The Effects of Price Regulation on Pharmaceutical Expenditure and Availability*

Mika Kortelainen Jaakko Markkanen Markku Siikanen Otto Toivanen

February 13, 2023

Abstract

Quasi-experimental evidence on the effectiveness of price regulation policies in curbing pharmaceutical expenditure in markets with generic competition is scarce. We analyze widely utilized generic substitution and reference price policies using data from the Nordic countries. Constructing treatment and control groups by matching data across countries by active ingredients and employing modern difference-in-differences methods, we find that expenditure decreases by 40% moving from the laxest to the strictest regime. Prices decrease by less than expenditure: patient incentives to choose a cheaper product probably explain the difference. We find no adverse effects on pharmaceutical availability and small to non-existent quantity effects.

Keywords: *pharmaceutical expenditure, pharmaceutical pricing, generic competition, reference pricing, regulation*
JEL-Classification: *I11, I18, H51, L51, L65, C23*

*Mika Kortelainen: University of Turku, InFLAMES, VATT and Helsinki GSE. Email: mika.kortelainen@utu.fi. Jaakko Markkanen: Aalto U. and Helsinki GSE. Email: jaakko.m.markkanen@aalto.fi. Markku Siikanen VATT and Helsinki GSE. Email: markku.siikanen@vatt.fi. Otto Toivanen: Aalto U., Helsinki GSE and CEPR. Email: otto.toivanen@aalto.fi. We thank Pierre Dubois, Ben Handel, Nano Barahona, Tobias Salz and Thomas Rønde as well as seminar audiences at CBS, EARIE, Helsinki GSE, NHESG and TSE for comments and suggestions. We also thank Hertta Hankimaa and Hung Le for excellent research assistance, the Finnish Social Insurance Institution (Kela) and the Yrjö Jahnsson Foundation for financial support. The statements, findings, conclusions, views and opinions contained and expressed in this report are based in part on data obtained under license from the following IQVIA AG information services: IQVIA Pricing Insights and IQVIA MIDAS Quarterly Sales, 2007-2013. All Rights Reserved. The statements, findings, conclusions, views, and opinions contained and expressed herein are not necessarily those of IQVIA Inc. or any of its affiliated or subsidiary entities. All errors are ours.

1 Introduction

"*A vial of insulin costs drug companies $10 to make, but families pay an average of $683 a month for it. It's ridiculous.*" In April 2022, President Biden pointed out the high cost of diabetes treatment in the United States and urged Congress to implement new price cap regulations: "*That's why I'm fighting to cap the cost of this lifesaving drug at $35 a month.*"[1]. The Inflation Reduction Act implemented this change for Medicare beneficiaries in 2022 (Cutler 2022).

Global spending on medicine has doubled during the last 10 years, reaching 1.3 trillion dollars. As shown in Figure 1, spending on medicines has increased everywhere, with the US in the lead. It is therefore not surprising that most OECD countries have adopted various cost-containment policies. An important part of such policies are those targeting markets with generic competition.[2] Such policies are often needed to counter the low price-sensitivity of consumers which is due to generous public insurance and low consumer co-payments.[3] Consequently, various price regulation policies for prescription drugs are common, especially in Europe but as is clear from above, such policies have recently been promoted and adopted also in the US.[4] Despite the wide-spread utilization of price regulation policies, credible causal evidence on whether such policies have decreased pharmaceutical expenditure and how much is scant. The objective of this paper is to provide such evidence.

We investigate the effects of different Reference Pricing (RP) and Generic Substitution (GS) policies on pharmaceutical expenditure, availability, prices, and quantities in markets with generic competition using data from four Nordic countries. Generic markets are an important part of the overall pharmaceutical (pharmacy) market—16% of markets with close to one third of sales in our data and even much higher in the US.[5] We have chosen to study the Nordic countries for two primary reasons: First, they are examples of societies that provide generous public insurance against pharmaceutical expenditure. They have

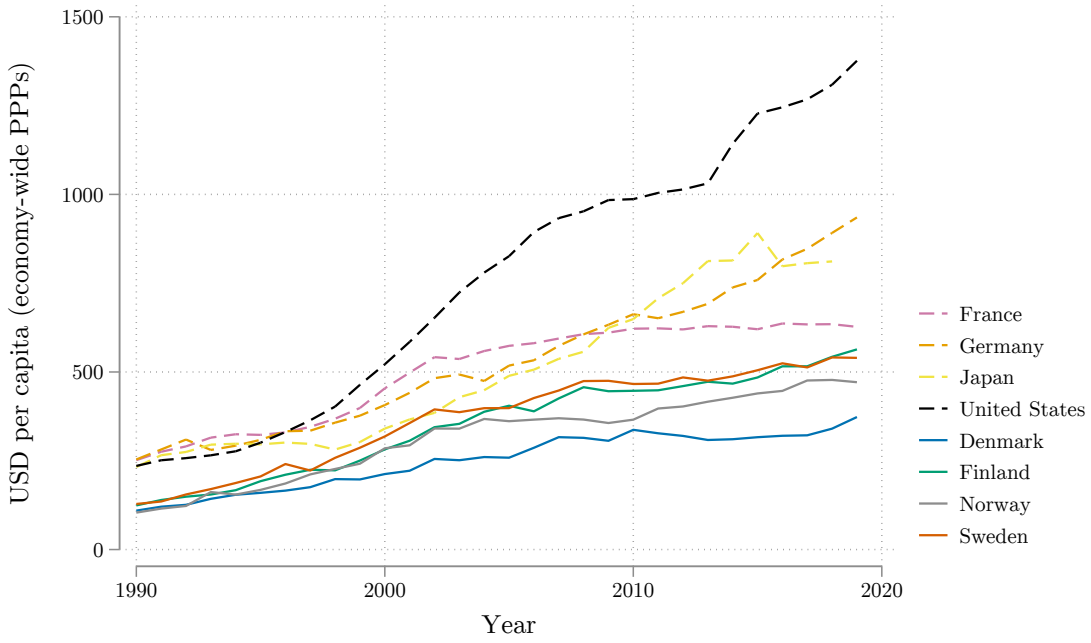
1. The statement was published on Twitter.

2. These are markets where the patent protection of the original (branded) drug has already lapsed, thus enabling competitor entry with products based on the same chemical molecule.

3. Note that besides taxpayer funded public programs, the same logic may well apply within privately provided health insurance, such as in the US.

4. See e.g., the Trump administration's Executive Order 13948 of September 13, 2020 on external reference pricing and the Biden administration's Executive Order 14036 of July 9, 2021 on promoting generic and biosimilar competition.

5. The rest are monopoly markets with patent protection, some of which may face competition through parallel imports.



OECD (2022), Pharmaceutical spending (indicator). doi: 10.1787/998febf6-en
(Accessed on 28 February 2022). Solid lines represent countires whose reforms are investigated.

Figure 1: Pharmaceutical expenditure in OECD countries

adopted several variants of Internal Reference Pricing (IRP) and External Reference Pricing (ERP) policies combined with GS during the 2000s, moving toward stricter regimes with greater financial incentives for patients. The main objective of these policies is to reduce pharmaceutical expenditure through (generic) competition and to increase consumer price sensitivity within generous reimbursement systems. Second, these countries are as homogeneous as groups of countries come, making the use of them as controls for each other appealing.

We depart from the existing literature in a number of ways. First, we use market-level data, thereby addressing the problem plaguing much of the literature that arises from the Stable Unit Treatment Value Assumption (SUTVA) in evaluating the causal effects of price regulation. When studying the prices of individual products—the approach of the vast majority of the existing literature[6]—this assumption translates into the untenable

6. The only exception we know of is one of the analyses in Brekke, Holmas, and Straume (2011).

assertion that the price of a product is not affected by the prices of its substitutes within the country. This assumption of no equilibrium effects runs against both theoretical models of competition and empirical evidence. Our key dependent variable is the market-level average expenditure per purchased dose (henceforth average expenditure), which encapsulates such equilibrium effects. Because our estimates capture the change in expenditure per dose, we directly analyze whether a reform has the effect intended by the regulator. Second, a central worry related to price regulation is its impact on product availability: unlike any of the existing papers on the effects of regulation on pharmaceutical prices, we also study the impact of regulation on product availability. Third, we design our control group by employing the same markets (active ingredients) in a neighboring country of similar appearance.[7] This is important in terms of both the quantity of data and the quality of the match between the treatment and control groups. The existing quasi-experimental literature has almost exclusively had to rely on different active ingredients within the same country in constructing the control group, or on products in the same group that entered the new regulatory regime at a different time.[8] Different active ingredients often have different price trends due to differences in regulation and competition or can be indirectly affected by the regulatory change ("spillover effects" or therapeutic competition), violating standard identification assumptions. In addition, our data allows us to study several regulatory reforms within a common framework, whereas the existing literature is concentrated on studying one reform at a time.

We find that price regulations decrease expenditure without affecting availability of products, with some policies being more effective than others: The move in Finland in 2003 from Voluntary Generic Substitution (VGS), the most producer-friendly regime in our data, to GS had no significant impact on the expenditure per dose sold. The Finnish move in 2009 from GS to IRP reduced expenditure by 13%.[9] Denmark reduced expenditure by some 5% moving from IRP very similar to that adopted in Finland in 2009 to ERP in 2000, and lost this gain when moving back to IRP 5 years later. Sweden, on the other hand, moved in 2009 from an IRP system that was stricter than that adopted by Finland the same year to a Product of the Month Auction (Auction-IRP) regime where each month essentially only the cheapest product in a given category is reimbursed. This

7. Our market definition is an active ingredient because reforms studied in this paper influence consumer choices within active ingredients.

8. Dubois and Lasio (2018) is a notable exception.

9. The reported magnitudes are based on our ATT-estimates displayed in Table 4. Naturally, one must keep in mind their confidence intervals.

change reduced expenditure by 27%. Similarly, Norway was able to reduce expenditure by 21% in 2005 when moving from GS to a system with government-dictated price cuts after the introduction of generic competition. Going from the laxest price regime (Finnish VGS) in our data to the strictest (Swedish Auction-IRP) we find an expenditure decrease of $(1 - 0.87 \times 0.73) \times 100\% = 36\%$ without change in availability or quantity, and an even larger one ($(1 - 0.75 \times 0.63) \times 100\% = 53\%$) when markets are weighted by their size.

We also study average posted prices per dose, which has been the key dependent variable in most of the existing literature. We observe that in half of our studied reforms the treatment effect estimates are systematically smaller in absolute value for prices than for expenditure, and in the other half of the reforms, this pattern is reversed. This shows that focusing on posted prices does not yield a good picture of the performance of the reforms. Based on our point estimates, the Finnish 2009 reform leading to a 13% reduction in average expenditure per dose lowers the average posted prices by only 5%; The Norwegian reform from GS-IRP to Step-Price (SP) that lead to an 21% reduction in average expenditure per dose had only 10% effect on average posted prices; and the Swedish Auction-IRP reform lowered average posted prices only by 4%, i.e., less than a sixth the effect it had on average expenditure per dose. When we compare our market-level average price analyses to package-level analyses found in most previous studies, we find that possible violations of SUTVA in the latter lead to minor differences in the results. Our shift in emphasis from posted (average) prices to expenditure turns thus to be more important than our change in the unit of analysis, at least for the data at hand.

The Swedish 2009 Auction-IRP reform demonstrates the likely mechanism behind the difference between the effects on average posted prices and on average expenditure. In Auction-IRP, the vast majority of consumers need to buy the cheapest available product to be reimbursed, but some consumers may be prescribed a more expensive product. By definition, all other products are priced higher than the product of the month, and some firms, the producer of the branded original drug in particular, may have an incentive to price their product very high to cream-skim locked-in customers. The average posted price in a market may thus remain relatively high, while at the same time the lowest price—the price of the product that most patients are dispensed and which therefore dominates expenditure—can be very low, thereby resulting in a large expenditure decrease.

Our paper contributes to three strands of literature. We contribute foremost to the literature on the effects of pharmaceutical price regulation on expenditure. The existing literature has mostly shied away from studying the effect of price regulation on expenditure,

the likely explanation being restrictions in data and research designs. When the question has been addressed, the existing literature has either used package or product level data on (changes in) average posted prices as a proxy for (changes in) pharmaceutical expenditure (e.g. Danzon and Chao 2000; Brekke, Holmas, and Straume 2011) or used a structural approach to evaluate the impact of price regulation on competition, welfare, and savings in public expenditure (Dubois and Lasio 2018; Maini and Pammolli 2022; Dubois, Gandhi, and Vasserman 2022). Our results cast doubt on the credibility of the first approach and provides a complementary approach to structural modeling, which requires carefully modeling the sometimes intricate details of price regulation. The only quasi-experimental analysis of the effect of a regulatory reform on pharmaceutical expenditure that we know of is Brekke, Holmas, and Straume (2011) who study the same regulatory change as Brekke, Grasdal, and Holmås (2009). Using within-country data on 24 Anatomical Therapeutic Chemical classification system (ATC)5 groups, 8 of which are treated, they find that the introduction of RP reduced expenditure in Norway by 30%.[10]

Dubois and Lasio (2018) is an important precursor of our study in that they also use multi-country data, although the difference-in-difference analysis with which they complement their structural modeling is carried out using package-level (= product-level) data, not market-level data on active ingredients. Our data and setting allow us to use a more direct approach to evaluate the causal effects of price regulation reforms on pharmaceutical expenditure. Although the structural approach of Dubois and Lasio (2018) has the potential to allow for an evaluation of the welfare effects of regulation, they are careful to point out that the complicated process of choosing a particular drug, involving the physician, the patient, and the pharmacist, makes the interpretation of traditional welfare measures difficult. The central feature of GS is that the regulator views such products as perfect substitutes from a medical point of view and therefore sees it justified to incentivize people to substitute to cheaper products.[11] In line with the regulators' approach, our main objective is to analyze whether the implemented regulations have led to a decrease in the amount of expenditure per dose. In this paper we shy away from the structural modeling because of the scale of the challenge: Depending on level of detail one would adopt in the modeling, our data contain 7–11 different regulatory regimes that we would have to model.

10. The emphasis in Brekke, Holmas, and Straume (2011) is in formulating a theory of how branded drugs can be priced differently from generic ones and testing the predictions of the model using product-level data.

11. However, there is plenty of evidence that patients prefer branded products to generic products; see e.g. Dubois and Lasio (2018).

Steering patients to choose generic and less expensive drugs matter for pharmaceutical expenditure also in markets where private insurance providers play an important role. A number of studies have investigated the effects of Medicare Part D and its incentive structures on drug prices and pharmaceutical expenditure. For example, Duggan and Scott Morton (2010) demonstrate that private insurers have been able to decrease prices for previously uninsured with incentive-based formularies, which encourage patients to choose generic and lower priced drugs.[12] Einav, Finkelstein, and Polyakova (2018) show complementary evidence that private insurance plans in Part D systematically set higher out-of-pocket prices (coinsurance rates) for drugs or classes associated with more elastic demand. Our results are in line with these observations: We find that successful regulatory tools in our setting combine direct interventions in the form of reduced maximum prices with increased patient incentives.

Second, to complement the analysis of expenditure and prices, we also evaluate the effects of price regulation policies on the availability of pharmaceuticals. A common concern and source of criticism of pharmaceutical price regulation is its possible adverse effect on pharmaceutical availability and innovation (see, e.g., Lakdawalla 2018). The literature on pharmaceutical shortages has documented that consolidation, fierce price competition, reimbursements, and low prices can increase pharmaceutical shortages (Yurukoglu, Liebman, and Ridley 2017; Stomberg 2016; Lee, Lee, Shin, and Krishnan 2021). However, we are not aware of any papers that study the effect of price regulation policies on pharmaceutical availability. Note that while the effect of pharmaceutical price regulation on innovation is an obvious concern in general (Acemoglu and Linn 2004; Yin 2008; Ornaghi 2009; Dubois, Mouzon, Scott-Morton, and Seabright 2015), the regulations we study are unlikely to have a first-order impact on pharmaceutical innovation, as they concern mainly markets with off-patent drugs.[13]

12. Duggan and Scott Morton (2011) find that there are also price reductions in the medium term.

13. To illustrate, consider the following back-of-the-envelope calculation: Let us assume a discount rate of 0.95 and that the inventor firm can enjoy patent protection for 10 or alternatively 15 years. These period lengths are motivated by how the patent system works and how long it takes to launch a pharmaceutical product after filing for a patent. Patent protection is usually 20 years from filing data of the patent application, but pharmaceutical patents are often granted a 5-year extension. It is well known that the time to market from patent filing can be long for pharmaceuticals, e.g. Lexchin (2021) reports an average time to market in Canada of 11.8 years. Keeping the annual profits constant, the Net Present Value (NPV) of the profits in year 11 is 5.95% and in year 15 0.21% of the NPV of the profits in the first year. Even the NPV of the sum of profits from year 11 to year 50 (by which time one might expect a superior substitute to have arrived, rendering profits zero) are modest at 12.93% of the NPV of first year profits under patent protection, and those from year 15 to year 50 even more so at 0.37%.

The largest literature we contribute to is on the effects of price regulation on pharmaceutical pricing. Several articles have used single-country data and quasi-experimental variation in price regulation to evaluate the effect of an individual reform (e.g. Pavcnik 2002; Brekke, Grasdal, and Holmås 2009; Brekke, Holmas, and Straume 2011; Herr and Suppliet 2017) on average posted prices. Perhaps the first such paper was Pavcnik (2002), who studied the change in German reimbursement rules. She uses the differential timing of the reform for oral antidiabetics.[14] Pavcnik (2002) finds that the posted prices decrease 10–36%. In contrast to the regulations we study, the German reference price was set by the regulator and was directly based neither on domestic nor foreign actual prices. Brekke, Grasdal, and Holmås (2009) study the Norwegian 2003 change from an ERP regime where the maximum price is the average of prices in a set of countries to an IRP system where the regulator sets a maximum reimbursement price as a function of past domestic prices. At the same time, pharmacies received strong monetary incentives for substitution. Their treatment group consists of six active ingredients, chosen because they cover a wide range of diseases and have a high sales volume. Using products not subject to the regime change as a control group, Brekke, Grasdal, and Holmås (2009) find that the prices posted for branded products decreased by 18–19% and those for generics by 7–8%. In contrast to these important papers, we use market-level average prices of active ingredients instead of package-level prices to encapsulate equilibrium effects of price regulation reforms in four countries and examine their relative effectiveness using, when appropriate, modern difference-in-difference and event-study methods (e.g. Callaway and Sant'Anna 2021; de Chaisemartin and D'Haultfœuille 2020).

The rest of the paper is structured as follows. In Section 2 we present the relevant institutions and regulatory regimes and motivate our choice of control countries. We also discuss the minor reforms and other institutional changes that take place during our observation periods. We introduce the data and our matching procedure in Section 3 and present our difference-in-difference approach in Section 4. We also discuss the timing of reforms and the choice of estimation periods in that Section. Section 5 is devoted to the presentation and discussion of the results. We present our main results using event study graphs, and a summary of the main and auxiliary results based on average treatment effects. We discuss most of our robustness analyses along the way, but conclude Section 5 with a discussion of our analysis of whether the reforms we study had an impact on markets

14. Pavcnik (2002) also performs a before-after analysis of antiulcerants due to lack of a suitable control group.

that were not directly affected. This analysis is of particular importance as it enables us to investigate our definition of individual markets. We offer our conclusions in Section 6.

2 Institutions and Regulatory Regimes

All Nordic countries[15] have a universal single-payer insurance system, also called the Beveridge model, in which all citizens receive insurance coverage through the state (Bhattacharya, Hyde, and Tu 2013). The system is financed by taxes and enrollment into the system is automatic and free. The government operates most hospitals and clinics and decides their locations. Publicly provided care is offered at very low or non-existent prices and patients do not face deductibles or premiums when using public services. There are some exceptions to this rule, prescription drugs being a notable one.

There are two distinct approaches to public reimbursement of pharmaceuticals in the Nordic countries: a needs-based and a product-specific calculation. In the needs-based system, used in Sweden and Denmark,[16] the level of reimbursement and the consumer's co-payment are tied and capped to the consumer's annual pharmaceutical spending. The share of reimbursement (co-payment) increases (decreases) as the consumer spends more on reimbursed pharmaceuticals. After crossing a legal threshold, the consumer is fully reimbursed. In addition, the state typically grants full reimbursement for certain drugs and vulnerable groups. In the product-based reimbursement system, used in Finland and Norway, public reimbursement varies product by product. The level of reimbursement (usually 40% to 100%) is based on the severity of the disease; however, annual consumer spending is capped as in the needs-based system. The crucial difference is that in the needs-based system, conditional on the price negotiations with the manufacturer, the government only decides whether a product receives reimbursement or not. In the product-specific reimbursement system, the government also decides on the level of reimbursement product by product.

We next define the different regulatory regimes found in our data and then describe

15. The Nordic countries consist of Denmark, Finland, Iceland, Norway and Sweden. Appendix A.1 displays the map of Nordic countries and gives some relevant descriptive statistics.

16. Denmark changed the calculation of the reimbursements in March 2000 from a product specific reimbursement system to a needs-based system through LOV nr 1045 af 23/12/1998. This created a cut-off point for a so-called "end of the year effect" where consumers increase their spending before their level of reimbursement decreases. Our results for market-level estimations confirm some findings in Simonsen, Skipper, Skipper, and Christensen (2021).

the regimes in place in different countries at different points in time, as well as the reforms that we analyze.

2.1 Regulatory Regimes

In official use, different regulatory regimes can share the same name in different countries. We use the following definitions and acronyms:

Definition 2.1. Voluntary Generic Substitution (VGS). Substitution to a cheaper interchangeable product is possible, but requires the active decision of the prescribing physician.

Definition 2.2. Generic Substitution (GS). Substitution to a cheaper interchangeable product must be offered to the consumer in the pharmacy. The medicines authority determines which products are substitutable.

Definition 2.3. Reference Pricing (RP). The consumer has to pay out of pocket the price difference between the price of the prescribed product and the price of the reference product if she declines generic substitution.

Definition 2.4. External Reference Pricing (ERP). The reference price is determined as a function of prices in both foreign and domestic markets (exogenous reference price).

Definition 2.5. Internal Reference Pricing (IRP). The reference price is determined as a function of domestic prices only (endogenous reference price).

Definition 2.6. Step-Price (SP). A reference price system after generic entry in which the government enforces gradual and predetermined price decreases to the maximum reimbursed price.

Definition 2.7. Product of the Month Auction (Auction-IRP). An internal reference price system where reimbursement is only granted for the prescribed product and the winner of the monthly auction. The lowest bid in the auction determines the reference price.

Regulatory policies often consist of a combination of GS and some form of RP, but sometimes only one or the other is used. For example, in the early 2000s, Finland and Norway adopted GS systems without RP. On the other hand, the Swedish GS system has

always been coupled with RP. Even when not explicitly stated (as in SP), these regulations are complemented by a maximum price regulation of varying degrees of severity in all other Nordic countries but Denmark.

Some regulations come from European Union (EU) law, which sets the principles for market entry and the freedom of movement of goods in the single market. Parallel trade, that is, imports of pharmaceuticals (irrespective of patent status) from a low-priced Member State to high-priced Member States, is protected.[17] Other types of pharmaceutical regulation, such as public reimbursement, price regulation and regulation of the distribution of pharmaceuticals are left to individual Member States. However, EU can place some restrictions on national regulators.[18]

2.2 Summary of Reforms

We now summarize the relevant regulatory regimes in place in the four countries during our observation period and explain which regime changes we study. We provide detailed information on each of the regimes in Appendix A.3.6.

Figure 2 shows the regimes that are in place and the reforms (= changes in regime) that take place during our observation period, organized by country and chronologically. We exclude two reforms from our analysis: The Norwegian 2001 reform that combines pharmacy market liberalization and GS reform is excluded because we cannot separately identify their effects on the outcomes. The Norwegian 2003 reform introducing the so-called Index Pricing[19] is excluded because 1) it directly influenced only eight markets (active ingredients), and thus we cannot perform market-level analyses, and 2) given the timing of reforms in the other countries, we cannot form a good control group. We analyze the 2005 Norwegian reform using data on pharmaceuticals not included in the Index Price regulation.

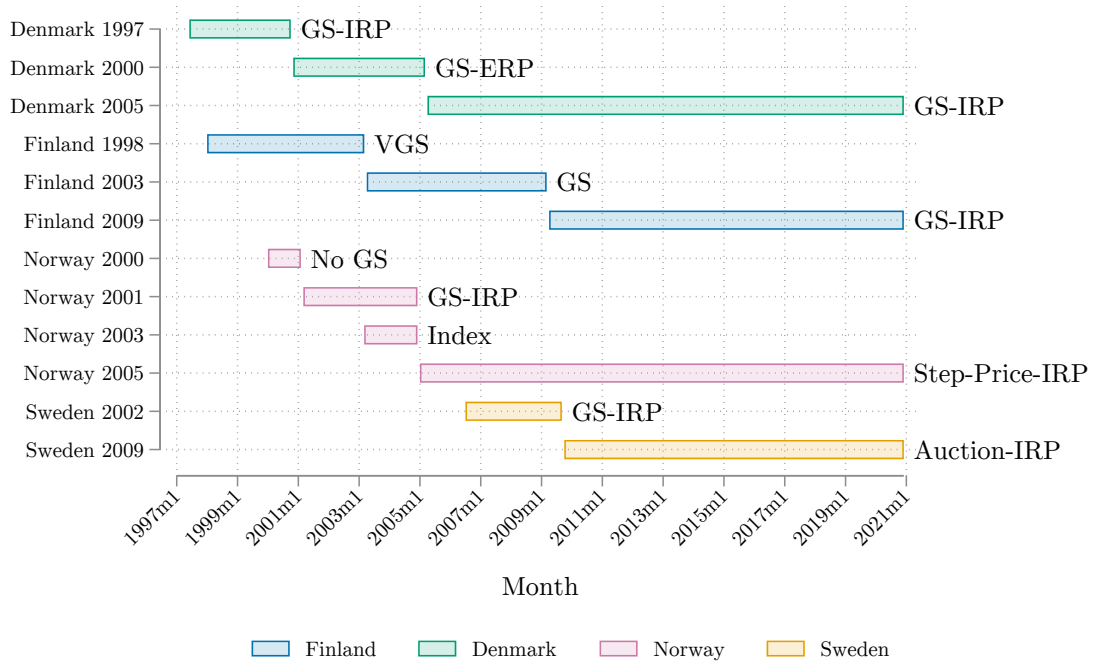
A more illuminating way to group the reforms is through the increasing strictness of the price regulation regimes, shown in Figure 3. The laxest regime in our data is the VGS regime of Finland, followed by the GS regime in the same country. Neither put much

17. The the Treaty on European Union, Articles 34 and 36, provide the legal basis. For reference, see the precedent of the Court of Justice of the European Union in Pfizer Inc. v Eurim-Pharm GmbH. (1981).

18. An example is the maximum processing time for reimbursement decisions: 180 days for new pricing and reimbursement decisions, 90 days for review of an application to increase prices. See Directive 89/105/EEC.

19. The Index Price system was an IRP system where the reference price was calculated as a sales-weighted average of producer prices by each reference price group; for a review of the Index Price system, see Brekke, Grasdal, and Holmås (2009) and Brekke, Holmas, and Straume (2011).

Figure 2: Timeline of reforms



Timeline of reforms by country and data availability.

pressure on firms to lower their prices nor gave consumers incentives to substitute toward cheaper products. GS coupled with internal or external reference pricing, as adopted by all four countries at some point, constitutes a clear tightening of the regulatory regime by giving consumers incentives to substitute toward cheaper products. Finally, while it is not clear whether the Norwegian SP is stricter than the Swedish Auction-IRP, they were both further steps toward ever tighter price regulation and, in the case of Auction-IRP, greater induced price sensitivity of consumers.

We find it useful to organize the reforms into three categories. The first category consists of the Finnish 2003 and 2009 GS and IRP reforms, i.e., moves toward IRP. These reforms demonstrate how well a simple substitution reform without financial incentives can reduce average expenditure and what happens when financial incentives in the form of RP are introduced to consumers. The second category of reforms consists of different ways of implementing RP: the Danish 2000 ERP and 2005 IRP reforms, which allow us to analyze

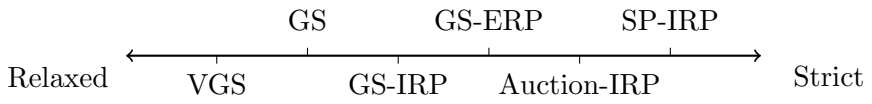


Figure 3: Reform strictness scale

Table 1: Treatment and Control Countries

Year	Treatment Country	Reform	Control Country	Control Regime
2003	Finland	VGS → GS	Denmark	ERP
2009	Finland	GS → IRP	Norway	SP
2000	Denmark	IRP → ERP	Finland	VGS
2005	Denmark	ERP → IRP	Finland	GS
2005	Norway	GS → SP	Finland	GS
2009	Sweden	GS → Auction-IRP	Norway	SP

Notes: IRP = Internal reference pricing, ERP = External reference pricing, VGS = Voluntary Generic substitution, GS = Generic substitution, SP = Step-price, Auction-IRP = Product of the Month Auction.

the benefits of using ERP as opposed to IRP. The third category of reforms consists of moves from the "regular" IRP to stricter versions of it. The Norwegian 2005 and Swedish 2009 reforms channel demand either to the cheapest product (Sweden 2009) or to products which are priced using predetermined pricing rules set by the regulator (Norway 2005).

We have collected the reforms analyzed into Table 1, ordered in increasing strictness of the regimes. The table shows the treatment country in question; the reform, i.e., the regulatory regimes before and after the reform; the country the markets of which act as the control group; and the regulatory regime in the control country during the estimation period. We next discuss the choice of the control country.

2.3 Choice of Control Countries

An important decision is the choice of a control country. We sought to identify a country where no major regulatory changes occur in the years right before and right after a given reform. Figure 2 reveals that one or two countries are available as control countries for the reforms we study.

We use Denmark as the control country when we study the Finnish 2003 reform, moving from VGS to GS. The Danish regime at that time was ERP. The second reform is the

Finnish change from GS to IRP in 2009. We use Norway as a control country; it was using SP at the time. Denmark is available as an alternative control country: those results, reported in Appendix A.9, are in line with main results reported below.[20]

We then move on to analyze the Danish 2000 switch from GS-IRP to GS-ERP using Finland as the control country. The Finnish regime at that time was VGS. We continue to use Finland as the control country when we study the Danish reversal from ERP to IRP in 2005. Using Denmark and Finland as control countries for each other is unlikely to affect our results for a number of reasons. First, the overlap between the different analyses in the time dimension is minor. Second, as we demonstrate below, the effects of the reforms stabilize rather quickly. Third, in our analysis of pre-trends, we do not find worrying signs. Furthermore, our different estimation samples consist neither of exactly the same markets (because the number of markets with generic competition increases over time due to patent expirations) nor of exactly the same products (due to generic entry): the overlap in products is usually $< 20\%$ and always $< 30\%$ (see Table 12 in Appendix A.4).

The fifth reform is the Norwegian change from GS-IRP to SP in 2005. Figure 2 reveals that the country with a stable regulatory regime is Finland, where GS was in place at that time. We discard 8 treated Norwegian markets due to the Index Price regulation implemented in 2003 in Norway because otherwise the pre-period market institutions would not be the same for all Norwegian markets.[21]

The only Swedish reform that we study is the implementation of the product of the month system in 2009. Sweden transitioned from GS with IRP to a first-price sealed bid procurement auction.[22] Similar to the Finnish 2009 reform, we use Norway as the control country in the main analysis and perform a robustness test substituting Denmark for Norway, and using both Denmark and Norway in Appendix A.9. Just like in the case of the Finnish 2009 reform, the results using different control countries are in line with each other.

20. We also report results with both Norway and Denmark as the control group in Appendix A.9.

21. We do not analyze the Index Price reform because the small number of treated markets (8) does not leave room to study market level outcomes.

22. The actual system is more complicated and includes tightening maximum reimbursement prices. See Appendix A.3.6.

2.4 Minor Reforms and Institutional Changes

Most of our analyses are affected by minor reforms or institutional changes that occur before or after the studied reform. These minor reforms can happen in the treated country or in the control country. Finland's 2003 (GS) reform is the only reform that we study that does not have any small regulatory changes occurring during the study period. In Appendix Section A.3.6 we detail all minor reforms we have identified. Here we list the minor reforms that mechanically influence our results and explain how we take these into account.

Finland implemented a 5% price cut to the maximum wholesale prices (price caps) of all reimbursed pharmaceuticals in January 1, 2006.[23] The effect of the mechanical wholesale price cut is clearly visible in our event studies for the Denmark 2005 (ERP to IRP) reform and Norwegian 2005 (GS-IRP to SP) reform in 2005, especially when price outcomes are studied. In both reforms, Finland is used as a control group. We mitigate the issues arising from the price cuts by limiting the post-periods in the main result event studies to time periods before the price cut and report results using a longer post-treatment period in Appendix Section A.10.

The second set of minor reforms are the maximum price cuts that happen before and after the Swedish 2009 Auction-IRP reform. A few months before the implementation of the reform, Sweden cut the maximum prices of the originator product by 65% from the price level that prevailed 12 months before patent expiration. [24] This price cut occurs close to the start of the Auction-IRP regime and the price cut potentially influences all markets in our sample: we identify the joint effect of the price cut and the regulatory change.[25] In 2011 Sweden changed the maximum wholesale price regulation rules, and this led to a 35% reduction in the price cap. The latter Swedish price cap change is less problematic than the first cut because the Auction-IRP reform directs market demand toward the cheapest product which rarely is priced below the price cap.

The third minor reform that mechanically influences our results is the change in the Danish reimbursement system in 2000. This policy change gave consumers an incentive to stockpile products before rules regarding annual reimbursement expense calculation were changed. This change in the reimbursement system happened 9 months before the

23. Declared in the commencement order of 885/2005.

24. This price cut was initially proposed by the Swedish pharmaceutical industry.

25. We study markets with generic competition and the price cut was imposed to markets where the patent has expired.

Danish 2001 (GS-IRP to GS-ERP) regime change. The results of an auxiliary event study presented in Appendix A.12 show that the demand shock arising from the anticipation of the reform is transitory and pricing is not influenced by the change.

3 Data and Matching

3.1 Sales and Reform Data

We use data from four different data providers on monthly revenues and quantities of drugs purchased by community pharmacies. Our data set covers the Nordic countries, excluding Iceland. The data sets contain information on the sales value and volume of each pharmaceutical package sold in the respective country. Sales values are defined in pharmacy purchase prices and volumes in Defined Daily Dosages (DDD) for each respective active ingredient according to the ATC.[26] To capture potential equilibrium (SUTVA) effects or to allow for them, we deviate from the literature and aggregate our sales and quantity data to the active ingredient (-month) level. The country-specific data sets also provide rich information on product characteristics. We supplement our sales data with rich regulatory information obtained from market regulators and directly from the relevant legislation. Table 11 in Appendix A.4 shows the data source for each country and reform.

We use wholesale prices. The wholesale price is the price a pharmacy pays for the product when the product is purchased from the wholesaler.[27] There are two reasons to use wholesale prices: First, the regulations target wholesale prices. Second, with one exception, the retail price in each country is determined using a mechanical formula based on the wholesale price. The only exception is the Norwegian 2005 SP regime, where only an upper bound on the retail price is based on the wholesale price. We show how price formulas work in the Nordic pharmaceutical market in Appendix Table 8.

Our main outcomes are (logarithms of) average expenditure at the market level $\left(\frac{\sum P \times Q_{Package}}{\sum Q_{Doses}}\right)$ (where P stands for the price of a package, $Q_{Package}$ for the number of packages of a given type sold and Q_{Doses} for the size of the daily dose) and availability, which we measure both by the number of product names and (in a robustness analysis)

26. The ATC system classifies active ingredients according to their therapeutic, pharmacological, and chemical properties. The classification groups active substances into five different categories. Active substances in the fifth category have the same active ingredients and are considered equivalent for the treatment of the same disease.

27. The monthly wholesale price is calculated by dividing the monthly sales of a given product by the monthly number of packages sold.

by the number of packages. The first one measures the expenditure per sold dose, i.e., the cost of treatment; the second measures the number of available products.

Our secondary outcome variables are the main outcome of the previous literature, the average wholesale price $\left(E\left[\frac{P}{Q_{Doses}}\right]\right)$, and the total quantity ($\sum Q_{Doses}$), in contrast to most of the literature, we measure them at the market level. The former measures the average price of a package on the pharmacy shelf but does not take into account which products are actually bought. The latter allows us to analyze whether the reforms affect the amount of pharmaceuticals consumed.

Prices and sales are measured in nominal national currencies.[28] Nominal values are used because in all Nordic countries price regulations work with nominal prices. As the sample periods are relatively short (2–4 years) and from an era of low inflation, differential inflation trajectories should not cause bias.

3.2 Sample Matching

Our empirical strategy is based on comparing the pharmaceutical retail markets of a country subject to a reform (treatment country) with identical retail markets in another Nordic country (control country) before and after each reform.[29] We match the markets by active ingredient (i.e., ATC5 level).[30] The matching process proceeds in four steps: i) we discard non-prescription pharmaceutical products (over-the-counter pharmaceuticals (OTC) products) and the hospital market for pharmaceuticals; ii) we identify the markets that are affected by the reform in question in the treated country; iii) we find the same markets in the control country; and iv) we drop non-treated markets, treated markets without a match, and matched markets where generic competition starts during the pre-period. Our estimation samples thus include different products and package sizes in the treatment and control markets.

Our matching process leads to the exclusion of some treated markets. A treated market

28. Sales data from Finland is in euros, because the switch from FIM to EUR occurs during our sample (2002). The reason for not converting prices to the same currency is the possibility of exchange rate shocks. Exchange rate shocks are problematic when data from Norway or Sweden are used, because these currencies are not linked to the euro like the Danish krone. A visual inspection of the data showed that this is a real concern. We demonstrate in Appendix A.5 how exchange rates evolve within our sample periods.

29. Pharmaceuticals used in inpatient care (hospital market) are excluded from the analysis, because competitive bidding is used in these markets and our data-set does not contain hospital prices.

30. Appendix A.4.1 Table 12 illustrates package level match rates between treatment and control countries before each studied reform. At the package level our match rates are less than 30% and this suggests that matching based on active ingredients is better modeling choice.

is excluded from the sample if the control country does not have the corresponding market, or if generic competition begins during the reform pre-period. A control country may not have the same markets as the treated country due to differences between countries in the markets served through the pharmacy sector, how OTC drugs are classified, or because some small markets may have experienced entry in one country but not yet in the other. The matching process also discards treated markets where generic competition has started during the reform pre-period, as otherwise we could confound changes caused by the reform with changes caused by increased competition through patent expiration and generic entry.

Table 2 illuminates how the estimation samples cover the pharmaceutical market, excluding the hospital sector. In Panel A, we describe how our matching process progresses from the number of existing markets to the number of markets included in each estimation sample. Panel B shows the same information in terms of share of sales.

The first Column in Panel A gives the number of markets in the treated country in the pre-period while the second Column shows the number of markets with generic competition. For example (see Row 1), there were 881 (ATC5) markets in Finland in 2003 during the reform pre-period, 120 of which had generic competition. Out of these, 100 (Column 3) are affected by the change from VGS to GS. Column 4 reveals that 90 of these 100 markets have experienced generic entry before the pre-period and are therefore included in the estimation sample. Thus, we end up with 82 matched markets (Column 5) after having discarded markets due to pre-period entry and due to the markets not existing in Denmark.

The difference between Columns 3 and 4 in Table 2 is informative about the exogeneity of reform timing with respect to markets becoming competitive. If regulators design new regulation policies while taking into account how patents expire, we might see that many markets become competitive during the reform pre-period (the difference between Columns 3 and 4 in Panel A) or the market size of markets with generic entry during pre-period is substantial (the difference between Columns 3 and 4 in Panel B). Table 2 shows that all reforms have markets that become competitive during the pre-period, but the number and the economic significance of these markets is relatively low in comparison to all treated markets.

In Panel B, Column 2 reveals that 20%–35% of pharmaceutical sales in the treated countries come from markets with competition, the rest coming from monopoly markets. Columns 3 and 4 show the sales share of the treated markets and markets with generic entry before the pre-period. The sales share of the treated markets varies from a low of 12% (Norwegian GS to SP reform in 2005) to a high of 33% (Finnish VGS to GS reform

in 2003). The sales share of unmatched markets (the difference between Columns 4 and 5) is small, as expected.[31] The only exception is the Norwegian 2005 reform, where the large decrease can be explained by the fact that we need to discard 8 markets that had been exposed to Index Price regulation in 2003.

A feature of the markets that affects the coverage of our analysis is the share of pharmaceuticals distributed through hospitals rather than through pharmacies. The share of pharmacy sales is close to or above 80% in all other countries in our sample but Denmark where pharmacy sales are round 70% in the early 2000s, but decrease to somewhat less than 50% by 2012.[32]

4 Empirical Strategy

4.1 Research Design

The primary obstacle in identifying the effects of price regulation policies on product market outcomes based on single-country data is that regulations are either rather broad, covering almost all markets, or targeted, covering markets of special interest.[33] As a consequence, non-regulated products are typically not similar enough to regulated products within a given country to form a plausible control group. The most prominent example is that price regulation policies related to GS can be applied to markets with generic competition. Products that remain outside of regulation presumably come from markets without competition. This leads to comparisons in which the treatment and control group products are in different stages of their product life cycle and the products under comparison come from different drug markets. In difference-in-difference analyses major dissimilarities between treatment and control groups can threaten the main identification assumptions, such as parallel trends condition.

The second major challenge in evaluating the effects of (price) regulation in markets with differentiated products using the differences-in-differences approach is SUTVA, which rules out equilibrium effects. The existing quasi-experimental literature on pharmaceutical market price regulation reforms has not considered SUTVA seriously: the studied outcomes have mostly been some variants of package or trade-name prices and used control

31. We display the number of observations all estimation samples in Appendix Table 13.
32. See Appendix Figure 10.
33. The Norwegian 2003 Index Price regulation is an example of the latter, targeting 8 markets deemed to be especially important.

group from the same market (Brekke, Holmas, and Straume 2011; Brekke, Grasdal, and Holmås 2009; Herr and Suppliet 2017). This means that previous studies impose the often unrealistic assumption that prices of products within the same market are independent of each other. In practice, price regulation reforms can also indirectly affect prices and sales of unregulated products, creating biases in within-country comparisons.

To address these shortcomings, we base our empirical strategy on cross-country comparisons at the market level (ATC5) between two Nordic countries rather than within country comparisons using product-level data.[34] This approach allows comparisons between identical markets in different countries.

Also our identification strategy is based on assumptions. We assume that there are no major pricing spillovers between countries that would change due to a reform, and that the prices and sales of pharmaceuticals in a given ATC5 market are comparable between countries.[35] The motivaation for choosing to use data from the Nordic countries is that they are rather homogeneous, thus reducing unobserved heterogeneity between the treatment and control groups. In addition to the aforementioned assumptions, we need to maintain further assumptions on pre-trends that depend on the employed estimator.

4.2 Difference-in-Difference Estimators

We estimate difference-in-difference models. Our empirical approach allow us to include market-country-specific fixed effects to account for level differences between markets and time-fixed effects to account for unobserved aggregate time trends and shocks.

Our primary approach is to use the standard Two-Way Fixed Effects (TWFE) estimator.[36] When the parallel trend assumption behind the TWFE estimator fails or is less credible or when the reform in question is implemented in a staggered fashion, we use the estimator proposed by Callaway and Sant'Anna (2021) which imposes less strict parallel

34. Some ATC5 classes are potential substitutes. To verify the robustness of our results to this, we also study the spillover effects of regulations between different ATC5 as a robustness test in Section 5.4. We find no significant spillover effects.

35. Pricing spillovers are possible in the European pharmaceutical market, because many countries have incorporated the ERP system to their institutional setup. Furthermore, Nordic countries use the system and other Nordic countries as a benchmark. We argue that pricing spillovers are not a problem in our setting because we study markets where generic competition has started before our sample period, and through descriptive statistics we demonstrate that the number of products, which could in theory be influenced by ERP in the control country, is surprisingly small. We show this in Appendix A.4.1.

36. We use the estimator proposed by Correia (2016) to absorb the fixed effects at the market or product level.

trends assumptions and allows for staggered treatment adoption.

Our base estimation equation has the following form:

$$y_{i,t,c} = \alpha_{i,c} + \lambda_t + \sum_{\tau \neq -1} \beta_\tau \text{Reform}_\tau + \epsilon_{i,t,c} \tag{1}$$

where $y_{i,t,c}$ represents the monthly market level average expenditure, number of product names, average price or quantity i in country c at time t. The subscript i denotes a market in our analysis of average expenditure, number of product names and quantity, and a package in our price analysis. $\alpha_{i,c}$ denotes the country observation unit-specific fixed effect which controls the country-specific time-invariant factors that influence the outcome. Reform_τ are relative time-to-treatment indicators which are set to 1 for treated markets if period t is τ periods from the start of treatment. The coefficients of interest (β_τ) denote the average change between time τ and the last period before treatment in markets exposed to treatment, relative to control markets. When we estimate the average effect of the reform, we augment equation (1) by replacing $\sum_{\tau \neq -1} \beta_\tau \text{Reform}_\tau$ with $\beta_{att} \text{Reform}_\tau$. This allows us to interpret β_{att} as the average impact of the reform on the treated units.

The estimation equation for the Callaway and Sant'Anna (2021) estimator has the following form when the comparisons are based on never treated units (control country) and treated units (treatment country) without conditioning on control variables X:

$$ATT^{unc}_{nev}(g, t; \delta) = \mathbb{E}\left[Y_t - Y_{g-\delta-1} | G_g = 1\right] - \mathbb{E}\left[Y_t - Y_{g-\delta-1} | D_{t+\delta} = 0\right] \tag{2}$$

G_g denotes the time period when unit i becomes treated; $D_{t+\delta}$ is an indicator of whether i has been treated at time $t + \delta$; Y_t is the outcome in period t; and $Y_{g-\delta-1}$ is the outcome in period $g - \delta - 1$, where g denotes the period when i becomes treated and δ denotes the number of anticipation periods. The expression for $ATT^{unc}_{nev}(g, t)$ clearly resembles the Average Treatment Effect on the Treated (ATT) in the canonical case of two periods and two groups. We use aggregation formulas derived in Callaway and Sant'Anna (2021) when we estimate the effects averaged over time.

The average effect of participating in the treatment for units in group g is identified by taking the path of outcomes (that is, the change in outcomes between the most recent period before they were affected by the treatment and the current period) actually experienced by that group and adjusting it by the path of outcomes experienced by the control

group (Callaway and Sant'Anna 2021). Under the maintained parallel trends assumption, discussed in detail in De Chaisemartin and d'Haultfœuille (2023), this latter path is the path of outcomes that units in group g would have experienced if they had not participated in the treatment.

The main benefit of using the Callaway and Sant'Anna (2021) (CS) estimator instead of the TWFE approach is that it is robust to treatment effect heterogeneity related to the timing of treatment. This avoids issues related to negative weighting (Goodman-Bacon 2021). Furthermore, the required parallel trends assumption is weaker in comparison to TWFE, since the CS estimator imposes fewer restrictions on the evolution of potential outcomes during the pre-period (see Callaway and Sant'Anna (2021) for discussion).

We cluster standard errors at the ATC5 level using a wild bootstrap procedure.[37] This clustering scheme allows dependencies within each market and is preferred over a block bootstrap because the number of clusters varies between 15–126 depending on the examined treatment.

4.3 Timing of Reforms and Choice of Estimation Periods

An important part of estimating causal effects of reforms is the timing of pre- and post-treatment periods. Each reform has an actual start date, which is public information, but it is possible that due to anticipation, companies or consumers react to the reform before the reform is implemented (Alpert 2016). Failing to take that into account could bias the estimates. Difference-in-difference estimators allow anticipation to exist, but the start of the anticipation period must be known (Callaway and Sant'Anna 2021).

Our reform timing is in most cases based on the date when the national parliament in question accepted the law imposing the new price regulation. The benefit of using the law acceptance date compared to the actual introduction of the law is that it circumvents anticipation concerns, and this date comes from the legislative process. Some reforms were implemented without changes to the legislation (e.g. Sweden 2009); in these cases we rely on other sources to pin down the timing.

Table 3 shows the duration of each sample period and our timing choices in relation to the timing of the reform. Column 1 shows the sample period and Column 2 displays the length (end-start) of the sample period in months. In selecting sample periods, we need

37. For our TWFE estimates, we use the method developed in Roodman, Nielsen, MacKinnon, and Webb (2019) in the estimation of the confidence intervals. Our Callaway and Sant'Anna (2021) estimations use the Mammen (1993) method.

to limit overlap between consecutive reforms and also at the same time guarantee that the post-reform period is long enough. Reforms can overlap each other because between 2000 and 2005 4 reforms out of 6 start and for years 2000-2006 data are only available from Denmark, Finland, and Norway.[38] The shortest sample period is 24 months and the longest 54.

Column 3 in Table 3 shows the actual start dates of the reforms, and Column 4 shows the reform timing used in our analysis. The duration of the anticipation period is reported in Column 5. Most of the studied reforms have a staggered implementation, i.e., different ATC5 markets are affected by the reform at a different point in time and the same anticipation length is applied to all cohorts within a given reform.[39]

When presenting the event study results, we separate the anticipation period estimates in orange from the blue pre-treatment period and the green post-treatment period point estimates. We do not include anticipation periods in the calculation of ATTs.

38. Swedish data starts in 2006.

39. The 2000 and 2005 reforms in Denmark do not have a staggered implementation, because the studied reform influences all markets immediately.

Table 2: Matching Descriptive Statistics

	All Markets	Generic Competition	Treatment Markets	Pre-Study Generic Competition	Matched Markets
	(1)	(2)	(3)	(4)	(5)
Panel A: Number of ATC 5 markets					
Finland 2003	881	120	100	90	82
Finland 2009	896	133	133	124	109
Denmark 2000	815	110	68	64	62
Denmark 2005	822	150	114	100	96
Norway 2005	1,016	171	26	22	15
Sweden 2009	1,013	146	146	138	129
Panel B: Share of total pharmacy sales, %					
Finland 2003	100.00	34.66	33.69	28.08	26.55
Finland 2009	100.00	29.26	29.26	25.57	23.86
Denmark 2000	100.00	26.32	23.51	21.30	21.14
Denmark 2005	100.00	33.52	32.22	26.24	25.67
Norway 2005	100.00	28.91	12.24	8.03	3.64
Sweden 2009	100.00	20.82	20.82	19.22	18.47

[1] All Markets = number of/market share of ATC 5 markets in pre-period; Generic Competition = number of markets/market share of markets with generic competition during the observation period ; Treatment Markets = number of markets/market share of markets where the new regulation is implemented; Pre-Study Competition = number of markets/market share of markets in which generic competition started before our observation period; Matched Markets = number of/market share of successfully matched markets.

[2] Outcome data source: DLI-MI (1999–2013), Farmastat (2004–2013), Fimea (1999–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

Table 3: Sample Periods and Reform Timings

Reform Name	Sample Period	Sample Period Lenght	Reform Start	Reform Timing	Anticipation Length
	(1)	(2)	(3)	(4)	(5)
Finland 2003	2001m7–2004m7	36m	2003m4	2003m1	4m
Finland 2009	2008m2–2011m1	36m	2009m4	2009m1	4m
Denmark 2000	1999m11–2001m11	24m	2000m11	2000m11	0m
Denmark 2005	2003m12–2005m12	24m	2004m12	2004m12	0m
Norway 2005	2004m1–2005m12	24m	2005m1	2004m9	4m
Sweden 2009	2008m4–2012m10	54m	2009m12	2009m10	2m

Notes: Sample Period = Sample period used in empirical analyses; Sample Period Length = Lenght of sample period used in empirical analyses; Reform Start = Public information on when reforms started; Reform Timing = months used for treatment analyses; Anticipation Length = difference between Reform Start and Timing.

5 Results

We present our results in four parts. First, we show event study graphs of our main analyses of expenditure and availability, reform by reform. We start from the less strict reforms and progress to the stricter ones. We then summarize these results by showing and discussing the average treatment effects. Third, we turn to our secondary outcomes, i.e., average prices and quantity, by presenting ATT estimates for these outcomes.[40] We conclude with an analysis of whether the reforms affected close-by markets not directly affected by the reform in question. This analysis serves as a robustness check on our decision to define markets at the ATC5 level.

5.1 Event Study Analysis of Main Outcomes

5.1.1 Event Study Part I: GS and IRP

As GS started in Finland on April 1st 2003, we set the base period to January 2003. The results are shown in Figure 4a: they suggest that immediately after implementation there was a 7% decrease in average expenditure (left-hand top Figure), but the effect decreases in magnitude and becomes statistically insignificant as time passes. The point estimates on availability (left-hand bottom Figure) are positive but very imprecise.

Our analysis of the Finnish 2009 IRP reform provides different results. The GS system was expanded with IRP in April 2009; we set the base period to January 2009. We find in Figure 4b (right-hand top Figure) that the adoption of IRP decreased average expenditure. The effect is initially modest, but increases and then stabilizes. Expenditure decreased by 16% a year after the implementation of the reform. The point estimates on availability (right-hand bottom Figure) vary, and one may detect a negative trend. All point estimates are, however, positive and noisily estimated and thus do not support the idea that this reform would have decreased availability.

There is a strong case to be made for why GS alone had only a limited effect. First, GS simply expanded the choice set of consumers when they shop in pharmacies. Although a generic alternative might have had a lower price than the original branded product, deciding against substitution did not affect the level of reimbursement or the co-payment faced by the consumer, thus mitigating the reform's effect on consumers. There were only limited incentives to accept substitution for fully reimbursed products and for consumers who had

40. We present the event study results for these outcome variables in Appendix A.12.

exceeded their annual aggregate co-payment cap. To test this assertion we performed a subgroup analysis based on the different reimbursement categories of the products, the results of which are provided in Appendix A.6, Figure 12.[41] As suspected, we find that package-level price decreases are the largest and statistically significant for products that enjoy only basic levels of reimbursement. The point estimates for products with the full level of reimbursement of 100% are close to zero. Therefore, the Finnish GS reform did not succeed in decreasing the prices of products that enjoyed the most generous public subsidies. On the other hand, the adoption of IRP led to price decreases and savings also for products with full public reimbursement—the same products that were less affected by the earlier GS reform in 2003.

The results of the two Finnish reforms should be interpreted jointly. Unlike the regulatory regimes in the other Nordic countries, the Finnish GS regime was unique in the sense that at first it did not include any kind of reference pricing. Our results imply two important conclusions. First, the GS reform expanded the choice set of consumers because they were no longer tied to the decision made by the prescribing physician. However, the increase in competition had no significant effect on the expenditure per dose. Second, we find that GS with IRP decreased average expenditure quite substantially. The result indicates that the inclusion of consumer incentives is important in this type of a context. Our package-level subgroup analysis further suggests that for reimbursed products, price decreases are achievable when GS is tied to RP for reimbursed products.

5.1.2 Event Study Part II: ERP

This subsection examines the Danish experiment with ERP between the years 2000–2005, going from ERP to IRP in 2000 and back in 2005. We provide the first difference-in-difference analysis of ERP.[42] Our results are shown in Figures 5a and 5b. Introduction of ERP leads to an average 5% decrease in average expenditure per dose. The point estimates for availability are quite stable at roughly -2%, but statistically insignificant.

41. This subgroup inquiry is based on product (package) level analysis instead of market level analysis, because the reimbursement statuses are defined at the package level. This analysis is thus subject to SUTVA violations, but the idea is to demonstrate the weakness of the GS policy design when it is applied without reference price regulation instead of estimating unbiased causal effects in different reimbursement categories.

42. The baseline in the first reform (the adoption of the ERP) is set to November 2000 while our baseline in the latter reform is coded to the month when the legislation repealing ERP was passed in the Danish Folketing (Parliament), that is, December 2004.

The results are highly symmetric when studying the change in 2005 from ERP to IRP: we find that average expenditure increased by roughly 4%, while the change in availability is not significant.

Our results suggest that the Danish transition from IRP to ERP in generic markets lead to substantial savings in expenditure per dose, which implies that the price levels of off-patent pharmaceuticals were higher in Denmark than in its reference countries. The Danish 2005 reform, where ERP was replaced again by IRP, increased average expenditure.[43] Unlike in the other Nordic countries, the Danish price regulation system does not impose strict price caps for pharmaceutical products and thus prices can increase as an unintended consequence of the regulation.

The results of the removal of the ERP system in Figure 5b are visible only a few months after the implementation of the reform. This probably is due to the fact that firms slowly re-optimize their prices and since the difference-in-difference estimator measures the changes in price relative to the control group, the positive treatment effect is most likely due to the fact that prices stopped decreasing or decreased less in Denmark than in the control country Finland. We have also estimated the effects beyond the 12 lags shown in Figure 5b finding that the positive effects dynamically increase.[44] However, these estimates of the lags $t > 12$ are biased upward because of a mandatory price decrease in the control country (Finland) at the start of 2006. The cut was set at exactly 5% and applied to the reimbursement price caps (which were most often not binding for generics).[45] Taking this into account, our results imply increases in average expenditure of roughly 12% up to 20 months after the legislation was passed in the Danish Parliament.[46]

Our conclusions of the effects of the Danish reform of 2005 differ from those of Kaiser, Mendez, Rønde, and Ullrich (2014). They used a before-after setup (without control group) that did not take into account the negative trend in pharmaceutical prices.[47]

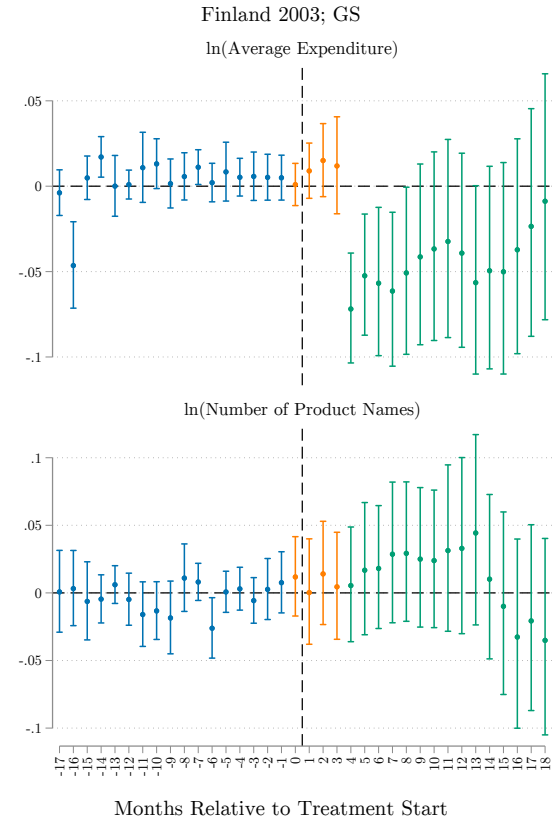
43. We study the effects of ERP transitions for non-competitive markets in Appendix A.7 and find weaker but similar asymmetry also for the non-competitive markets.

44. These results are also shown in Appendix A.8.

45. We have confirmed this in email exchange with the regulator, because legislation does not specify which price was cut.

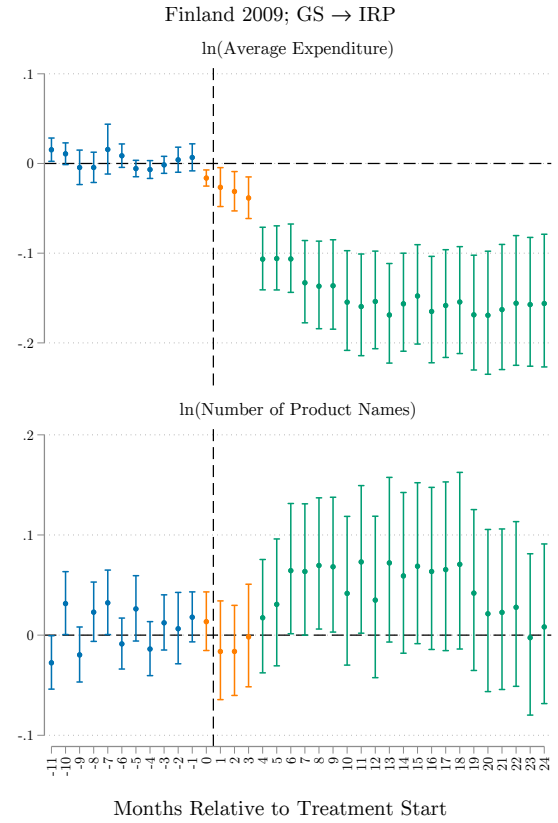
46. See Appendix Figure 23a for these event study plots.

47. We provide a replication of the before-after setup in Appendix Figure 14.



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Denmark used as a control group. Outcome data source: DLI-MI and Fimea.

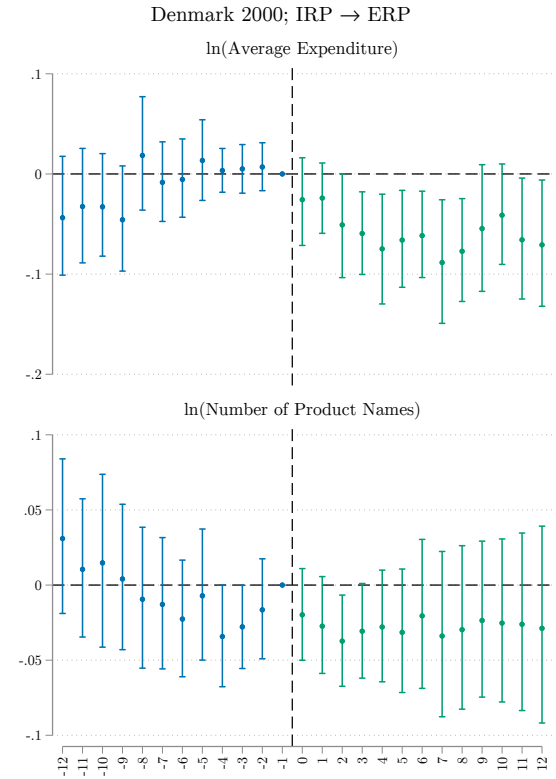
(a) Finland 2003



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Norway used as a control group. Outcome data source: Farmastat and Fimea.

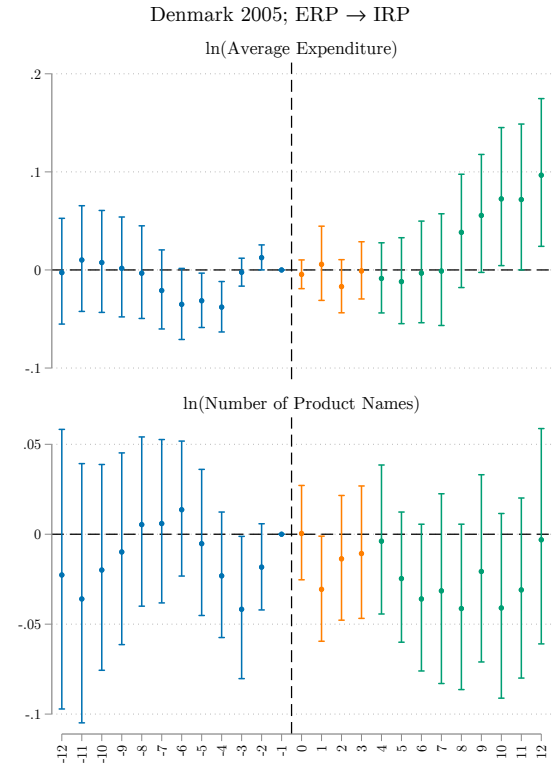
(b) Finland 2009

Figure 4: Part I: Main Results



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(a) Denmark 2000



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(b) Denmark 2005

Figure 5: Part II: Main Results

5.1.3 Event Study Part III: Stricter IRP Variants

We now turn to the Norwegian 2005 SP and the Swedish 2009 IRP auction reforms. These reforms are labeled "stricter" because they use polar ways to determine the reference price in regulated markets and imposed either very strict maximum price rules or steep incentives for customers to choose the cheapest product. In the SP reform pre-specified government rules determine the evolution of the reference price after patent expiry. In the auction system, the winner of the first price auction receives close to the entire market demand for each month. The exceptions to the rule are 1) patients who buy what the physician prescribed instead of the auction winner, ii) patients to whom the physician explicitly prescribes a different product than the cheapest and forbids substitution, and iii) buyers of the winner of the previous month's auction get reimbursed in the first half of the month. The auctions were and are conducted at the package level, which means that within an ATC5 market, there are several winners. An analysis of these policies offers new insights on how to design pharmaceutical price regulation because these reforms can be used to complement existing IRP regulations.

We start with the Norwegian SP reform. The baseline in our estimations is set to September 2004 when the reform was introduced in the Norwegian Parliament. Our results, shown in Figure 6a (top-left Figure), reveal that average expenditure per dose decreased by approximately 21%. The number of product names is not affected (bottom-left Figure): the point estimates are positive but imprecisely measured. Follow-up period is only 15 months long since Finland (control country) imposes a price cut in January 2006. In Appendix section A.10 we present results for the Norwegian 2005 reform using a post-period of 20 months.

The SP regulatory environment in Norway assigns the same price cap for both the original patented product and its generic alternatives, and the original price cap is based on an average of the prices of the original products in other European Economic Area (EEA) countries. Usually, the price cap is binding or close to binding for the branded manufacturer. Because the SP system forced a gradual decrease in this price cap, the price decrease can be expected to be the largest for products for which the price cap was binding.[48]

The reform also required pharmacies to have at least one product at or below the step-price (reference price) in stock. In the vertically integrated Norwegian retail market,

48. We show the SP rule for our observation period in Appendix Table 9.

wholesale prices for pharmaceuticals are in part just internal prices of the pharmacy chains. After the adoption of the SP model, the pharmacy chains had little incentive to sell generics below the maximum wholesale price (the price cap) or the maximum retail price (the price cap with the retail margin).

The results of the Swedish Auction-IRP reform are reported in Figure 6b. The baseline in the estimations is set to November 2009 when pharmaceutical companies were asked to submit their bids for the first time. Our estimates suggest that the reform led to statistically significant decreases between 20%–34% in average expenditure per dose. In line with our previous results, we find (bottom-right Figure) that even a reform with a large effect on expenditure seems to have no discernible effect on the availability of products.

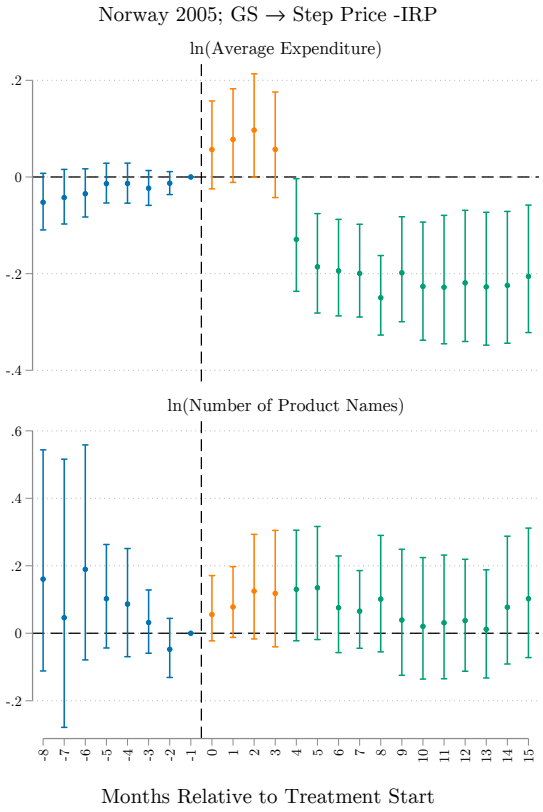
The importance of the results for the Swedish auction system should not be underestimated. By strongly restricting the set of products for which consumers are reimbursed, the competition for these consumers increases substantially. In practice, the winner of the monthly auction can expect to gain a very large share of the market. Our results suggest that the Swedish reform works very well from the point of view of curtailing expenditure—the main objective of pharmaceutical price regulation. However, one should keep in mind that the auction format was introduced simultaneously with a tightening of the maximum wholesale price regulation. Our reduced form approach does not allow us to disentangle the effects of the auction format and the tightening of the maximum wholesale price regulation. Finally, as we demonstrate in the following, the established estimation strategy of studying average prices would have provided a severe underestimate of the effectiveness of this regulation.

5.2 ATT Results of Main Outcomes

We use ATT estimates to summarize our results even if this may be a somewhat conservative choice in light of the event-study results presented above which often show increasing effects over time.[49]

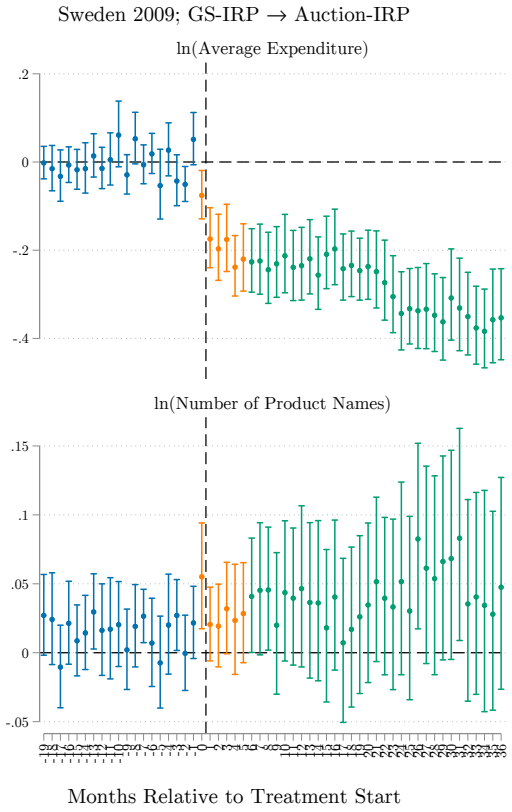
We have collected the ATT results into Table 4 where each Column is a different reform (in the same order as discussed above) and each Row is dedicated to a specific outcome variable. We show unweighted regression results in Panel A; these are comparable to the event study results discussed above. We have also run estimations where the individual

49. We explain in Section 4 how the ATTs are derived for the TWFE and Callaway and Sant'Anna (2021) estimators.



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: Farmastat and Fimea.

(a) Norway 2005



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Norway used as a control group. Outcome data source: Farmastat, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(b) Sweden 2009

Figure 6: Part III: Main Results

markets are weighted by their share of the treatment country pharmacy sales of prescription pharmaceuticals;[50] these are presented in Panel B.

We summarize first the results on average expenditure. The ATT results (first Row of Panel A) confirm our event study results and show significant expenditure decreases going from GS to IRP in Finland in 2009 and from IRP to either SP in Norway in 2005 or to Auction-IRP in Sweden in 2009. Also the results on the Danish reforms mirror nicely the event-study results, albeit the point estimates are lower. Comparing the first rows of Panels A and B shows that the weighted results suggest even larger expenditure decreases for the Finnish and Swedish 2009 reforms, as well as the Danish ones.

All in all, our results suggest four main conclusions: First, that giving consumers incentives to choose cheaper drugs is essential in the settings we study. Second, the results on the Danish reforms suggest that at least in a Nordic context, ERP delivers larger savings than regular IRP, albeit smaller than the stricter Norwegian SP and Swedish Auction-IRP regulations. Third, using strict maximum wholesale price rules (SP) and combining them with an auction-setting and steep consumer incentives (Auction-IRP) produces further significant savings. Depending on whether one uses the unweighted or weighted results, we find that going from the laxest regime - Finland's VGS—to the strictest—Sweden's Auction-IRP—produces savings of 37% to 55%.

Our fourth main result is that none of the reforms, even those that considerably decreased expenditure per dose, seem to have had any effect on product availability. Our results thus suggest that one of the feared cost of stricter regulation, decreased availability, is not warranted. Our results should however not be taken as definitive evidence that price regulation has no effect on pharmaceutical availability. Our analysis is limited to at most 36 months post reform. Also, our analysis does not cover all aspects of availability. We focus only on markets that are included in the main analysis, and issues related to product entry delays are beyond the scope of our analysis (see Maini and Pammolli 2022).

5.3 ATT Results on Price and Quantity

We have collected our ATT estimates of the effect of the reforms on price and quantity also into Table 4. Turning first to the price results (Row "Average Price") where the dependent variable is the period-specific arithmetic average of prices per DDD (i.e, measured at the market level), we find coefficients that are clearly smaller in absolute value than the

50. We use sales in the months -12 to -6 in each estimation period to calculate the weights.

Table 4: Average Treatment Effects

	Part I		Part II		Part III	
	Finland 2003	Finland 2009	Denmark 2000	Denmark 2005	Norway 2005	Sweden 2009
Panel A: Main Estimations						
Average Expenditure	-0.03	-0.13*	-0.05*	0.04	-0.21*	-0.27*
	[-0.07, 0.01]	[-0.18, -0.08]	[-0.09, -0.01]	[-0.01, 0.09]	[-0.29, -0.12]	[-0.34, -0.20]
Number of Product Names	0.01	0.04	-0.02	-0.01	-0.01	0.04
	[-0.03, 0.05]	[-0.02, 0.10]	[-0.06, 0.02]	[-0.05, 0.03]	[-0.15, 0.15]	[-0.00, 0.09]
Average Price	-0.04	-0.05	-0.07*	0.07*	-0.10	-0.04
	[-0.12, 0.04]	[-0.09, -0.00]	[-0.12, -0.01]	[0.02, 0.12]	[-0.18, -0.00]	[-0.11, 0.04]
Number of Doses	0.01	0.04*	0.00	0.07*	0.04	0.12*
	[-0.04, 0.07]	[0.01, 0.07]	[-0.04, 0.04]	[0.03, 0.12]	[-0.00, 0.09]	[0.02, 0.22]
Wholesale Price	-0.05	-0.10*	-0.09*	0.05	-0.11*	-0.06*
	[-0.11, 0.02]	[-0.14, -0.07]	[-0.13, -0.05]	[-0.02, 0.12]	[-0.20, -0.01]	[-0.11, -0.01]
Panel B: Weighted Estimations						
Average Expenditure	0.03	-0.25*	-0.05*	0.04	-0.21*	-0.37*
	[-0.13, 0.23]	[-0.34, -0.15]	[-0.09, -0.01]	[-0.01, 0.09]	[-0.29, -0.12]	[-0.49, -0.22]
Number of Product Names	-0.00	0.12	-0.02	-0.01	-0.01	0.06
	[-0.06, 0.05]	[-0.02, 0.29]	[-0.06, 0.02]	[-0.05, 0.03]	[-0.15, 0.15]	[-0.02, 0.15]
Average Price	0.04	-0.16*	-0.07*	0.07*	-0.10	0.06
	[-0.11, 0.22]	[-0.23, -0.08]	[-0.12, -0.01]	[0.02, 0.12]	[-0.18, -0.00]	[-0.24, 0.50]
Number of Doses	-0.06	0.07*	0.00	0.07*	0.04	0.14
	[-0.14, 0.03]	[0.04, 0.10]	[-0.04, 0.04]	[0.03, 0.12]	[-0.00, 0.09]	[-0.05, 0.37]
Wholesale Price	0.04	-0.17*	-0.09*	0.05	-0.11*	-0.07
	[-0.19, 0.32]	[-0.23, -0.10]	[-0.13, -0.05]	[-0.02, 0.12]	[-0.20, -0.01]	[-0.14, -0.00]

[1] Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). Outcome data source: DLI-MI (1999–2013), Farmastat (2004–2013), Fimea (1999–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

[2] * = statistically significant at the 95% confidence level. 10000 replications for ATC-5 wild bootstrapped standard errors.

estimated expenditure effects for the Finnish 2009 reform (switching to IRP from GS), the Norwegian 2005 reform (switching to SP from IRP) and the Swedish 2009 reform (switching from IRP to Auction-IRP). The estimated price effects for the Norwegian and Swedish reforms are at best marginally statistically significant. As a detail, it is noteworthy that the weighted ATT for the 2009 Swedish reform is positive albeit insignificant while the expenditure results suggest double-digit savings per does. We find price effects that are larger than the expenditure effects for the two Danish reforms only.

These results reveal a common pattern in that the price change is larger in absolute value than the expenditure change. The most likely explanation for this is that these reforms affect not only prices, but also increase consumers' financial incentives to choose a cheaper product.

Sweden's Auction-IRP system is a good example of this effect. The Auction-IRP is in practice a monthly auction where the policymaker procures the pharmaceuticals included in the tax-funded funded social insurance. Consumers can only be reimbursed for the product for which they receive the prescription or for the substitutable product that was offered for the lowest price. This leads to a stark form of the general point we have made earlier: average market prices do not matter; what matters is average expenditure per dose sold. In the Swedish case, where almost all sales are channeled to the cheapest product in the market, the prices of the other products become almost irrelevant. The estimation strategy employed in the existing studies and by us in the analysis of average posted price does not take this substitution effect into account. The conclusion we draw is that when studying price regulation it is not advisable to use prices as a proxy for expenditures when consumer incentives are also affected.

To enable a comparison to the existing literature, we have also estimated price effects using package-level data; these results are reported in Rows 'Wholesale Price'. The difference to the Average Price results are mostly modest, suggesting only small bias due to possible violation of SUTVA: the one exception is the 2009 reform in Finland moving from GS to IRP. The small differences in the two price effect estimates are comforting given the prevalence of package-level analyses in the literature. However, these estimates are as different from the expenditure results as are the Average Price results.

Finally, we find mostly very mild quantity effects across all reforms (Row 'Number of Doses' in Table 4. The reform with the greatest effect is the Swedish 2009 reform replacing IRP with Auction-IRP which we find to increase quantity by 12%. The Norwegian 2005 SP reform and the Finnish 2009 IRP reform also seem to have increased consumption by

4% while simultaneously reducing expenditure per dose by 21% and 13%. As the above three reforms are the ones leading to largest expenditure decreases per dose, these results do suggest non-zero price elasticities for prescription pharmaceuticals.

5.4 Spillovers Between Pharmaceutical Markets

The existing literature has examined the spillovers of regulation to markets not directly affected by the reform. We have defined markets at the active ingredient, i.e., ATC5 level, but there are diseases that are treated with pharmaceuticals from more than one ATC5 class. Since not all (ATC5) markets used in the treatment of a given disease are necessarily subject to a given reform, it is possible that a reform nonetheless indirectly affects those markets that are not directly affected. Spillovers of this type are called *therapeutic competition*, and in some studies, the effect of therapeutic competition on prices has been found to be economically significant (Brekke, Grasdal, and Holmås 2009; Brekke, Holmas, and Straume 2011). To test our market definition, we execute an analysis of whether such effects exist in our data.

For our analysis of spillovers, the treatment group consists of markets in the treatment country that share the same ATC4 class but are not directly affected by the reform. The control group consists of the same ATC5 markets in the control country. We use the same estimation methods as in the main analysis. Here we report only the weighted and non-weighted ATTs.[51] The Danish reforms (2000 and 2005) are excluded from the spillover analysis because these reforms influenced all products in the Danish pharmaceutical market.[52]

The main result is that the estimated spillover effects on expenditure (first Row) are small in absolute magnitude, negative in sign and statistically insignificant. The weighted results (first Row in Panel B) mostly confirm the outcomes, though now also the Finnish move from GS to IRP in 2009 produces a statistically significant negative impact, but the estimate is just one fifth of the direct effect.

The effects on availability (number of products) are consistently estimated to be very small in magnitude and quite precise: the exception is the Swedish 2009 move to Auction-IRP where we find a large positive effect on availability.

Turning to the secondary outcomes, our market- (Row "Average Price") and package-

51. In the future we will add event studies to the Appendix.
52. The results for the monopoly Danish markets are discussed and reported in the Appendix section A.7.

Table 5: Average Treatment Effects (Spillover Samples)

	Part I		Part III	
	Finland 2003	Finland 2009	Norway 2005	Sweden 2009
Panel A: Main Estimations				
Average Expenditure	-0.01	-0.03	-0.03	0.04
	[-0.04, 0.03]	[-0.06, 0.01]	[-0.10, 0.05]	[-0.02, 0.10]
Number of Product Names	-0.01	0.00	-0.02	0.04
	[-0.03, 0.02]	[-0.02, 0.04]	[-0.08, 0.04]	[-0.00, 0.08]
Average Price	-0.01	-0.03	0.00	0.05
	[-0.05, 0.02]	[-0.17, 0.13]	[-0.05, 0.07]	[-0.04, 0.14]
Number of Doses	0.08*	0.02	0.13*	-0.07
	[0.01, 0.16]	[-0.09, 0.14]	[0.01, 0.26]	[-0.14, 0.02]
Wholesale Price	0.00	-0.02	-0.04	0.02
	[-0.01, 0.01]	[-0.04, 0.00]	[-0.12, 0.04]	[0.00, 0.04]
Panel B: Weighted Estimations				
Average Expenditure	0.02	-0.05*	-0.03	0.09
	[-0.00, 0.04]	[-0.08, -0.03]	[-0.10, 0.05]	[-0.04, 0.25]
Number of Product Names	-0.01	-0.03	-0.02	0.11
	[-0.06, 0.05]	[-0.10, 0.03]	[-0.08, 0.04]	[-0.00, 0.24]
Average Price	0.04	-0.07	0.00	0.06
	[-0.01, 0.09]	[-0.18, 0.05]	[-0.05, 0.07]	[-0.05, 0.18]
Number of Doses	0.11*	0.06	0.13*	-0.09
	[0.07, 0.14]	[-0.05, 0.17]	[0.01, 0.26]	[-0.20, 0.03]
Wholesale Price	0.02*	-0.04*	-0.04	0.05*
	[0.02, 0.03]	[-0.06, -0.02]	[-0.12, 0.04]	[0.01, 0.10]

[1] Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). Outcome data source: DLI-MI (2007–2013), Farmastat (2004–2013), Fimea (2007–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

[2] * = statistically significant at the 95% confidence level. 10000 replications for ATC-5 wild bootstrapped standard errors.

level (Row "Wholesale Price") price estimations deliver mostly small and insignificant estimates. We do find though that the package level results which violate SUTVA are more often statistically significant; so much so that one would draw different conclusions than when using market level average prices at least for the Swedish 2009 Auction-IRP reform and possibly also for the Finnish 2009 IRP reform. Our estimates are smaller in absolute value than the spillover effects on prices reported in the literature (Brekke, Grasdal, and Holmås 2009) both overall and especially regarding the statistically significant price effects.

Changes in physician behavior can create links between demand in regulated and non-regulated markets: Cost-conscious physicians might start to write more prescriptions for pharmaceuticals where competition has lowered prices, and this could decrease sold quantities in non-regulated markets if products. The estimated average spillover effect on sold quantity varies between -4% and 1%, but none is statistically significant. The estimates for the two Finnish reforms are, however, close to statistical significance. These results support our decision to define the relevant market at the ATC5 active ingredient level.

6 Conclusions

We investigate the causal effects of different price regulation policies on pharmaceutical expenditure and product availability in the Nordic pharmaceutical markets facing generic competition. Such policies are important because pharmaceutical spending has been increasing and because public and private health insurance schemes have in many countries reduced or even removed the price sensitivity of citizens.

We combine product-level price, quantity, and sales information with extensive information on different regulation policies and market institutions that were in place 1999–2010 and employ modern difference-in-difference methods to analyze the effects of several reforms.

The regimes in our data can for the most part be ordered by the strictness of the price regulations and steepness of the financial incentives of patients to choose a cheaper drugs at the pharmacy. We find that several reforms decrease public expenditure considerably: Moving from the least strict regulatory regime in our data to the strictest reduced expenditure by at least 36% and possibly by 53%. The effects on expenditure were greater than those on prices. This is explained by the fact that the reforms also introduced stronger financial incentives for patients to choose cheaper drugs within the same ATC5 group. This implies that the existing literature that heavily relies on estimating the effect of regula-

tions on prices may have underestimated the effectiveness of price regulations in curbing expenditure. Despite the large effects on expenditure, the reforms did not have an adverse effect on product availability, nor any meaningful effects on quantity. Our results suggest that regulations that combine maximum price regulation in markets with generic competition with steep patient incentives to facilitate competition are a powerful tool to decrease pharmaceutical expenditure without having to compromise availability.

References

Acemoglu, Daron, and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *The Quarterly Journal of Economics* 119 (3): 1049–1090.

Alpert, Abby. 2016. "The anticipatory effects of Medicare Part D on drug utilization." *Journal of Health Economics* 49 (September): 28–45.

Bhattacharya, Jay, Timothy Hyde, and Peter Tu. 2013. *Health Economics.* Macmillan International Higher Education.

Brekke, Kurt R., Astrid L. Grasdal, and Tor Helge Holmås. 2009. "Regulation and pricing of pharmaceuticals: Reference pricing or price cap regulation?" *European Economic Review* 53, no. 2 (February): 170–185.

Brekke, Kurt R., Tor Helge Holmas, and Odd Rune Straume. 2011. "Reference pricing, competition, and pharmaceutical expenditures: Theory and evidence from a natural experiment." *Journal of Public Economics* 95, no. 7 (August): 624–638.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* 225, no. 2 (December): 200–230.

Correia, Sergio. 2016. *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator.* Technical report.

Cutler, David. 2022. "Medicare Enters the Pharmaceutical Purchasing Business." *Journal of American Medical Association Health Forume* 3 (9): 255–285.

Danzon, Patricia M., and Li-Wei Chao. 2000. "Does Regulation Drive out Competition in Pharmaceutical Markets?" *The Journal of Law and Economics* 43, no. 2 (October): 311–358.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110, no. 9 (September): 2964–2996.

De Chaisemartin, Clément, and Xavier d'Haultfœuille. 2023. "Two-Way Fixed Effects and Differences-in-Differences with Heterogeneous Treatment Effects: A Survey." *Econometrics Journal.*

Dubois, Pierre, Ashvin Gandhi, and Shoshana Vasserman. 2022. *Bargaining and International Reference Pricing in the Pharmaceutical Industry.* Working Paper 30053. National Bureau of Economic Research, May.

Dubois, Pierre, and Laura Lasio. 2018. "Identifying Industry Margins with Price Constraints: Structural Estimation on Pharmaceuticals." *American Economic Review* 108, no. 12 (December): 3685–3724.

Dubois, Pierre, Olivier de Mouzon, Fiona Scott-Morton, and Paul Seabright. 2015. "Market size and pharmaceutical innovation." *The RAND Journal of Economics* 46 (4): 844–871.

Duggan, Mark G, and Fiona Scott Morton. 2010. "The Effect of Medicare Part D on Pharmaceutical Prices and Utilization." *American Economic Review* 100 (1): 590–607.

———. 2011. "The Medium-Term Impact of Medicare Part D on Pharmaceutical Prices." *American Economic Review* 101 (3): 387–92.

Einav, Lirav, Amy Finkelstein, and Maria Polyakova. 2018. "Private Provision of Social Insurance: Drug-Specific Price Elasticities and Cost Sharing in Medicare Part D." *American Economic Journal: Economic Policy* 10(3):122–153.

Goodman-Bacon, Andrew. 2021. "Difference-in-differences with variation in treatment timing." *Journal of Econometrics* 225, no. 2 (June): 254–277.

Herr, Annika, and Moritz Suppliet. 2017. "Tiered co-payments, pricing, and demand in reference price markets for pharmaceuticals." *Journal of Health Economics* 56 (December): 19–29.

Kaiser, Ulrich, Susan J. Mendez, Thomas Rønde, and Hannes Ullrich. 2014. "Regulation of pharmaceutical prices: Evidence from a reference price reform in Denmark." *Journal of Health Economics* 36 (July): 174–187.

Lakdawalla, Darius N. 2018. "Economics of the Pharmaceutical Industry." *Journal of Economic Literature* 56, no. 2 (June): 397–449.

Lee, Junghee, Hyun Seok (Huck) Lee, Hyoduk Shin, and Vish Krishnan. 2021. "Alleviating Drug Shortages: The Role of Mandated Reporting Induced Operational Transparency." *Management Science* 67, no. 4 (April): 2326–2339.

Leopold, Christine, Sabine Vogler, Aukje K Mantel-Teeuwisse, Kees de Joncheere, Hurbert GM Leufkens, and Richard Laing. 2012. "Differences in external price referencing in Europe — A descriptive overview." *Health policy* 104 (1): 50–60.

Lexchin, Joel. 2021. "Time to market for drugs approved in Canada between 2014 and 2018: an observational study." *BMJ Open* 11 (7).

Maini, Luca, and Fabio Pammolli. 2022. "Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market." *American Economic Journal: Microeconomics Forthcoming.*

Mammen, Enno. 1993. "Bootstrap and Wild Bootstrap for High Dimensional Linear Models." *The Annals of Statistics* 21 (1): 255–285.

Ornaghi, Carmine. 2009. "Mergers and innovation in big pharma." *International Journal of Industrial Organization* 27 (1): 70–79.

Pavcnik, Nina. 2002. "Do Pharmaceutical Prices Respond to Potential Patient Out-of-Pocket Expenses?" *The RAND Journal of Economics* 33 (3): 469–487.

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon, and Matthew D. Webb. 2019. "Fast and wild: Bootstrap inference in Stata using boottest." *The Stata Journal* 19, no. 1 (March): 4–60.

Simonsen, Marianne, Lars Skipper, Niels Skipper, and Anne Illemann Christensen. 2021. "Spot price biases in non-linear health insurance contracts." *Journal of Public Economics* 203 (November): 104508.

Stomberg, Christopher. 2016. "Drug Shortages, Pricing, and Regulatory Activity." In *Measuring and Modeling Health Care Costs,* 323–348. University of Chicago Press, October.

Yin, Wesley. 2008. "Market incentives and pharmaceutical innovation." *Journal of Health Economics* 27 (4): 1060–1077.

Yurukoglu, Ali, Eli Liebman, and David B. Ridley. 2017. "The Role of Government Reimbursement in Drug Shortages." *American Economic Journal: Economic Policy* 9, no. 2 (May): 348–382.

A Appendix

The subsections of the Appendix are ordered as they are referred to in the main text.

A.1 Nordic Countries, Reimbursement systems and Pharmacy Markup-rules

This subsection provides an overview of the four Nordic countries included in this study and gives details on the reimbursement systems in use, as well as the pharmacy markup-rules in place.

Country overview. Figure 7 shows the Nordic countries on a map and Table 7 displays some relevant descriptive statistics of the four countries. All countries except Norway are EU member states. All four countries belong to the EEA, meaning that Norway also follows many EU regulations. Finland is the only Nordic country without her own national currency, having adopted the euro in 2002. In 2007 Sweden had the largest population, which was more than 9 million, while Norway's population of 4.6 million was the smallest. The percentage of population aged 65 years and older was also the highest in Sweden and the lowest in Norway. In 2007, GDP per capita was the highest in Norway and the lowest in Finland. Sweden had the largest pharmaceutical market with total sales of more than 2.7 billion euros in 2007, while Norway had the smallest market with sales of 1.46 billion euros. At 8.5%, the Swedish pharmaceutical market was also the largest relative to GDP. In Finland and Denmark, the pharmaceutical market represented approximately 6.3% of GDP, and in Norway 3.3%.

Reimbursement systems. Table 6 summarizes the structure of reimbursement systems in Nordic countries. The countries are divided into product-based (Finland and Norway) and consumption-based (Denmark and Sweden) reimbursement systems. Although their reimbursement systems are quite similar, individual countries have different reimbursement rates and annual out-of-pocket ceilings. The Finnish reimbursement system is the least generous, because the smallest reimbursement rate is 42% and the annual out-of-pocket cost ceiling is 610 euros, almost three times higher than in Norway or Sweden.

Table 7: Nordics Descriptive Statistics

	Population	Population aged 65 and above, %	EU Member	EEA Country	Currency	GDP per capita	Market Size	GDP share, %
Finland	5.3	16.5	Yes	Yes	Euro	29,900	1,873	6.27
Sweden	9.1	17.4	Yes	Yes	Swedish krona	32,300	2,774	8.59
Denmark	5.4	15.5	Yes	Yes	Danish krone	30,800	1,951	6.33
Norway	4.7	14.6	No	Yes	Norwegian krone	44,200	1,461	3.31

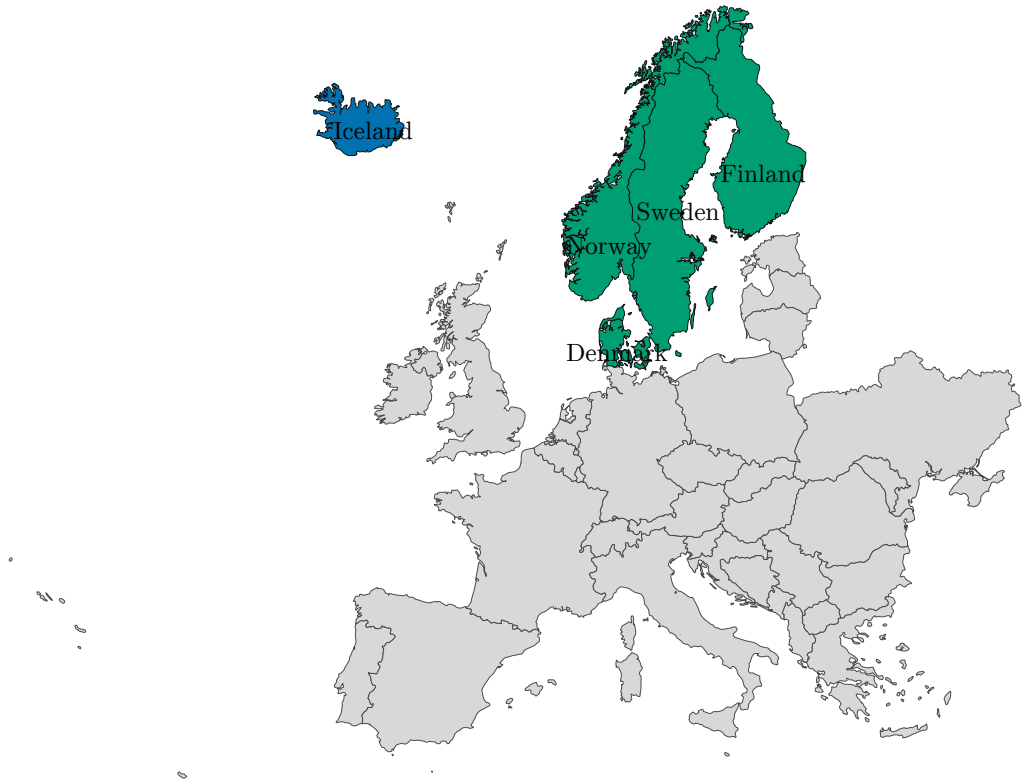
Notes: The values are from 2007 and population is expressed in millions. The second column displays the percentage of total population that were aged 65 and above. The EU member column indicates whether a country is an European Union member state and the EEA country column indicates whether the countries belong to the European Economic Area. The currency column shows which currency is used in each country. GDP per capita is expressed in euros (PPS). Market size is expressed in millions of euros and is calculated as the sum of sales using pharmacy purchase prices (wholesale prices) in 2007. Market share denotes the share that the pharmaceutical market forms of the country's total GDP. Outcomes data source: DLI-MI, Farmastat, Fimea, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007).

Table 6: Reimbursement Systems in the Nordics

	Reimbursement % **	Annual out-of-pocket ceiling	Out-of-pocket threshold	Time period for annual ceiling	Reference countries (2012)	Type of referencing	Annual reimbursement expenditure
Panel A: Product specific							
Finland	Basic: 42% Lower special: 72% Higher special: 100%	610 EUR **	N/A	calendar year	EEA (excl. Croatia) + UK	directional	1142 EUR***
Norway	Standard: 64% Serious contagious diseases: 100%	205 EUR**	N/A	calendar year	avg. of 3 lowest countries	direct	11480 NOK**
Panel B: Consumption based							
Sweden	901-1700 SEK: 50% 1701-3300 SEK: 75% 3301-4300 SEK: 90% 4301 SEK: 100%	194 EUR **	900 SEK**	calendar year	N/A	N/A	21500 SEK**
Denmark	0-480 DKK: 0% 480-1165 DKK: 50% 1165-2730 DKK: 75% > 2730 DKK: 85%	Only for chronically ill after 472 EUR **	480 DKK**	continuous 12 month period	N/A	N/A	11447 DKK***

Notes: Reimbursement (%) = Different reimbursement categories and reimbursement classes; Annual out-of-pocket ceiling = Annual limit for out-of-pocket expenditures; Out-of-pocket threshold = Threshold for out-of-pocket expenditure; Time period for annual ceiling = Time window where the out-of-pocket annual ceiling contributes; Reference countries (2012) = Countries that are used in external reference price calculations; * : 2005, ** : 2006, *** : 2007. Annual reimbursement expenditures are expressed in millions. Sources: PPRI, KELA and Leopold, Vogler, Mantel-Teeuwisse, Joncheere, Leufkens, and Laing (2012).

Figure 7: Nordic Countries in Europe



Colored countries denote the Nordic siblings.
Green denotes the Nordic countries included in this study.

Pharmacy mark-ups. All countries except Norway have a mathematical formula for the pharmacy mark-up, that is, pharmacies do not decide retail prices. Table 8 shows how these formulas (we display formulas for the year 2009) convert the pharmacy purchase price (PPP) into the pharmacy retail price (PRP), which is the price from which reimbursements are calculated. The main takeaway from the table is that the retail price formulas transmit changes in pharmacy purchase prices to pharmacy retail prices.

Norway is a slight exception because the institutional setting allows pharmacies to charge a lower markup than what the formula presented in Table 8 would yield.

Table 8: Pharmacy Retail Price Formulas

Effective Period	Type	Register Price
	Denmark	
	Prescription drugs (DDK)	
26/03/2004 - 21/03/2007	<= 30	PPP + 0.61×(0.6×PPP+1.8 DKK)
	30-60	PPP + 0.61×(0.4×PPP+7.8 DKK)
	> 60	PPP + 0.61×(0.2×PPP+19.8 DKK)
	Finland	
	Prescription drugs (€)	
1/1/2003 - 1/1/2014	0-9.25	1.5 × PPP + 0,50 €
	9.26-46.25	1.4 × PPP + 1,43 €
	46.26-100.91	1.3 × PPP + 6,05 €
	100.92-420.47	1.2 × PPP + 16,15 €
	> 420.47	1.125 × PPP + 47,68 €
	Norway	
	Prescription drugs (NOK)	
1/1/2001 - 1/1/2009	0-200	1.08 × PPP
	> 200	1.05 × PPP
	Sweden	
	Prescription drugs (SEK)	
	0-75	PPP × 1.20 + 31.25
15/7/2009 - 1/11/2009	> 75-300	PPP × 1.03 + 44.00
	> 300-6000	PPP × 1.02 + 47
	> 6000	PPP + 167.00

Notes: Effective Period = Period when the retail price formula was in use; Type = Price range where the retail price formula applies; Register Price = How list price is determined from the wholesale price.

A.2 Branded, generic and parallel imported pharmaceuticals

There can be three types of products in a given ATC5 category by origin: the unique branded (original patented) product that was (is) protected by a patent;[53] generic products that feature the same molecule as the original drug, but are most of the time produced by different firms than the branded drug (brand manufacturers sometimes have their own generic products, too); and third, so-called parallel imported products, which are manufactured by the producer of the branded drug, but originally sold to a different geographic

53. Parallel imports may take place while patent protection is in place.

market (= EU Member State), bought there and shipped to the country in question by an intermediary company (parallel importer).

Figure 8 illustrates how three different substitutable products (the original patented product, the parallel import, and the generic copy) look. The packages are from Finland and all three products contain the same active ingredient venlafaxine (ATC5: N06AX16), which is an antidepressant. If the patient received a prescription for the branded product, substitution can be made for the products at the bottom of the figure.

The branded product is placed at the top of Figure 8 and has a unique product name "Efexor". The generic product is at the bottom left, and the parallel imported product is at the bottom right. All packages provide information on package size (98 tablets), strength (150 mg) type of product, product id (Nordic Article Number (VNR)) and the company that sells the product. All packages contain detailed instructions related to pharmaceutical use and information on possible adverse effects on the use of the product.

Figure 8: Example of a branded (top), generic (bottom left) and parallel (bottom right) imported pharmaceutical

A.3 EU Regulations, Price Regulation Regimes and Minor Price Regulation Reforms

Here we provide more details on the regulatory institutions regarding market entry at the European A.3.1 level, and then details for the price regulation that we study in each of the four countries in our data: Denmark (A.3.2), Finland (A.3.3), Norway (A.3.4) and Sweden (A.3.5). We close the subsection with a discussion of minor price regulation reforms (A.3.6).

A.3.1 Relevant EU Regulations

We briefly describe the regulatory process for a given pharmaceutical product to be allowed to enter the market in an EU Member State. There are two routes: obtaining market authorization and (after that has been granted), so-called parallel imports.

Obtaining market authorization. There are four distinct processes through which a product can receive market authorization for sale in the European common market and in its Member States. Three of these processes, namely, the centralized, decentralized, and mutual recognition processes, are based on legislation passed by the EU. The fourth option, national market authorization, is regulated by the Member States themselves. In the centralized procedure, authorization is granted by the European Medicines Agency (EMA) through which the authorization is valid in the European Economic Area (EEA).[54] In the decentralized process, a company simultaneously applies for market authorization in more than one Member State through the respective national authorities, on the condition that the product has no market authorization in any of the Member States. The decentralized process is led by one of the Member States, and other national authorities provide assistance in the process. In the mutual recognition process, a company applies for market authorization for a product that has already been approved in at least one Member State.

Parallel imports. Parallel imports are a feature of European pharmaceutical markets. The market share of parallel imported products varies from country to country, but the possibility of parallel imports from within the EU exists in all EU Member States and banning them is illegal.

54. The EEA covers the EU Member States and Iceland, Liechtenstein, and Norway.

A.3.2 Denmark

May 1997–Oct. 2000: GS and IRP. In 1997, Denmark adopted mandatory substitution of generics on top of an existing RP system for generics.[55] This regime corresponds to our definition of a GS system with IRP. The Danish system required pharmacies to substitute to the cheapest interchangeable available product unless the price differential was (roughly) less than 5%.[56] The prescribing physician could still opt out of substitution for medical reasons. If a consumer did not buy the reference-priced product, he/she was required to pay the price difference between the products out of pocket.

Nov. 2000– Dec. 2004: GS and ERP. Denmark switched from generic GS with IRP to GS with ERP in November 2000. Reference prices were calculated using prices in other European countries.[57] If a product was sold only in Denmark or the domestic price was lower than the price calculated using the other European prices, the price in Denmark was used as the reference price.

The implementation process of ERP on the Danish market had already started in 1998 when manufacturers of new pharmaceutical substances (defined by market entry after April 1, 1997) were required to inform the Danish government of their prices in other European countries.[58] The process was finalized in November 2000 when the Danish government stopped the reimbursement of all products that exceeded their European average prices.[59] While the use of ERP was included in the legislation in summer 2001, the regulator started applying ERP already in November 2000.[60] We use November 2000 as the date of the reform.

Jan. 2005– : GS and IRP. ERP lasted until January 1, 2005, when it was replaced by IRP.[61] In the new regime, the reference price was again the lowest domestic price within a substitution group. The government also abolished the ERP of patented pharmaceuticals.

There are two other institutional changes that occur in Denmark during our study period that are not directly related to the reforms studied. The first is the overhaul of the reimbursement system. In March 2000, the Danish government adopted a new reim-

55. See BEK nr 308 af 06/05/1997 §36–§37.

56. This "price corridor" in Denmark has remained mostly the same since 1996. See BEK nr 724 af 01/08/1996 §37.

57. EU-15 excluding Greece, Luxembourg, Spain, and Portugal.

58. The government would then use this price information to cap the public reimbursement to the average of the lowest two prices.

59. As stated in LOV nr 1031 af 23/11/2000 §7j.

60. See LOV nr 495 af 07/06/2001 §7d.

61. See LOV nr 1431 af 22/12/2004 §7d.

bursement model in which the fixed product-specific reimbursement level was replaced by a system in which the patient's reimbursement level was non-linearly calculated based on spending (see Simonsen, Skipper, Skipper, and Christensen 2021). The other change is a price freeze agreement between the Danish government and an association of pharmaceutical manufacturers. We explain these changes in more detail in Appendix A.3.6.

A.3.3 Finland

Up to March 2003: VGS. Throughout the 1990s, the Finnish VGS required prescribing physicians to actively opt-in to allow GS to occur. In practice, GS and prescription of generics was almost non-existent.[62]

April 2003–March 2009: GS. The Finnish government adopted mandatory GS in April 2003.[63] In the new regime, pharmacies were required to deliver one of the products at or close to the cheapest price.[64] The reimbursement of a consumer was not affected if she decided against substitution; the monetary incentives to substitute were small in drug categories with high reimbursement rates. Unlike Finland, other countries combined substitution policies with financial incentives for the patient.

April 2009– : GS and IRP. To address the incentive problems related to GS and high reimbursement rates, Finland adopted IRP in April 2009.[65] Reference pricing was applied to products that were publicly reimbursed and to which at least one generic substitute was available. The reference price in a substitution group is the highest reimbursed retail price. During our sample period, the reference price was defined as the cheapest retail price within the reference price group +1.5€(retail price less than 40€) or 2.5€(retail price greater than 40€). Reference prices were updated quarterly. If the price of the purchased product exceeds the reference price, the consumer is reimbursed on the basis of the reference price and pays the price difference out of pocket. Parallel imports were not included in the system until 2017.[66] GS continued to be applied for non-reimbursed and parallel imported products.

62. See the government proposal HE 165/2002 vp, page 6.
63. See 80/2003 §57b.
64. Pharmacies were required to offer substitution if the prescribed product was either 2€(retail price less than 40€) or 3€(retail price more than 40€) more expensive than the cheapest product in the substitution group.
65. See Chapter 6 §18-§23.
66. See 1100/2016 Chapter 6 §18. Before this, parallel imports could be included in reference price groups if other generics were on the market. After the 2017 change, this requirement was lifted. In practice, this allowed RP to start even during the patent period.

In addition to the above major reforms, Finland has implemented minor reforms in the 2000s. The first minor reform in 2006 imposed that the price cap for new entrants should be 40% lower than the cap of the original product. The second reform was a 5% price cap cut on all reimbursement drugs. These reforms are explained in more detail in Appendix A.3.6.

A.3.4 Norway

March 2001–2005: GS. Norway adopted a GS policy and liberalized the pharmacy sector simultaneously in 2001.[67] Prior to the 2001 reform Norway had an ERP system.[68] Thus, the GS system with ERP elements is the baseline regulatory regime for subsequent reforms in Norway.

If the consumer substituted to the cheapest alternative in that regime, she had to pay the difference in price between the cheapest alternative and the chosen product out of pocket.[69] The Norwegian GS did not explicitly require pharmacies to substitute for a cheaper alternative available; instead, pharmacies are incentivized to offer GS.[70]

Jan. 2005– : GS and the Step-Price regime. Norway implemented a major change to the GS system in 2005 by introducing the current SP system. After generic entry has taken place, the maximum reimbursement price (now called as the Step-Price) gradually decreases.[71] The base level for the price is established as the maximum allowed retail price at the time of generic entry. If a consumer decides not to buy the product priced at the Step-Price, she is required to pay the difference in price out of pocket. The

67. We do not study the effects of this substitution reform because the effects of the reform cannot be separated from the effects of pharmacy market liberalization. For further information, see LOV-2000-06-02-39.

68. The maximum reimbursement price was the average of the three lowest prices of the original patented product in the other EEA countries.

69. In comparison to the Finnish GS, the Norwegian regime provided financial incentives, and the Finnish policy did not offer incentives. See LOV-2000-06-02-39 for further information.

70. Originally, if pharmacies sold a product whose wholesale price was below the maximum wholesale price, they could keep 50% of the difference between the retail price and the maximum retail price. See FOR-2001-12-17-1537 §12-3. Generic alternatives received the same maximum pharmacy purchase price as the original manufacturer. The difference was calculated from the product's maximum wholesale price with the maximum retail markup and the actual retail price, which was also subject to the maximum markup rule. Between 2003 and 2005 eight active ingredients were subject to IRP (called the index price). These active ingredients are excluded from our estimation sample for the Norwegian 2005 reform; for a review of the index price system, see Brekke, Grasdal, and Holmås (2009) and Brekke, Holmas, and Straume (2011).

71. The Norwegian Medicines Agency determined when generic entry has taken place. In practice, it requires that the generic product be available in pharmacies.

Table 9: The Step Price Schedule

Starting from	Step-Price Calculation					
01/01/2005	**<100 Mill. NOK 12 months before**		**>= 100 Mill. NOK 12 months before**			
	1. Generic competition	-30%	1. Generic competition	-30%		
	2. 6 months after	-40%	2. 6 months after	-50%		
	3. 12 months after	-50%	3. 12 months after	-70%		
01/01/2007	**<100 Mill. NOK 12 months before**		**>= 100 Mill. NOK 12 months before**		**Cut rate**	
	1. Generic competition	-30%	1. Generic competition	-30%	**Simvastatin**	-85%
	2. 12 months after	-55%	2. 12 months after	-75%		
	3. Final cut if sales >100 Mill. NOK	-85%	3. Final cut if sales >100 Mill. NOK	-85%		

Notes: This table provides the two first Stepped Price rules from Norway. The starting price for calculating the Stepped-Price is the price cap of the original at the start of generic competition. See FOR-2004-12-17-1712 and FOR-2006-12-01-1327 for further details.

first price cut occurs at the beginning of generic competition, followed by further cuts after 6 months and 12 months.[72] The magnitude of the price cuts is related to the total sales prior to generic entry: During our sample period, the first price cut was -30%, the second between 40–50%, and the third between 50–70%.[73] The Step-Price acts as a reference price whose future development is known and fixed by the government. The reform also required pharmacies to keep at least one product in stock at or below the reference price.

Step Price-IRP Schedule. Table 9 shows how the SP regulation worked during our observation period. SP regulation uses predetermined rules to set the price where reimbursement is paid, instead of competition determining the reimbursement price. The price formulas for SP regulation start from the onset of generic competition, and the formula depends on the size of the market before the generic competition started. Table 9 also shows that the steps of price decreases change over time. In the price formulas valid from January 1, 2005, the largest price decrease was 70% but this was changed to 85% in formulas starting January 1, 2007.

A.3.5 Sweden

Nov. 2002–2009: GS and IRP. Sweden adopted GS in November 2002.[74] The system included IRP from the beginning by requiring pharmacies to substitute for the cheapest substitutable product available. Unlike other Nordic countries, patients were reimbursed

72. Appendix section A.3.4 shows the price cut timing in the Step-Price system.

73. See FOR-2004-12-17-1712.

74. See Lag (2002:160) om läkemedelsförmåner m.m. §21. Before 2002, Sweden used IRP without GS. In practice, this meant that the government issued mandatory price decreases as a function of the lowest price of substitutable products.

only for the prescribed product or the product to which the pharmacy offered substitution: This means that if a patient wanted to buy another product (without the decision of the prescribing physician), she would pay the full price (not the price difference between the chosen and the cheapest product) out of pocket. A notable factor in the Swedish GS system was the fact that all pharmacies in the country were operated at the time by the government-owned monopoly Apotek Ab until 2009, when the pharmacy sector was liberalized.

Dec. 2009– : Auction-IRP. Following the liberalization of the pharmacy sector in 2009, a new interpretation of the law was adopted: The cheapest product would be determined at the national level. This led to the establishment of the current "Product of the Month Auction" system, where pharmaceutical manufacturers issue monthly prices (bids) within a given package size and a substitution group. Winners are called products of the month. Consumers can in practice only choose between the prescribed product and the product of the month, although for the first two weeks of each month, the legislation allows pharmacies to also substitute to the winning product of the previous month. The winner and the previous winner thus have high market shares. The government also declares secondary and third alternatives to the winner in case the initial winner has problems in supplying the market.

During our sample period, Sweden also implemented minor price regulation reforms that are related to price caps and the mechanics of the Auction-IRP system. Price caps were subject to one-time cuts in 2009, and later price cap rules within substitution groups were changed.[75] The Auction-IRP system was reformed in 2011 by redefining substitutable products, and in 2012 the backup winners were included in the regulation. These minor reforms are explained in more detail in Appendix A.3.6.

Auction-IRP timing. Figure 9 shows how auction timing works in the Swedish 2009 Auction-IRP reform. In the Auction-IRP system, bids for prices are submitted before they become effective. If a bid is submitted during Month 1, the bid is revealed to all participants during Month 2, and the price is effective during Month 3. Another important feature of the timing of the Auction-IRP is that winning the auction provides benefits only for one month at a time. Regulation allows the previous month's winning product to be dispensed two weeks into the next month. This is represented by the curly brackets denoting the effective prices in Figure 9.

75. It important to note, that the Swedish Pharmaceutical industry proposed the 2009 price cut to regulator.

Figure 9: Auction-IRP Timing

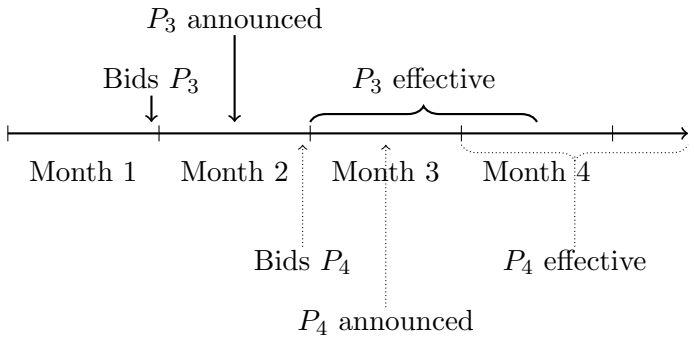


Table 10: Minor Price Regulation in the Nordics 2001-2012

Country	Year	Reform Type	Studied Reform(s)
Denmark	2000–	Reimbursement system overhaul	Denmark 2000
Denmark	2001–	Price Caps	Denmark 2000
Norway	2003–2004	IRP for 8 active ingredients	Norway 2005
Finland	2006	5% Price cap cut for reimbursed products	Denmark 2005, Norway 2005
Finland	2006–	Price cap rule for generic entrants	Norway 2005, Finland 2009
Sweden	2009	Mandatory price cap cut	Sweden 2009
Sweden	2011–	Mandatory price caps in substitution groups	Sweden 2009
Sweden	2011–	Redefinition of substitutable products in Auction-IRP	Sweden 2009
Sweden	2012	Back-up winners in Auction-IRP system	Sweden 2009

Notes: Country = Country where the minor reform happened; Year = When the minor reform happened; Reform Type = Minor reform type; Studied Reform(s) = Reforms that are studied in the paper, where the minor reform happens during the sample period.

A.3.6 Minor Price Regulation Reforms

During the periods that our estimation samples cover, Nordic countries implemented reforms that we categorize as minor. These reforms create changes, e.g., in the way pharmaceuticals are priced and reimbursed.[76] We have collected the minor reforms into Table 10: There are two minor reforms in Denmark, two in Finland, one in Norway, and four in Sweden during our observation periods.

Denmark. During our sample periods, the Danish regulator made price cap or price "freeze" agreements with pharmaceutical firms represented by the Danish Association of the

76. Changes in the reimbursement rates, reimbursement ceilings and OTC deregulation policies (pricing and distribution) are excluded from the table. OTC-deregulation policies are excluded because we study prescription drugs.

Pharmaceutical Industry (LIF). Not all firms present in the Danish market are represented by LIF, and this leads to a situation where price caps were not imposed on all products. In these agreements, the Danish regulator and the firms agree that a market price from a certain date acts as the price cap for a period of time. These price agreements were in place during our observation periods.[77]

In March 2000, the Danish government adopted a new reimbursement model where the fixed product-specific reimbursement level was replaced by a system where the patients' reimbursement level was non-linearly calculated based on spending (see Simonsen, Skipper, Skipper, and Christensen 2021). This reimbursement system change happens in the pre-period of the Danish 2000 reform. The reimbursement reform gave incentives to persons who already exceeded their annual pharmaceutical cost limit to stock pharmaceuticals, because after the reform they faced 100% coinsurance. We see this effect as a pre-period increase in Figure 4a when the outcome studied is the market size in DDD. The change in the reimbursement system does not affect average expenditure because pricing did not respond to the change.

Finland. In 2006 Finland implemented two minor reforms related to pharmaceutical pricing. The first reform was a 5% price cap cut for reimbursed pharmaceuticals, and the second was the price cap rule for generic products. The price cap cut reduced the maximum price of the reimbursed product and led to a decrease in wholesale and retail prices for the products for which the price cap was binding.[78] These price cuts can indirectly influence the evaluation of the Danish 2005 ERP reform and the Norwegian SP reform, because we use Finland as the control group. We deal with this issue by constraining the sample period to the time before the price cut. In the Appendix section A.10, we present results where the sample period is not constrained by the price cut.

The second Finnish reform in summer 2006 was the formalization of how price caps of the generic entrants are calculated when the first generic product enters the Finnish market. This reform formalized that generic products are accepted into the reimbursement system only if they are priced at least 40% lower than the cap of the originator product. If a company does not accept this proposed cap, the product can enter the market, but it is not eligible for public reimbursement. This regulation change does not complicate

77. Price cap agreement signed on 19.3.2019 states that the first price cap agreement was signed in 2006, but WP version of Kaiser, Mendez, Rønde, and Ullrich 2014 mentions that the price agreements between LIF and the Danish government were already implemented in 2001.

78. See 885/2005 for additional details.

our empirical analyses like the implemented price cut, because the markets we study had generic entry before our observation period.

Norway. The only minor change in price regulation in Norway during our sample period was the IRP-experiment (Index-Price) for eight active ingredients (=ATC5 categories). This policy was in place 2003–2004. The Index-Price policy was an IRP variant similar to the Finnish 2009 policy.[79] This means that the Index-Price policy change occurs during the pre-period of the SP reform. To ensure that all treated markets have the same pre-period regulation regime, we discard the markets where index-price regulation was implemented. Brekke, Grasdal, and Holmås (2009) report that the Index-Price policy was shut down because the policy did not achieve the desired amount of cost savings and price reductions.

Sweden. Minor Swedish reforms during our sample period are related to the implementation of the price cap and minor changes in the contents of the Auction-IRP regulation. Sweden introduced a mandatory one-time price cap cut for originator products in markets with substitutable products and generic competition in 2009.[80] The unique feature of this price cut is that it was proposed by Läkemedelsindustriföreningen (Lif), the trade association for the research-based pharmaceutical industry in Sweden.[81] The cut was a 65% decrease from the price of the originator product that prevailed 12 months before the expiration of the patent.[82] The price cap decrease was planned so that after the price cut the originator price cannot be lower than the cheapest comparable generic product. According to the Swedish National Board of Health and Welfare, the price cut decreased annual expenditure by 400 million SEK, or approximately 40 million euros.[83]

79. Index price at GIP (producer price) level was calculated as the total turnover value for all products in the index price group for the period, divided by the total quantity sold during the period. The index price was determined at the producer level (GIP), to which a 10% maximum profit was added for the benefit of the wholesalers. The final index price was obtained by adding the maximum pharmacy mark-up to the index price at the PPP (pharmacy purchase price) level. The final index prices were in PRP (pharmacy retail price). See Brekke, Holmas, and Straume (2011) and Brekke, Grasdal, and Holmås (2009) for more details.

80. See Price cut announcement for additional details.

81. The Swedish price cut resembles the Danish price freeze agreements that are based on the negotiations between pharmaceutical industry and the government.

82. For products that experienced patent expiration before October 2002, the price cut is either calculated from the price that was applied on September 2001 or price that was applied 12 months before the patent expiration.

83. See Läkemedelsförsäljningen i Sverige – analys och prognos for additional details. The problem with this cost savings estimate is that the price reduction occurs almost simultaneously with the implementation of Auction-IRP.

The price cap regulation was changed in 2011 and the regulation contains two phases.[84] In the first phase, generic competition has not started within a substitution group and the price cap is defined as the maximum price in the substitution group. This price cap is defined as the initial price cap. In the second phase, price cap decreases are triggered by (generic) competition. Mandatory price caps were imposed if four months had passed since generic competition had started in the substitution group and at least one product within the substitution group is priced 30% lower than the initial price cap. When these conditions are met, the price cap for all products in the substitution group is reduced by 35% of the initial price cap. This regulation change means that a decrease in the price of one product triggers a decrease in the price cap for all products in the substitution group.

In 2011 Sweden changed the definition of substitutable pharmaceutical products to the package level.[85] Prior to this change, substitution would be made to the cheapest package within the substitution group, which favored small packages. After the definition change Auction-IRP would be considered within the same package sizes of the substitution groups set by the Swedish Medicinal Products Agency.

In 2012 Sweden changed the Auction-IRP regulation to allow multiple winners in the auction.[86] The reason for the change was to allow pharmacies to substitute to the backup products (second or third cheapest product in the auction) if the auction winner has problems supplying the market.

A.4 Data Sources and Sample Statistics

Data sources. Our data sources are detailed in Table A.4.

Number of observations. Our sample sizes are detailed in Table 13. Panel A displays market-level statistics by reform, and Panel B displays the same for product-level outcomes. The samples are balanced by design for market-level analyses. Product-level sample sizes are well balanced between the treatment and control groups in different reforms with two exceptions. In the Norwegian 2005 reform, the control group is much larger than the treatment group, and in the Swedish 2009 reform, the treatment group is larger than the control group.

84. TLVFS 2009:4
85. TLVFS 2009:4
86. TLVFS 2009:5

Table 11: Sales Data Coverage

	Years	Source
Panel A: Sales Data		
Finland	1998–2017	FIMEA
Sweden	2006Q2–2017	IQVIA
Denmark	1991–2017	DLI-MI
Norway	2000–2018	Farmastat
Panel B: Reform Data		
2000 Denmark	1999–2005	Legislation
2003 Finland	2003–2009	FIMEA+Legislation
2005 Denmark	2003–2007	Legislation
2005 Norway	2003–2007	NOMA+Legislation
2009 Finland	2009–2015	PPB+Legislation
2009 Sweden	2005–2013	TLV+Legislation

Notes: FIMEA = Finnish Medicines Agency; PPB = (Finnish) Pharmaceutical Pricing Board; NOMA = Norwegian Medicines Agency; TLV = (Swedish) Dental and Pharmaceutical Benefits Agency.

Table 12: Package-level Matching Rates

	Treatment (1)	Control (2)	Union (3)	Union-% w.r.t treatment (4)
Finland 2003	1654	1936	369	22.31
Finland 2009	2393	1393	392	16.38
Denmark 2000	1551	1098	250	16.12
Denmark 2005	2183	2146	454	20.80
Norway 2005	331	484	93	28.10
Sweden 2009	2962	1626	493	16.64

[1] This table lists the package level match rates between the treatment and control countries in all estimations. Outcome data source: DLI-MI (1999–2013), Farmastat (2004–2013), Fimea (1999–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

Table 13: Number of Observations and Clusters

	Part I		Part II		Part III	
	Finland 2003	Finland 2009	Denmark 2000	Denmark 2005	Norway 2005	Sweden 2009
Panel A: Market Level						
Number of Observations	5843	7590	2842	6716	1110	12824
Number of Clusters	80	106	59	118	15	123
Panel B: Product Level						
Number of Observations	96384	109744	59152	118949	24780	200232
Number of Clusters	82	109	62	129	15	129
Panel C: Market Level (Spillover and Monopoly)						
Number of Observations	4040	4555	24153	29654	3014	8126
Number of Clusters	73	74	727	688	44	121
Panel D: Product Level (Spillover and Monopoly)						
Number of Observations	19739	16739	122473	135892	37684	64901
Number of Clusters	75	76	968	930	44	139

[1] This table presents the number of observations and number of bootstrap cluster by each estimation (reform).

[2] Panel A gives market level statistics from Average Expenditure estimations. Panel B gives product level statistics from Wholesale Price estimations. Other outcomes might have slightly different values due to missing values.

[3] Outcome data source: DLI-MI (1999–2013), Farmastat (2004–2013), Fimea (1999–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

A.4.1 Additional Descriptive Statistics

Share of identical products in reform comparisons. Nordic countries that use ERP-policies include other Nordic countries in their ERP-baskets, and this can facilitate regulation spillovers or externalities between treatment and control countries. In Table 12 we calculate how large a share of products (packages) sold in the treatment country is also sold in the control country.

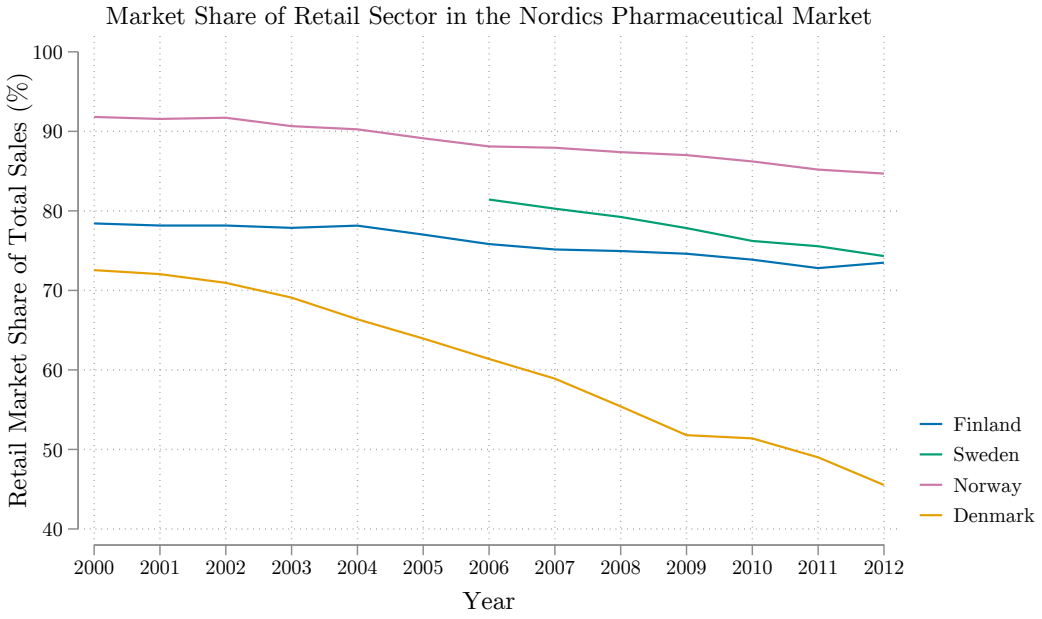
Table 12 shows the number of unique packages and the number of identical packages in estimation samples by each reform during the reform pre-period. Column 1 shows the unique number of products in the treatment country, and column 2 shows the same for the control country. Column 3 displays the number of identical unique packages that are found both from treatment and control countries, and Column 4 shows how large share of the treatment country packages are present in both countries during pre-period. The overlap between products being sold in both countries during the pre-period varies between 16% and 28%. ERP-policies used in the Nordics compare prices at the package-level and small product overlap means that ERP is not likely to invalidate our cross-country research design.

Role of the hospital market. Pharmaceuticals are distributed through pharmacies and hospitals in the Nordic countries. We concentrate on the pharmacy market: Figure 10 shows the share of pharmaceuticals sold through pharmacies (shares are calculated using wholesale prices).[87] We find that the share of pharmaceuticals distributed through pharmacies has been quite stable in Finland, Sweden, and Norway during our observation period. However, in Denmark the share of pharmaceuticals distributed through pharmacies decreased during our observation period from around 70% to less than 50%.

A large hospital share of pharmaceutical sales can be problematic in our cross-country matching procedure because it is possible that a given ATC5 market in Denmark has only hospital market sales, leading to unmatched markets. The difference between Columns 4 and 5 of Table 2 in the main text illustrates the number (Panel A) and economic significance (Panel B) of unmatched markets. All comparisons in which Denmark is used as a control group have unmatched markets, but the economic significance of these markets in the

87. The Nordic hospital pharmaceutical market works through competitive bidding. Unfortunately, we do not have access to bids and therefore we need to rely on wholesale prices while calculating market shares. This leads to a situation where the market shares presented in Figure 10 are the upper bound of the actual market share.

Figure 10: Aggregate Pharmacy Market Share



The development of the retail market share in the Nordics pharmaceutical market from 2000 to 2012.
Data source: DLI-MI, Farmastat, Fimea, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2000–2012).

treated country is small (1%–2% of the sales of the pharmacy market).[88]

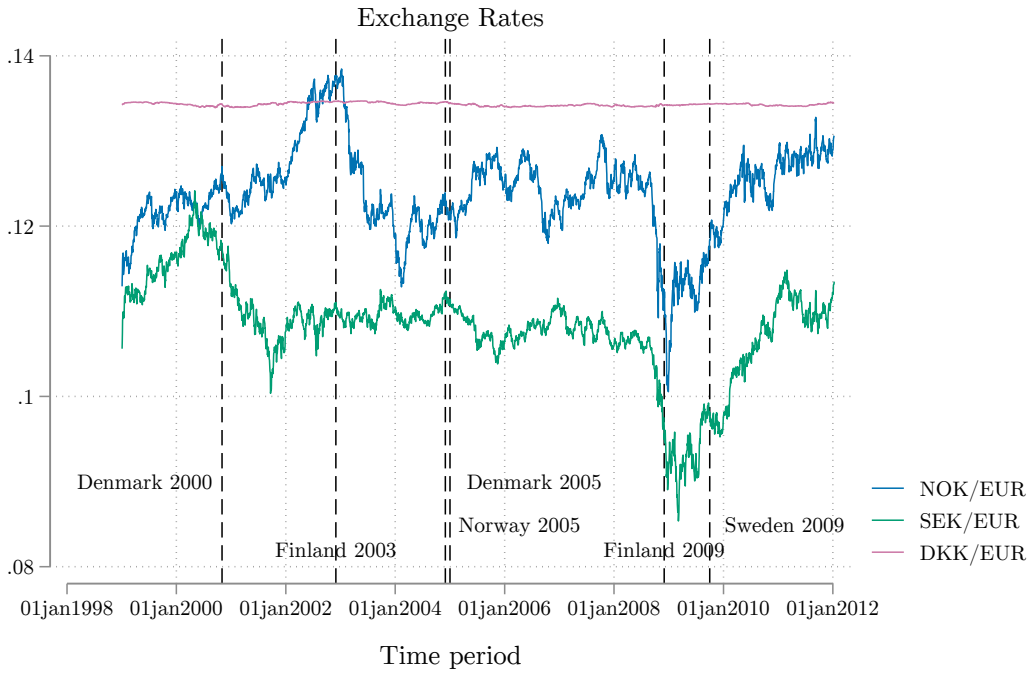
A.5 Exchange Rate Shocks

We use domestic currencies in our analyses. The rationale for this is that sudden changes in exchange rates can bias our results. This is illustrated in Figure 11 which plots the NOK–EUR, SEK–EUR and DKK–EUR exchange rates and the start dates of the reforms we study.

Figure 11 shows that the DKK–EUR exchange rate evolves differently. This follows from the fact that during the study period, the Danish Krone (DKK–EUR) is linked to the Euro. It is evident from the figure that some reforms start close to sudden and extreme changes in the exchange rate, such as the Sweden 2009 reform. The 2009 fluctuations in exchange rates were induced by the financial crisis. If the analyses were done using outcomes converted to the same currency, the exchange rate movements would influence

88. Overall the economic significance of unmatched markets is small with respect to all reforms we study.

Figure 11: Exchange Rate Shocks



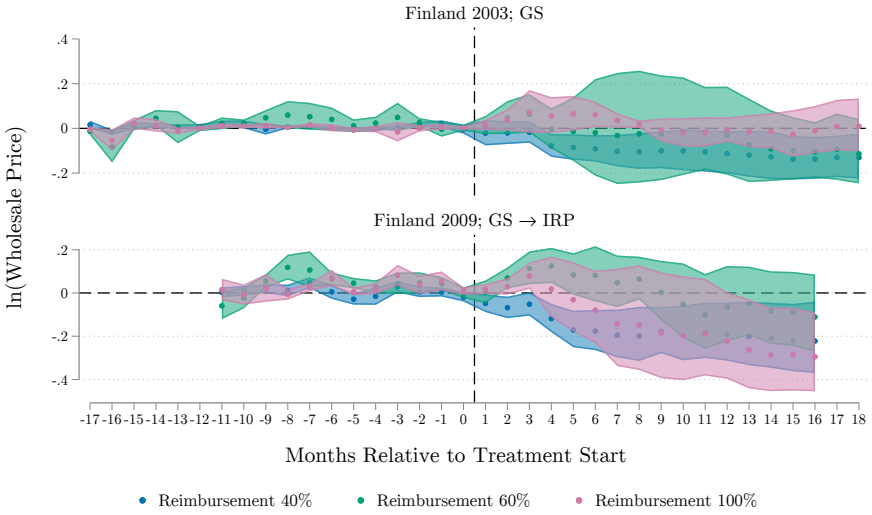
Data source: European Central Bank.

the results because our differences-in-differences specifications cannot be augmented to separate exchange rate movements from the reform effects.

A.6 Reimbursement Rates and Finnish 2003 and 2009 Reforms

The main text presented results for the Finnish 2003 and 2009 reforms. The results showed quite clearly that the 2009 IRP reform was much more effective than the GS reform of 2003 in reducing pharmaceutical expenditure. The main explanation for this difference is that in the 2003 reform consumer choices did not influence the reimbursement consumer received. This meant that a consumer with full reimbursement (100%) has no incentive to substitute to cheaper products. 2009 reform tied consumer reimbursement to the cheapest products in the substitution group, giving consumers an additional incentive to substitute to cheaper products. In this subsection, we examine how the effects of regulation depend on the reimbursement rate.

Figure 12: Finland 2003 and 2009 by Reimbursement Status



Estimator: Callaway and Sant'Anna (2020). 999 replications for ATC-5 wild bootstrapped standard errors. Denmark used as a control group. Outcome data source: DLI-MI and FIMEA.

The upper part of Figure 12 shows the results for the 2003 GS reform and the lower part for the 2009 GS-IRP reform. Both panels present results for three reimbursement sub-samples (40%), 60%) and 100%).[89]

The top panel of Figure 12 clearly shows that the negative price effect is driven by products with the 40% reimbursement rate, because the treatment effect for higher rates is zero. These results help to rationalize why the 2003 reform delivered only modest savings. Average market-specific spending per DDD did not decrease much because product prices did not respond to the reform in all reimbursement categories. The bottom panel of Figure 12 shows that in the 2009 IRP reform, all reimbursement categories show decreasing prices due to the implemented reform. These price results are also in line with the expenditure results shown along the main results. Expenditure per DDD decreased substantially with the IRP reform, and an explanation for the decrease is that the average price in all categories decreased due to the reform.

89. These sub sample regressions are estimated using product specific data instead of market level data as in the main analysis. This change helps to show whether incentives related to reimbursements explain the differences between the two reforms or not.

A.7 Denmark 2000 and 2005: Monopoly Markets

The main analysis showed results on Danish 2000 and 2005 reforms for competitive markets. These reforms were designed in a way that reform can also influence non-competitive markets, and this subsection presents the effects of ERP switches on monopoly markets. The analysis of monopoly markets is an important addition to the discussion of how ERP-like regulatory measures work. The structure of the analysis and sample matching is the same as before; the only change is that now the focus is on markets where (generic) competition has not started yet. The results are displayed in an event study format and are also summarized as simple ATT measures.

Event studies presented in Figures 13a and 13b show that the ERP policy changes had some short-term effects on average expenditure and no effect on pharmaceutical availability. The results follow the same patterns as the results for the competitive markets in the main text. The only difference between competitive and monopoly market results is that the monopoly results are more imprecise, and the effect size seems to decrease in absolute terms over time during the follow up period. Table 14 shows that on average average expenditure and average price decreased both by -4% during the reform of 2000 and average expenditure and prices increased statistically insignificantly by 1% and 2% during the reform of 2005.

The main takeaway from the results presented in this subsection is that ERP policies have the ability to influence the pricing and sales of pharmaceuticals during the period when the market is not subject to generic competition. This means that ERP can be used to augment simple price cap regulation when price competition-based regulation cannot be used. However, it is important to note that implementation of ERP policy could also have adverse effects on reference countries, because firms could have an incentive to increase prices or delay entry in order to dilute the benefits of using ERP (Dubois, Gandhi, and Vasserman 2022; Maini and Pammolli 2022). The results from a (small) Nordic country might not be directly applicable to a larger country because it is possible that ERP implementation in a small geographical market might not cause large adverse effects compared to a situation where ERP is implemented in larger European countries and markets.

A.8 Denmark 2005: Additional Results

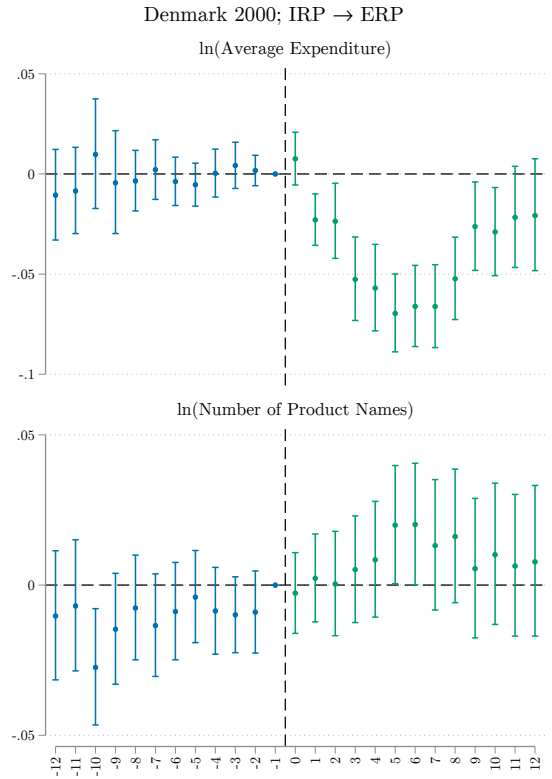
Denmark 2005 switch from ERP to IRP has already been studied in the earlier literature, but our results are quantitatively different from what Kaiser, Mendez, Rønde, and Ullrich 2014 finds. These results imply that the change from ERP to IRP decreases the prices,

Table 14: Average Treatment Effects (Monopoly Samples)

	Part II	
	Denmark 2000	Denmark 2005
Panel A: Main Estimations		
Average Expenditure	-0.04*	0.01
	[-0.06, -0.02]	[-0.00, 0.03]
Number of Product Names	0.02	-0.00
	[0.00, 0.04]	[-0.02, 0.02]
Average Price	-0.04*	0.02
	[-0.06, -0.02]	[0.00, 0.04]
Number of Doses	-0.04	0.00
	[-0.08, 0.01]	[-0.04, 0.05]
Wholesale Price	-0.06*	0.02*
	[-0.07, -0.04]	[0.01, 0.03]
Panel B: Weighted Estimations		
Average Expenditure	-0.04*	0.01
	[-0.06, -0.02]	[-0.00, 0.03]
Number of Product Names	0.02	-0.00
	[0.00, 0.04]	[-0.02, 0.02]
Average Price	-0.04*	0.02
	[-0.06, -0.02]	[0.00, 0.04]
Number of Doses	-0.04	0.00
	[-0.08, 0.01]	[-0.04, 0.05]
Wholesale Price	-0.06*	0.02*
	[-0.07, -0.04]	[0.01, 0.03]

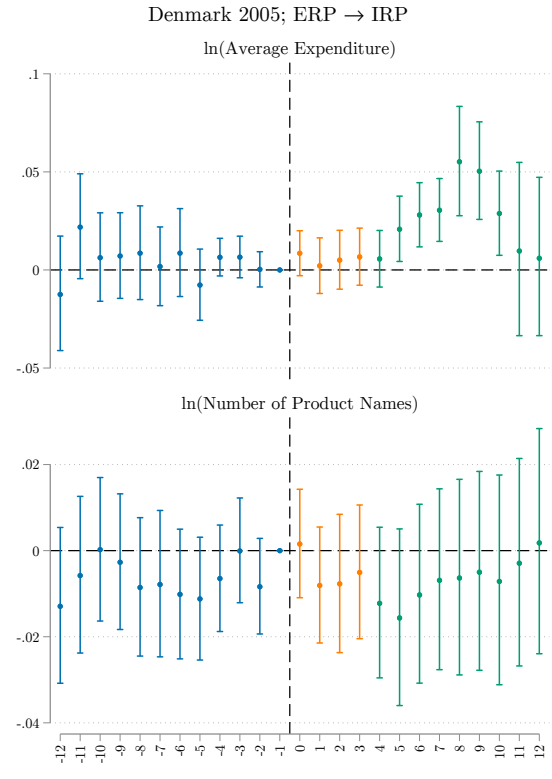
[1] Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). Outcome data source: DLI-MI (1999–2006) and Fimea (1999–2006).

[2] * = statistically significant at the 95% confidence level. 10000 replications for ATC-5 wild bootstrapped standard errors.



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(a) Denmark 2000



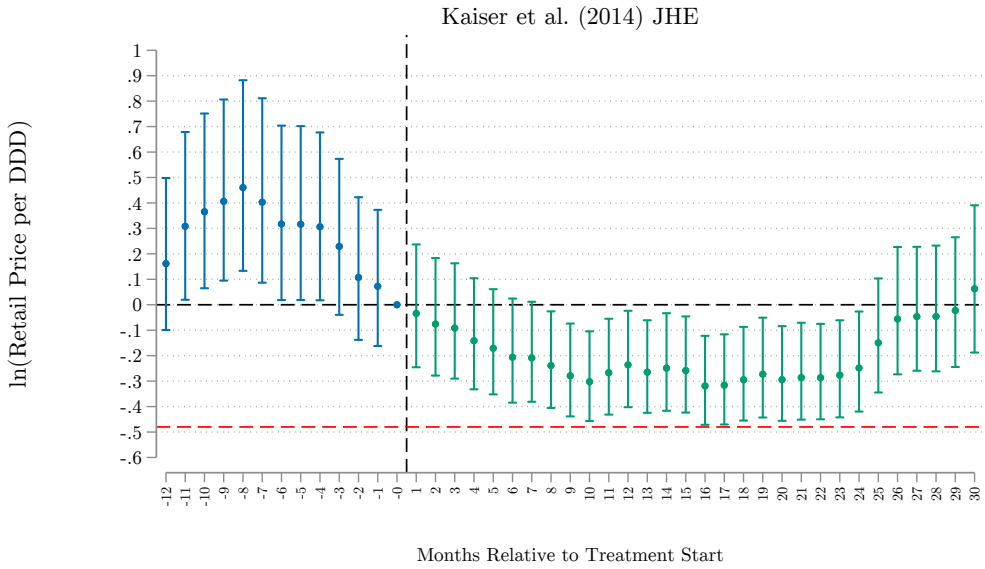
Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(b) Denmark 2005

Figure 13: Part II: Monopoly Markets

but our results show that the prices, in fact, increase as a result of the change. Figure 14 illustrates why the negative price effect result does not identify the true effect of the reform.

Figure 14: Replication of Kaiser, Mendez, Rønde, and Ullrich (2014)



Regression 1: ln(Retail Price per DDD) on reform dummy and 12 seasonal monthly dummies.
The red dashed line indicates the ATT.
Regression 2: ln(price) on all month dummies.

Figure 14 replicates the price analysis in Kaiser, Mendez, Rønde, and Ullrich (2014) using the same markets and the first regression specification presented in the article[90]. This specification regresses log prices on seasonal monthly dummies and a post-reform dummy. Danish 2005 reform influenced all Danish markets, so it is not possible to construct control groups within the country, and the authors can only use a simple differences-estimator. The coefficient on the reform dummy is plotted as the red dashed line in Figure 14. The figure also displays the coefficient of a more flexible specification, where the reform dummy and seasonal monthly dummies are replaced with calendar-time fixed effects. These fixed effects are plotted in the figure with confidence intervals. There is already a clear decreasing price trend in the Danish statin market before the implementation of the reform (dashed vertical line). This suggests that the price results reported in Kaiser, Mendez, Rønde, and

90. See Table 4, column one on page 180.

Ullrich (2014) are a mixture of actual reform effects and competition effects unrelated with the reform.

A.9 Finland 2009 and Sweden 2009 with Alternative Control Groups

In Section 2 we show how reform timings go in the Nordic pharmaceutical market by each country. From Figure 2 we see that for most reforms there is only a one country that can serve as the control group, but Finnish and Swedish 2009 reforms have also the possibility to use Denmark as the control group.[91] In this subsection we illustrate how robust our unweighted and weighted main results are with respect to the used control country. We summarize our results by estimating ATTs (see Table 15) and we illustrate how reform effects evolve over time by estimating event study regressions.

We find that main results presented in Table 15 are qualitatively the same irregardless of the used control group. There are some differences in estimate sizes, but almost in all cases the point estimates from the model with the alternative control group fall within confidence intervals of the original estimates. The most notable exception is Finnish 2009 reform (presented in Figure 15) where the unweighted results using Denmark as a control group yield larger effect size as the main results when Average Expenditure is used as the outcome. When results are weighted (Figure 16), we see that the difference between main results and results using the alternative control group is not as large as with the unweighted results.

A.10 Extended Sample Period for Denmark 2005 and Norway 2005 Reforms

In the main text, we showed results for Denmark and Norway 2005 reforms with a short post-reform period. The reason for this choice was the price cut implemented in the control country Finland in January 2006. This shock in the control country cannot be "controlled away" in our framework, and the shock directly influences our results. Now in figures 23a and 24b we show event study results in a longer post-reform time window. Column A shows results for the main outcomes, and Column B for the spillover (Norway 2005) or monopoly (Denmark 2005) results. Solid-green event study estimates represent the results already

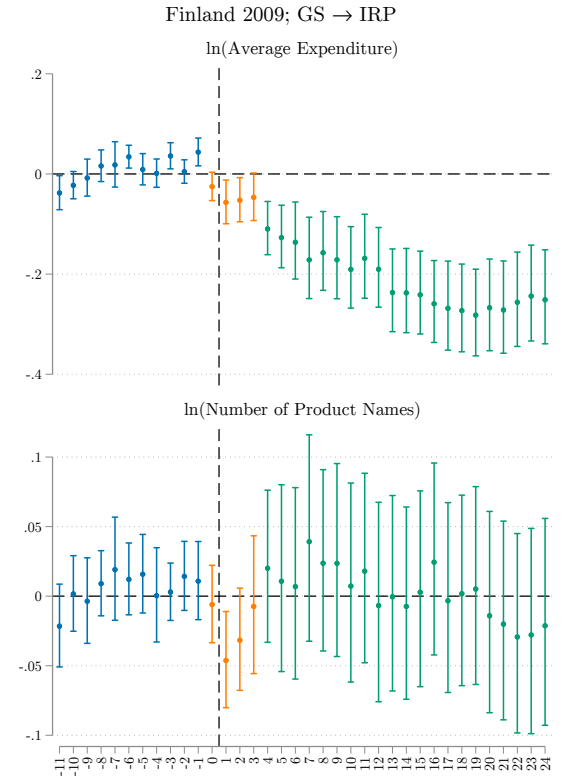
91. We could also use Sweden as a control group for Finland in the 2009 reform for one quarter due to the differential reform timing in 2009. We don't perform this analysis, because the follow up period is short compared to other reforms.

Table 15: Average Treatment Effects (Control Group Comparison)

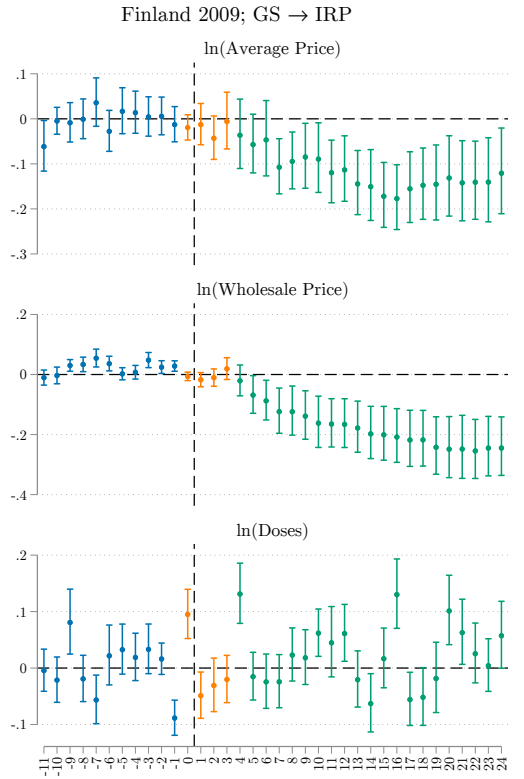
	Finland 2009			Sweden 2009		
	FIN-NOR	FIN-DEN	FIN-NOR-DEN	SWE-NOR	SWE-DEN	SWE-NOR–DEN
Panel A: Main Estimations						
Average Expenditure	-0.19*	-0.13*	-0.15*	-0.29*	-0.27*	-0.29*
	[-0.26, -0.12]	[-0.18, -0.08]	[-0.20, -0.10]	[-0.35, -0.22]	[-0.34, -0.20]	[-0.35, -0.22]
Number of Product Names	-0.00	0.04	0.02	0.03	0.04	0.04*
	[-0.05, 0.05]	[-0.02, 0.10]	[-0.03, 0.08]	[-0.01, 0.08]	[-0.00, 0.09]	[0.01, 0.08]
Average Price	-0.11*	-0.05	-0.07*	-0.12*	-0.04	-0.08*
	[-0.16, -0.04]	[-0.09, -0.00]	[-0.12, -0.02]	[-0.19, -0.05]	[-0.11, 0.04]	[-0.14, -0.01]
Number of Doses	0.02	0.04*	0.03*	0.04	0.12*	0.10*
	[-0.02, 0.05]	[0.01, 0.07]	[0.01, 0.06]	[-0.04, 0.13]	[0.02, 0.22]	[0.01, 0.19]
Wholesale Price	-0.15*	-0.10*	-0.13*	-0.11*	-0.06*	-0.09*
	[-0.21, -0.08]	[-0.14, -0.07]	[-0.17, -0.08]	[-0.17, -0.05]	[-0.11, -0.01]	[-0.14, -0.04]
Panel B: Weighted Estimations						
Average Expenditure	-0.22*	-0.25*	-0.23*	-0.32*	-0.37*	-0.36*
	[-0.35, -0.06]	[-0.34, -0.15]	[-0.31, -0.14]	[-0.45, -0.17]	[-0.49, -0.22]	[-0.48, -0.21]
Number of Product Names	-0.05	0.12	0.07	0.08	0.06	0.08
	[-0.17, 0.08]	[-0.02, 0.29]	[-0.07, 0.21]	[-0.01, 0.17]	[-0.02, 0.15]	[0.00, 0.16]
Average Price	-0.20*	-0.16*	-0.18*	-0.05	0.06	0.00
	[-0.27, -0.12]	[-0.23, -0.08]	[-0.24, -0.12]	[-0.34, 0.37]	[-0.24, 0.50]	[-0.26, 0.37]
Number of Doses	0.07*	0.07*	0.08*	0.04	0.14	0.11
	[0.01, 0.14]	[0.04, 0.10]	[0.02, 0.13]	[-0.10, 0.20]	[-0.05, 0.37]	[-0.06, 0.31]
Wholesale Price	-0.23*	-0.17*	-0.20*	-0.12*	-0.07	-0.10*
	[-0.34, -0.10]	[-0.23, -0.10]	[-0.28, -0.11]	[-0.19, -0.04]	[-0.14, -0.00]	[-0.16, -0.04]

[1] Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). Outcome data source: DLI-MI (1999–2013), Farmastat (2004–2013), Fimea (1999–2012), IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

[2] * = statistically significant at the 95% confidence level. 10000 replications for ATC-5 wild bootstrapped standard errors.

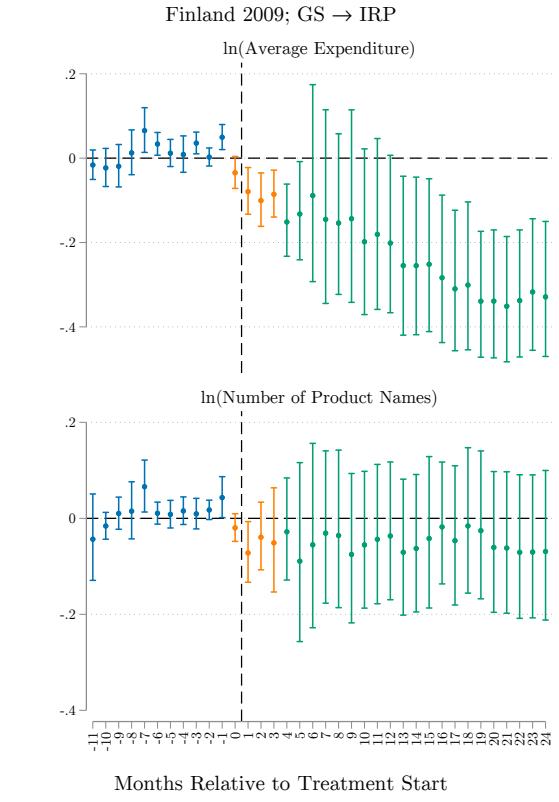


(a) Finland 2009 – Main



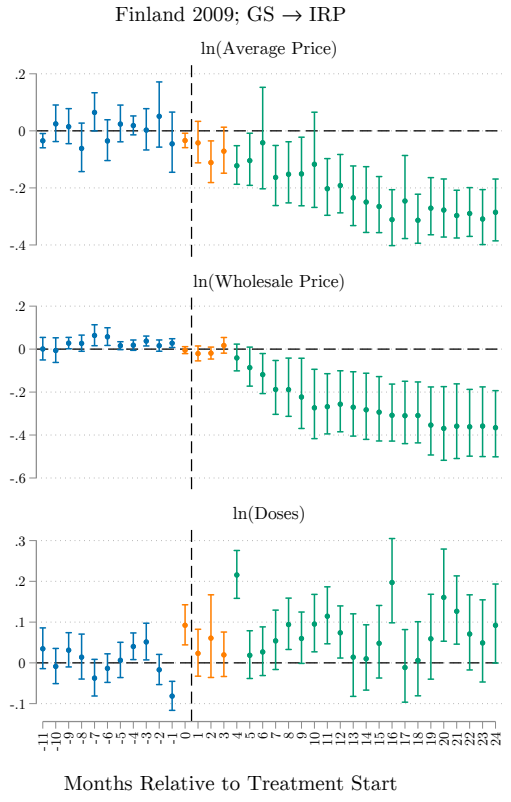
(b) Finland 2009 – Secondary

Figure 15: Finland 2009 Results with Denmark as the control group



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI and Fimea.

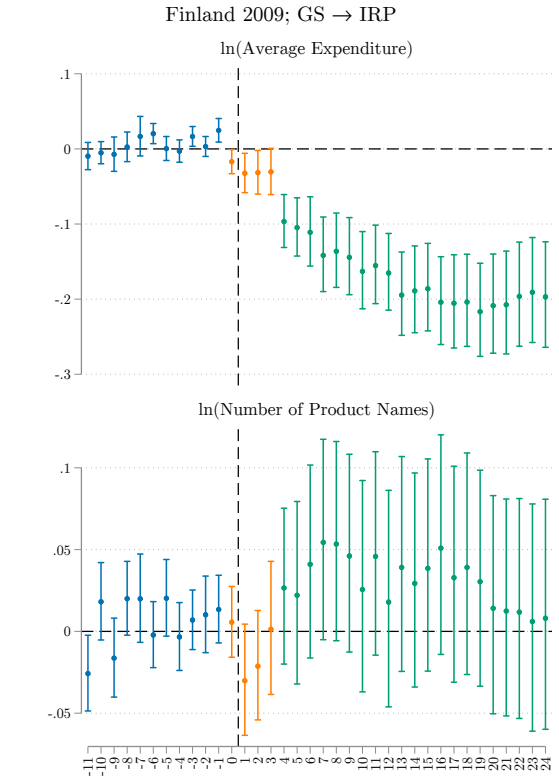
(a) Finland 2009 – Main



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI and Fimea.

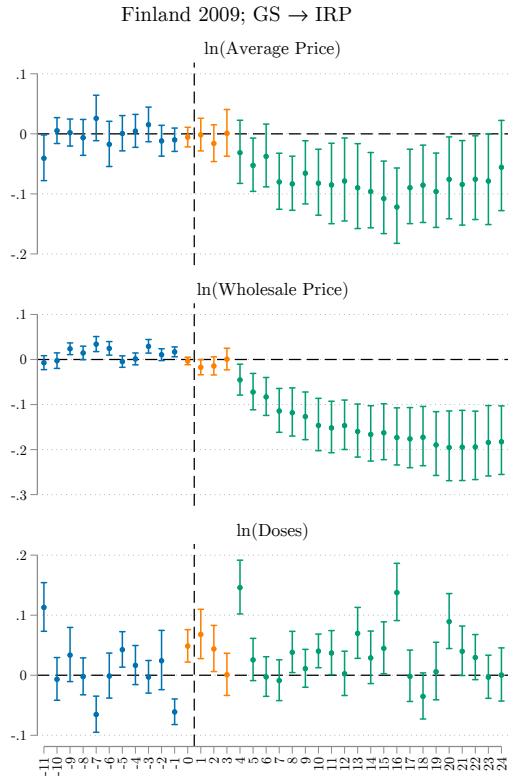
(b) Finland 2009 – Secondary

Figure 16: Finland 2009 Weighted Results with Denmark as the control group



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Denmark Norway used as a control group. Outcome data source: DLI-MI, Farmastat and Fimea.

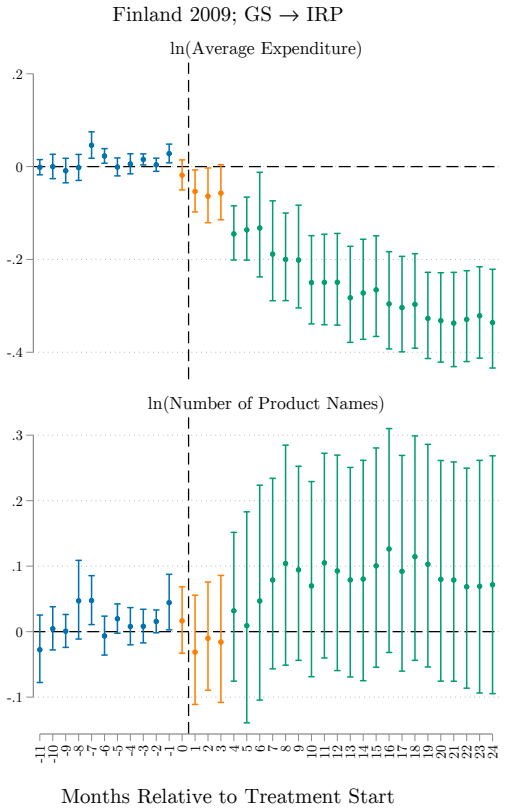
(a) Finland 2009 – Main



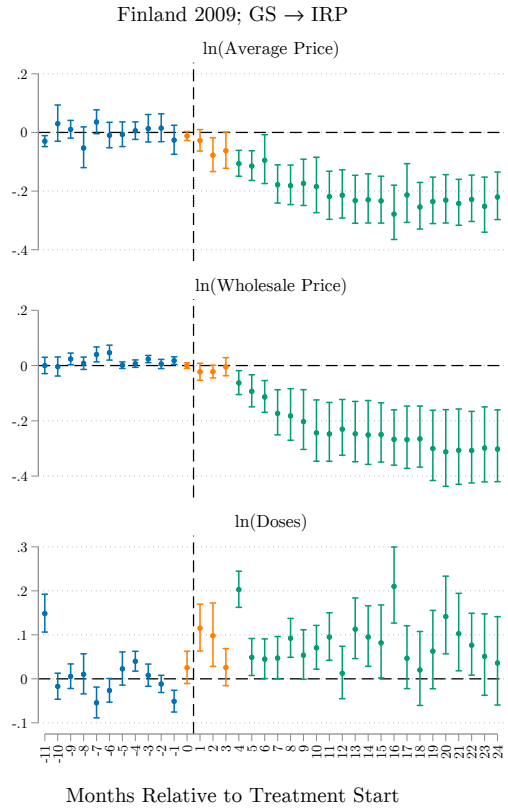
Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Denmark and Norway used as a control group. Outcome data source: DLI-MI, Farmastat and Fimea.

(b) Finland 2009 – Secondary

Figure 17: Finland 2009 Results with Norway and Denmark as the control group

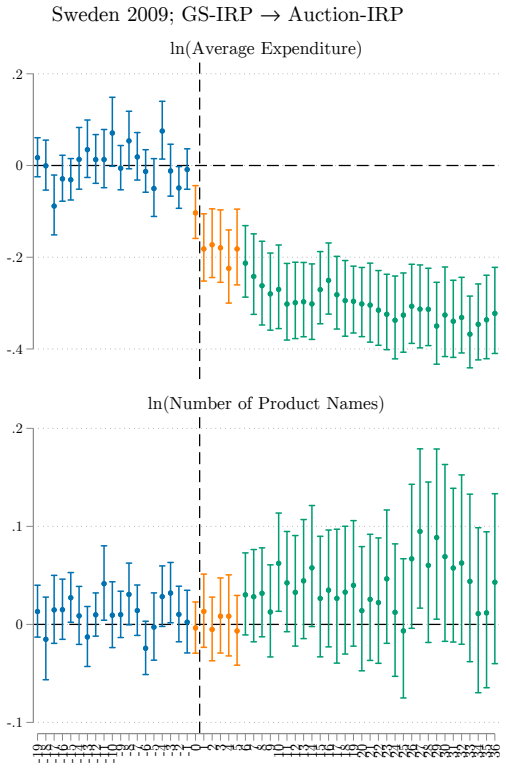


(a) Finland 2009 – Main



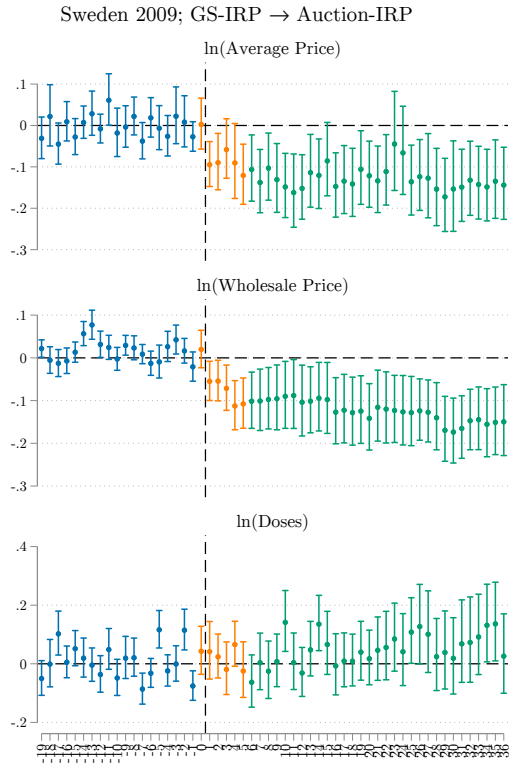
(b) Finland 2009 – Secondary

Figure 18: Finland 2009 Weighted Results with Norway and Denmark as the control group



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Outcome data source: DLI-MI, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

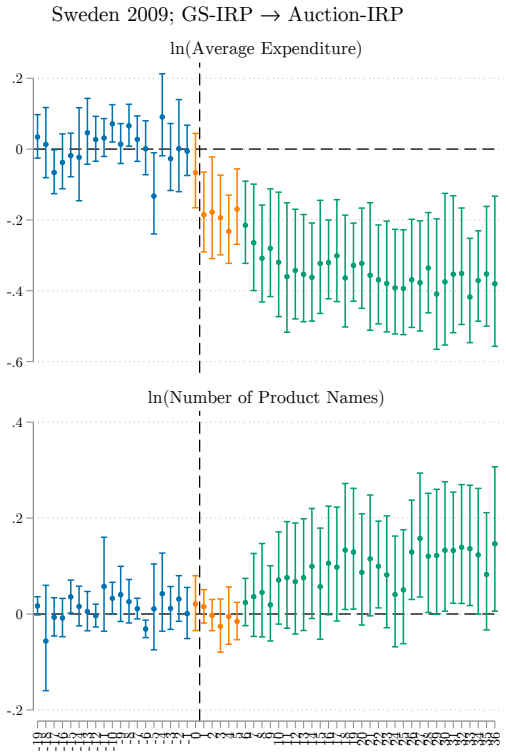
(a) Sweden 2009 – Main



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Outcome data source: DLI-MI, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

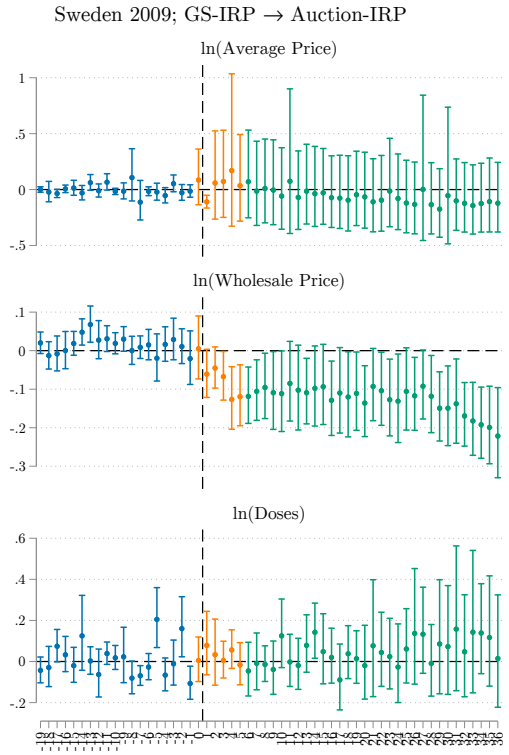
(b) Sweden 2009 – Secondary

Figure 19: Sweden 2009 Results with Denmark as the control group



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

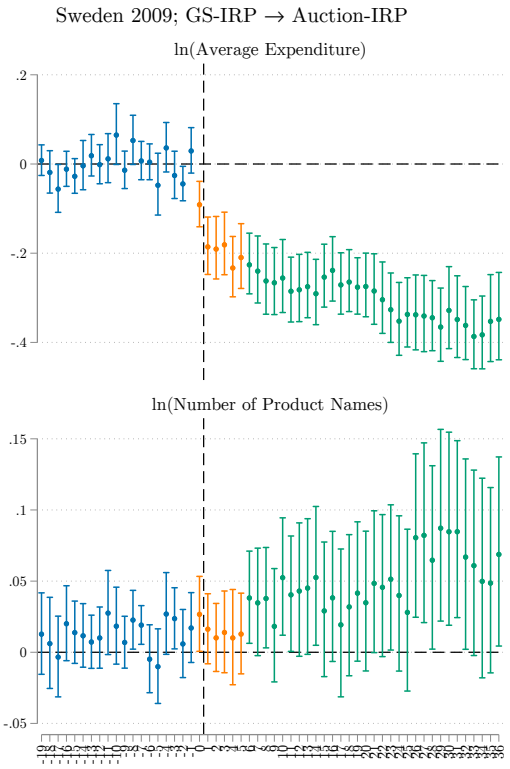
(a) Sweden 2009 – Main



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(b) Sweden 2009 – Secondary

Figure 20: Sweden 2009 Weighted Results with Denmark as the control group



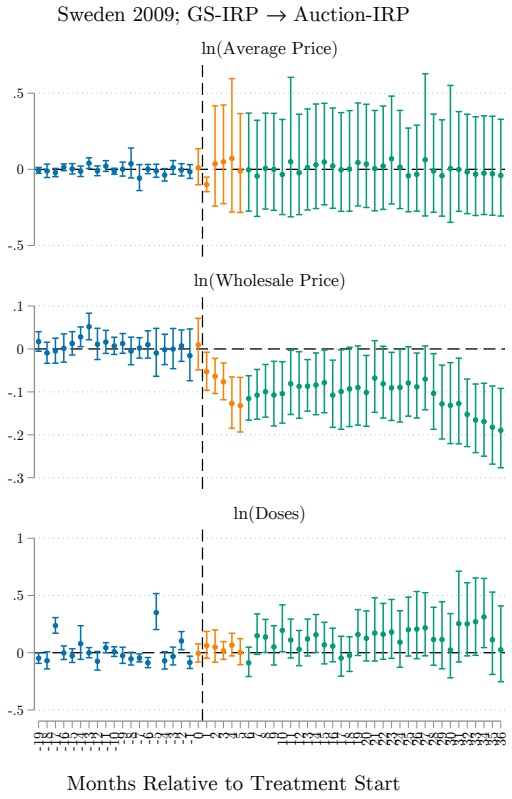
Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange c denotes anticipation period. Denmark and Norway used as a control group. Outcome data source: DLI-MI, Farmastat, MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(a) Sweden 2009 – Main

Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark and Norway used as a control group. Outcome data source: DLI-MI, Farmastat, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

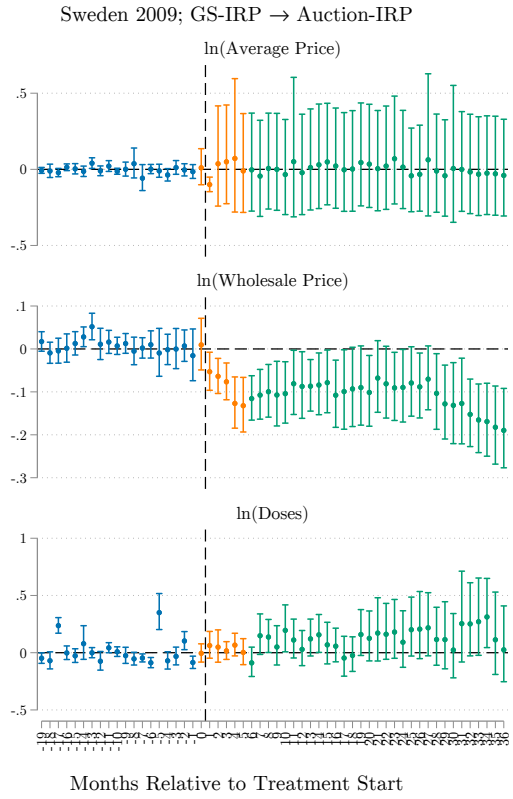
(b) Sweden 2009 – Secondary

Figure 21: Sweden 2009 Results with Norway and Denmark as the control group



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange c denotes anticipation period. Denmark and Norway used as a control group. Each ATC-5 weighted by it's share of tota sample sales in the treatment country. Outcome data source: DLI-MI, Farmastat, IQVIA MIDAS Quarterly Sales and MIDAS (2007–2013).

(a) Sweden 2009 – Main



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark and Norway used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI, Farmastat, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(b) Sweden 2009 – Secondary

Figure 22: Sweden 2009 Weighted Results with Norway and Denmark as the control group

shown in the main text, and the light-green estimates are the time periods added to the study period. The most notable changes in the event study coefficient sizes occur when monopoly markets or spillovers are studied. The reason for this is that these markets are the markets where the price cut had the largest effect on the wholesale price, and therefore the impact of the price cut is seen in the figures.[92]

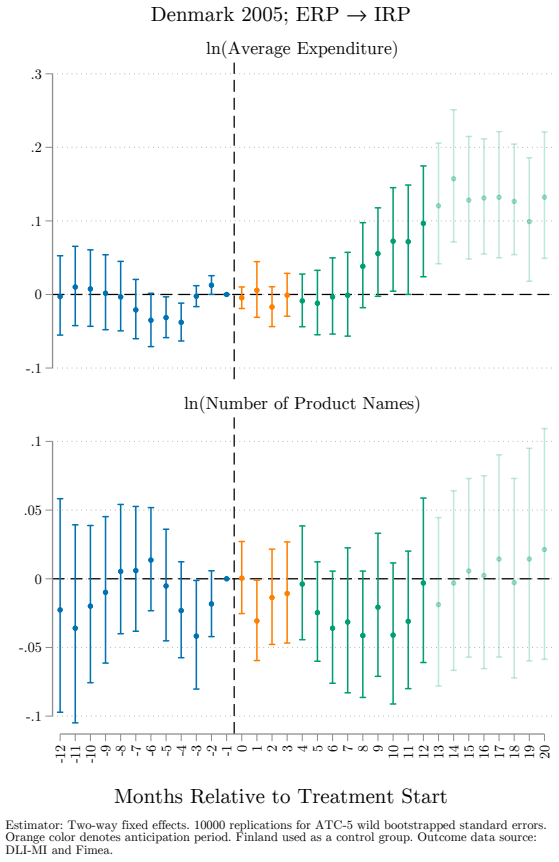
A.11 Weighted Results

In this subsection we display our weighted main result event studies by each studied reform. We estimate weighted versions of our main results, because we want to be sure that our findings are not driven by markets with small economic significance. Studied markets are weighted by their share of the treatment country pharmacy sales of prescription pharmaceuticals. Sales from periods -12 to -6 are used in constructing the weights. We derive constant weights from the pre-period, because otherwise the studied reform would also influence the weights we use.

Part I Weighted Event Study Results: Figure 25 displays weighted main results for Finnish 2003 and 2009 reforms. Unweighted results (presented in Figure 4 display a small statistically insignificant Average Expenditure reduction for Finland 2003 reform, but when weights are used this negative effect disappears. Weighted results for the Finland 2009 reform have the same trajectory as unweighted results, but the point estimates have larger size.

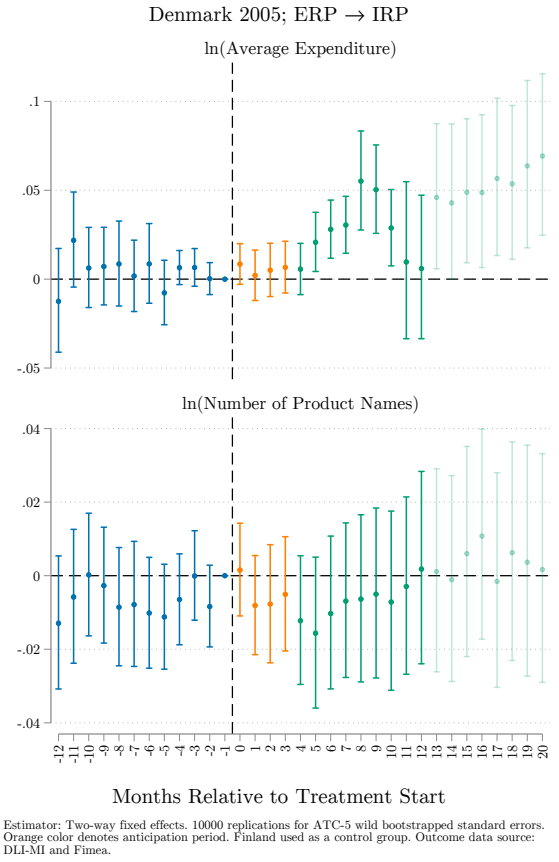
Part II Weighted Event Study Results: Figure 26 shows weighted results for Danish 2000 and 2005 reforms. Weighted results for Denmark 2000 are almost identical with the unweighted results, but weighting changes the interpretation with respect to Denmark 2005 reform. Unweighted results presented in Figure 5 shows that Average Expenditure weakly increases as a result of the studied reform. However, weighted results show that the reform impact on Average Expenditure is much smaller in size than in the case of unweighted results. With the weighted results the most likely result interpretation is that the reform had only negligible impact on the Average Expenditure. Weighting seems not to have any significant impact on product availability outcome (Number of Product

92. The price cut was imposed on the price caps and in competitive markets large share of products are priced under the price cap and a 5 % reduction in the cap does not have a large impact on firm pricing. In monopoly markets or markets included in our spillover analyses, the price cap cut can have a full 5% decrease in wholesale prices because products in these markets do not face competition and are priced to the cap.



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

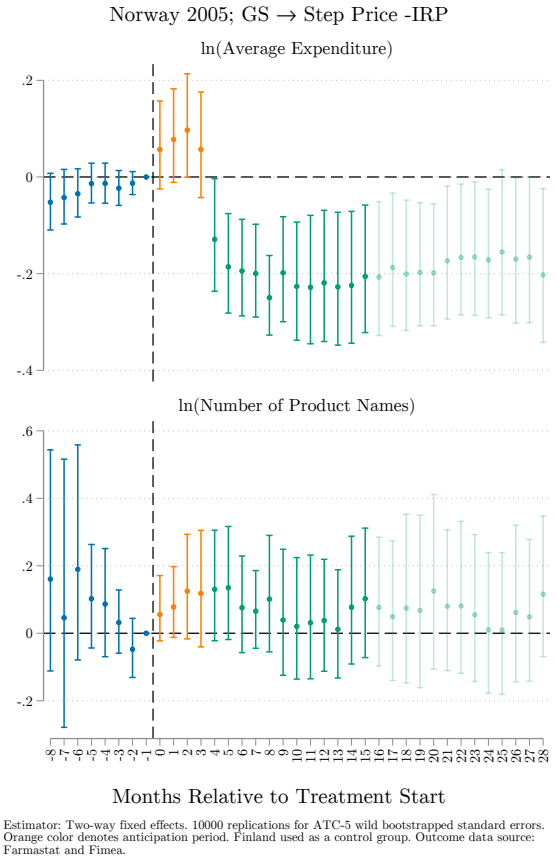
(a) Denmark 2005 – Longer



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

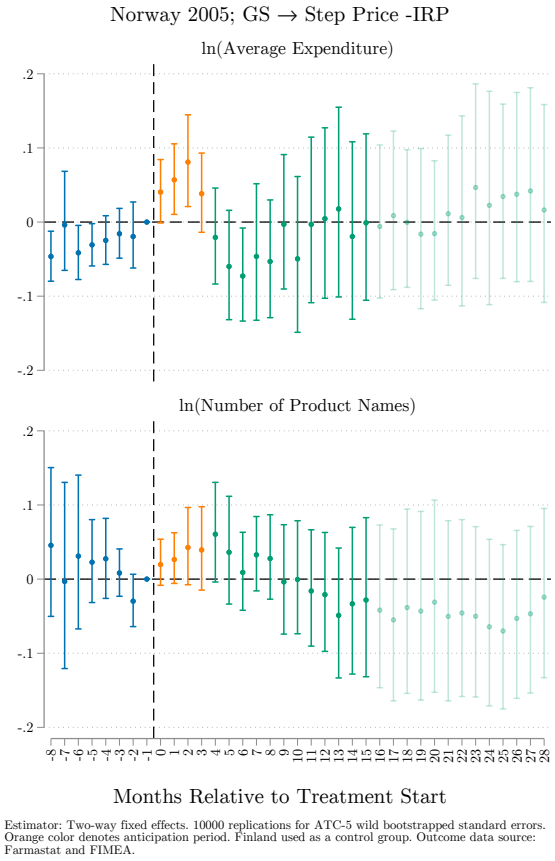
(b) Denmark 2005 Monopoly – Longer

Figure 23: Denmark 2005 Results with Post 2006 coefficients.



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: Farmastat and Fimea.

(a) Norway 2005 – Longer



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: Farmastat and FIMEA.

(b) Norway 2005 Spillover – Longer

Figure 24: Norway 2005 Results with Post 2006 coefficients.

Names) for neither of the reforms.

Part III Weighted Event Study Results: Figure 27 shows weighted results for Norway 2005 and Sweden 2009 reforms. Weighted and non-weighted results are similar to each other for Norway 2005, but weighting has an impact on Sweden 2009 reform results. Weighted and non-weighted expenditure (Average Expenditure) results share the same profile, but the weighted results have larger estimate size. The weighted availability results have larger effect size than the unweighted results, but the effect size difference is not substantial.

A.12 Event Study Results for Secondary Outcomes

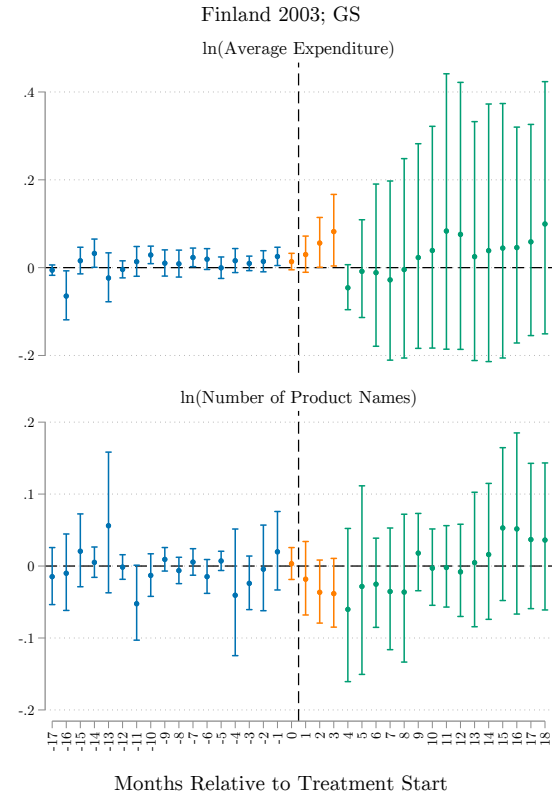
In this section we present event study results for our secondary outcomes (Average Price, Wholesale Price and Doses) by each reform in the same order as results are presented in the main text table 4.

Part I Event Study Results: Figure 28 collects unweighted event study results for Finland 2003 and 2009 reforms. Results show that market level price effect (Average Price) has a smaller effect size in absolute value than the package level price (Wholesale Price) for both Finnish reforms. It is interesting to see that in Finland 2009 reform the average price converges to zero in the end of the follow up period and the package level price converges to 11%.

Part II Event Study Results: Figure 29 collects unweighted event study results for Denmark 2000 and 2005 reforms. Event study results for both Danish reforms follow the same patterns as in the case of the Finnish reforms, the wholesale price estimate size is larger in absolute value than the market-level price response. The surge in the Denmark 2000 quantity (Doses) results is a result of the change in the Danish reimbursement system. This change had no effect on pricing, because neither of the price measures reacts to the change in the reimbursement system.[93]

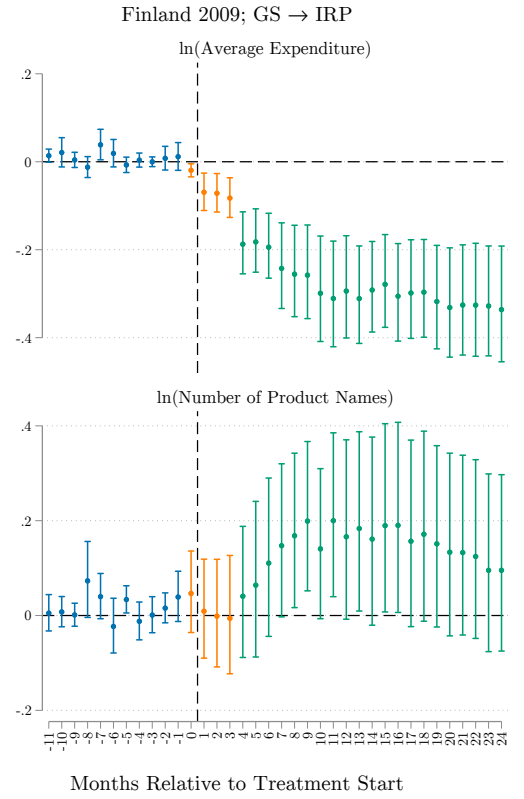
Part III Event Study Results: Figure 30 collects unweighted event study results for Norway 2005 and Sweden 2009 reforms. Norway 2005 reform shows again the previous finding that package level prices (Wholesale Price) can yield a different results than market level price (Average Price) when consumer choice reforms are studied, but we find almost identical price effect results for the Sweden 2009 reform. Interestingly, with package level prices we find almost constant and immediate price decrease whereas with market level

93. Appendix section A.3.6 describes the Danish reimbursement system change in detail.



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI and Fimea.

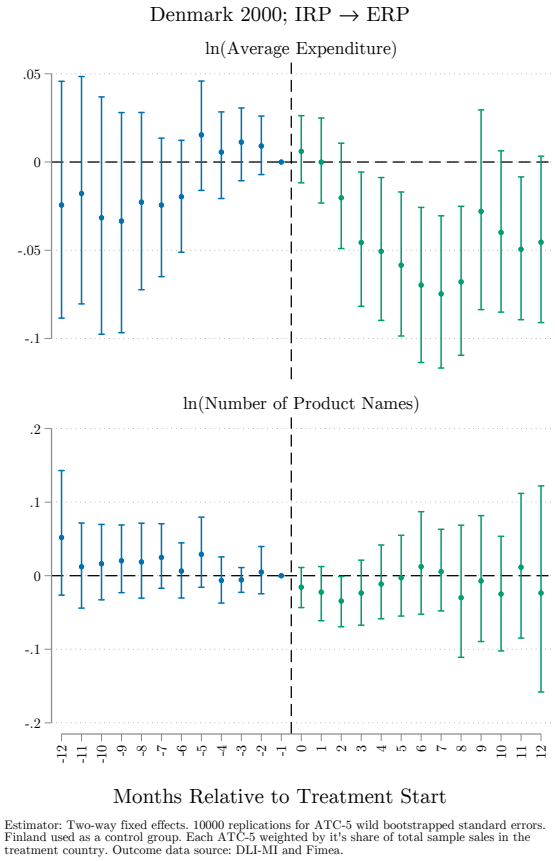
(a) Finland 2003 – Weighted



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Norway used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: Farmastat and Fimea.

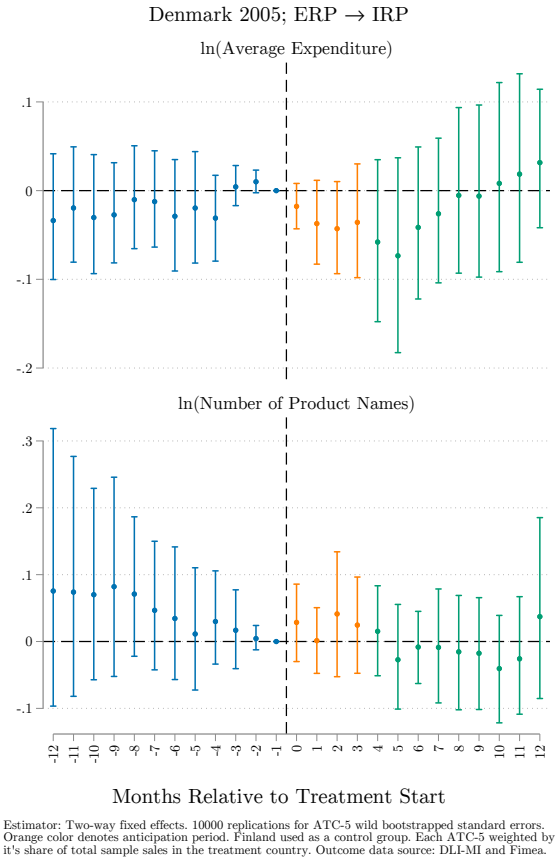
(b) Finland 2009 – Weighted

Figure 25: Weighted Part I Main Outcome Results



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Finland used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI and Fimea.

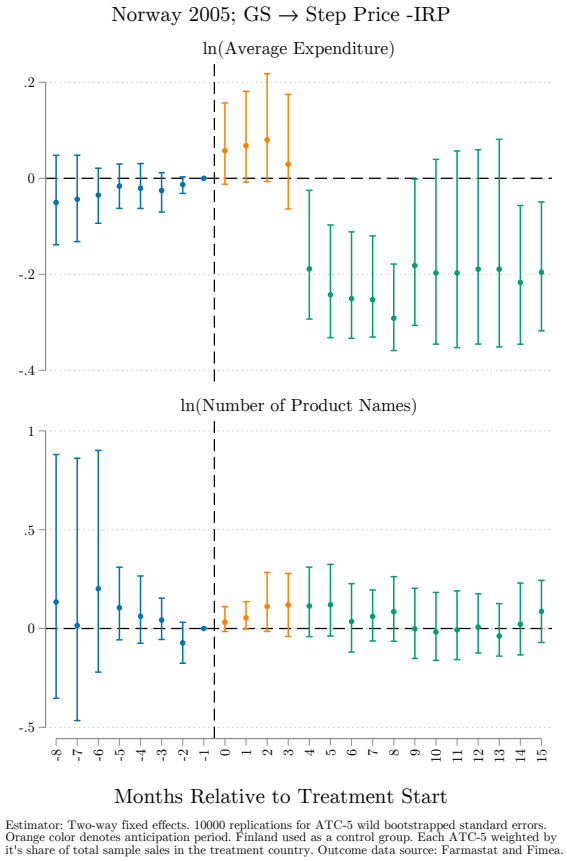
(a) Denmark 2000 – Weighted



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: DLI-MI and Fimea.

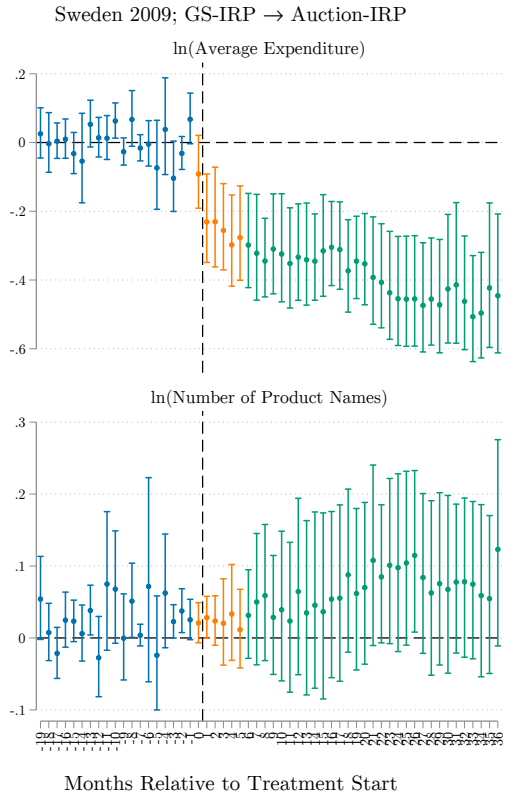
(b) Denmark 2005 – Weighted

Figure 26: Weighted Part II Main Outcome Results



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: Farmastat and Fimea.

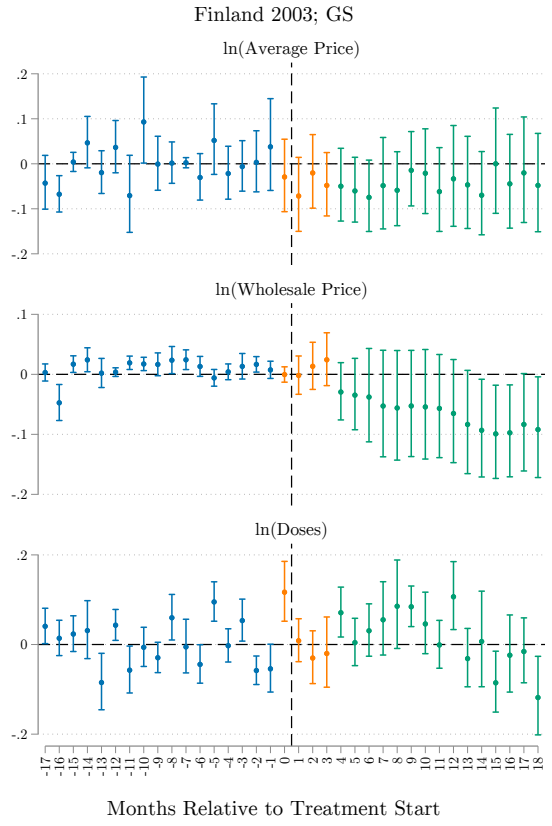
(a) Norway 2005 – Weighted



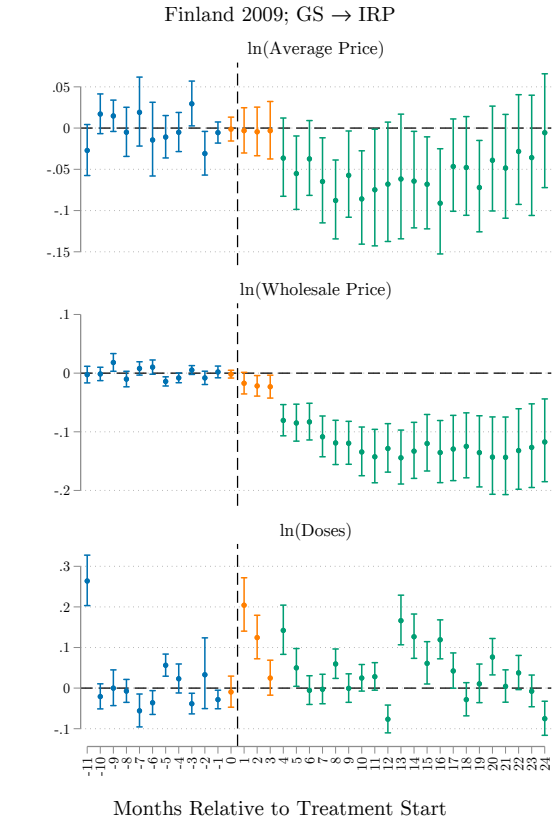
Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Norway used as a control group. Each ATC-5 weighted by it's share of total sample sales in the treatment country. Outcome data source: Farmastat, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(b) Sweden 2009 – Weighted

Figure 27: Weighted Part III Main Outcome Results



(a) Finland 2003

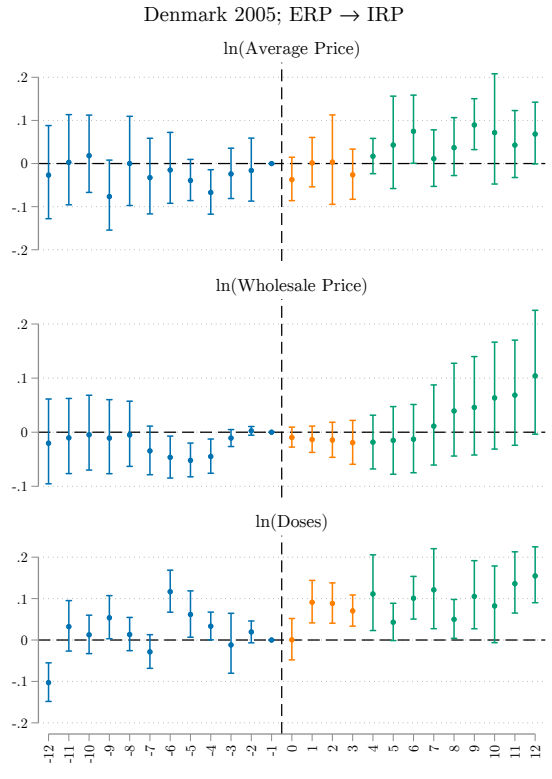


(b) Finland 2009

Figure 28: Part I: Secondary Results

Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(a) Denmark 2000

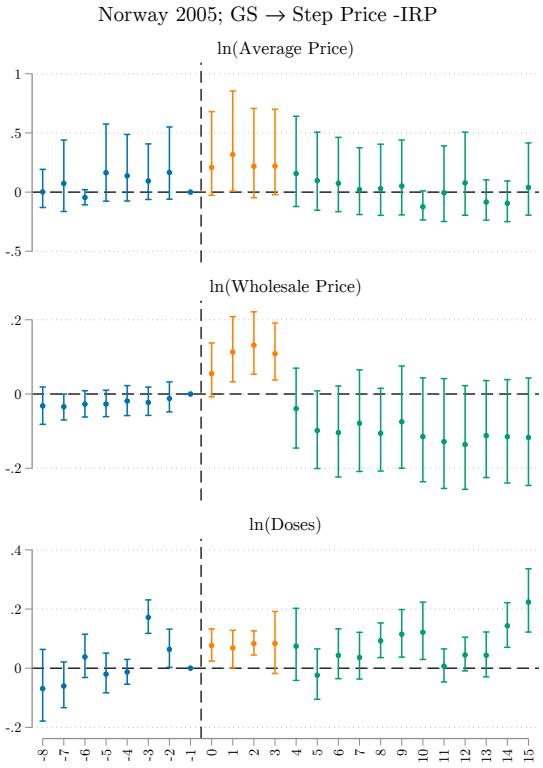


Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and Fimea.

(b) Denmark 2005

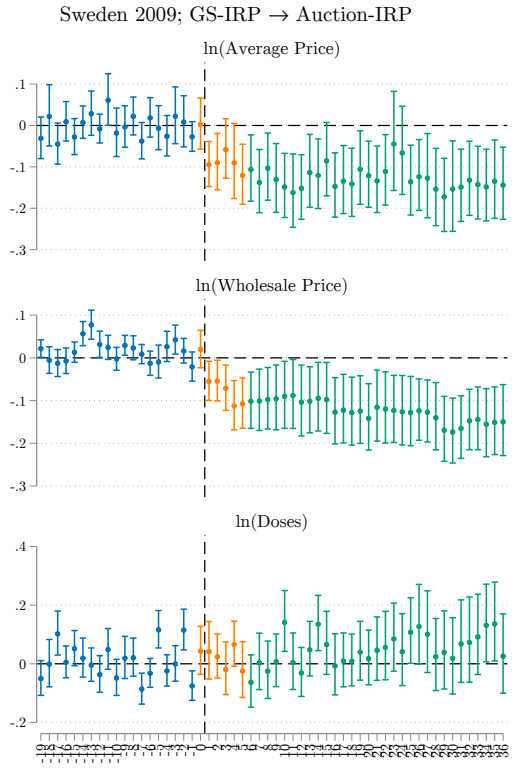
Figure 29: Part II: Secondary Results

prices we find slightly downward sloping price effect. For Sweden 2009 price results using package level prices vary less than the market level price. The most likely explanation for this is the regulation which gives the market demand for the cheapest product.



Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: Farmastat and Fimea.

(a) Norway 2005



Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Outcome data source: DLI-MI, IQVIA MIDAS Quarterly Sales and IQVIA MIDAS (2007–2013).

(b) Sweden 2009

Figure 30: Part III: Secondary Results