The Effects of Price Regulation on Pharmaceutical Expenditure and Availability*

Mika Kortelainen Jaakko Markkanen Markku Siikanen Otto Toivanen

4th November 2022

Abstract

There is little quasi-experimental evidence on the effectiveness of different price regulation policies in curbing pharmaceutical expenditure. Using data from the Nordic countries, we analyze how widely used generic substitution and reference price policies influence market-level pharmaceutical expenditure, prices, and availability of pharmaceuticals. Constructing treatment and control groups by matching data across countries by active ingredients and employing recent difference-in-difference methods, we find that several price regulation reforms decrease pharmaceutical expenditure considerably and proportionally more than average prices. Different regulatory policies seem not to affect pharmaceutical availability in the markets we study.

Keywords: *pharmaceuticals, regulation, generic competition, reference pricing*
JEL-Classification: *I11, H51, L51, C23*

*Mika Kortelainen: University of Turku, InFLAMES Research Flagship Center, VATT Institute for Economic Research, Helsinki Graduate School of Economics. Email: mika.kortelainen@utu.fi. Jaakko Markkanen: Aalto University, Helsinki Graduate School of Economics. Email: jaakko.m.markkanen@aalto.fi. Markku Siikanen VATT Institute for Economic Research, Helsinki Graduate School of Economics. Email: markku.siikanen@vatt.fi. Toivanen: Aalto University School of Business, Helsinki Graduate School of Economics and KU Leuven. Email: otto.toivanen@aalto.fi. We thank thank Hertta Hankimaa, Hung Le, Ella-Kaisa Luoma, and Helena Rantakaulio for excellent research assistance, the Finnish Social Insurance Foundation (Kela) and the Yrjö Jahnsson Foundation for financial support and participants in several seminars for comments. 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, GMD Scope, 2006–2017 IQVIA AG. 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 costs of diabetes treatment in the United States and called on 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]

Global spending on medicine has doubled in the last 10 years, reaching 1.3 trillion dollars. Spending on medicines has increased everywhere with the US in the lead (see Figure 1). Therefore, it is not surprising that most OECD countries have adopted various cost-containment policies. Often, such policies are needed to counterbalance the fact that societies have reduced or even removed the price sensitivity of citizens by providing public insurance (e.g. through small co-payments or co-insurance).[2] Various price regulation policies for prescription drugs are common, especially in Europe, but as our starting quote makes clear, such policies have also recently been promoted in the US.[3] Despite the widespread utilization of price regulation policies, there is little credible causal evidence on whether such policies have decreased pharmaceutical expenditure or achieved their main objectives. The contribution of this paper is to provide this evidence.

We investigate the effects of different Reference Pricing (RP) and Generic Substitution (GS) policies on pharmaceutical expenditure, prices, and availability using data from four Nordic countries. We have chosen these countries for two primary reasons: First, the Nordic countries are examples of societies that provide generous public insurance against pharmaceutical expenditure. Consequently, these countries have adopted several variants of Internal Reference Pricing (IRP) and External Reference Pricing (ERP) policies combined with GS during the 2000s. The main objective of these policies is to reduce pharmaceutical expenditure through (generic) competition and to increase consumer price sensitivity within a generous reimbursement system. Second, they are as homogeneous as groups of countries go, making the use of them as controls for each other appealing.

We depart from the existing literature first, by employing the same markets (active ingredients) in a country of similar appearance as our control group.[4] The previous quasi-experimental literature has mostly had to rely on different active ingredients within the same country to construct a control group, or on products in the same group that entered the new regulatory regime at a different time.[5] Different active ingredients often have dif-

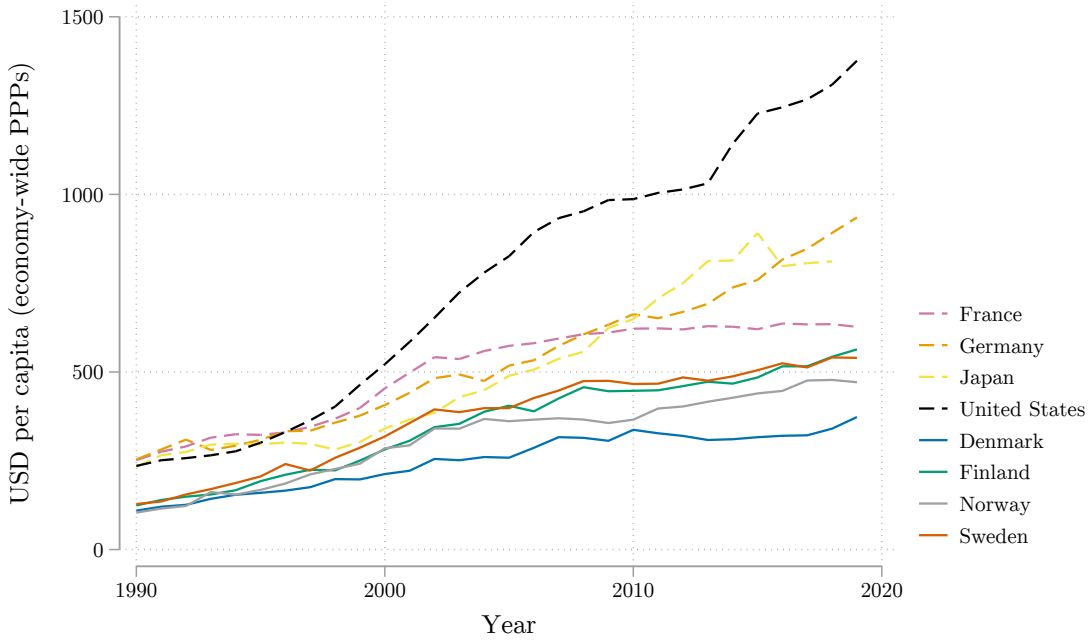
1. The statement was published on Twitter.

2. Note that besides taxpayer funded health programs, the same logic may well apply within private provision of health insurance, as e.g., employed in the United States.

3. 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.

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

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



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

ferent price trends due to differences in regulation and competition and are likely indirectly affected by the regulatory change. Second, we use market-level outcomes, thereby solving the problem plaguing much of the literature that arises from the Stable Unit Treatment Value Assumption (SUTVA). When studying the prices of individual products, this assumption translates into the untenable assertion that the price of a product is not affected by the prices of its substitutes. This assumption of no equilibrium effects runs against both theoretical models of competition and empirical evidence. Third, our key dependent variable is the average expenditure per purchased dose (henceforth average expenditure), which allows us to directly interpret our findings in terms of changes in expenditure per dose. To our knowledge, the existing literature (with only one exception; see below) has looked at the effect of regulatory changes on average posted prices, that is, not taking into account which products are actually sold as opposed to being on sale. As documented below, changes in posted prices do not translate one-to-one into changes in expenditure because different regulatory regimes have differing incentives for choosing the cheaper or cheapest alternative. Fourth, a central worry related to price regulation is its impact on product availability: unlike most papers on the effects of regulation on pharmaceutical

prices, we also study the impact of regulation on product availability.

We find that price regulations decrease expenditure, 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 expenditure per sold dose. The Finnish move in 2009 from GS to IRP reduced expenditure by 25%.[6] Denmark reduced expenditure by some 8% moving from IRP very symmetrical 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 to a Product of the Month Auction (Auction-IRP) regime where each month essentially only the cheapest product in a given category was distributed. This change reduced expenditure by 20%. Similarly, Norway was able to reduce expenditure by 20% in 2005 when moving from GS to a system with government-dictated price cuts after the introduction of generic competition.

We also study average posted prices per dose, which is the key dependent variable in almost all existing literature on the effects of pharmaceutical price regulation. Here, the results are quite different. Based on our point estimates, the Finnish reform leading to a 25% reduction in average expenditure per dose lowers the average posted prices by only 15%; The Danish reform from IRP to ERP that lead to an 8% reduction in average expenditure per dose had no effect on average posted prices; and the Swedish product-of-the-month reform lowered average posted prices only by 10%, i.e., half the effect it had on average expenditure per dose.

The Swedish reform of 2009 demonstrates the difference between average posted prices and average expenditure, i.e., average purchase prices per dose. In the product-of-the-month regime, the vast majority of consumers need to buy the cheapest available product to be reimbursed, but some consumers may be prescribed the more expensive products. By definition, all other products are priced higher than the product of the month, and some firms may have an incentive to price their product very high to skim-cream-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.

Finally, our analysis of the effects of the different regulatory reforms on product availability shows that none of the reforms we study, despite the large effects on expenditure, has an impact on the number of products and product variants sold on the market.

Our article contributes to three strands of literature. The first is the literature on the effects of price regulation on pharmaceutical pricing. Several earlier articles have used single-country data and quasi-experimental variation in price regulation to evaluate the effect of an individual policy (Pavcnik 2002; Brekke, Grasdal and Holmås 2009; Brekke, Holmas and Straume 2011; Herr and Suppliet 2017) on average posted prices. Perhaps the

6. The reported magnitudes are based on our point estimates. Naturally, one needs to keep their confidence intervals in mind.

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.[7] Pavcnik (2002) finds that the posted prices decrease 10–36%. In contrast to 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. The treatment group consists of six active ingredients, chosen because they cover a wide set 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 of generics by 7–8%. In contrast to these important papers, we study various different reforms of price regulation in Nordic countries and examine their relative effectiveness in reducing both pharmaceutical expenditure and posted prices.

Second, we contribute to the literature on the effects of pharmaceutical regulation on expenditure. The existing literature has mostly shied away from studying the effect of price regulation on expenditure, with the likely explanation being restrictions in data and research designs. When the question has been addressed, the existing literature has used (changes in) average posted prices as a proxy for (changes in) pharmaceutical expenditure (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 2021). Our result casts 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 reduced-form analysis 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), but extend the analysis in a number of ways. Using within-country data, they find that the introduction of RP lowered expenditure in Norway by 30%.

Dubois and Lasio (2018) is an important precursor of our study in that it also uses multi-country data, although the difference-in-difference analysis that they use to complement their structural modeling is carried out using package-level (= product) data, not market-level data. 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 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

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

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[8] and therefore sees it justified to incentivize people to substitute for cheaper products. In line with the regulator's approach, our main objective is to analyze whether the implemented regulations have led to a decrease in the amount of expenditure per dose.

Third, 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. By evaluating short-term effects on pharmaceutical availability, we provide new evidence on the effect of regulation on availability. Note that the regulations we study are unlikely to have a first-order impact on pharmaceutical innovation, as all but one of them mainly concern markets with off-patent drugs.

Fourth, our work illustrates how quasi-experimental methods and, in particular, modern event study and difference-in-difference estimators can be utilized to evaluate the impacts of market- or country-specific large reforms. Recent developments in difference-in-differences methods emphasize the use of well-defined control groups in both staggered and non-staggered designs, as well as the potential spillover effects of major policies on the control group (e.g. Callaway and Sant'Anna 2020; de Chaisemartin and D'Haultfœuille 2020). Using the same pharmaceutical market in another country as a control group, our approach allows regulation policies to have spillover effects within the country of interest and allows us to investigate market-level effects of the reforms, thus capturing equilibrium (SUTVA) effects within country. Our setting allows us to evaluate the identification assumptions of each research design and to implement various robustness analyzes.

The rest of the paper is structured as follows. In Section 2 we present the relevant institutions and regulatory regimes. We introduce the data and our matching procedure in Section 3 and present our difference-in-differences estimation approach in Section 4. Section 5 is devoted to the presentation and discussion of the results. We offer our conclusions in Section 6.

8. There is plenty of evidence that patients prefer branded products to generic products; see e.g. Dubois and Lasio (2018).

2 Institutions and Regulatory Regimes

All Nordic countries[9] 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, co-insurance, or premiums when using public services. There are some exceptions to this rule, prescription drugs being a notable one. Next, we briefly explain how pharmaceutical reimbursement systems are operated in the Nordic countries and describe the existing price controls that affect private and public spending on pharmaceuticals.

There are two distinct approaches to public reimbursement of pharmaceuticals in the Nordic countries: a needs-based calculation and a product-specific calculation of reimbursement. In the needs-based system, used in Sweden and Denmark,[10] the level of reimbursement and the consumer's copayment are tied and capped to the consumer's annual spending used on pharmaceuticals. The share of reimbursement (copayment) increases (decreases) as the consumer spends more on reimbursed pharmaceuticals. After crossing some legal threshold, the consumer is fully reimbursed. However, 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 tied to 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, the reimbursement, conditional on the price negations 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. Thus, the government's decision in the latter is in two parts: first, the government decides whether to reimburse, and second, by how much.

Two regulatory regimes can share the same name even in official use, while operating differently. To avoid confusion, we use the following definitions:

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

9. The Nordic countries consist of Denmark, Finland, Iceland, Norway and Sweden.

10. 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).

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

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

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

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

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

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.

Most often, regulatory policies consist of a combination of GS and some form of RP, but sometimes only one or the other is in use. 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.

Some regulations or regulatory limitations come from the law European Union (EU), which sets the principles for market entry and the freedom of movement of goods in the single market. For example, parallel trade, that is, imports of pharmaceuticals (irrespective of patent status) from a low-priced member state to high-priced member states, are protected by EU law.[11] 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, the EU can place some restrictions on national regulators, an example of which is the maximum processing time for reimbursement decisions[12].

2.1 Denmark

May 1997–Oct. 2000: GS and IRP. In 1997, Denmark adopted the mandatory substitution of generics on top of an already existing RP system for generics.[13] This regime

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

12. 180 days for new pricing and reimbursement decisions, 90 days for review of an application to increase prices. See Directive 89/105/EEC.

13. See BEK nr 308 af 06/05/1997 §36 - §37.

corresponds to our definition of a GS system with IRP. The Danish system required pharmacies to substitute for the cheapest interchangeable product available unless the price differential was (roughly) less than 5%.[14] The prescribing physicians could still opt out of substitution on medical grounds. If a consumer did not choose to buy the reference-priced product, then the consumer was required to pay the price difference between the reference-priced product and the purchased product fully 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 from other European[15] prices and domestic prices. If a product was sold only in Denmark or the domestic price was lower than the other European prices, the price in Denmark was used to calculate 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.[16] The process was finalized in November 2000 when the Danish government stopped the reimbursement of all products that exceeded their European average prices.[17] While the use of ERP was included in the legislation in summer 2001, the regulator started to apply ERP already in November 2000.[18] Thus, we use November 2000 as the date of the change of the regulation regime in our analyses.

Jan. 2005– : GS and IRP. ERP lasted until January 1, 2005, when it was replaced by IRP.[19] 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 periods 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 reimbursement 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 a co-op of pharmaceutical manufacturers. Appendix A.2 explains these changes in more detail.

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

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

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

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

18. See LOV nr 495 af 07/06/2001 §7d

19. See LOV nr 1431 af 22/12/2004 §7d

2.2 Finland

Up to March 2003: VGS. Throughout the 1990s, Finland allowed VGS and generic prescribing by physicians. The system required prescribing physicians to actively opt-in to allow GS to occur. In practice, GS and prescription of generics was almost non-existent.[20]

April 2003–March 2009: GS. Following the experiences of other Nordic countries, the Finnish government adopted mandatory GS starting from April 2003.[21] In the new regime, pharmacies were required to deliver one of the products at or close to the cheapest price.[22] The reimbursement of a consumer was not affected if she decided against substitution; the monetary incentives to substitute were small in drug categories with higher reimbursement rates. This feature makes the Finnish substitution policy unique in comparison to other studied Nordic reforms, because other countries combined substitution policies with financial incentives.

April 2009– : GS and IRP. To address the incentive problem related to GS and high reimbursement rates, the Finnish government adopted IRP in April 2009.[23] Reference pricing was applied only to products that were publicly reimbursed and had at least one generic substitute available. The reference price is the highest retail price from which reimbursement can be paid.[24]. If the consumer purchased a product more expensive than the reference priced product, the consumer was reimbursed on the basis of the reference price, and the consumer needed to pay the difference between the reference price and the chosen product fully out of pocket. Parallel imports were not included in the system until 2017.[25] Due to the reference price eligibility rules, the policy in 2009 did not stop the GS policy because markets not eligible for the reference price policy remained eligible for GS because GS was applied for non-reimbursed and parallel imported products.

During the 2000s, Finland also implemented minor reforms related to pharmaceutical pricing in the Finnish market. The first minor reform is how the price cap of a reimbursed generic product is determined when a new generic enters the market in 2006. The regulation change imposed that the price cap of new entrants should be 40% lower than the cap of the originator product. The second reform is the 5% price cap cut on all reimbursement drugs. These reforms are explained in more detail in the Appendix A.2.

20. See the government proposal HE 165/2002 vp, page 6.

21. See 80/2003 §57b.

22. Pharmacy was required to offer substitution if 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.

23. See Chapter 6 §18-§23.

24. During our sample period reference price was defined as the cheapest retail price within reference price group +1.5€(retail price less than 40€) or 2.5€(retail price more than 40€). Reference prices were updated quarterly.

25. See 1100/2016 Chapter 6 §18. Prior to this, parallel imports could be included to reference price groups if other generics were on the market. After the 2017 change, this requirement was lifted. In practice, this made it possible for RP to start even during the patent period.

2.3 Norway

March 2001–2005: GS. In 2001, Norway simultaneously adopted a GS policy and liberalized the pharmacy sector.[26] Prior to the 2001 reform Norway had adopted a system similar to ERP, where the government calculates the maximum reimbursement price as an average of the three lowest prices of the original patented product from a basket consisting of other European Economic Area (EEA) countries.[27] Thus, the GS system with ERP elements is the baseline regulatory regime for subsequent reforms in Norway.

Norwegian GS policy provided financial incentives for the consumer to substitute to cheap products.[28] If the consumer substituted to the cheapest alternative, the consumer had to pay the price difference between the cheapest alternative and the chosen product out of pocket.[29] Norwegian GS does not explicitly require pharmacies to substitute for a cheaper alternative available, but instead pharmacies are incentivized to offer GS. 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.[30]

Jan. 2005– : GS and the Step-Price regime. Norway implemented a major change to the GS system in 2005 by introducing the current system, called the Step-Price (SP) system. After the generic entry has taken place, the maximum reimbursement price (the Step-Price) gradually decreases.[31] The base level for the price was established as the maximum allowed retail price (maximum wholesale price and maximum retail margin) at the time of generic entry. If a consumer decides not to buy the product priced at Step-Price, he is required to pay the difference between Step-Price and the actual price entirely out of pocket. The first price cut takes place at the start of generic competition, followed by further cuts after 6 months and 12 months.[32] The magnitude of the price cuts is related to the total sales prior to generic entry: During our sample period, the first price cut is

26. In this paper we do not study the effects of this substitution reform, because the reform effects cannot be separated from the pharmacy market liberalization. See for further information LOV-2000-06-02-39

27. In practice, this means that most of the time the price cap is binding only to the branded. To provide perspective, the Danish price cap for 2000 - 2004 was calculated from the pooled prices of all substitutable products.

28. See LOV-2000-06-02-39 for further information.

29. In comparison to Finnish substitution policy, Norwegian provided financial incentives and the Finnish policy did not offer incentives.

30. See FOR-2001-12-17-1537 §12-3. Generic alternatives receive 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 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/or Brekke, Holmas and Straume (2011).

31. The Norwegian Medicines Agency determined when generic entry had taken place. In practice, it required that the generic product be available in pharmacies.

32. Appendix section A.7 shows the price cut timing in the Step-Price system.

-30%, the second between 40–50%, and the third between 50–70%.[33] 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.

2.4 Sweden

Nov. 2002–2009: GS and IRP. Sweden adopted GS in November 2002.[34] 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 only for the prescribed product or the product for which the pharmacy offered substitution to. 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.

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. The winners are called the products of the month. Consumers can in practice only choose between the prescribed product and the product of the month, though 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.[35]

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 by changing price cap rules within substitution groups. 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 the Appendix A.2.

2.5 Summary of Reforms

We now summarize the changes in the price-regulation regimes that we study and explain how we select our control countries.

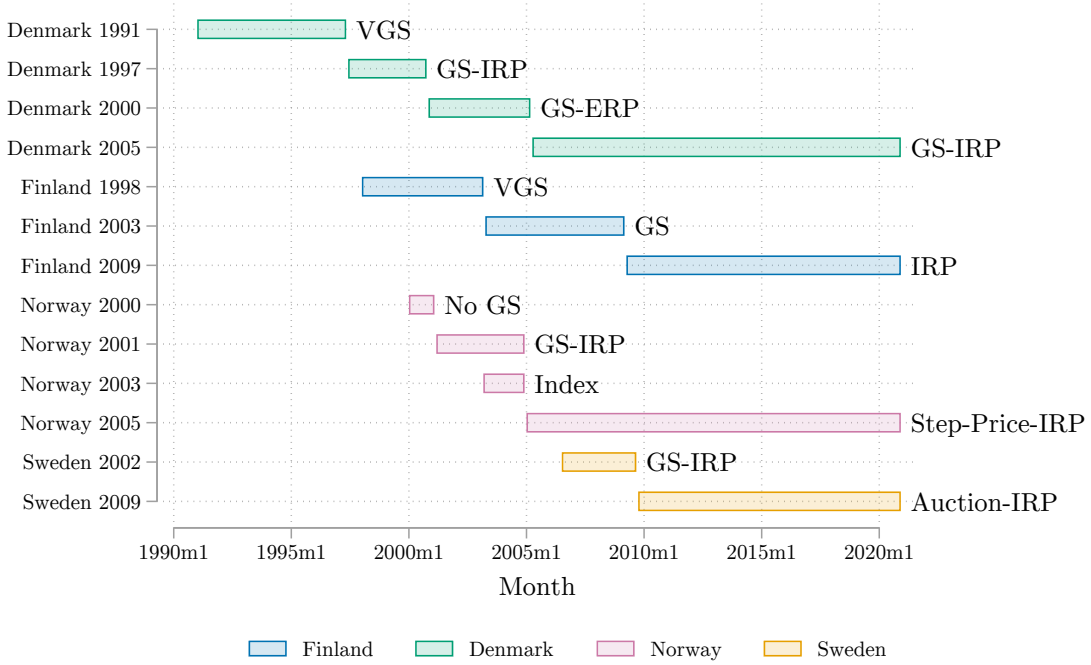
We have collected the regulatory regimes and the 9 regulatory changes in Figure 2. We observe 2 regulatory regimes for Sweden, 3 for Finland, and 4 for both Denmark

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

34. 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

35. Appendix section A.7 displays the timing in the Auction-IRP system.

Figure 2: Timeline of reforms



Timeline of reforms by country and data availability.

and Norway. We exclude three of the nine reforms from this paper: The Danish 1997 change from VGS to GS is excluded because we cannot form a well-defined external control group. The Norwegian 2003 reform is excluded because it directly influenced only eight markets (active ingredients), and thus we cannot perform market-level analyses. Finally, the Norwegian 2001 reform of pharmacy market liberalization and GS reform are left outside the scope of this article because the GS policy cannot be identified separately from market liberalization. All reforms occur over a nine-year period.

The reforms we analyze affect consumer choice and can be divided into three different groups based on the content of the reforms. The first category are the Finnish 2003 and 2009 GS and IRP reforms. These reforms demonstrate how well a simple substitution reform without financial incentives can reduce average expenditure and what happens when financial incentives are included to the substitution policy in a form of reference price regulation. The second category of reforms is the Danish ERP and IRP reforms. These reforms highlight the benefits of using ERP policies. The third category of reforms is the "extreme" versions of IRP regulation. We label Norwegian 2005 and Swedish 2009

Table 1: Treatment and Control Countries

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

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

reforms as extreme because these reforms either hand out market demand for the cheapest product (Sweden 2009) or market demand is targeted for products which are priced using predetermined pricing rules set by the regulator (Norway 2005).

We have collected the reforms analyzed into Table 1. The table contains information on the treatment country in question; the regulatory regime before and after the change; the control country that we have chosen to use; and the regulatory regime in the control country during the estimation period.

An important decision is the choice of a control country. As Figure 2 shows, the regulatory environment in the Nordics changed a lot during 2000–2009. Our main criterion in choosing a control country is to identify a control country where no major regulatory changes occur in the years before and after the reform in question. Figure 2 helps to show that for each of the reforms we study, there is essentially only one country available that can act as a control group.

For the Danish 2000 reform, switching from GS-IRP to GS-ERP we use Finland as the control country. The Finnish regime at that time was VGS.[36] The countries change roles when we study the Finnish reform, replacing VGS with GS in 2003. In the following, we show that the concern that the time-varying effects of the earlier Danish reform could contaminate the control period for the later Finnish reform does not materialize. Furthermore, our different estimation samples consist neither of the same markets (because the number of markets with generic competition increases over time due to patent expirations) nor of the same products (due to generic entry).

The same countries switch roles once more when we study the Danish reversal from ERP to IRP in 2005. Similarly to the previous case, we checked whether the time-varying effects of the Finnish reform would lead us to misinterpreting the findings of the analysis of the 2005 Danish reform. This worry turns out to be unfounded.

36. We were not able to obtain Swedish sales data for 2000–2006. This would have provided us with a new control group for comparisons, where we currently use Denmark as a treatment group.

The fourth reform that we analyze is the Norwegian change from GS-IRP to SP in 2005. Figure 2 reveals that the country with a stable regulatory regime (and with which we have data) is Finland, where GS was in place at the time. We discard 8 treated markets (= Norway) due to the index price regulation implemented in 2003 in Norway.[37] We discard markets transitioning from index price to SP because otherwise the pre-period market institutions would not be the same for all Norwegian markets.

The second Finnish reform that we study is the change from GS to IRP, implemented in 2009. Denmark is again used as a control country.

The only Swedish reform studied in this paper is the implementation of the product of the month system in 2009. Sweden transitioned from GS with IRP to a system similar to a first-price sealed bid procurement auction. We use Denmark, where IRP was used at the time, as the control country.

Some of our reform comparisons are affected by other minor reforms occurring either in the pre- or post-period of the studied reform. Most notably, Finland implemented a 5% price cut to the wholesale prices of all reimbursed pharmaceuticals in January 1st 2006,[38] the effect of which is clearly visible in our event studies for the Danish and Norwegian reforms in 2005 where Finland is used as a control group. Although the first-order effects of the price cut are known precisely, we cannot rule out effects on the equilibrium price levels, and thus the affected coefficients should be interpreted conservatively. Another reform influencing the interpretation of our results is the change in the Danish reimbursement system in 2000. This change influences pharmaceutical consumption, and this is clearly visible when sold quantities and sales are investigated using event studies in Denmark 2000 reform.

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.[39] Panel A in Table 2 shows the country-specific source for each data set and the data coverage for each country. The data sets contain information on the sales value and volume of each pharmaceutical package sold in the respective country. The sales values are defined in pharmacy purchase prices and volumes Defined Daily Dosages (DDD) for each respective active incredient according to the Anatomical Therapeutic Chemical

37. 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.

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

39. Data from Iceland was not available for us to purchase.

Table 2: 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.

classification system (ATC).[40] In our analysis, we aggregate our sales and quantity data at the package level to the active ingredient level. The country-specific datasets also provide rich information on product characteristics.

We supplement our sales data with rich regulatory information that is obtained from market regulators and directly from the relevant legislation. Panel B in Table 2 shows the source and coverage of the reform data for each reform studied. The reform information covers time periods before and after the implementation of the reform in question, both for the treated and for the control country.

We use wholesale prices in our analyzes. The wholesale price is the price a pharmacy pays for the product when the product is purchased from the wholesaler. The monthly wholesale price is calculated by dividing the monthly sales of a given product by the monthly number of packages sold. There are two reasons for using wholesale prices: First, the regulations relate to 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. Appendix section A.6 illustrates how price formulas

40. The ATC system classifies active incredients 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

work in the Nordic pharmaceutical market.

All main outcome variables are market-level variables, because unlike package- or trade name-level outcomes, with market-level outcomes we can satisfy SUTVA in our differences-in-differences identification strategy.[41] Our main outcomes are logarithms of average expenditure at market level $\left(\frac{\sum P \times Q_{Package}}{\sum Q_{Doses}}\right)$, average wholesale price $\left(E\left[\frac{P}{Q_{Doses}}\right]\right)$, and total quantity $(\sum Q_{Doses})$. The average expenditure outcome measures how much market-level expenditure is for each dose sold. The average price measures the average wholesale price per DDD, and the total quantity measures the size of the market measured in DDDs. These outcomes decompose the changes in average expenditure into price and quantity responses.

As secondary outcomes, we also study package-level wholesale prices (per DDD) and market-level drug availability measures. Package-level wholesale price analysis allows us to compare the results of the previous literature with ours. Market-level availability outcomes are the number of unique sold packages and sold trade names.

The price and sales variables are measured in nominal national currencies.[42] The reason for not converting prices into the same currency is the possibility of exchange rate shocks, which could bias and distort our findings. Exchange rate shocks are problematic in analyzes where data from Norway or Sweden are used, because these currencies are not linked to the euro like the Danish krone.[43] Nominal values are used because in all Nordic countries price regulation works with nominal prices. The sample periods in our empirical analyzes are relatively short (2–4 years), so 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.[44] We match the markets between countries based on the active ingredient defined by the ATC 5 code. The matching process contains the following four steps: i) we discard non-prescription pharmaceutical products (over-the-counter pharmaceuticals (OTC) products); ii) we identify the markets that are affected by the reform in question in the treated country c; 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. This matching strategy produces samples with identical pharmaceutical markets while avoiding restrictions on market structure and product variety within the markets.

41. The next section explains the identification strategy in detail.

42. Sales data from Finland is represented in euros, because the switch from FIM to EUR occurs during our sample (2002).

43. Visual inspections of the data showed that this is not only a potential, but also a real concern. The appendix section A.8 demonstrates how exchange rates evolve within our sample periods.

44. Pharmaceuticals used in inpatient care (hospital market) are excluded from the analysis, because competitive bidding is used in these markets and our dataset does not contain tendered prices hospitals pay for their products.

Table 3: Matching Table

	All Markets	Generic Competition	Treatment Markets	Pre-Period Generic Entry	Pre-Study Competition	Unmatched Markets	Discarded Markets	Matched Markets
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A: Number of ATC 5 markets								
Denmark 2000	817	111	68	4	64	2	6	62
Finland 2003	903	142	103	11	92	9	20	83
Denmark 2005	822	150	115	14	101	4	18	97
Norway 2005	925	170	18	3	15	0	3	15
Finland 2009	912	139	139	15	124	13	28	111
Sweden 2009	1120	144	144	6	138	12	18	126
B: Share of overall pharmacy sales, %								
Denmark 2000	100.00	26.33	23.51	2.20	21.30	0.16	2.36	21.14
Finland 2003	100.00	38.93	35.44	5.10	30.34	2.07	7.17	28.27
Denmark 2005	100.00	33.53	32.54	5.98	26.56	0.57	6.55	25.99
Norway 2005	100.00	28.91	6.57	2.93	3.64	0.00	2.93	3.64
Finland 2009	100.00	33.67	33.67	7.69	25.98	1.30	8.99	24.68
Sweden 2009	100.00	21.99	21.99	1.16	20.83	1.06	2.22	19.77

Notes: All Markets = All markets in during pre-period ; Generic Competition = Markets with generic competition ; Treatment Markets = Markets where regulation is implemented ; Pre-Period Generic Entry = Generic entry happens during reform pre-period ; Pre-Study Competition = Generic competition started before our observation period ; Unmatched Markets = Market cannot be match to a control country ; Discarded Markets = Treated markets dropped from the sample ; Matched Markets = Successfully matched markets.

Our matching process may lead to the exclusion of some treated markets. A treated market is excluded from the sample if the control country does not have the corresponding market, or 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 retail or hospital sector, how OTC drugs are classified, or because some small markets may have experienced entry in one country but not in the other country. For example, the size of the hospital market in Denmark is larger than in the other Nordic countries, leading to a situation where Denmark could have fewer retail markets. 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 3 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 the estimation sample. Panel B shows the same information in terms of sales share.

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, there were a total of 817 (ATC 5) markets in Denmark in 2000, 111 of which had generic competition. Out of these, 68 (Column 3) are affected by the change from IRP to ERP. As can be seen from column 4, four of these 68 markets experience generic

Table 4: Number of Observations

Reform	Treatment	Control	Total
Denmark 2000	1537	1537	3074
Finland 2003	3032	3032	6064
Denmark 2005	4126	4126	8252
Norway 2005	555	555	1110
Finland 2009	3218	3218	6436
Sweden 2009	7490	7490	14980

Notes: Market-level observations by treatment status.

entry during the pre-period and are therefore excluded from the estimation sample, leaving 64 markets (Column 5) to be potentially matched with the same markets in the control country (Finland in this case). Column 6 tells us that two of these markets were not found in Finland. Thus, we end up with 62 matched markets (Column 8) after having discarded a total of 6 as reported in Column 7 (4 due to pre-period entry and 2 due to the markets not existing in Finland).

In Panel B, a comparison of Columns 1 and 2 shows that a minority of pharmaceutical sales in treated countries come from markets with competition (22%–39%) and the rest come from monopoly markets. Columns 3, 4 and 5 show the sales share of the treated markets, of markets with generic entry during the pre-period, and of markets with generic competition before the observation period. The sales share of the treated markets varies from a low of 6.6% (Norway 2005, GS to SP) to a high of 35.4% (Finland 2003, VGS to GS).[45] We lose between 1% and 8 % of sales due to generic entry taking place in the pre-period. The sales share of unmatched markets is small, as expected. In general, the samples discarded from the sample markets with a total sales share 2%–9%, and our estimation samples cover between 3.6% (Norway 2005) and 28.3% (Finland 2003) of total pharmaceutical sales, measured in pre-period sales.

In general, our matching procedure leads to a sample containing identical markets (ATC 5 classification) in the treatment and control countries, but allows for differences with respect to product variety and market structure. We do not exclude firms and products that appear only in control or in treated markets because national generic producers and parallel importers can and often do have large market shares. We also allow product entry and exit during the estimation samples, since those might have large price effects.

3.3 Descriptive Statistics

Table 4 gives the number of observations by treatment status. We perform analysis using a market (ATC 5) in a given time period as our observation unit.[46] The number of observations used in market-level analyzes varies between 555 and 7 490 observations per country.[47] Our largest estimation sample pertains to the Swedish 2009 reform (IRP - Auction-IRP), and the smallest to the Norwegian 2005 reform (IRP-SP).

We investigate the size of the different markets in the treatment and control countries in the Appendix Figure 1 which shows retail market ATC 1 level per capita sales in the pharmacy market we study, separately for below and above median sales per capita markets for the treatment and control country. A recurring finding is that Danish sales per capita are systematically lower than in comparison countries. The most likely explanation for this is the Danish practice, in which the hospital sector distributes a larger share of pharmaceuticals than in any other Nordic country.

4 Empirical Strategy

4.1 Research Design

The primary obstacle to identifying the effects of pharmaceutical price regulation policies on product market outcomes based on single-country data is that regulated products are typically not similar enough to non-regulated products within a given country. 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 probably come from markets without competition. This leads to comparisons in which treatment and control group products are in different stages of their product life cycle and the products under comparison come from different drug markets. Furthermore, in practice, price regulation reforms can also indirectly affect sales and prices of unregulated products, creating biases in within-country comparisons.

The second major challenge in evaluating the effects of (price) regulation in markets with differentiated products using the differences-in-differences approach is the SUTVA-assumption, which rules out all equilibrium effects (e.g. regulated products decrease price and as a response also nonregulated products change their prices). Previous reduced-form literature on pharmaceutical market price regulation reforms has not considered the SUTVA seriously, because often the studied outcomes have been some variants of package or trade-name prices (Brekke, Holmas and Straume 2011; Brekke, Grasdal and Holmås 2009;

45. Norway 2005 is an outlier, because reform started with a few markets and the reform expansion happens outside the study period.

46. Appendix subsection A.4 also displays the number of observations for supplementary analyzes where package-level outcomes are used instead of market-level outcomes.

47. The strong balance between treatment and control groups comes from the sample matching rules.

Herr and Suppliet 2017). This means that previous studies impose the often unrealistic assumption that products within the same market are independent of each other.

To address these shortcomings, we base our empirical strategy on cross-country comparisons at the market level (ATC 5) between Nordic countries rather than within country comparisons using product-level data. This approach allows comparisons between identical markets in different countries. Thus, the markets in comparison share similar products and demand conditions, except for the implementation of the reform.

Our identification strategy, too, is based on assumptions. First, we assume that there are no major pricing spillovers between countries (which would change due to a reform), and second, that the prices and sales of pharmaceuticals in a given ATC 5 market are comparable between countries.[48] The choice to use data from Nordic countries is that they are rather homogeneous, thus reducing unobserved heterogeneity between the treatment and control groups.

4.2 Difference-in-Differences Estimation

We estimate the effect of different price regulation policies on our primary outcomes of average expenditure, average prices, and sold quantities using different difference-in-differences strategies. Our empirical models 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 or shocks.

Besides our main analysis, we also study how implemented regulations influence short-term pharmaceutical availability and whether studied price regulation policies have spillovers to nonregulated markets. Price regulation policies can have adverse effects, and in the context of Nordic markets, the adverse effects are more likely to appear in the form of decreasing pharmaceutical availability in regulated markets[49] Previous studies have studied spillovers arising from price regulation policies, and spillovers have had economic significance (Brekke, Grasdal and Holmås 2009; Brekke, Holmas and Straume 2011). We also study what happens to markets outside the scope of the studied reform when regulation is implemented to the regulated markets, because the existence of between-market spillovers is an important factor in designing future regulation policies.

Our primary approach is using the standard Two-Way Fixed Effects (TWFE) estim-

48. 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 Nordic countries use 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. This is shown in Appendix XXX.

49. Based on the small global market share, Nordic markets do not drive pharmaceutical innovation and that is the reason why we leave innovation issues outside the scope of this study.

ator.[50] When the parallel trend assumption behind the TWFE estimator fails, we use a modern version of differences-in-differences methods, namely the estimator proposed by Callaway and Sant'Anna (2020) which imposes less strict parallel trends assumptions.

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, average price or quantity i in country c at time t. $\alpha_{i,c}$ denotes the fixed effects specific to the country observation unit controlling 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 the time 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 between markets exposed to change in price regulation relative to the same change between markets in the control country. When we estimate the average effect of the reform, we augment equation 1 by replacing $\sum_{\tau \neq -1} \beta_\tau \text{Reform}_\tau$ the term with $\beta_{att} \text{Reform}_\tau$. This allows us to interpret β_{att} as the average impact of the reform studied on the treated units.

The estimation equation for the Callaway and Sant'Anna (2020) estimator has the following form when the comparisons are based on never treated units (control country) and treated units (observations from the country that implements the reform) without conditioning on control variables X:

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

where 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. When we estimate the effects averaged over time, we use aggregation formulas derived in Callaway and Sant'Anna (2020).

The average effect of participating in the treatment of units in the 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 a comparison group (Callaway and Sant'Anna 2020). Under the maintained parallel trends assumption, discussed in detail in the next subsection, this latter path is the path of outcomes that units in group g would have experienced if they had not participated in the treatment.

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

The main benefit of using the Callaway and Sant'Anna (2020) estimator instead of the TWFE approach is that it is robust to the of the treatment effect heterogeneity related to the timing of treatment. This avoids issues related to negative weighting (Goodman-Bacon 2021). In addition, the required parallel trends assumption is weaker than when using TWFE.

We cluster standard errors at the ATC 5 level using a wild bootstrap procedure.[51] 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 Parallel Trends Assumptions

The differences-in-differences estimators we use in this paper rely on different versions of the parallel trends assumption. When the TWFE estimator is used, the assumption takes the following form:

Assumption 1 (TWFE parallel trends)**.** For each $g \in G$ and each $t \in \{2, ..., T\}$, $E[Y_t(0) - Y_{t-1}(0)|G_g = 1] = E[Y_t(0) - Y_{t-1}(0)|C = 1] = E[Y_t(0) - Y_{t-1}]$ *a.s.*

Assumption 1 implies that the parallel trends between the treatment and control group should hold for all groups and time periods. We impose Assumption 1 when we investigate the following reforms: Denmark 2000, Denmark 2005 and Norway 2005. In the case of the Finnish and Swedish reforms, we observe that Assumption 1 probably does not hold, since treatment and control group prices are diverging already during the pre-period. Therefore, we switch from TWFE to Callaway and Sant'Anna (2020) for these two reforms, since the required parallel trend assumption is weaker.

Assumption 2 (Callaway and Sant'Anna (2020) parallel trends based on the "never treated" group.)**.** Let δ define the anticipation interval. For each $g \in G$ and each $(s, t) \in \{2, ..., T\}$ such that $t \geq g - \delta$, $E[Y_t(0) - Y_{t-1}(0)|G_g = 1] = E[Y_t(0) - Y_{t-1}(0)|C = 1]$ *a.s.*

The crucial difference between the assumptions 1 and 2 is that the latter assumes only parallel trends for $t \geq g - \delta$, not for all $t \in T$. This means that Callaway and Sant'Anna (2020) does not restrict parallel trends in the pre-treatment period at all (Marcus and Sant'Anna 2020). This convenient property of the Callaway and Sant'Anna (2020) estimator allows us to investigate reforms when standard TWFE diagnostics suggest violation of the parallel trends assumption.

51. For our TWFE estimates, we use the method developed in Roodman, Nielsen, MacKinnon and Webb (2019) in the calculation of the confidence intervals. Our Callaway and Sant'Anna (2020) estimations use the Mammen (1993) method.

Table 5: Reform Timing

Reform	Start	Timing	Anticipation length	Source
Finland 2003	2003m4	2003m1	4m	
Finland 2009	2009m4	2009m1	4m	
Denmark 2000	2000m11	2000m11	0m	
Denmark 2005	2004m12	2004m12	0m	
Norway 2005	2005m1	2004m9	4m	
Sweden 2009	2009m12	2009m10	2m	

Notes: Start = Public information on when reforms started; Timing = months used for treatment analyses; Anticipation = difference between Start and Timing.

4.4 Reform Timing

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, firms or consumers react to the reform before the reform is implemented (Alpert 2016). If firms change pricing or consumers stock pharmaceuticals before reform implementation, then the results will be biased, and the direction of the bias depends on firm and consumer behavior. Modern differences-in-differences estimators allow anticipation to exist, but the start of the anticipation period must be known (Callaway and Sant'Anna 2020).

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

Table 5 shows how reform anticipation is modeled in our analyzes. Column 1 shows the actual reform start date, and Column 2 shows the date used in the analyzes. Columns 3 and 4 display the duration of the anticipation period and the legislative source for the reform start date used. Most of the reforms studied have staggered implementation, and the same anticipation length is applied to all reform cohorts.[52]

In the event study analyzes presented in Section 5, we separate the anticipation period in orange from the (blue) pre-treatment period and (green) post-treatment period point estimates. When summarizing our event-study results as reform-specific ATTs, we do not include anticipation periods in the ATT calculation.

52. Denmark 2000 and 2005 reforms do not have staggered implementation.

5 Results

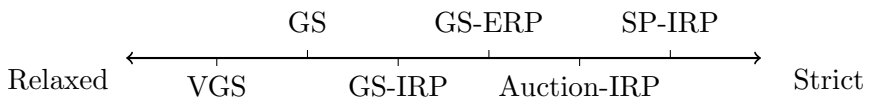


Figure 3: Reform strictness scale

We start this section by showing our main results in the event study framework. We divide the discussion of our results into three parts, starting from the less strict reforms and progressing to the stricter type of reforms. Our measure of regulation strictness is shown in Figure 3.[53] The interpretation and justifications are as follows: VGS is less strict than GS because it leaves the decision of substitution solely to the prescribing physician. GS is less strict than GS-IRP because it does not affect the level of consumer reimbursement. ERP stricter than IRP because it requires companies to submit price data outside the domestic market to the authorities. However, Auction-IRP-IRP is stricter than ERP since the government procures only one product for reimbursement. The SP system is the strictest type of regulation because the government itself decides the evolution of the reference price.

The first pair of results discussed in this section considers the Finnish transition from VGS to GS in 2003, followed by the adoption of GS-IRP in 2009. These two reforms allow us to separate the effects of GS from RP. The second pair is formed from the Danish experiment with ERP between the years 2000 and 2005. The transition from an IRP system to ERP and back allows comparison of these two types of reference pricing. Finally, our third pair considers the strictest types of reference pricing, namely the Swedish Auction-IRP-IRP and the Norwegian SP-IRP system.

For each reform, our main outcome variables are log average prices, log expenditure per DDD and sold quantities in terms of DDD. After examining the event study results, we also report the average treatment effects by reform and outcome. After the main results, we investigate whether the reforms studied have any adverse effects on product availability. We conclude this section by examining the size of between-market spillovers arising from the examined reforms. The results for our secondary outcome measures are provided in the Appendix.

5.1 Part I: GS and IRP

GS started in Finland on 1 April 2003; we set the base period in our estimations to January 2003. The results are shown in Figure 4a. We find no statistically significant effects for average prices in our event study results. However, the results suggest that immediately after implementation there was a modest 7% decrease in average expenditure, but the

53. The strictness scale is based on legislation and discussions with the market regulators.

effect decreases in magnitude and becomes statistically insignificant in later periods. The results for the quantities (in terms of doses) follow a shotgun pattern, varying from positive coefficients to negative coefficients both in the pre- and post-periods, implying inelastic consumer demand in the markets in our estimation sample. Based on our results, we conclude that there is no strong evidence that simple GS without any RP leads to savings in terms of expenditure.

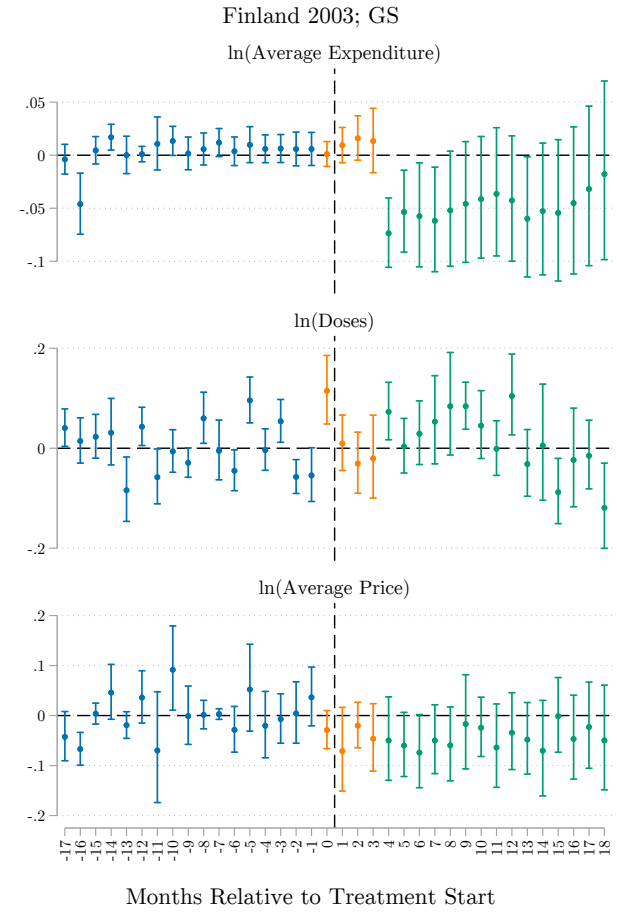
Unlike the GS system, our analysis of the 2009 IRP reform provides different results. The GS system was expanded with IRP in April 2009. The base period of our estimation is set to January 2009. Based on our event study results, as shown in Figure 4b, we find that the adoption of IRP decreased the average expenditure by 25% after a year of reform implementation. As expected, our results for the average prices per dose are more moderate: a decrease between 15%–20% by the end of the sample period, a result that is also statistically significant. Like before, we do not find significant effects on the total quantities.

There is a strong case to be made for why GS alone has only a limited effect on prices, aggregate quantities, and consumption. First, GS simply expands the choice set of consumers when they shop in pharmacies. Although a generic alternative might have a lower copayment than the original branded product, deciding against the substitution does not affect the level of reimbursement or the copayment faced by the consumer, thus mitigating its effects on consumers. For fully reimbursed products or consumers who have exceeded their annual aggregate co-payment cap, there are only limited incentives to accept substitution. To test this theory, we also perform a subgroup analysis based on the different reimbursement categories of the products, the results of which are provided in Appendix A.10 Figure 16.[54] 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 which were less affected by the earlier GS reform in 2003.

The results of the two Finnish reforms should be interpreted jointly. Unlike 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 widened the choice set of consumers because they were no longer tied to the decision made by their prescribing physician. However, the increase in competition had no significant effect on the expenditure per dose. Second, we show that GS with

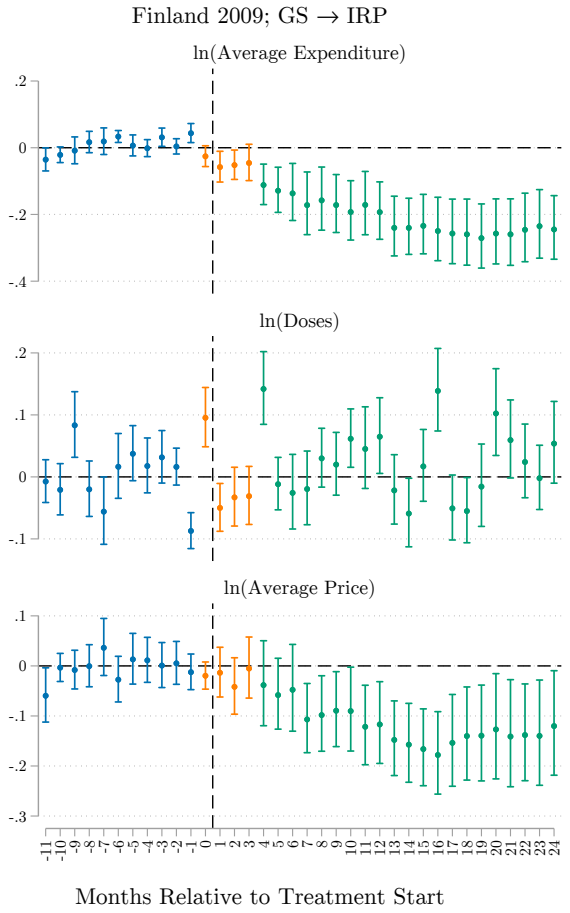
54. This subgroup inquiry is based on product (package) level analysis instead of market level analysis, because the reimbursement statuses are defined on the package level. This analysis is prone to general equilibrium violations, but the idea is to demonstrate the weakness of the GS policy design when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

IRP decreased average expenditure per dose by more than 20%. The result indicates that the inclusion of consumer incentives has an economically significant effect, since average expenditure decreases are 2–3 times greater than in the case of Finnish 2003 GS regulation. Our package-level subgroup analysis further suggests that for reimbursed products, price decreases are achievable when GS is tied to RP for reimbursed products.



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. Denmark used as a control group. Outcome data source: DLI-MI and FIMEA.

(b) Finland 2009

Figure 4: Part I: Main Results

5.2 Part II: ERP

This subsection examines the Danish experiment with ERP between the years 2000–2005. These results are important because this is the first study in which the effects of ERP switches are evaluated using the differences-in-differences approach.[55] Our results are shown in Figures 5a and 5b. When ERP is implemented, we find statistically significant results that, on average, average expenditure per dose decreased by 8%, average wholesale price decreased by 10% in Denmark compared to the control country Finland. Surprisingly, we don't find any statistically significant demand effects. The results are highly symmetric when studying the change in 2005 from ERP to IRP: we find that average expenditure increased by roughly 10%, average price by 10% and demand seems to increase almost by 5%.[56]

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 in off-patent markets in Denmark were higher than in its reference countries. Surprisingly, Danish 2005 reform, where ERP was replaced again by IRP leads to increases in average expenditures and in average prices.[57] Danish price regulation system does not impose strict price caps for pharmaceutical products, and this makes it possible that prices can increase as an unintended consequence of the implemented regulation policy.

The results of abolishing the ERP system presented in Figure 5b are visible only a few months after the implementation of the reform. We point out that this probably is due to the fact that firms slowly re-optimize their prices and since the difference-in-differences 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.[58] We have also estimated the effects beyond the 12 lags shown in Figure 5b showing that the positive effects dynamically increase.[59] However, these estimates of the lags $t > 12$ are upwards biased because of a mandatory price decrease in the control country (Finland) at the start of 2006. The cut was set to exactly 5% and

55. The baseline in the of 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.

56. A noticeable result in the quantity and expenditure analysis is the sharp increase and immediate decrease at lags nine and eight in the pre-period Denmark 2000 reform. These lags correspond to March and April 2000, a time when Denmark transferred from a product-specific reimbursement system to a needs-based reimbursement system. The reform created an incentive for patients to bring forward their prescription purchases before their reimbursement levels were reset. Our results suggest that, on average, consumers increased their purchases by roughly 40% in terms of doses. This was followed by decreases in purchases by up to a quarter of a year later. Surprisingly, pharmaceutical companies did not take advantage of the sharp increase in demand by temporarily increasing their prices. Our estimation of these lags replicates and confirms the results of Simonsen, Skipper, Skipper and Christensen (2021).

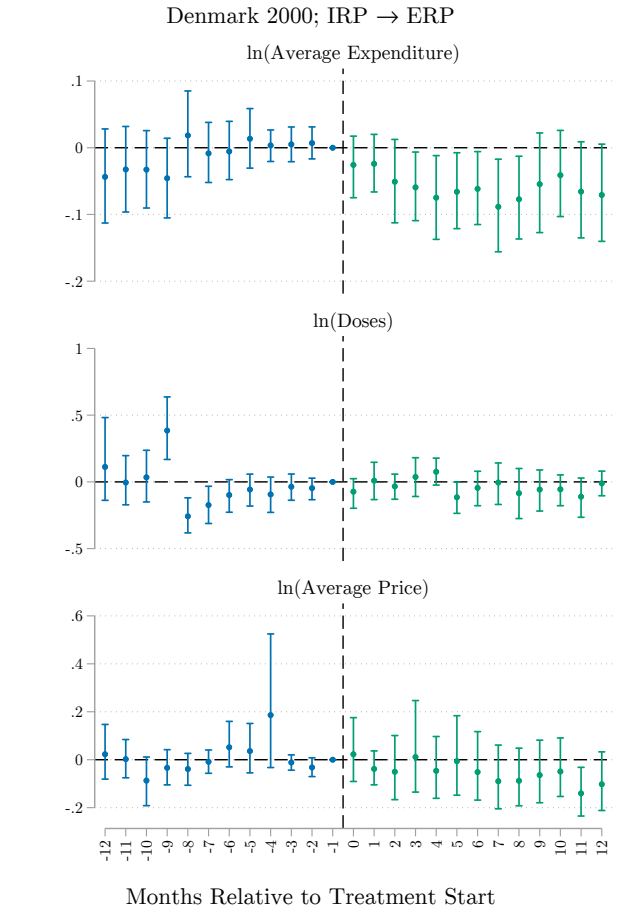
57. In the Appendix A.11 we study the effects of ERP transitions for non-competitive markets, and we find weaker but similar asymmetry also for the non-competitive markets.

58. Figure 20 in Appendix subsection A.12 shows this.

59. These results are also shown in Appendix A.12.

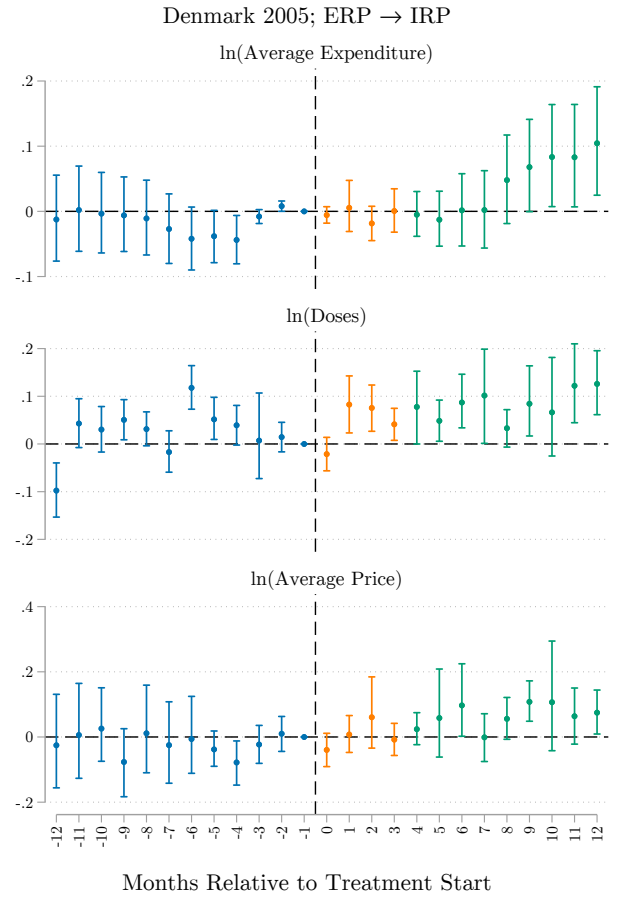
applied to all publicly reimbursed products. Taking this into account, our results imply increases in average expenditure roughly above 10% up to 30 lags after the legislation was passed in the Danish Parliament.

In particular, our analysis on the Denmark 2005 reform comes to a different conclusion on the effects of the reform than the previous study by Kaiser, Mendez, Rønde and Ullrich (2014) who studied the same reform. In their study, the authors employed a simple before-after setup for the reform price effects that does not take into account the general negative trend in pharmaceutical prices at the time in the studied markets, and the authors concluded that the switch from ERP to IRP decreased pharmaceutical prices. Our replication of the before-after setup is provided in Appendix Figure 19. We demonstrate the negative trend in the treatment and control group markets in Figure 20, which implies that the results of their study reflect the decrease in the price trend rather than the effect of the change from ERP to IRP. This highlights the benefits of using differences-in-differences research design compared to a simple before-after comparison.



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

5.3 Part III: Extreme IRP Variants

This subsection discusses the results of the Norwegian 2005 SP reform and the Swedish 2009 IRP auction reform. These reforms are labeled extreme IRP reforms because these reforms use polar ways of determining the reference price in regulated markets. SP reform is categorized as extreme because pre-specified government rules determine the evolution of the reference price after patent expiry, and in the auction system, the winner of the first price auction receives the entire market demand for each month. Analysis of these policies that affect pricing and expenditure offers new aspects on how to design future price regulation measures, as these reforms can be used to complement existing IRP regulations.

We start the discussion of the extreme type IRP's with the Norwegian SP reform. The baseline in our estimations is set to September 2004 when the reform was introduced in the parliament of Norway. Our results are shown in Figure 6a. They show that, on average, the implementation of the Norwegian SP reform decreased average expenditure per dose by approximately 20%. We find no effect on average wholesale prices, showing that simple price measures could not capture all relevant aspects of the implemented reform.

The regulatory environment in Norway assigns the same price cap for both the original patented product and the generic alternatives, and the price cap is calculated as an average of the prices of the original products in other 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, it can be expected to be the largest for products for which the price cap was binding.[60] The reform also required pharmacies to have at least one product at or below the reference price (SP) in stock. In the vertically integrated Norwegian retail market, wholesale prices for pharmaceuticals are in part just internal prices of the pharmacy chains. After the adoption of the SP model, the pharmacy chain had little incentive to sell generics in stock below the maximum wholesale price (the price cap) or the maximum retail price (the price cap with the retail margin).

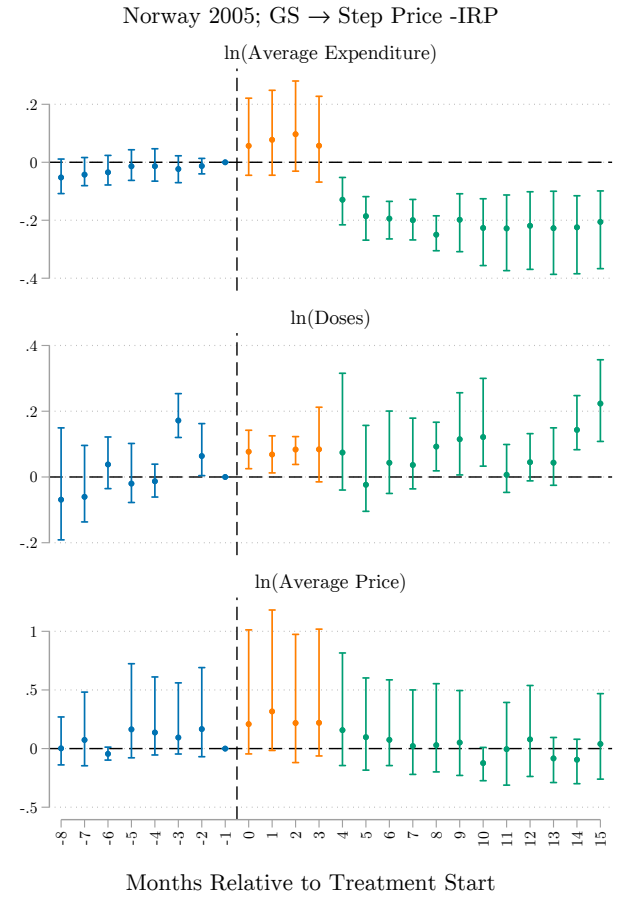
The results of the Swedish Auction-IRP reform are reported in Figure 6b. The baseline in the estimations is set to November 2009 when companies were asked to submit their bids for the first time. Our estimation parameters suggest that average wholesale prices decreased on average by approximately 10%. Not surprisingly, the effects on average expenditure are larger in magnitude. We find that the reform led to a statistically significant decreases between 10%–20% in average expenditure per dose. This implies, as hypothesized, that the reform decreased the minimum prices within substitutable products more than the overall average price within each active ingredient market. We find no statistically significant results for quantity, although most point estimates in the post-period are positive and have economically significant size.

As discussed in Section 2, Sweden's Auction-IRP system is in practice a monthly auction where the policymaker *procures* the pharmaceuticals included in the tax-funded funded social insurance. This is due to the fact that consumers can only be reimbursed for the

60. Appendix table 12 shows the Stepped Price rule for the time period we use in our analysis.

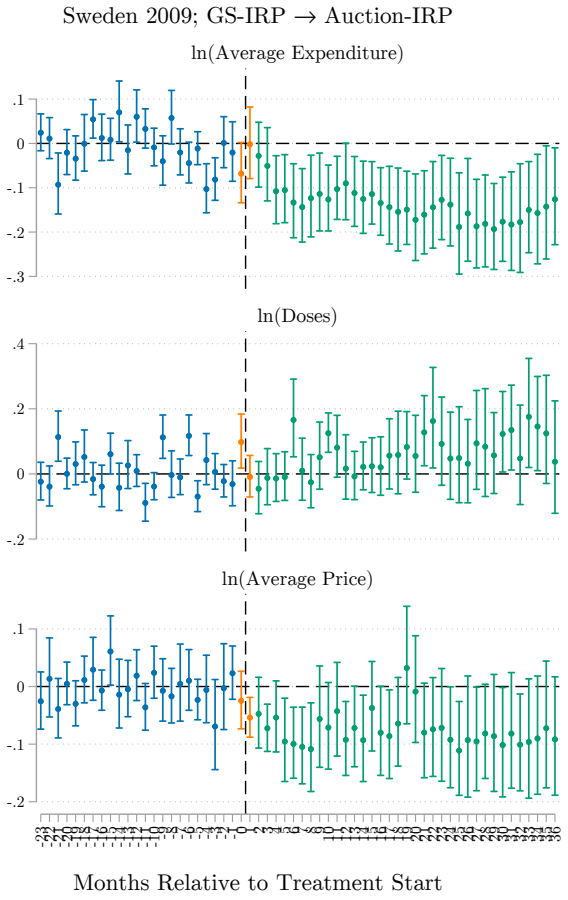
product they receive the prescription or the substitutable product decided by the government in a monthly bidding process. This leads to a stark form of the general point we have made earlier: average market prices do not matter; what matters is average expenditure, that is, the average expenditure per dose. In the Swedish case, if all or almost all sales are channeled to the cheapest product in the market, the prices of the other products become (almost) irrelevant, yet the estimation strategy employed in the existing studies and by us in the product-level price analyses does not take these into account, and this dilutes the estimated price effects.

The importance of the results for the Swedish auction-like system should not be underestimated. By directly restricting the choice set of consumers within the public reimbursement system to the prescribed product (which, in practice, is often the original patented product) and the cheapest generic substitute, the generic competition is at its extreme. In practice, the winner of the monthly auction can expect to gain almost the entire market. Our results suggest that the Swedish reform works very well from the point of view of curtailing expenditure—the objective of pharmaceutical price regulation. This fact emphasizes the contribution of our approach, as the established estimation strategy in the empirical literature would have missed this important result.



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 and IQVIA MIDAS.

(b) Sweden 2009

Figure 6: Part III: Main Results

5.4 ATT Results

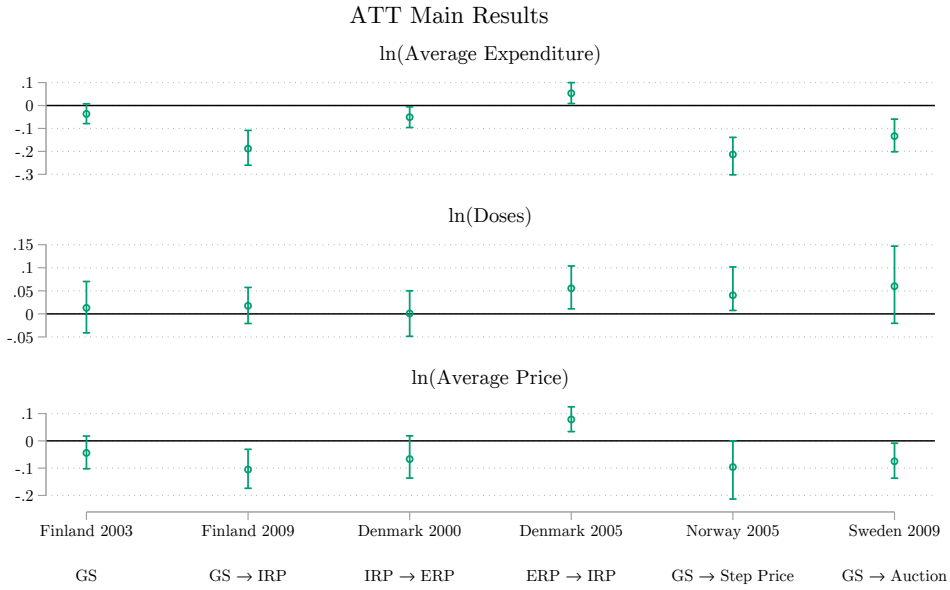
We use the ATT to summarize the information presented in our event study estimations.[61] Figure 7 displays the reform ATTs by outcome and reform with three main results. First, the reference price regulation has the tendency to decrease average expenditure per dose. The only clear outlier is the Denmark 2005 reform, where switching from ERP to IRP increased the expenditure per dose. An important result of policy design is the difference between Finland 2003 GS and 2009 IRP reform results. The Finnish 2009 IRP reform added reference pricing on top GS. The results indicate that the average expenditure effect of the 2009 reform is almost three times larger in absolute value compared to the less strict 2003 reform. The Finnish reforms that increase competition without reference price regulation are significantly less effective in creating savings in terms of average expenditure.

Second, the reference price calculation method has great significance for the savings created. Our symmetric results of the Danish reforms show that ERP worked better in Denmark than the IRP systems in decreasing prices and average expenditure per dose. We argue that this is due to the special characteristics of the Danish ERP system: Unlike many other ERP systems, the Danish system was actually a generalization of an RPS. This is due to the fact that the ERP system was also in practice an IRP system in markets where the domestic price level was below the foreign price levels. Thus, at the extensive level, our results suggest that ERP is a more efficient version of RP than IRP in decreasing average expenditure.

However, the calculation of the reference price also has significance at the intensive margin. Norway 2005, Finland 2009 and Sweden 2009 are different variants of IRP. The regulation in Norway imposed predetermined price decreases; the Swedish regulation commits market demand to the cheapest product through an auction-based mechanism; and the Finnish reform is a milder version of the Swedish one. Despite the fact that the point estimates of the expenditure effects are close to each other, it is crucial that the estimates are relative measures and do not consider levels (e.g. in prices and expenditure). Although the point estimates for the Finnish IRP reform and the Swedish Auction-IRP system have the same size, the latter was only a change on how the already existing reference prices were calculated. Therefore, Sweden was able to induce more savings by adopting a new mechanism to calculate the reference price.

Our third main result is that, on average, demand is rather inelastic. In all reforms examined, the quantity seems to increase as prices decrease. The extensive social security programs in the Nordics explain this finding at least partially. These programs try to ensure that people get the medications they need regardless of their income level, and this may explain why quantity responses are quite small compared to price or expenditure outcomes.

61. Section 4 on identification strategy explains how the ATTs are derived in the case of TWFE and Callaway and Sant'Anna (2020) approaches.



Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). replications for ATC-5 wild bootstrapped standard errors. Outcome data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

Figure 7: ATTs by outcome and reform

5.5 Price Regulation Policies and Pharmaceutical availability

Previous sections showed that price regulation policies implemented in the Nordic pharmaceutical market have decreased average market-level expenditure. A common concern and source of criticism of pharmaceutical price regulation policies are their possible adverse effects on pharmaceutical availability and competition. Previous 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). Previous literature on shortages implies that our earlier findings could indicate that pharmaceutical shortage issues could increase due to the adopted reforms.

We study the effects of price regulation policy changes on pharmaceutical availability in the short run by re-estimating event study models presented in Section 5 with availability outcomes. We use the number of packages (defined by their Nordic Article Number (VNR) code) and the number of product names sold in each market i in the country c at the time t as our measures of pharmaceutical availability. We focus on all markets studied in the earlier analysis and separately also on markets with injectable drugs.[62] Previous literature

62. Market belongs to the injectable sample if sold product types contain injections or infusions.

on pharmaceutical shortages has pointed out that sterile injectable drugs are prone to have availability problems due to the nature of drug manufacturing, storage, and use conditions (Yurukoglu, Liebman and Ridley 2017).

Our conclusion is the same for all reforms, both availability outcomes and samples: we find no economically and statistically significant average effect on pharmaceutical availability in the Nordic markets. Most significantly, the point estimates are negative only for the number of packages for the Danish 2000 ERP adoption, and even in that case we find no effects on the number of product names in the market. This result is especially striking in the case of the Swedish 2009 Auction-IRP reform, where in practice the winner of the auction could expect to gain most of the sales in the market. The event study results and ATT-tables are listed in the Appendix subsection A.13.

Yet, we provide some descriptive evidence of long-term net entry of new products and firms into the Nordic market structure. Figure 24 shows the number of firms in all the markets studied (ATC 5 active ingredients) in the Nordic countries.

However, our results should not be taken as definitive evidence that price regulation has no effect on pharmaceutical availability in the Nordics. Our analysis is limited at most to 36 lags or three years of the post-reform period, and thus we are unable to study the long-term effects of the studied reforms. Our analysis does not cover all aspects of the availability theme. We focus only on markets that are included in the main analysis, and issues related to entry delays are beyond the scope of our analysis (Maini and Pammolli 2022).

5.6 Spillovers Between Pharmaceutical Markets

Previous sections analyzed the impact of price regulation on regulated markets. The previous literature has also examined the spillovers of regulation to outside markets. Since not all markets (ATC 5 active ingredients) used in the treatment of certain diseases are included in the given regulation, it is possible that outcomes change also in the non-regulated markets when regulation influences pricing in the treated markets. Spillovers like these are labelled as *therapeutic competition*, and in some studies, the magnitude of the effect of therapeutic competition on prices has been economically significant (Brekke, Grasdal and Holmås 2009; Brekke, Holmas and Straume 2011).

The sample matching used in the analysis of spillovers follows the same ideas as in the main analysis, but now the sample contains only markets which are not included in the analyzed reform and have a matching market in the control country. We use the same estimation methods as in the main analysis. The Danish reforms (2000 and 2005) are excluded from the spillover analysis because these reforms influence all products in the Danish pharmaceutical market.[63] This section reports only the reform ATTs and the Appendix subsection A.14 reports the same results in an event study form.

63. The results for the monopoly Danish markets are discussed alongside Danish main results and results are provided in the Appendix.

The main result from the spillover analysis is that the effects are much smaller than what previous studies have found (Brekke, Grasdal and Holmås 2009). Our point estimates for the ATTs are close to zero in terms of magnitude. Furthermore, these estimates are also statistically precise zeros. This implies that even if certain illnesses can be treated with different active ingredients, pricing decisions in the non-treated markets are not greatly influenced by pricing in the treated markets. In comparison to the average expenditure results, the quantity results are more interesting. Changes in physician behavior can create links between demand in regulated and non-regulated markets. Cost-conscious physicians might start to write more prescriptions from markets where competition has lowered prices, and this could decrease sold quantities and sales in non-regulated markets if products have the same treatment response. The estimated average spillover effect on sold quantity varies between 0% and 7%, but no spillover effect is statistically significant. This result also supports the idea that our definition of the relevant market at the ATC 5 active ingredient level is narrow in the context of pharmaceutical markets. However, the ATT estimate for the quantities sold in the Swedish Auction-IRP reform is close to statistically significant, and the event study graph suggests an increase in the quantities sold after two years of reform implementation.

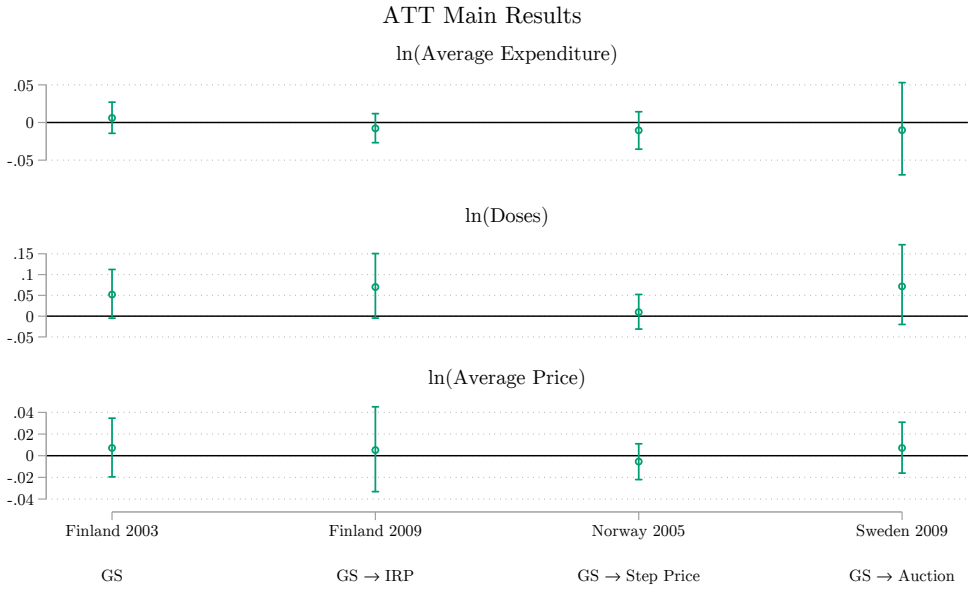
The main take-away from the analysis presented in this subsection is simple. Direct regulation measures are much more effective than indirect measures in decreasing pharmaceutical prices and expenditure in the countries we study. These results help to design new ways to reduce pharmaceutical expenditure by providing evidence that indirect effects of price regulation may be insufficient to create cost savings and provide affordable medicines. The results also build trust on our parallel trends assumptions because the cross-country non-treated markets also seem to follow the same paths in terms of expenditure and prices as shown in the event study graphs in the Appendix A.14.

6 Conclusions

We investigate the causal effects of different price regulation policies on market outcomes in the Nordic pharmaceutical markets. Such policies are important because it is empirically known that pharmaceutical spending has been increasing and because of the institutional fact that, in many countries, public (or private) health insurance schemes have 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 implemented during 1999–2010 in the Nordic countries and employ modern difference-in-difference methods to analyze the effects of the reforms. We provide new evidence on how different variants of widely used regulations influence pricing and whether they yield savings in pharmaceutical expenditure.

Interestingly, we found that several price regulation reforms decrease public expenditure considerably and proportionally more than average pharmaceutical prices. Despite these



Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). replications for ATC-5 wild bootstrapped standard errors. Outcome data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

Figure 8: Spillover Effects by Reform

effects on expenditure and prices, we did not find that regulation affects pharmaceutical availability in the Nordic markets.

References

Alpert, Abby. 2016. 'The anticipatory effects of Medicare Part D on drug utilization' [in en]. *Journal of Health Economics* 49 (September): 28–45. (Cited on page 24).

Bhattacharya, Jay, Timothy Hyde and Peter Tu. 2013. *Health Economics.* Macmillan International Higher Education. (Cited on page 7).

Brekke, Kurt R., Astrid L. Grasdal and Tor Helge Holmås. 2009. 'Regulation and pricing of pharmaceuticals: Reference pricing or price cap regulation?' [In en]. *European Economic Review* 53, no. 2 (February): 170–185. (Cited on pages 4, 5, 11, 20, 21, 37, 38, 44).

Brekke, Kurt R., Tor Helge Holmas and Odd Rune Straume. 2011. 'Reference pricing, competition, and pharmaceutical expenditures: Theory and evidence from a natural experiment' [in en]. *Journal of Public Economics* 95, no. 7 (August): 624–638. (Cited on pages 4, 5, 11, 20, 21, 37, 44).

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2020. 'Difference-in-Differences with multiple time periods' [in en]. *Journal of Econometrics* (December). (Cited on pages 6, 22–24, 35).

Correia, Sergio. 2016. *Linear Models with High-Dimensional Fixed Effects: An Efficient and Feasible Estimator.* Technical report. (Cited on page 22).

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. (Cited on page 5).

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. 'Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects' [in en]. *American Economic Review* 110, no. 9 (September): 2964–2996. (Cited on page 6).

Dubois, Pierre, Ashvin Gandhi and Shoshana Vasserman. 2021. *Bargaining and international reference pricing in the pharmaceutical industry.* Technical report. Technical report. (Cited on page 5).

———. 2022. *Bargaining and International Reference Pricing in the Pharmaceutical Industry.* Working Paper 30053. National Bureau of Economic Research, May. (Cited on page 62).

Dubois, Pierre, and Laura Lasio. 2018. 'Identifying Industry Margins with Price Constraints: Structural Estimation on Pharmaceuticals' [in en]. *American Economic Review* 108, no. 12 (December): 3685–3724. (Cited on pages 2, 5, 6).

Goodman-Bacon, Andrew. 2021. 'Difference-in-differences with variation in treatment timing' [in en]. *Journal of Econometrics* (June). (Cited on page 23).

Herr, Annika, and Moritz Suppliet. 2017. 'Tiered co-payments, pricing, and demand in reference price markets for pharmaceuticals' [in en]. *Journal of Health Economics* 56 (December): 19–29. (Cited on pages 4, 20).

Kaiser, Ulrich, Susan J. Mendez, Thomas Rønde and Hannes Ullrich. 2014. 'Regulation of pharmaceutical prices: Evidence from a reference price reform in Denmark' [in en]. *Journal of Health Economics* 36 (July): 174–187. (Cited on pages 30, 43, 63, 64).

Lakdawalla, Darius N. 2018. 'Economics of the Pharmaceutical Industry' [in en]. *Journal of Economic Literature* 56, no. 2 (June): 397–449. (Cited on page 6).

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. (Cited on pages 6, 36).

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. (Cited on page 47).

Maini and Fabio Pammolli. 2022. 'Reference Pricing as a Deterrent to Entry: Evidence from the European Pharmaceutical Market' [in en]. *American Economic Journal: Microeconomics Forthcoming*. (Cited on pages 5, 37, 62).

Mammen, Enno. 1993. 'Bootstrap and wild bootstrap for high dimensional linear models'. *The annals of statistics* 21 (1): 255–285. (Cited on page 23).

Marcus, Michelle, and Pedro H. C. Sant'Anna. 2020. 'The Role of Parallel Trends in Event Study Settings: An Application to Environmental Economics'. *Journal of the Association of Environmental and Resource Economists* 8, no. 2 (October): 235–275. (Cited on page 23).

Pavcnik, Nina. 2002. 'Do Pharmaceutical Prices Respond to Potential Patient Out-of-Pocket Expenses?' *The RAND Journal of Economics* 33 (3): 469–487. (Cited on pages 4, 5).

Roodman, David, Morten Ørregaard Nielsen, James G. MacKinnon and Matthew D. Webb. 2019. 'Fast and wild: Bootstrap inference in Stata using boottest' [in en]. *The Stata Journal* 19, no. 1 (March): 4–60. (Cited on page 23).

Simonsen, Marianne, Lars Skipper, Niels Skipper and Anne Illemann Christensen. 2021. 'Spot price biases in non-linear health insurance contracts' [in en]. *Journal of Public Economics* 203 (November): 104508. (Cited on pages 7, 9, 29, 44).

Stomberg, Christopher. 2016. 'Drug Shortages, Pricing, and Regulatory Activity'. In *Measuring and Modeling Health Care Costs,* 323–348. University of Chicago Press, October. (Cited on pages 6, 36).

Yurukoglu, Ali, Eli Liebman and David B. Ridley. 2017. 'The Role of Government Reimbursement in Drug Shortages' [in en]. *American Economic Journal: Economic Policy* 9, no. 2 (May): 348–382. (Cited on pages 6, 36, 37).

A Appendix

A.1 Regulation in the European Common Market

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 the legislation passed by the EU. The fourth option, national market authorization, is regulated by the Member States themselves. In the centralized procedure, the authorization process is applied through the European Medicines Agency (EMA) through which the authorization is valid in the European Economic Area (EEA).[64] In the decentralized process, a company simultaneously applies for market authorization from more than one Member State through the 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.

A.2 Minor Price Regulation Reforms in the Nordics

Section 2 of the main text describes the reforms that we study in this paper. During the periods our estimation samples cover, Nordic countries implement reforms which we categorize as minor. These reforms create changes in the way pharmaceuticals are priced, reimbursed, or how regulation mechanics are tweaked.[65] Table 6 shows two minor reforms from Denmark, one from Norway, two from Finland, and four from Sweden.

Denmark. During our sample periods, the Danish regulator makes price cap or price "freeze" agreements with pharmaceutical firms represented by the Danish Association of the Pharmaceutical Industry (LIF). Not all firms present in the Danish market are represented by LIF, and this leads to a situation where price caps are not imposed on all products. The idea of these agreements is that the Danish regulator and firms agree that a market price from a certain date acts as the price cap for the time agreement is valid. These price agreements were in place during the periods in which we use Danish sales data in our analyses.[66]

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'

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

65. Minor 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 our study focuses on prescription drugs.

66. 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.

Table 6: Minor Price Regulation in the Nordics 2001-2012

Country	Year	Reform Type
Denmark	2001–	Price Caps
Denmark	2002–	Reimbursement system overhaul
Norway	2003–2004	IRP for 8 active ingredients
Finland	2006	5% Price cap cut for reimbursed products
Finland	2006–	Price cap rule for generic entrants
Sweden	2009	Mandatory price cap cut
Sweden	2011–	Mandatory price caps in substitution groups
Sweden	2011–	Redefinition of substitutable products in Auction-IRP
Sweden	2012	Back-up winners in Auction-IRP system

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 Denmark 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. It is interesting to see that the change in the reimbursement system does not influence average expenditure or average prices.

Norway. The only minor change in price regulation in Norway during our sample period was the IRP-experiment (Index-Price) for a subset of eight active ingredients. This policy was in place between 2003–2004. Index-Price policy was an IRP variant similar to the Finnish 2009 policy.[67] This means that the Index-Price policy change happens during the pre-period of SP reform. In our analyses, we discard the markets where index-price regulation was implemented to ensure that all treated markets have the same pre-period regulation regime. Brekke, Grasdal and Holmås 2009 mentions that Index-Price policy was shutdown because the policy did not achieve the desired amount of cost savings and price reductions.

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 mechanically led to a situation where wholesale and

67. 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 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; Brekke, Grasdal and Holmås 2009 for more details.

retail prices decreased when the price cap was binding.[68] These price cuts can indirectly influence the Danish 2005 ERP reform and the Norwegian SP reform, because in these comparisons Finland is used as the control group. We deal with this issue by constraining the main text sample period to the time before the price cut and in the appendix 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 price 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 our empirical analyses like the implemented price cut, because the markets we study have had generic entry before our observation period.

Sweden. Minor Swedish reforms during our sample period are related to price cap implementation and minor tweaks in the contents of the Auction-IRP regulation. In 2009 Sweden introduced a mandatory one-time price cap cut for originator products in markets with substitutable products and generic competition.[69] 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.[70] The cut was a 65% decrease from the price of the originator product which prevailed 12 months before the expiration of the patent.[71] The price cap decrease was planned so that after the price cut the originator price cannot be lower than the cheapest comparable generic product. The Swedish National Board of Health and Welfare has documented that the price cut decreased annual expenditure by 400 million SEK.[72]

The rice cap regulation was changed in 2011 and the regulation contains two phases.[73] 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 started in the substitution group and at least one product within the

68. See 885/2005 for additional details.

69. See Price cut announcement for additional details.

70. The Finnish 2006 price cut was proposed by the market regulator, but the Swedish cut resembles the Danish price freeze agreements that are based on the negotiations between pharmaceutical industry and the government.

71. 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.

72. 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.

73. TLVFS 2009:4

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 price for 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.[74] Prior to this change, substitution would be made towards the cheapest package within the substitution group, which favored small packages. After the definition change Auction-IRP would be considered within 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.[75] Idea behind the change was that pharmacies can substitute to the backup products (second or third cheapest product in the auction) if the auction winner has problems in supplying the market.

A.3 Nordic Countries

This subsection provides an overview of the four Nordic countries included in this study. Figure 9 shows these Nordic countries on a map, and Table 7 displays the descriptive statistics of the four countries. As seen from the table, all countries except Norway are EU member states. All four countries however belong to the EEA. 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 over 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. The highest GDP per capita in 2007 was the highest in Norway, while Finland's GDP per capita was the lowest. Sweden had the largest pharmaceutical market with a turnover of more than 2.7 billion euros in 2007, while Norway had the smallest market with a size of 1.46 billion euros. The Swedish pharmaceutical market also had the highest contribution to total GDP, as it represented more than 8.5% of Sweden's GDP in 2007. In Finland and Denmark, the pharmaceutical market accounted for roughly 6.3% of their GDP, respectively, and in Norway it only accounted for 3.31% of GDP in 2007.

74. TLVFS 2009:4
75. TLVFS 2009:5

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 in PPP in 2007. Market share denotes the share that the pharmaceutical market forms of the country's total GDP. Sources: Nordic Statistics, OECD, **main data (was not sure how to mark this)**

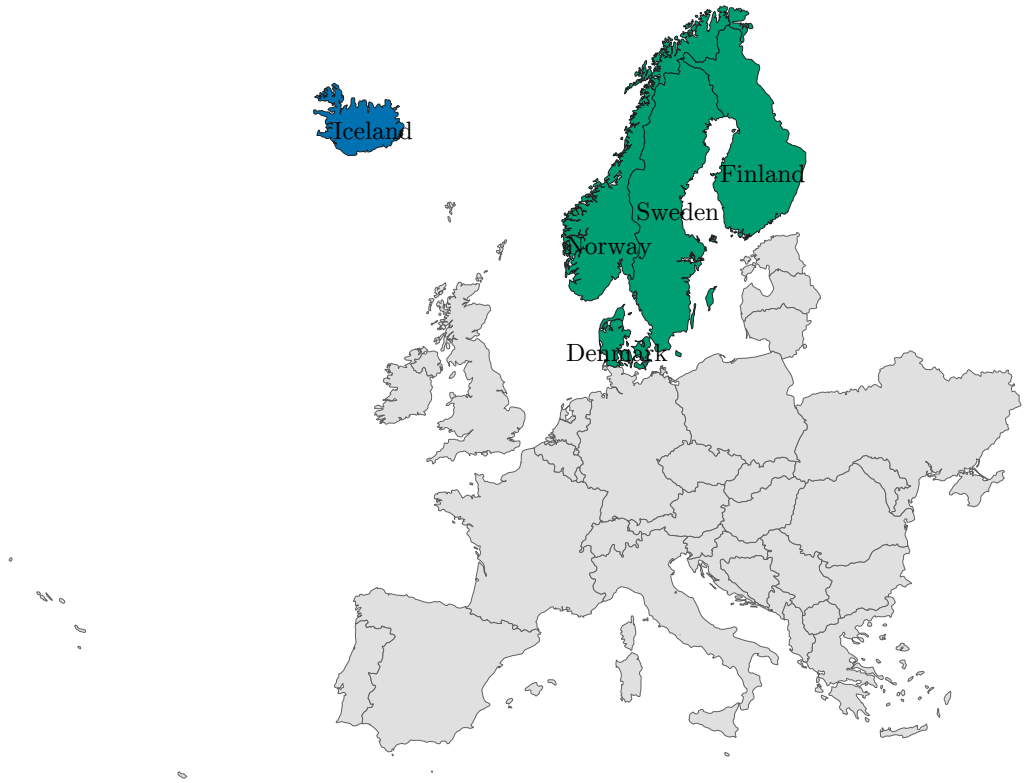
Table 8 summarizes the structure of reimbursement systems in Nordic countries. Countries are divided into product (Finland and Norway) and consumption (Denmark and Sweden) based reimbursement systems. Despite similar reimbursement systems, countries have different reimbursement rates and annual out-of-pocket ceilings. Finnish reimbursement system is the least generous, because the smallest reimbursement rate is 42 % and the annual out-of-pocket cost ceiling is 610 euro, which is almost three times higher than in Norway or Sweden.

Table 8: Reimbursement Systems in the Nordics

	Reimbursement % **	Annual out-of-pocket ceiling	Out-of-pocket threshold	Time period for annual ceiling	Reference countires, 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: * : 2005, ** : 2006, *** : 2007. Annual reimbursement expenditures are expressed in millions. Sources: PPRI, KELA and Leopold, Vogler, Mantel-Teeuwisse, Joncheere, Leufkens and Laing (2012).

Figure 9: Nordic Countries in Europe



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

A.4 Descriptive Statistics on the Nordic Countries

Figure 10 illustrates how three different substitutable products (the original patented product and two generic copies) look like. The sample packages are from Finland and the products contain the active ingredient venlafaxine (ATC5: N06AX16), which is an antidepressant. In the Finnish GS policy, these packages can be substituted for each other in the pharmacy because the products belong to the same substitution group. If the patient received a prescription for the brand name product, then substitution can be made for the products in the bottom of the figure.

The brand-name product is placed at the top of Figure 10 and the brand-name product has a novel product name "Efexor", the generic product is on the bottom left, and a parallel imported product is bottom right. All packages provide information on package size (98 tablets), strength (150 mg) product type, product id (VNR) and the firm that sells the product. The packages contain pharmaceuticals and detailed instructions related to pharmaceutical use and information on possible adverse effects on the use of the product.

Figure 10: Example of a brand (top), generic (bottom left) and parallel (bottom right) import pharmaceutical



A.5 Additional Descriptive Statistics

The main text presented descriptive statistics of market-level estimations and samples. In this Appendix subsection, we provide additional descriptive statistics and results using package-level prices. Table 9 shows the number of observations used in these analyzes. The number of observations used in the analysis of package-level prices is shown in Panel A, and the number of observations used in main text market level analyzes are shown in Panel B.[76] In product level analyzes, the number of observations varies from 24,809 to 28,723 observations, and in market level analyzes between 555 and 7,490 observations. The largest estimation sample is the Swedish 2009 reform (IRP-Auction-IRP), and the smallest is the Norwegian 2005 reform (IRP-SP).

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

76. Market-level information was already reported in the main text.

Table 9: Number of Observations

	Treatment	Control	Total
Panel A: Product			
Denmark 2000	33661	25577	59238
Finland 2003	49753	53088	102841
Denmark 2005	74038	66625	140663
Norway 2005	9543	15266	24809
Finland 2009	58965	54054	113019
Sweden 2009	162791	118932	281723
Panel B: Market			
Denmark 2000	1537	1537	3074
Finland 2003	3032	3032	6064
Denmark 2005	4126	4126	8252
Norway 2005	555	555	1110
Finland 2009	3218	3218	6436
Sweden 2009	7490	7490	14980

Notes: Product = Product-level observations, Market = Market-level observations.

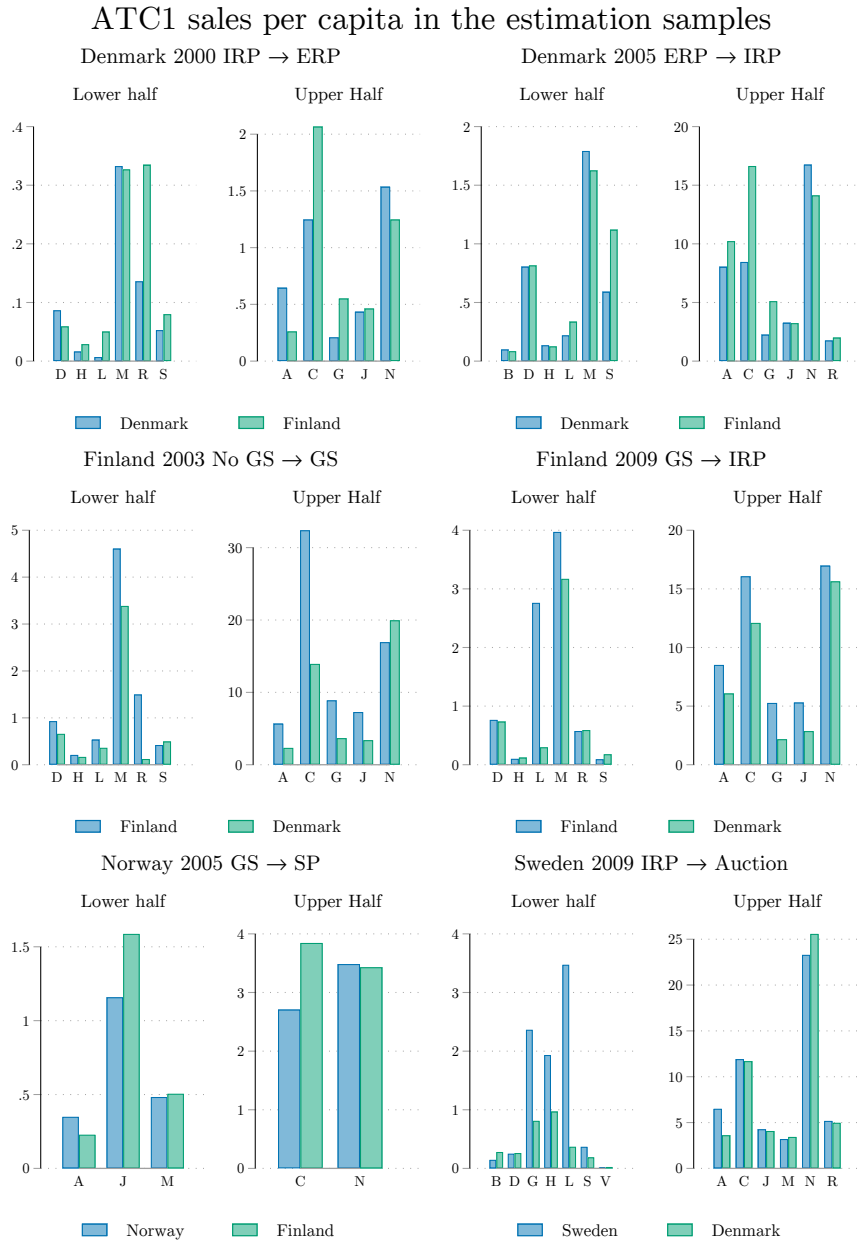
group is larger than the control group. For market-level analyzes, the samples are balanced by design.

Figure 11 shows sales per capita in the retail market at the ATC 1 level in the pharmacy markets investigated, separately for the markets below and above the median sales per capita for the treatment and control country. A recurring finding is that Danish sales per capita are systematically lower than in the comparison countries. The most likely explanation for this is the Danish practice, in which the hospital sector distributes a larger share of pharmaceuticals than in other Nordic countries. This is further illustrated in Figure 12, which shows differences between countries in hospital market share for Nordic countries. Market shares are based on monetary values and wholesale prices.[77] This comparison shows that the hospital market share has been quite stable in Finland, Sweden, and Norway during the sample periods. However, Denmark is a clear exception, because hospital market share is increasing during the sample period. During 2000–2012, the market share decreased from around 70 % to less than 50 % in Denmark.

Large increases in the hospital market share can be problematic in our cross-country

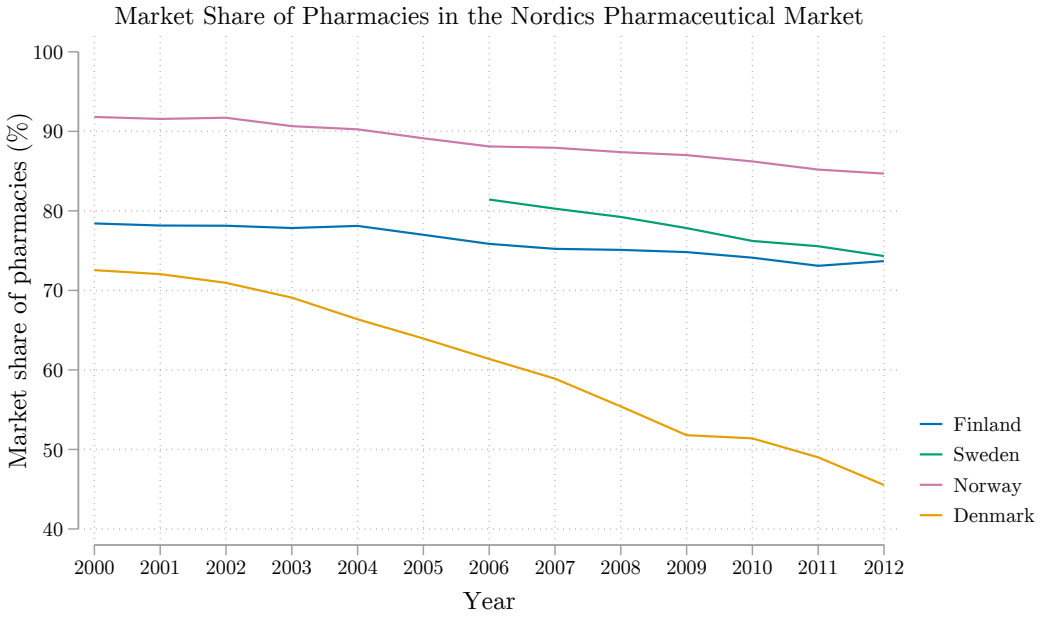
77. The Nordic hospital pharmaceutical market works through competitive bidding. Unfortunately, we do not have access to bids, and therefore we need to rely on wholesale prices while calculating market shares. This leads to a situation where the market shares presented in Figure 12 are the upper bound of the actual market share.

Figure 11: Estimation Sample Sales per capita



Aggregate sales per capita in the sample preperiod by ATC1 level and country, divided by the median. Treatment group denoted with darker grey. Data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

Figure 12: Aggregate Pharmacy Market Share



The development of the market share of pharmacies in the Nordics pharmaceutical market from 2000 to 2012. Data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

matching procedure because it is possible that a given market in Denmark has only hospital market sales, whereas in other Nordic countries sales come from hospital and retail markets. This leads to a situation where our matching procedure yields unmatched markets. Table 3 column 6 shows the number of unmatched markets and the main finding is that all comparisons where Denmark is used as a control group (Finland 2003, Finland 2009, Sweden 2009) have larger number of unmatched markets than comparisons where Denmark is not used as a control group (Denmark 2000, Denmark 2005, Norway 2005). This leads to a situation where 9-13 markets (ATC codes) are dropped, but the economic significance of these markets in the treated country is small (around 1.06 % - 2.07% of the turnover of the retail market).

Table 10 shows how the number of matched markets evolves in Denmark and Finland when reforms are implemented in repeated fashion. Panel A displays how the number of markets is divided between Denmark 2000 and 2005 reforms and Panel B the same for Finland 2003 and 2009 reforms. Panel C compares how many same markets 2009 reforms in Finland and Sweden have in common.

Panel A of Table 10 shows that all markets included in the 2000 reform in Denmark are also included in the 2005 reform and the 2005 reform has 33 markets that were not

Table 10: The Number of Markets in Consecutive Reforms

(1)	(2)	(3)	(4)	(5)
Panel A: Denmark 2000 (IRP → ERP) and Denmark 2005 (ERP → IRP)				
Only in Denmark 2000	Only in Denmark 2005	Both reforms	Total of Denmark 2000	Total of Denmark 2005
0	33	62	62	95
Panel B: Finland 2003 (VGS → GS) and Finland 2009 (GS → IRP)				
Only in Finland 2003	Only in Finland 2009	Both reforms	Total of Finland 2003	Total of Finland 2009
23	51	60	83	111
Panel C: Finland 2009 (GS → IRP) and Sweden 2009 (GS → Auction)				
Only in Finland 2009	Only in Sweden 2009	Both reforms	Total of Finland 2009	Total of Sweden 2009
35	50	76	111	126

Notes: The first and the second columns document the number of exclusive markets included to the estimation samples. The third column displays the number of markets that appear in both samples and columns four and five show how many markets are included to the samples.

included in the earlier 2000 reform. The most likely reason for this is that these markets experience generic entry after 2000 reform implementation. Panel B shows that Finnish reforms have the pattern that both 2003 and 2009 reforms have markets that do not appear in both reforms. The reason for this is that the 2003 GS reform was designed in a way that patent-protected markets with parallel imports can be included in the reform, but these cannot be included in the 2009 reform when the reform started. Finnish 2009 reform has markets that were missing from 2003 because new markets experience generic competition between 2003 and 2009.

Panel C in Table 10 shows differences in the 2009 Finnish and Sweden reform samples in how many identical markets are included. Interestingly, both countries have markets exclusively included in their respective reform sample, which are missing from the other country. Reform design and differences in market structures can explain these differences.

A.6 Pharmacy Retail Price Formulas

This subsection illustrates how pharmacy markup formulas work in the Nordic pharmaceutical market. Table 11 shows how the formulas that were in use during 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 retail

price formulas transmit changes in pharmacy purchase prices to pharmacy retail prices.

Norway is a slight exception compared to other countries because the institutional setting allows pharmacies to charge a lower markup than what the formula presented in table 11 would allow. This allows that in Norway, pharmacy purchase price changes can transmit differently into pharmacy retail prices than in other investigated countries.

Table 11: 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)	
15/7/2009 - 1/11/2009	0-75	PPP × 1.20 + 31.25
	> 75-300	PPP × 1.03 + 44.00
	> 300-6000	PPP × 1.02 + 47
	> 6000	PPP + 167.00

A.7 Step Price-IRP Schedule and Auction-IRP Timing

Swedish 2009 Auction-IRP and Norwegian 2005 SP price regulation reforms have a unique timing structure that can influence how the outcomes studied evolve within these regulations. Table 12 shows how the SP regulation worked during the time period of our estimation sample. SP regulation uses predetermined rules to set the price where reimbursement is paid, instead of that competition in the market would determine the 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 12 also shows that the steps of price decreases change over time. In price formulas valid from 1.1.2005, the largest price decrease was 70% but this was changed to 85% in formulas starting January 1, 2007.

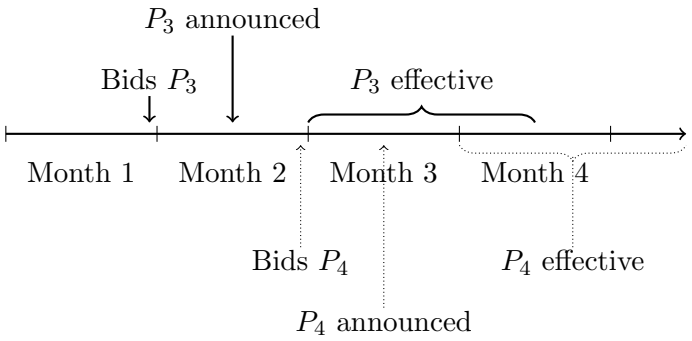
Table 12: The Step Price Schedule

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.

Figure 13 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. This timing structure creates lag, because prices do not become effective immediately. 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 that the previous month's winning product can be dispensed two weeks into the next month. This is represented by the curly brackets denoting the effective prices in Figure 13.

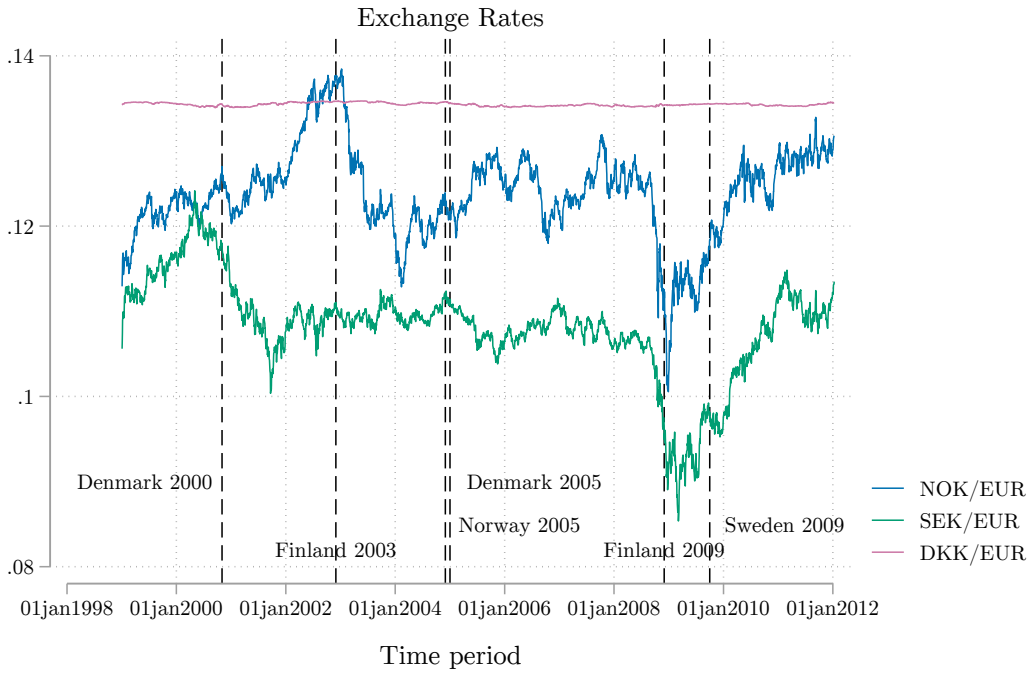
Figure 13: Auction-IRP Timing



A.8 Exchange Rate Shocks

In our empirical analysis, we use domestic currencies instead of converting monetary units into the same currency (e.g., Euro). The rationale for this is that sudden changes in exchange rates can bias our results. This is further illustrated in Figure 14 which plots the

Figure 14: Exchange Rate Shocks

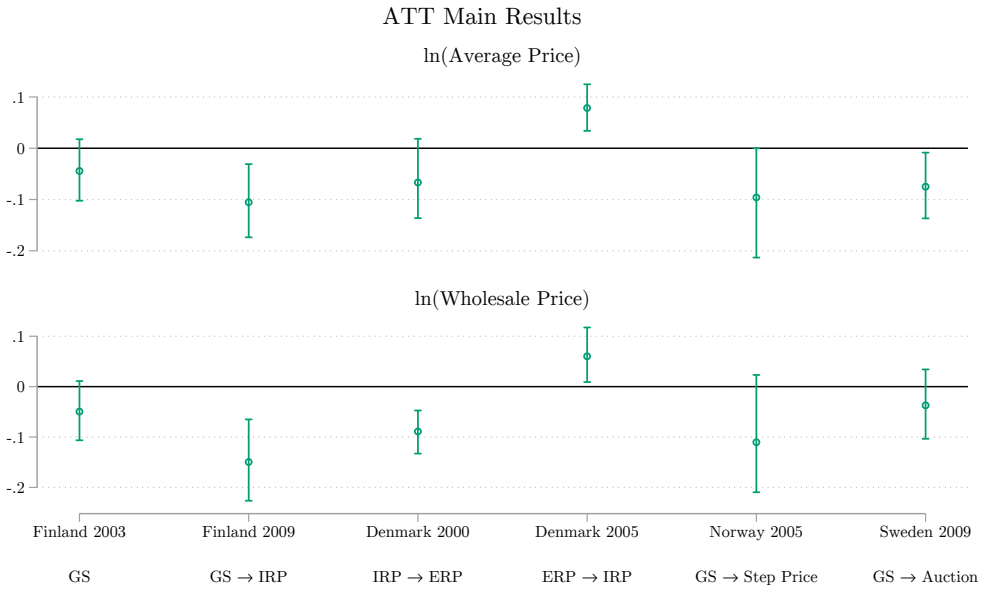


Data source: European Central Bank.

NOK–EUR, SEK–EUR and DKK–EUR exchange rates and start dates of the reforms we study.

Figure 14 shows that the DKK–EUR exchange rate evolves differently from the rest (SEK–EUR and NOK–EUR) and this follows from the fact that during the study period, the Danish Krone (DKK–EUR) is linked to the Euro. All analyzes in which Finland and Denmark are compared to each other yield the same results regardless of the choice of outcome currency. However, this is not the case for the rest of the countries used in our analyzes. Norwegian Krone (NOK) and Swedish Krone (SEK) are not pegged to Euro, and this is seen in Figure 14 as a sawtooth pattern. The figure also shows that some reform starts are timed close to sudden and extreme changes in the exchange rate, such as the Sweden 2009 reform. If the analysis were done using outcomes converted to the same currency, then exchange rate movements would influence the results because our differences-in-differences specifications cannot be augmented in a way that exchange rate effects could be separated from the reform effects.

Figure 15: Main Result ATTs With Different Levels of Aggregation



Estimator: Two-way fixed effects and Callaway and Sant'Anna (2020). replications for ATC-5 wild bootstrapped standard errors. Outcome data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

A.9 Package vs Market-level Price Results

The main results are based on market-level analysis, but a large share of the previous literature dealing with reforms in the pharmaceutical markets uses less aggregated outcome data (e.g. tradename or package-level). In this section, we re-estimate our price regressions with package-level data in order to compare our approach to empirical approaches used in the previous literature.

Figure 15 shows the results using the market-level average price (top panel) and the package-level price (bottom panel) as the outcome. The key result is that the differences between the two specifications are not drastic when the results are compared reform-by-reform. Furthermore, there is no clear ranking on which outcome variable produces the largest estimates in absolute terms. It is important to note that, because of the wide confidence intervals, we cannot reject the hypothesis that the results from the two specifications are different from each other.

The largest differences between the results presented in Figure 15 come from the Norwegian 2005 SP and Swedish 2009 Auction-IRP reforms. In these two reforms, the economic significance of the differences between the two specifications is substantial. In the case of the Auction-IRP reform, the analysis focusing on posted prices at the package level does not give a correct view of the effects of the reform, since the policy diverts the demand

to the cheapest product. Due to this institutional feature, a policymaker should be more interested in the prices of purchased products than in the prices of all products available in the market. This fact explains why the reform treatment effect size almost doubles when the average price (top panel) is compared to the price at the product level (bottom panel). In the Norwegian reform, the treatment effect from product specific prices is around 30% larger than the effect obtained by using the average price as the outcome. This result suggests that in Norway, the prices of packages with limited amounts of sales decreased more than the prices of products with large sales. Switching costs or consumer state dependence can partially explain this behavior. The alternative explanation for this result is that this happens by random chance because the sample size in the market-level analyzes contains only 15 different markets.

A.10 Reimbursement Rates and Finnish Reforms

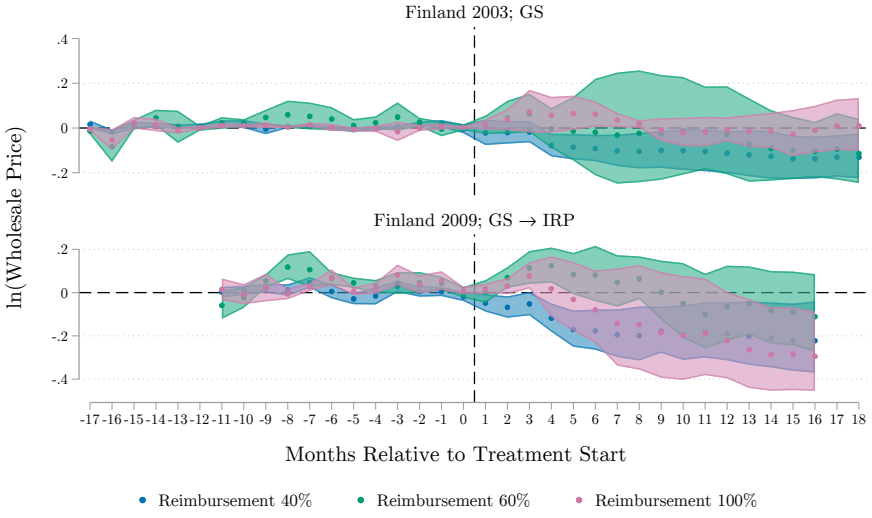
The main text presented results for the Finnish 2003 and 2009 reforms. The results showed quite clearly that the 2009 IRP reform was much more effective than the GS reform of 2003 in reducing pharmaceutical expenditure. The main explanation for this difference is that in the 2003 reform consumer choices 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.

The upper part of Figure 16 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%).[78]

The top panel of Figure 16 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 16 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.

78. 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.

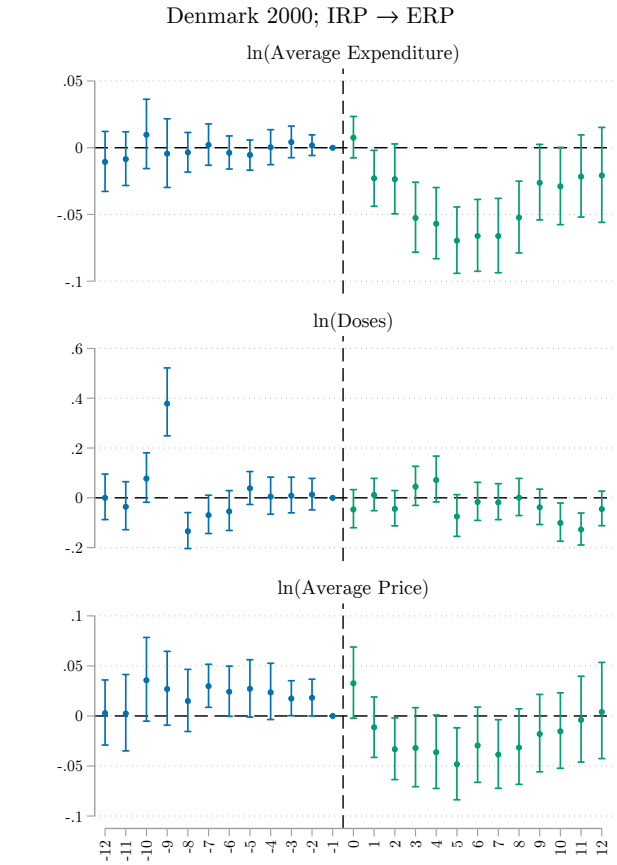
Figure 16: Finland 2003 and 2009 by Reimbursement Status



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.

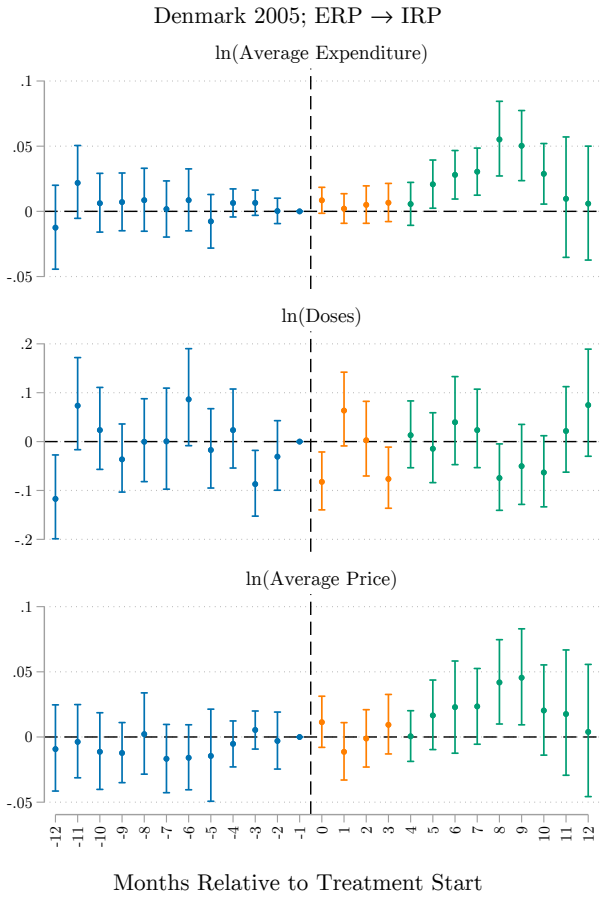
A.11 Denmark 2000 and 2005: Monopoly Markets

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. 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.



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

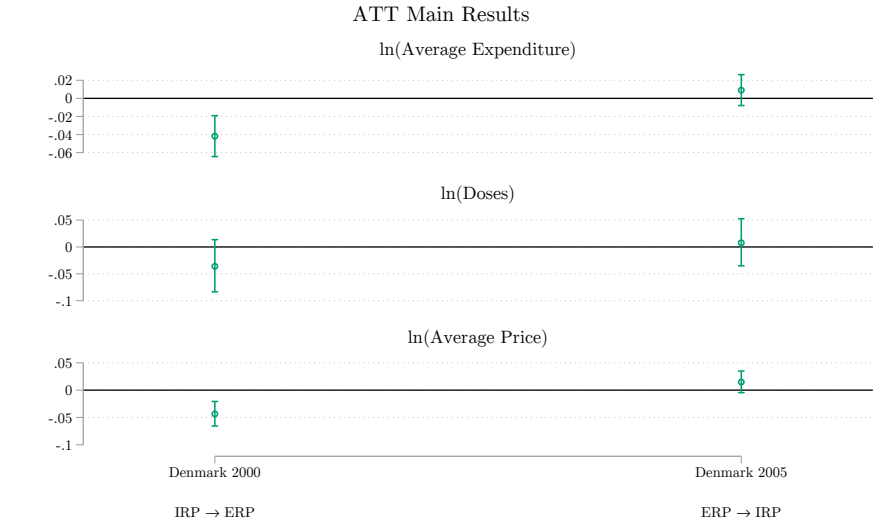


Estimator: Two-way fixed effects. 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Finland used as a control group. Outcome data source: DLI-MI and FIMEA.

(b) Denmark 2005

Figure 17: Part II: Monopoly Markets

Figure 18: Denmark 2000 and 2005

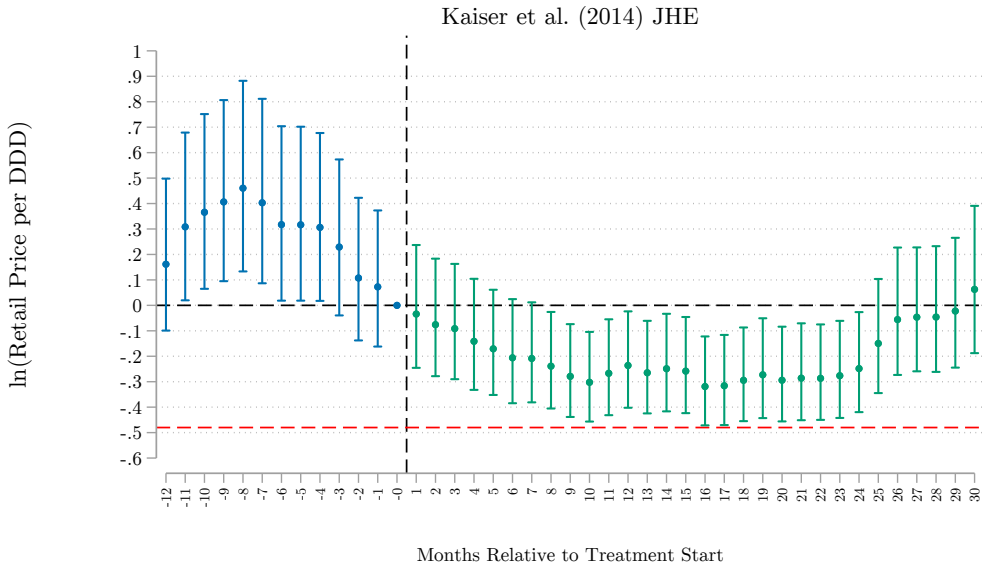


Estimator: Two-way fixed effects. replications for ATC-5 wild bootstrapped standard errors.
Outcome data source: DLI-MI and FIMEA.

Event studies presented in Figures 17a and 17b show that the ERP policy changes had some short-term effects on average expenditure and prices. 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 reform effects seem to decrease in absolute terms over time. Figure 18 shows that on average average expenditure and prices decreased by -4% and -5% during the reform of 2000 and average expenditure and prices increased statistically insignificantly by 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 might also have adverse effects for reference countries, because firms might have an incentive to increase prices or delay entry in order to dilute the gains of using ERP (Dubois, Gandhi and Vasserman 2022; Maini and Pammolli 2022). 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.

Figure 19: 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.

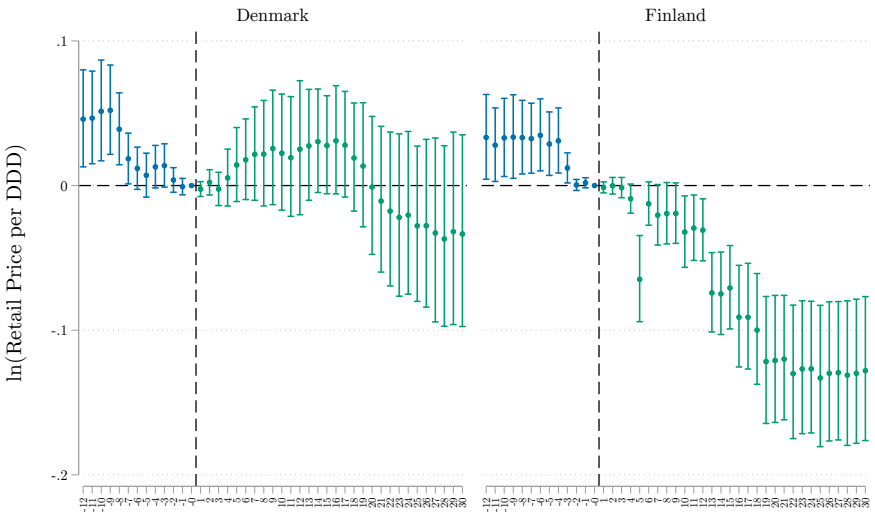
A.12 Denmark 2005: Additional Results

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 prices, but our results show that prices, in fact, increase due to the change. Figure 19 illustrates why the negative price effect result does not identify the true effect of the reform.

Figure 19 replicates the price analysis in Kaiser, Mendez, Rønde and Ullrich (2014) using the same markets and the first regression specification presented in the article[79]. 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 19. 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

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

Figure 20: Replication of Kaiser, Mendez, Rønde and Ullrich (2014) - Both Countries



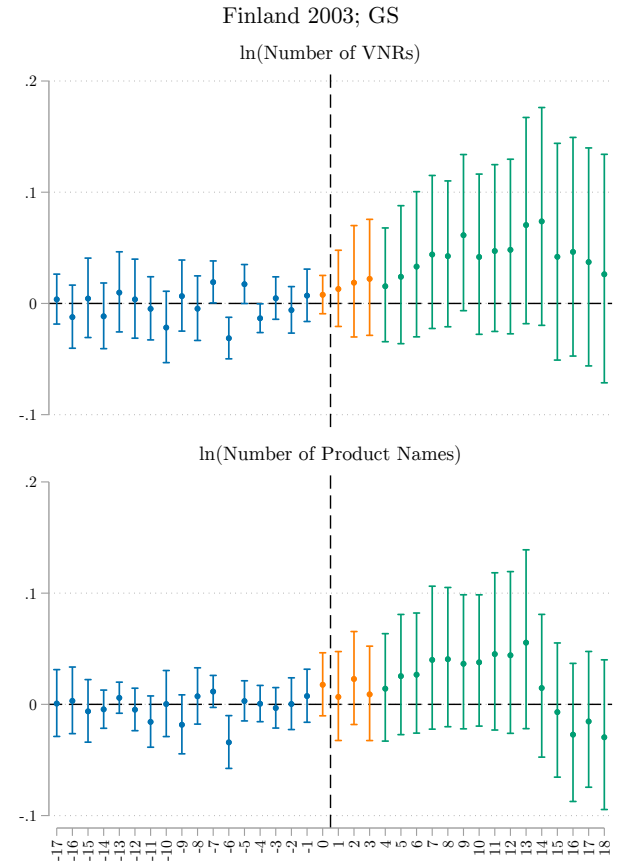
Estimator: Two-way fixed effects. ln(Retail Price per DDD) is regressed on time and product fixed effects.

vertical line). This suggests that the price results reported in Kaiser, Mendez, Rønde and Ullrich (2014) do not measure the actual reform effect on prices.

Figure 20 illustrates the benefits of cross-country comparison in reform evaluation. This example uses the same markets for both countries that are used in the main analysis of this paper. Figure plots calendar-time fixed effects coefficients from separate TWFE regressions for Finland and Denmark to show how average price trends evolve near Danish 2005 reform implementation. Both countries have a decreasing price trend before the reform in Denmark, and when countries are compared, we see that prices continue to decrease in Finland and the decreasing trend stops in Denmark after the reform, and this explains our main result regarding the Denmark 2005 reform.

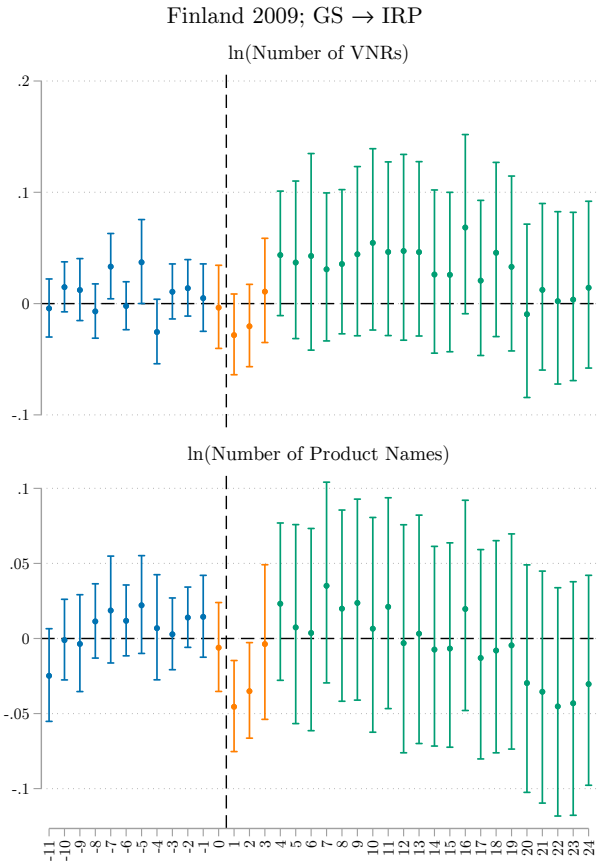
A.13 Regulation and the Availability of Pharmaceuticals

This section shows event study results related to availability analyses presented in the main text. The structure and analysis methods are as in the main text. The pharmaceutical availability is studied using the number of packages and tradenames sold in each market as an outcome. Event study analysis shows that there are no clear parallel trends assumption violations in the availability analyses we do.



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.

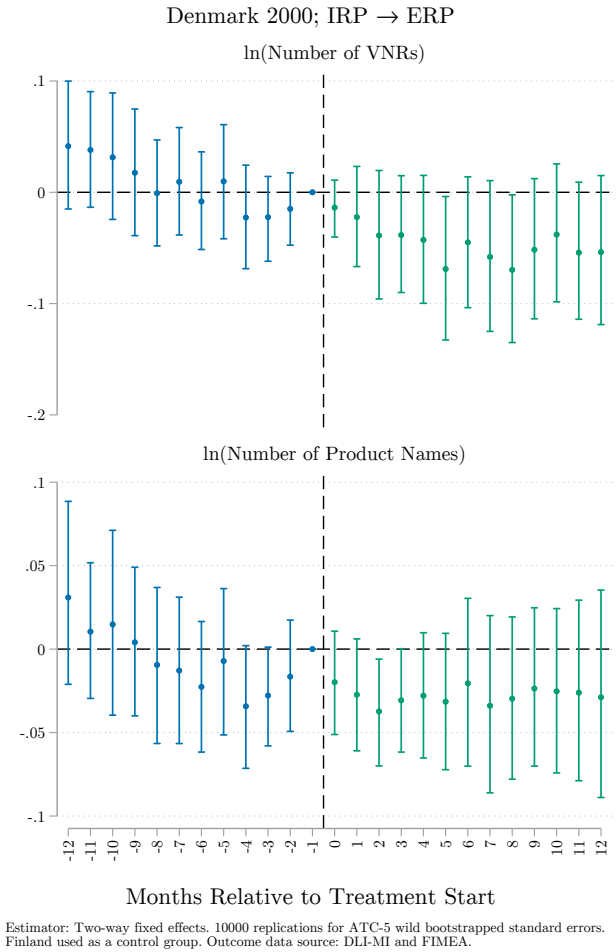
(a) Finland 2003 Pharmaceutical Availability



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.

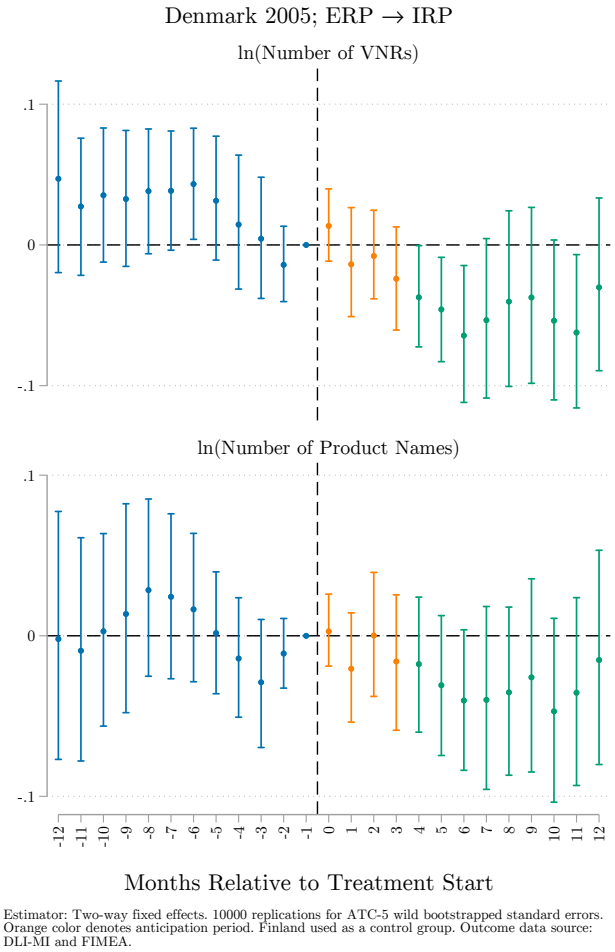
(b) Finland 2009 Pharmaceutical Availability

Figure 21: Part I: Availability



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.

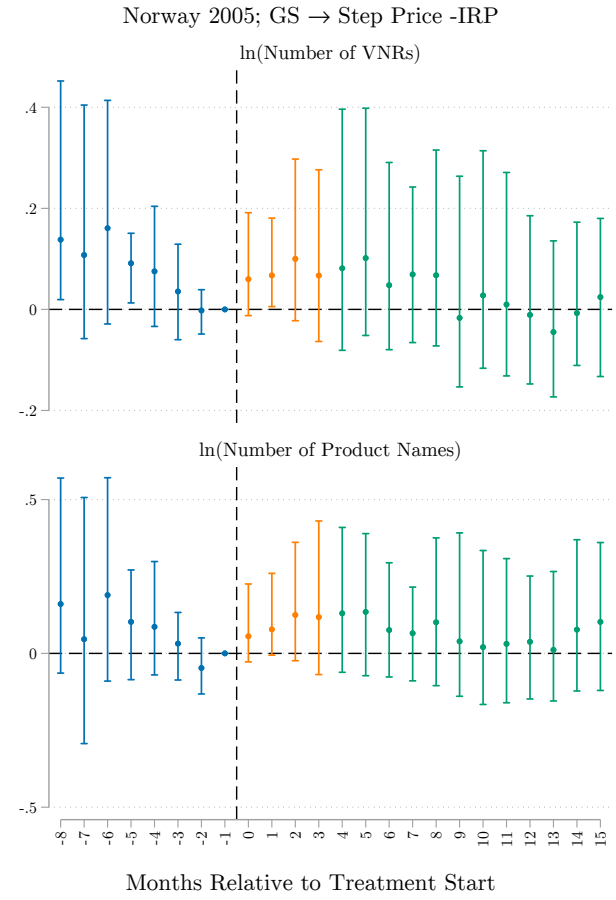
(a) Denmark 2000 Pharmaceutical Availability



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.

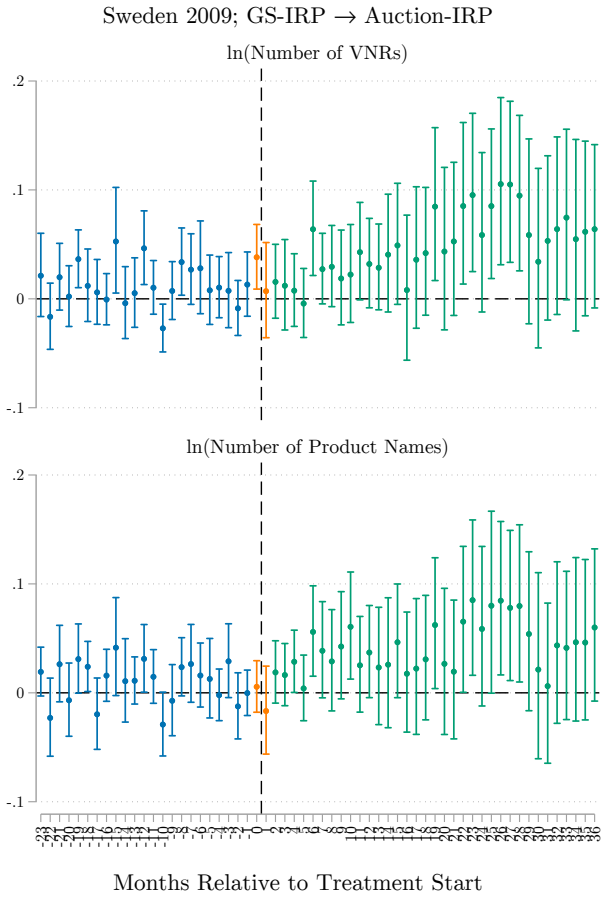
(b) Denmark 2005 Pharmaceutical Availability

Figure 22: Part II: Availability



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 Pharmaceutical Availability

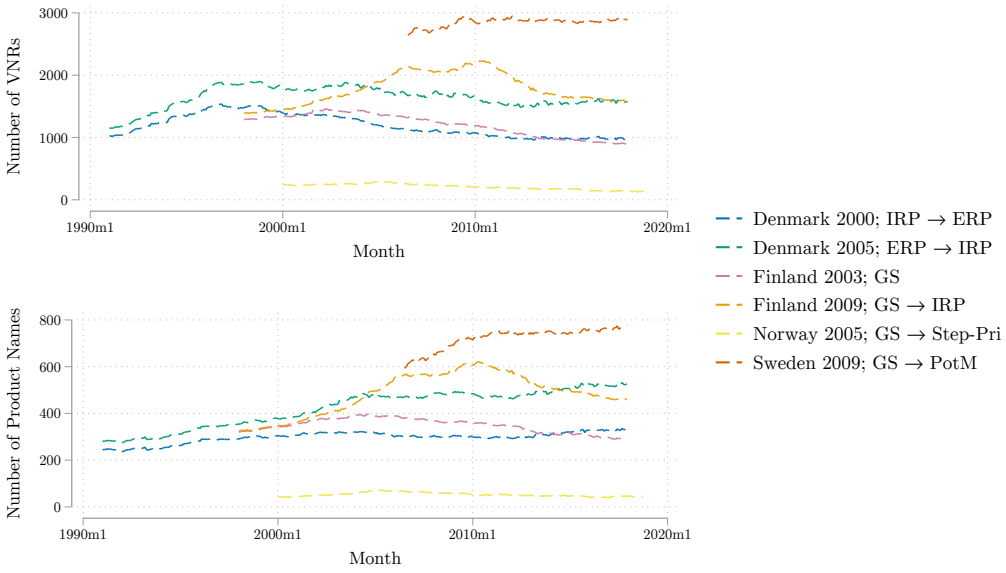


Estimator: Callaway and Sant'Anna (2020). 10000 replications for ATC-5 wild bootstrapped standard errors. Orange color denotes anticipation period. Denmark used as a control group. Outcome data source: DLI-MI and IQVIA MIDAS.

(b) Sweden 2009 Pharmaceutical Availability

Figure 23: Part III: Availability

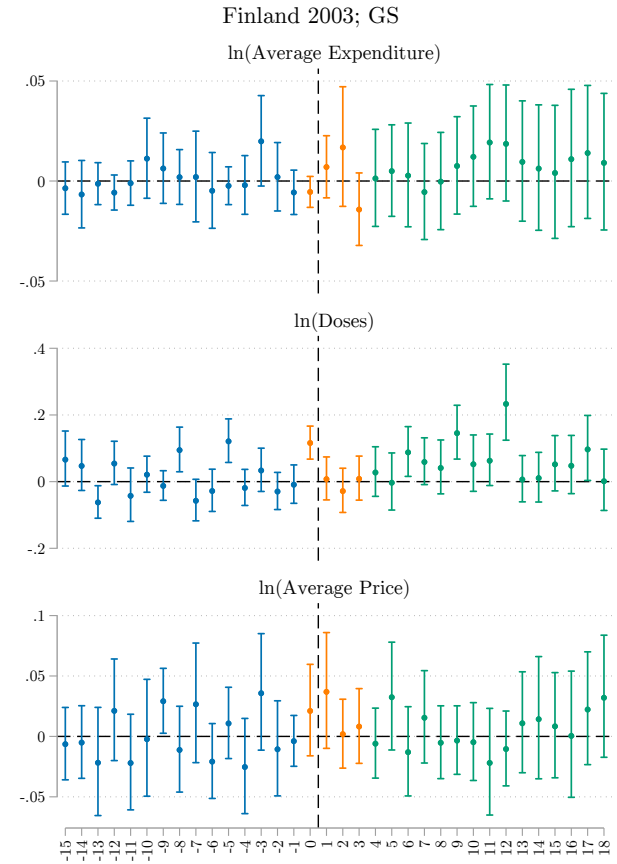
Figure 24: Long Run Pharmaceutical Availability



Number of packages and product names in markets included in at least one estimation sample. Data source: DLI-MI, FIMEA, Pharmastat and IQVIA MIDAS.

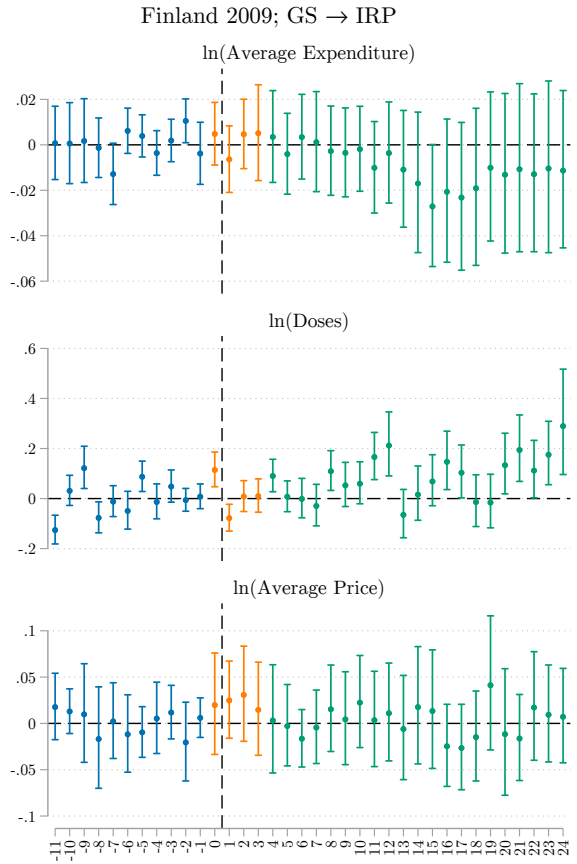
A.14 Additional Results for the Spillover Analyses

This section shows event study results related to between market spillover analyzes presented in the main text. The structure of the analysis, the results, and the methods follow the main text. Event study analysis shows that there are no clear parallel trends assumption violations in these analyses.



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.

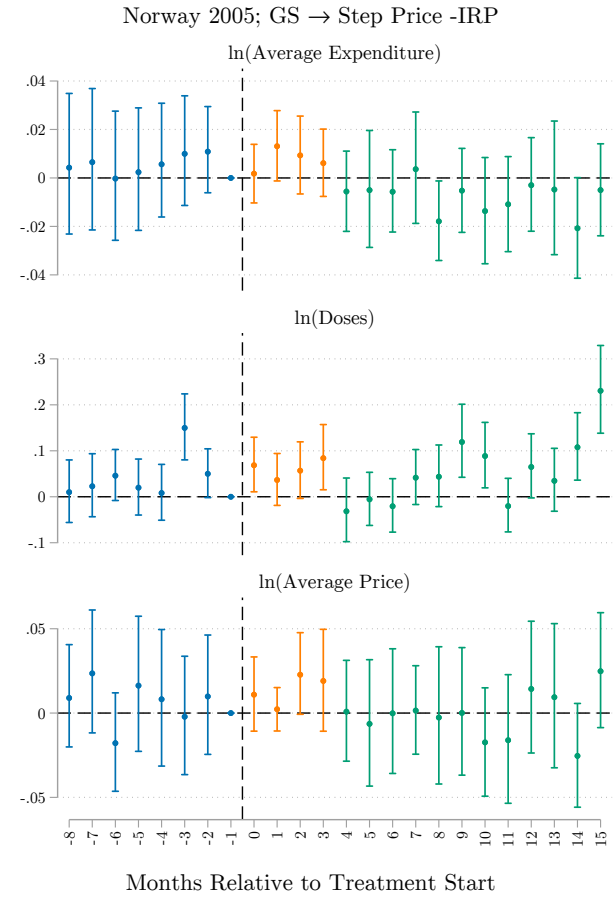
(a) Finland 2003



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

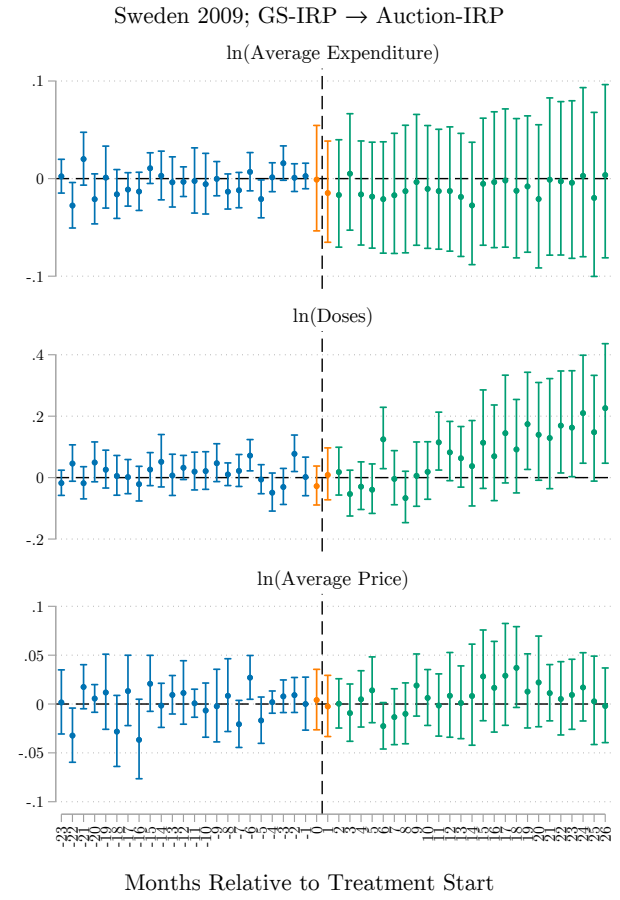
(b) Finland 2009

Figure 25: Part I: Spillover 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. 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 and IQVIA MIDAS.

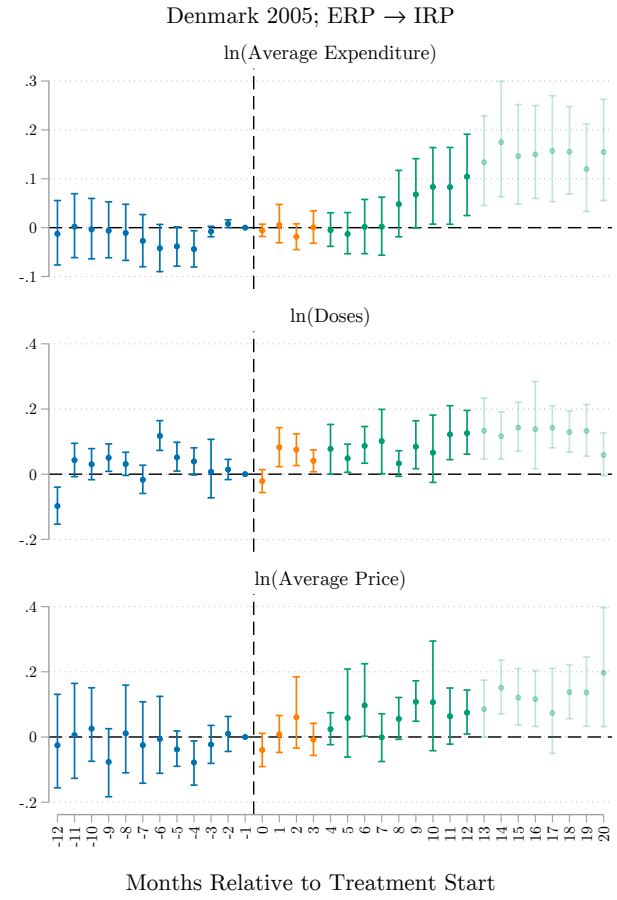
(b) Sweden 2009

Figure 26: Part II: Spillover Results

A.15 Extended Sample Period for Denmark 2005 and Norway 2005 Reforms

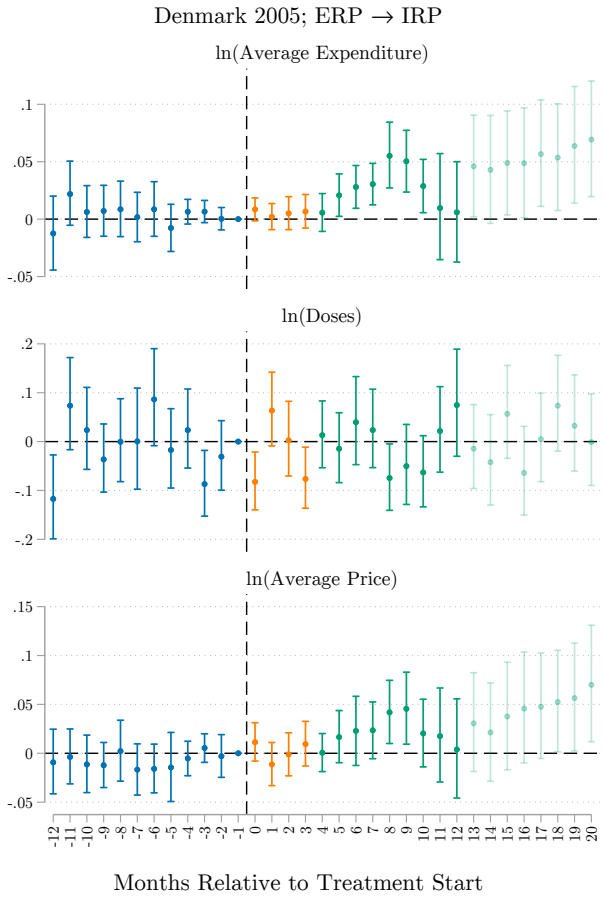
In the main text, we showed results for Denmark and Norway 2005 reforms with a short post-reform period. The reason for this choice was the price cut implemented in the control country Finland in January 2006. This shock in the control country cannot be "controlled away" in our framework, and the shock directly influences our results. Now in figures 27a and 28b 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 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.[80]

80. 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 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.

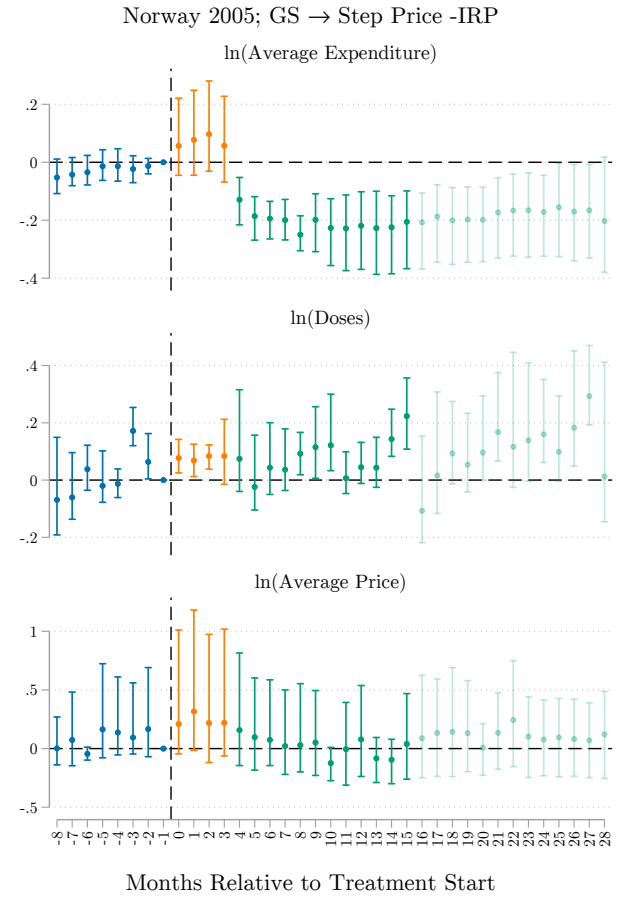
(a) Denmark 2005 – Longer



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.

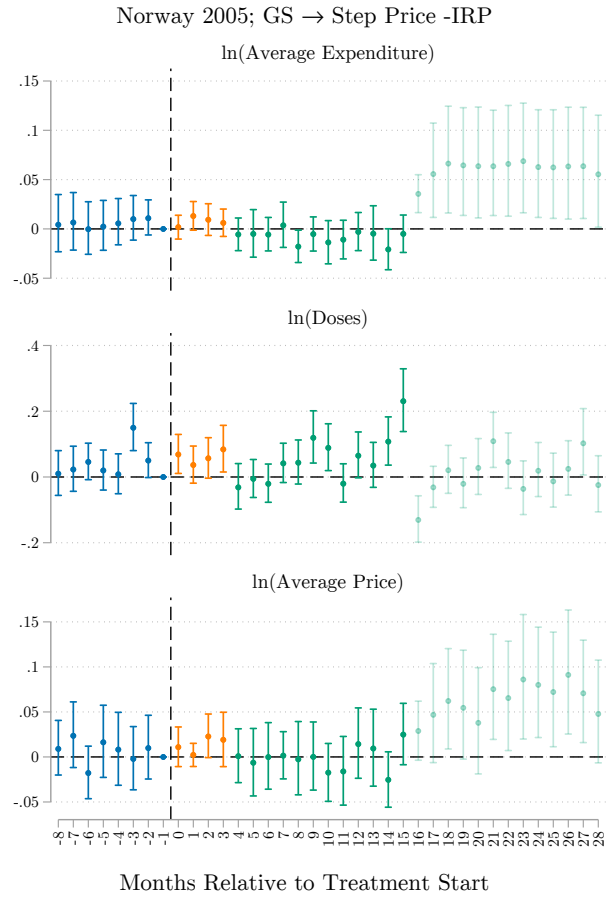
(b) Denmark 2005 Monopoly – Longer

Figure 27: 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 28: Norway 2005 Results with Post 2006 coefficients.