TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

COST-SHARING AND ACCESS IN PRIMARY HEALTH CARE

Tapio Haaga

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. E OSA – TOM. 113 | OECONOMICA | TURKU 2024

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

COST-SHARING AND ACCESS IN PRIMARY HEALTH CARE

Tapio Haaga

TURUN YLIOPISTON JULKAISUJA – ANNALES UNIVERSITATIS TURKUENSIS
SARJA – SER. E OSA – TOM. 113 | OECONOMICA | TURKU 2024

University of Turku

Turku School of Economics
Department of Economics
Economics
Doctoral Programme of Turku School of Economics

Supervised by

Professor Mika Kortelainen
Turku School of Economics

Professor Janne Tukiainen
Turku School of Economics

Professor Petri Böckerman
Jyväskylä University School of Business and Economics

Reviewed by

Associate Professor Gustav Kjellsson
Gothenburg University

Associate Professor Matti Sarvimäki
Aalto University

Opponent

Associate Professor Matti Sarvimäki
Aalto University

ISBN 978-951-29-9668-1 (PRINT)
ISBN 978-951-29-9669-8 (PDF)
ISSN 2343-3159 (PRINT)
ISSN 2343-3167 (ONLINE)
Painosalama Oy, Turku, Finland, 2024

UNIVERSITY OF TURKU
Turku School of Economics
Department of Economics
Economics
HAAGA, TAPIO: Cost-Sharing and Access in Primary Health Care
Doctoral dissertation, 268 pp.
Doctoral Programme of Turku School of Economics
March 2024

ABSTRACT

In developed countries, population is aging, demand for primary health care (PHC) rising, and shortage of healthcare professionals growing. Increased cost-sharing can ease this challenge, yielding more revenue to finance the services and most likely reducing healthcare expenditure in the short term by decreasing service use. However, increased cost-sharing may have negative side effects on inequality as low-income individuals may face a higher barrier to access. It could also worsen population health, if missed visits lead to unmet health needs.

This doctoral thesis studies empirically whether moderate copayments in Finnish public PHC causally affect the use of PHC services and whether low-income individuals respond more strongly to copayments than the rest. The thesis consists of four coauthored essays that are based on comprehensive Finnish administrative register data.

The first essay analyzes the effects of introducing a 10-euro copayment for nurse visits on PHC use of adults in 2014–2019. Nurses increasingly examine and treat primary care patients, but there is little evidence on the impacts of cost-sharing for nurse visits. We fill this gap by using a staggered difference-in-differences (DID) design and state-of-the-art estimators. Our results show that the copayment reduced nurse visits by 9–12% during a one-year follow-up. We find heterogeneity by income in absolute terms, but not in relative terms. Thus, the resulting barrier to access may have been higher for low-income households. The effects on general practitioner (GP) use were negative but small, and statistical significance varies depending on the specification. Unlike much of the earlier Nordic literature, we estimate the effects for the whole adult population and not just for adolescents at a specific birthday. We also use a detailed pre-analysis plan (PAP), which has been rare in nonexperimental economics.

The second essay examines the impacts of copayments of 14–21 euros on GP visits at the 18th birthday, when previously exempted adolescents become subject to copayments in many municipalities. The study uses an age-based regression discontinuity (RD) design and variation across municipalities in whether the copayment changes discontinuously at the 18th birthday (RD-DID). We find that GP visits decreased by 4–5% in the copayment municipalities relative to the comparison municipalities at the 18th birthday. The reductions were largest for the bottom 20% of the equivalized family disposable income distribution, but surprisingly also larger than average for the top 50%. Compared to earlier related Nordic studies, our effect estimates are smaller and the heterogeneity by income level is weaker. We contribute

by combining recent data from 2011–2019 with a design that also has comparison areas and by using state-of-the-art methods for RD designs with a discrete running variable.

The third essay analyzes whether abolishing a 14-euro copayment for GP visits in Helsinki, the capital, increased the number of GP visits among adults. Using a DID design and data from 2011–2014, we find that the abolition was associated with only a small increase in GP visits ($+0.04$ visits annually, or $+4.4\%$, for all adults). The increase was driven by low-income adults ($+0.06$ visits, or $+4.5\%$, at the bottom 40%). The setting is challenging for inference due to only one treated cluster and a finite number of comparison clusters. Although our point estimates are rather robustly positive, conclusions regarding the statistical significance are sensitive to how we account for clustering.

The fourth essay reports the results of a randomized controlled trial that examines the effectiveness of an informational campaign that reminded citizens aged 55 and above about the importance of early detection and treatment of health conditions. The campaign, which took place during the second year of the COVID-19 pandemic, also informed of a policy change that abolished the copayment for nurse visits in PHC. We found no evidence of either the intervention in general or the information on the copayment abolition having increased PHC utilization. Furthermore, we found no evidence of treatment effect heterogeneity. These findings suggest that informational outreach programs may not be effective in inducing curative PHC visits in a gatekeeping system. They are also a healthy reminder that not all nudges always work.

The thesis essays and our related policy report (in Finnish) make an important contribution to national policy discussions. The Finnish setting is characterized by moderate copayments, gatekeeping, and relatively tight supply. Our findings suggest that copayments have a moderate or small effect on PHC utilization in Finland. Copayments also seem to have non-trivial effects on inequality as low-income individuals respond more strongly in terms of visits, although we do not find such heterogeneity in relative terms. Besides these results, the thesis contributes by advancing the use of good research practices, such as the sharing of all replication codes and the use of detailed pre-analysis plans. I believe that both practices should be adopted much more widely, especially by researchers evaluating the impacts of public policies.

KEYWORDS: Cost-sharing, copayment, barriers to access, primary health care, public primary care, healthcare utilization, causal inference, pre-analysis plan

TURUN YLIOPISTO
Turun kauppakorkeakoulu
Taloustieteen laitos
Taloustiede
HAAGA, TAPIO: Cost-Sharing and Access in Primary Health Care
Väitöskirja, 268 s.
Turun kauppakorkeakoulun tohtoriohjelma
Maaliskuu 2024

TIIVISTELMÄ

Kehittyneissä maissa väestö vanhenee ja perusterveydenhuollon kysyntä kasvaa, mutta lääkäreistä ja hoitajista on pulaa. Asiakasmaksujen korottaminen on eräs tapa helpottaa terveydenhuoltoon kohdistuvaa painetta. Maksukorotuksilla voidaan rahoittaa palveluja ja todennäköisesti laskea menoja lyhyellä tähtäimellä palvelujen käytön vähentyessä. Maksukorotuksilla voi olla myös haitallisia sivuvaikutuksia eriarvoisuuteen, sillä ne voivat olla erityisesti pienituloisille este palvelujen saamiselle. Myös terveystulemat voivat heikentyä, jos lääketieteellisesti arvokkaita käyntejä jää toteutumatta.

Tässä väitöskirjassa tarkastellaan empiirisesti, onko kohtuulliseksi katsotuilla asiakasmaksuilla syy-seuraussuhdetta perusterveydenhuollon käyttöön Suomessa. Toinen tutkimuskysymys on, vaikuttavatko asiakasmaksut palvelujen käyttöön erityisesti pienituloisten keskuudessa. Väitöskirja koostuu neljästä artikkelista, jotka perustuvat laajoihin kansallisiin yksilötason rekisteriaineistoihin.

Ensimmäisessä artikkelissa tarkastellaan, miten 10 euron hoitajakäyntimaksun käyttöönotto vaiheittain vuosina 2014–2019 vaikutti perusterveydenhuollon käyttöön aikuisilla. Hoitajien rooli potilaiden tutkimisessa ja hoidossa on kasvanut, mutta aiempaa näyttöä hoitajakäyntimaksujen vaikutuksista ei juuri ole. Analyysissa käytetään vaiheittaisiin erotus erotuksissa -tutkimusasetelmiin (eng. *difference-in-differences*) soveltuvia menetelmiä. Tulosten mukaan asiakasmaksun käyttöönotto vähensi hoitajakäyntejä 9–12 % vuoden seuranta-ajalla. Pienituloiset reagoivat vahvemmin absoluuttisesti käynneissä mitattuna, mutta eroa ei havaittu suhteellisesti. Vaikutukset lääkärikäynteihin olivat negatiivisia mutta pieniä ja usein tilastollisesti ei-merkitseviä. Tulokset estimoidaan koko aikuisväestölle eikä vain nuorille tiettynä syntymäpäivänä, mikä poikkeaa valtaosasta aiempaa pohjoismaista kirjallisuutta. Uutuusarvoa tuo myös tarkan analyysisuunnitelman käyttö havaintoaineistoon perustuvassa tutkimuksessa (eng. *pre-analysis plan*).

Toisessa artikkelissa tutkitaan, vaikuttivatko 14–21 euron lääkärikäyntimaksut lääkärikäyntien määrään 18-vuotissyntymäpäivänä, jolloin alaikäisten vapautus asiakasmaksusta päättyy. Tutkimus perustuu regressioepäjatkuvuusmenetelmään (eng. *regression discontinuity design*) sekä vaihteluun siinä, mitkä kunnat perivät asiakasmaksua täysi-ikäistyneiltä nuorilta. Tulosten mukaan lääkärikäynnit vähenivät 18-vuotissyntymäpäivänä 4–5 % asiakasmaksua perineissä kunnissa verrattuna vertailukuntiin. Estimaatit vaikutuksista ovat suurimpia tulojakauman alapäässä, mutta yllättäen ne ovat keskiarvoa suurempia myös jakauman yläpäässä. Keskimäärin vaikutukset ovat pienempiä ja näyttö tuloryhmien välisistä eroista

heikompaa kuin aiemmissa nuoria koskevissa pohjoismaisissa tutkimuksissa. Artikkeli erottautuu aiemmista tutkimuksista käyttämällä tuoretta aineistoa vuosilta 2011–2019 asetelmassa, jossa on myös vertailualueita, ja käyttämällä uudempia ekonometrisia menetelmiä.

Kolmannessa artikkelissa tutkitaan, lisäsikö 14 euron lääkärikäyntimaksun poistaminen Helsingissä vuonna 2013 lääkärikäyntejä aikuisten keskuudessa. Erotus erotuksissa -tutkimusasetelmaan perustuen havaitsemme, että asiakasmaksun poisto oli yhteydessä vain pieneen nousuun käyntien määrässä ($+0.04$ käyntiä vuositasolla koko aikuisväestössä, eli $+4.4\%$). Tulosta ajoi tulojakauman alapää ($+0.06$ käyntiä tulojakauman alimmassa 40 %:ssa, eli $+4.5\%$). Vaikka piste-estimaatit ovat kautta linjan positiivisia, tulosten tilastollinen merkitsevyys on epäselvää ja riippuu klusterointitavasta.

Neljännessä artikkelissa tarkastellaan, vaikuttivatko satunnaisesti valittuihin kotitalouksiin lähetetyt informaatiokirjeet perusterveydenhuollon käyttöön. Koronapandemian toisen vuoden aikana toteutetussa kokeilussa muistutettiin yli 55-vuotiaita oikea-aikaisen hoidon tärkeydestä ja mahdollisuudesta hakea hoitoa terveysongelmiin terveyskeskuksesta. Osassa kirjeistä tiedotettiin, että asiakasmaksulain uudistuksen myötä terveyskeskuksen sairaanhoidollisista hoitajavastaanotoista oli tullut maksuttomia. Näyttöä ei saatu, että kirjeillä itsessään tai tiedolla hoitajakäyntien maksuttomuudesta olisi ollut havaittavaa vaikutusta sairaanhoidollisten hoitaja- tai lääkärikäyntien määrään keskimäärin, eikä havaittu näyttöä eroista etukäteen määriteltyjen ryhmien välillä. Tulosten perusteella informaatioon perustuvat tuuppaukset (eng. *nudge*) eivät välttämättä ole tehokkaita lisäämään perusterveydenhuollon käyntejä tilanteessa, jossa tarjonta on niukkaa ja saatavuutta rajoitetaan hoidon tarpeen arvioinnilla.

Artikkelit ovat osa toteuttamaamme jälkiarviointikokonaisuutta, joka tuottaa tietoa asiakasmaksujen vaikutuksista perusterveydenhuollon käyttöön Suomessa. Asiayhteydelle ominaista on kohtuulliseksi katsotut asiakasmaksut, hoidon tarpeen arviointi perusterveydenhuollossa sekä tarjonnan ja sen niukkuuden merkitys. Tulostemme mukaan asiakasmaksut vaikuttavat jonkin verran perusterveydenhuollon käyttöön. Asiakasmaksuilla on myös eriarvoisuuden näkökulmasta kiinnostavia vaikutuksia, sillä pienituloisten palvelukäyttö muuttuu käyntien määrässä mitattuna enemmän. Vastaavaa eroa ei kuitenkaan havaita suhteellisissa vaikutuksissa. Varsinaisten tulosten lisäksi väitöskirjan arvoa lisää eräiden hyvien tutkimuskäytäntöjen, kuten replikointikoodien avoimen jakamisen ja tarkkojen analyysisuunnitelmien käytön (eng. *pre-analysis plan*), edistäminen. Nähdäkseni molempien käytäntöjen soisi nähdä merkittävästi yleistyvän etenkin politiikkamuutosten vaikutusten jälkiarvioinnissa.

ASIASANAT: Asiakasmaksut, hoidon esteet, perusterveydenhuolto, julkinen perusterveydenhuolto, terveyspalvelujen käyttö, kausaalipäättely, analyysisuunnitelma

Acknowledgements

Achievements in life or research are not accomplished alone. The available research infrastructure, such as learning materials, data, econometric methods, and open-source software, is primarily created by others, having a great impact on what a single person can do. Even with great research infrastructure, a single person can do very little without the support of other people.

My excellent co-authors and supervisors have had an important role in conducting the research, as well as my development as a researcher. Petri Böckerman, Mika Kortelainen, and Janne Tukiainen were my supervisors and co-authors in three essays. Frequent discussions with Mika have been beneficial for generating ideas, improving the essays, and developing as a researcher. Petri is extremely accessible and fast to provide feedback and suggestions. His input has especially improved my writing. Janne can pose important questions that others may not detect, improving the essays from motivating them to the econometric methods. Lauri Sääksvuori and Jussi Tervola were co-authors in one essay and our policy report. Jussi has kindly helped me in several important ways during the years, and he is a very nice person indeed. I admire Lauri's ability and courage to pose honest, critical questions, which is an extremely valuable skill in any workplace.

I want to thank Matti Sarvimäki and Gustav Kjellsson for pre-examining the thesis and providing valuable feedback. Matti will also act as an opponent to this dissertation.

This thesis would not exist without the support of the Finnish Institute for Health and Welfare (THL). I started working on the dissertation topic at THL as a bachelor student in 2019 and wrote my master's thesis at THL in 2020. THL has provided excellent research infrastructure, including but not limited to the data, ever since. I will always be grateful to Mikko Peltola for his trust, kindness, and guidance at the start of this journey. In 2022, I again worked at THL with Jussi and Lauri. I want to thank Eeva Nykänen, Maria Vaalavuo, Merja Korajoki, Joonas Ollonqvist, and all the others at THL who I have had the pleasure to work with.

The second crucial source of support was the Turku School of Economics, where I have been working as a doctoral researcher since Summer 2022. It was Mika who encouraged me to apply, and it was an excellent choice. I am grateful to Heikki Kauppi for his trust, support, and an excellent research environment. The department has talented people, and administrative matters work out smoothly. I enjoyed the

weekly research seminar. Besides Heikki, Mika, and Janne, I want to single out Henri Salokangas, Mikko Hertzig, and Salomo Hirvonen for valuable discussions, and Jenni Heervä for all her help.

The research described in this thesis and our policy report has received financial support from the Finnish Ministry of Social Affairs and Health (STM) and Yrjö Jahnsson Foundation (research grant no. 20197209). I thank them and the members of the steering group set up by STM: Marjo Kekki, Mikko Meuronen, Tero Tyni, Timo Hujanen, and Mika. I express my gratitude to the regional PHC areas of Kymenlaakso, Päijät-Häme, and South Karelia and their representatives for cooperation in the informational outreach campaign, reported in the fourth essay.

I thank several people for feedback and comments on the essays or the thesis: Tanja Saxell, Markku Siikanen, Liisa T. Laine, Tuomas Markkula, Mikko Nurminen, Jukka Pirttilä to whom I also express gratitude for supervising my master's thesis, Mikko Peltola, Lauri, Jussi, Mika, Henri, Maria, Heikki, Joonas, and Karri Heikkinen. Besides, I thank seven anonymous referees and editors Owen O'Donnell, Heather Royer, Hendrik Schmitz, Trine Kjær, and Brigitte Dormond for considering and reviewing our work. I also thank all seminar participants who have provided comments related to this research and the organizers of the Finnish Health Economics PhD Student Workshop, the Helsinki GSE Labor and Public PhD Workshop, the THL CHESS Seminar, the TSER Seminar, the INVEST Workshop, and the Finnish Society for Health Economics.

Before the doctoral studies, I worked as a research assistant for Liisa T. Laine and Antti Ripatti at the PREDICT consortium, funded by the Strategic Research Council at the Academy of Finland (no. 340551; PI Liisa T. Laine). I thank Liisa and Antti for this stimulating time that led to the doctoral studies. With respect to future, I thank Oskari Nokso-Koivisto for his great work in building opportunities, as well as Mika, Lauri, Mikko, and Tanja.

I am grateful to the economics unit at the University of Helsinki, the Helsinki Graduate School of Economics, as well as the Lahden yhteiskoulu secondary school for precious memories, wonderful people, and excellent learning environments.

I thank friends, relatives, and godparents — your support means a lot to me. Most importantly, I owe an enormous debt of gratitude to my parents for their love, support, and an exceptional environment to grow and live happily. I am also extremely grateful to my partner, who has supported me whenever I needed it and made the last few years so good.

February 16th, 2024
Tapio Haaga

Table of Contents

Abbreviations

ACSC	Ambulatory care sensitive condition
ATE	Average treatment effect
ATT	Average treatment effect on the treated
CS	The Callaway and Sant'Anna estimator
DID	Difference-in-differences
DDD	Triple differences
ED	Emergency department
EHR	Electronic health records
GP	General practitioner
IID	Independent and identically distributed
MSE	Mean squared error
OLS	Ordinary Least Squares
PAP	Pre-analysis plan
PHC	Primary health care
PTA	Parallel trends assumption
RD	Regression discontinuity
RD-DID	Difference-in-discontinuities
RS	The Roth and Sant'Anna estimator
SC	Synthetic control
SE	Standard error
THL	Finnish Institute for Health and Welfare
TWFE	Two-way fixed effects
WCR	Restricted wild cluster bootstrap
WCU	Unrestricted wild cluster bootstrap

List of Original Publications

This dissertation is based on the following original articles, which are referred to in the text by their Roman numerals:

I Haaga, Tapio; Böckerman, Petri; Kortelainen, Mika; Tukiainen, Janne. Effects of Nurse Visit Copayment on Primary Care Use: Do Low-Income Households Pay the Price? Version 4, 2023; https://osf.io/skuv9/.

II Haaga, Tapio; Böckerman, Petri; Kortelainen, Mika; Tukiainen, Janne. Do Adolescents from Low-Income Families Respond More to Cost-Sharing in Primary Care? Version 2, 2023; https://osf.io/r83sg.

III Haaga, Tapio; Böckerman, Petri; Kortelainen, Mika; Tukiainen, Janne. Does Abolishing a Copayment Increase Doctor Visits? A Comparative Case Study. Version 2, 2023; https://osf.io/v7b5s.

IV Haaga, Tapio; Sääksvuori, Lauri; Tervola, Jussi. The Impact of an Informational Campaign on Primary Care Utilization among Older Citizens: Evidence from a Randomized Field Experiment. Version 2, 2023; https://osf.io/q72b5.

Revised versions of Article 1 and Article 3 have recently been published as [1] and [2], respectively. The original publications have been reproduced with the permission of the copyright holders.

1 Introduction

1.1 Cost-Sharing and Moral Hazard

Health care services aim to maintain or improve people's physical and mental well-being and prevent negative health shocks. Thus, access to health care can have a major impact on quality of life. The question of whether individuals consume less health care when they have to pay more for it has been an important theoretical and empirical topic in economics for decades [3]. A seminal study in this literature is the RAND Health Insurance Experiment, conducted in the U.S. in 1971–1986 [4].

In the health economics literature, "moral hazard" typically refers to whether higher out-of-pocket costs reduce healthcare utilization conditional on health [3]. Without empirical evidence, it is not obvious whether such moral hazard exists. Standard demand theory predicts that higher prices reduce demand. However, health care can be price inelastic if it is consumed only when needed. It is also possible that lower prices for preventive services lead to an increase in the use of preventive services, better health, and ultimately a reduction in the use of curative services and, possibly, healthcare costs.

Understanding the extent and nature of patients' response to out-of-pocket costs is important for the design of better health insurance, whether private or public. Health insurance balances between providing a high level of financial security against unexpected health shocks on the one hand, and sufficient incentives for consuming services only when needed and maintaining a healthy lifestyle on the other [3]. In developed countries, where the growing use of telemedicine has reduced other costs associated with seeking care, populations are aging, and the shortage for healthcare professionals increasing, cost-sharing is becoming increasingly important. Potentially, higher out-of-pocket costs may ease the pressure on healthcare systems, if patients seek less of the kind of care that is of low value medically. Cost-sharing may also impact inequality as out-of-pocket costs constitute a larger fraction of disposable income for low-income individuals, probably creating a greater barrier to access for them than for the rest of the population. Higher out-of-pocket costs may widen health inequalities as a potential consequence of income-related heterogeneity in the responsiveness to cost-sharing and missed, medically valuable care.

An extensive previous literature has found that out-of-pocket costs do reduce health care utilization [3; 4]. However, most of this literature is based on variation in

insurance coverage in the U.S. context. This thesis focuses on moderate copayments in a publicly funded primary health care (PHC) system, where there is gatekeeping at the entry. The copayments in question are transparent for patients, easy to bill, and their low level mitigates financial risks to patients but generates less revenue to healthcare providers. Still, adding a small price to a previously free health care service may have disproportionately large utilization effects [5]. These impacts may be especially important when the copayment is levied on preventive entry-level services that serve a gatekeeping purpose, such as many PHC services. Any needs-based prioritization by professionals is conditional on patients having contacted PHC in the first place, and it could be better if professionals make assessments about the need instead of patients alone.

This thesis studies empirically whether moderate copayments in Finnish public PHC causally affect the use of PHC services and whether low-income individuals respond more strongly to copayments than the rest.[1] Besides guiding healthcare use and collecting revenue, one of the Finnish policy objectives has been to ensure that copayments are moderate and do not threaten citizen's right to use these services. The thesis consists of four essays, each of which is based on comprehensive administrative register data. The empirical challenge in estimating causal relationships is that the individual-level counterfactuals of what would have happened had the exposure been different are never observed. This problem is circumvented by using mainstream research designs developed in the causal inference literature.

1.2 Previous Literature

There is a voluminous literature that examines the impacts of cost-sharing on health care utilization and drug consumption. To my knowledge, no extensive literature reviews or meta-analyses have been conducted during the last ten years. However, some reviews that focus on specific reforms, such as the Affordable Care Act under the Obama administration in the U.S., are available. Our policy report [6] lists 263 references by country from 2012–2022 that are relevant to this broader literature, focusing on (health) economics journals. That list is not systematic and may not be representative or extensive either (see the report for details). However, it is, to my knowledge, the most extensive reference list available that considers the literature from 2012–2022.

Of the 263 listed references, a clear majority (168) are based on U.S. data. A typical study based on U.S. data examines the impacts of a policy reform that aims to increase the share of people covered by health insurance. Getting covered decreases

[1] Such moderate copayments are also widely utilized in tax-funded public healthcare systems elsewhere, including the Nordic countries.

out-of-pocket costs as the insurer pays a share of the costs. Smaller out-of-pocket costs may increase the consumption of health services and drugs. The ultimate interest is often in the health effects of cost-sharing, but credible results of health effects are rare compared to results on the first-order utilization effects.[2] Most of the 263 studies are based on observational data with the aim of causal inference and use mainstream identification strategies, such as DID or RD. The consensus is that insurance coverage (or lower out-of-pocket costs) increase health care utilization [3]. It is safe to say that the studies from 2012–2022 use better data and designs than the older literature reviewed by [7]. This is visible in the use of large administrative registers (vs. small surveys) and in an increased emphasis on plausibly exogenous variation in cost-sharing and identification strategy.

Results from U.S. studies that focus mainly on the impacts of healthcare coverage may not easily generalize to public health insurance systems covering all citizens. In such systems, cost-sharing policies often represent moderate changes at the intensive margin. Table 1 lists studies that examine the impacts of copayments on healthcare utilization using data from the Nordic countries. The essays of this thesis are also included for comparisons. The studies of the table are discussed more extensively in our policy report [6] that also summarizes the Nordic evidence about the impacts of out-of-pocket costs on prescription drug consumption. Five studies use Swedish data, two use Norwegian, and two use Danish. Of these nine studies, all examine the impacts of copayments or copayment changes. Six focus on adolescents or children using either an age-based RD design or a DID-type strategy. Six studies examine GP visits in PHC as the outcome, while two focus on psychologist visits. Five studies analyze the heterogeneity of the effect estimates by income.[3]

The Nordic studies rather consistently suggest that people consume less health care services and prescription drugs when the spot price (the price at the point of consumption) is higher [6]. Of the nine studies from Sweden, Norway, and Denmark, only one reports no observable utilization effects [8]. In that study, the exposure is also the least intensive in a sense that the study examines modest intensive-margin changes of 33% to existing copayments, while in other studies copayments are introduced or abolished or significantly changed for a service or a subgroup. Five studies (three from Sweden and two from Norway) report that primary care GP visits or telemedicine consultations decreased when previously exempted children or adolescents start facing copayments [9; 10; 11; 12; 13]. One

[2] Mortality is rare, and detecting mortality effects may require drastic interventions, very large samples, and statistically powerful designs. In practice, researchers often stratify by diagnosis code or focus on ED visits, but it is unclear how well these measures proxy health effects.

[3] There is room for studies that examine the impacts of copayments for all adults or the elderly population, focus on whether individuals from low-income households respond more to cost-sharing, use credible research designs other than RD designs, or examine the impacts of not only GP visit copayments but also of other important services, such as nurse visits or telemedicine consultations.

study from Sweden reports intertemporal substitution effects close to the copayment change at the 85th birthday, but finds no persistent effects [14]. Two studies based on Danish data report that the use of psychologist visits increased considerably either in terms of service months or treatment initiations when these services became partly or totally covered by the state [15; 16]. Of the eight studies summarized by [6], seven find that people respond to changes in spot prices for prescription drugs.

Regarding the heterogeneity of the effect estimates, the findings are mixed. Two studies using Swedish data report heterogeneity by income in both absolute and relative terms: children or adolescents from low-income households respond more than the rest [9; 10]. However, another study from Sweden does not find heterogeneity in effects by income at the 85th birthday [14]. The findings on the use of psychologist visits in Denmark are also mixed: one study [15] does not find treatment effect heterogeneity by income, while another does [16]. The former estimates the effects at the 38th birthday or for those aged 18 to 37, while the latter does so for those aged 18 to 21.

Table 1. The Literature on the Effects of Copayments on Healthcare Use in the Nordic Countries.

Study	Context	Outcome	Reform or exposure	Design	Sample	Key findings
Nilsson & Paul (2018) [9]	SWE 1999–2006	Outpatient doctor visits (both GPs and specialists)	People under a given age are exempted from copayments (€10–15 for GP visits). This age threshold was lowered from 20 to 7 in 1999 and increased back to 20 in 2002.	RD + RD-DID: comparing outcomes just before and after the two age cutoffs of 7 and 20, and accounting for discontinuities in outcomes that existed also when there was no policy discontinuity.	Children and adolescents close to the 7th or the 20th birthday in Region Skåne.	Doctor visits increased by 5–10% when care was free, and such effects were found both at the 7th and the 20th birthday. The effects were driven by those from low-income families.
Johansson, Jakobsson & Svensson (2019) [10]	SWE 2014–2015	PHC GP visits	Children and adolescents under the age of 20 are exempted from copayments for GP visits (€10).	RD: comparing outcomes just before and after the 20th birthday.	Adolescents close to the 20th birthday residing in Region Västra Götaland.	GP visits decreased by 7% when copayments were charged. The reductions were larger for women (9%), low-income individuals (11%), and low-income women (14%).
Johansson, de New, Kunz, Petrie & Svensson (2023) [14]	SWE 2014–2018	PHC visits	Individuals aged 85 or more are exempted from copayments for PHC visits (levels vary by service).	Kinked donut RD: testing for trend breaks in PHC use just before and after the 85th birthday.	Individuals close to the 85th birthday residing in Region Stockholm or Region Västra Götaland.	Patients had less PHC visits shortly before the copayment elimination, and these visits were shifted until shortly after the elimination. However, no persistent increase nor heterogeneous effects by income in PHC use were found.

Continued on next page

Continued from previous page

Study	Context	Outcome	Reform or exposure	Design	Sample	Key findings
Ellegård, Kjellsson & Mattisson (2022) [11]	SWE 2012–2018	PHC in-person physician consultations and direct-to-consumer telemedicine (DCT) consultations.	Children and adolescents under the age of 20 are exempted from copayments (€25) for DCT consultations.	Fuzzy RD-DID. The first stage: comparing DCT consultations just before and after the 20th birthday in 2018 (DCT was common) and subtracting jumps that were observable also in 2012–2016 (when DCT was uncommon).	Adolescents close to the 20th birthday residing in Region Västra Götaland or Region Stockholm.	DCT consultations decreased by half at the 20th birthday. Around half of all DCT consultations represented additional demand, the rest replacing in-person physician visits.
Jakobsson & Svensson (2016) [8]	SWE 2011–2012	PHC GP visits	Region Värmland harmonized copayments in health services by increasing the GP visit copayment from c. €18 to €23 (+33%) and lowering the copayment for emergency department visits from c. €35 to €23 (−33%).	DID, DDD: comparing outcomes in Region Värmland, before and after the policy change, to the same evolution in Region Örebro (the neighbouring control region). The third difference in DDD subtracts a time-placebo estimate from the effect estimate.	All PHC GP visits in the study areas.	No effects on PHC GP visits were observed.
Magnussen Landsem & Magnussen (2018) [12]	NOR 2009–2014	PHC GP visits	Adolescents under 16 are exempted from GP visit copayments (c. €18) since 2010.	RD: comparing outcomes just before and after the 16th birthday.	All those who visited a GP in 2009–2014 at age 10 to 20.	GP visits decreased by 10–15% when copayments were charged. Those with an acute condition responded less strongly than the rest.
Olsen & Melberg (2018) [13]	NOR 2006–2013	PHC GP visits	Adolescents under 16 were exempted from GP visit copayments (c. €18) in 2010. Previously, only those under 12 were exempted. The policy thus changed for those aged 12 to 15.	SC: comparing outcomes of those aged 12 to 15, before and after the policy change, to the same evolution of a weighted average of other age groups. The weights are chosen to optimize pre-treatment fit.	All those fulfilling an age-based eligibility criterium. In the main analysis, individuals aged 0 to 20 are included.	GP visits increased by 22% for women and by 14% for men when care was free.
Ly Serena (2021) [15]	DNK 2000–2011	The probability of treatment initiations (the first treatment in at least a year)	The coverage of psychologist treatment for depression was increased from 0% to 60% for those aged 18 to 37 in 2008.	RD, DID: comparing outcomes just before and after the 38th birthday and the 18th birthday, and comparing the outcomes of those aged 18 to 37, before and after the policy change, to the same evolution of those aged 38 to 50.	Full population. In the RD analyses, only individuals close to the age cutoffs.	When the coverage was in place, psychologist treatment initiations increased by 60–90%. The study also considers a large set of other outcomes.
Kruse, Olsen & Skovsgaard (2022) [16]	DNK 2014–2020	Monthly propensity to use psychologist treatment	The copayment of psychologist treatment for anxiety and depression (€50) was abolished in 2018 for those aged 18 to 21.	DID: comparing the outcomes of those aged 18 to 21, before and after the policy change, to the same evolution of those aged 16 to 17 or 22 to 23.	All those aged 16 to 23.	The monthly propensity to use psychologist treatment increased by 75% when care became free. The use of outpatient psychiatric care and prescriptions of antidepressants moderately increased. The effects were higher for adolescents from low-income families.

Thesis essays

Continued on next page

Continued from previous page

Study	Context	Outcome	Reform or exposure	Design	Sample	Key findings
Haaga, Böckerman, Kortelainen & Tukiainen (2023a) [17]	FIN 2013–2019 2020–2022	PHC nurse and GP visits	The staggered adoption of copayments for nurse visits (€10) in 2014–2019 by municipalities. The copayment was abolished nationally in 2021.	DID: comparing outcomes in municipalities that adopted the copayment, before and after the adoption, to the same evolution in municipalities that did not adopt the copayment or adopted it later.	Those aged 25 or above from all over the country.	Nurse visits decreased by 9–12% during a one-year follow-up. Low-income individuals responded more in absolute terms, but not in relative terms. The estimates on GP use were negative but small, with varying statistical significance.
Haaga, Böckerman, Kortelainen & Tukiainen (2023b) [18]	FIN 2011–2019	PHC GP visits	Minors are nationally exempted from copayments for GP visits (€14–21). Some areas charge no copayment at all or exempt students.	RD, RD-DID: comparing outcomes just before and after the 18th birthday in areas with and without a policy discontinuity at the cutoff.	Adolescents close to the 18th birthday residing in large municipalities.	GP visits decreased by 4–5% when copayments were charged. The reductions were largest for the bottom 20% of the income distribution, but also larger than average for the top 50%.
Haaga, Böckerman, Kortelainen & Tukiainen (2023c) [19]	FIN 2011–2014	PHC GP visits	Helsinki, the capital, abolished the copayment for GP visits (€14) in 2013.	DID, SC: comparing the outcomes of individuals residing in Helsinki, before and after the abolition, to individuals residing in other large municipalities that did not change copayments.	Those aged 25 or above residing in Helsinki and other large municipalities.	GP visits increased by 4% on top of the extrapolated increasing pre-trend difference, and the increase was driven by low-income individuals. Statistical significance is inconclusive.
Haaga, Sääksvuori & Tervola (2023) [20]	FIN 2021–2022	PHC nurse and GP visits	Informational letters reminding on the importance of early detection and treatment of health conditions. Two thirds of the letters informed that the nurse visit copayment (€12) was abolished nationally just recently.	RCT: letters were randomly assigned to households. Every third household (47,398) received a letter, the rest forming a control group.	Those aged 55 or above residing in three target regions (Kymenlaakso, Päijät-Häme, and South Karelia).	No evidence that the intervention in general or the information on the copayment abolition would have increased PHC use. No evidence of treatment effect heterogeneity.

Recent years have seen the emergence of interesting hypotheses in the broader cost-sharing literature which go beyond the core hypotheses that explore whether people respond to cost-sharing and whether there is potential treatment effect heterogeneity. For instance, studies have examined whether people respond asymmetrically to symmetric increases and decreases in cost-sharing [21] or to framing very similar incentives in terms of gains versus losses [22; 23], and whether the prize zero affects service use discontinuously [5]. Another strand of literature exploits nonlinearities in health insurance contracts, such as annual out-of-pocket caps, deductibles, or time-varying coinsurance rates determined by prior consumption, in order to analyze how forward-looking or myopic consumers are [24; 25; 26; 27; 28; 29; 30]. To summarize, a plausible hypothesis is that individuals exhibit myopic behavior and are responsive to spot prices and that increases in cost-sharing have a stronger impact on healthcare use compared to equal-sized decreases.

1.3 What Can We (Really) Learn from the Previous Estimates?

It is reasonable to hypothesize that increased out-of-pocket costs reduce the use of health care services also in Finland. However, the available literature provides limited information on the extent to which this reduction in usage may occur.

In general, interpreting the *estimates* of published empirical literature in social sciences is challenging and should be done with great caution. This argument is based on 1) challenges in extrapolating causal inference results out of sample (external validity), 2) the likely selection of published findings (publication bias), 3) the common lack of transparency, especially in nonexperimental literature, on to what extent the data analysis workflow was confirmatory versus exploratory (inflated p-values), and 4) the constant improvement in econometric methods. I discuss these points below in the wider context of empirical social sciences and not related to the specific cost-sharing literature in health economics. They may be beneficial in reviewing confirmatory studies critically and motivate our use of a pre-analysis plan in two thesis essays.

Comparing different estimates is challenging. The estimated causal effect is by definition tied to a specific context, such as the combination of place, time, intervention, and local institutions. Additional assumptions are always needed when trying to extrapolate causal effect estimates out of sample. Such assumptions should be explicitly stated and discussed.

Price elasticity is a key concept in the cost-sharing literature. However, one should be cautious in summarizing the effects of nonlinear[4] cost-sharing to a single elasticity estimate [31]. Defining and estimating the numerator (change in quantity) and the denominator (change in price) of the elasticity is often nontrivial [31]. It is not obvious whether individuals respond myopically to the spot price, an expected end-of-year price, or something else under nonlinear pricing. Considering quantity changes, it may be reasonable to group separate health care contacts into episodes and assume that cost-sharing affects only the initiations of such episodes (extensive margin), but not what happens within the episode (intensive margin) [31].

However, a single price elasticity estimate does not illustrate how sensitive it is to changes in these assumptions and other choices made in data cleaning and construction. It may be impossible to identify which feature of a specific context causes a difference in elasticity estimates. For these reasons, I would be cautious to construct and use price elasticity estimates in cross-context comparisons.

Publication bias. Standard frequentist inference methods are valid if and only if publication probabilities do not depend on the point estimates in any way [32].

[4] In nonlinear pricing, prices change as a function of the (e.g., yearly) accumulated expenses. In Finnish PHC, nonlinearity stems from an annual out-of-pocket cap and from the fact that copayments for nurse and GP visits are charged only for the first three visits annually in most municipalities.

The properties of frequentist tests and estimators are considered under hypothetical replications of the data. By definition, these tests should not depend on the observed realizations of the data. If publication is selected based on findings, then the point estimates are biased and test sizes and confidence sets distorted [33].[5]

The understanding is that published research is indeed selected. For instance, Maximilian Kasy summarizes this in the Journal of Economic Perspectives: *"Published research is selected through a process that includes both researchers and journals, so that consumers of such research cannot, in general, assume that reported estimates are unbiased, either in their point estimates or their confidence intervals."* [33] For some important empirical evidence on publication bias in economics or in some of its subfields, see [34; 35; 36; 37; 38]. The form and extent of publication bias arguably differs across empirical fields and journals. Some researchers and journals may overvalue estimates that are statistically significant, of a given sign, or which they view surprising or plausible [33].

There are some methods to estimate the form and magnitude of selection if replication studies or meta-analyses are available [34]. In practice, it may be hard for a reader to assess the extent (or absence) of publication bias in a specific journal or a strand of literature, let alone in a single study. For many fields and journals, there are currently no studies available that could provide evidence that rejects the null hypothesis of no publication bias.[6] However, a more plausible null hypothesis is arguably publication bias, and its refutation should require proof.[7]

Without a reasonable estimate of the size of the publication bias, the published point estimates can be misleading. For instance, the replicated effect sizes of all between-subject laboratory experiments published in two top five economics journals between 2011 and 2014 tend to be of the same sign as the original ones but smaller, the mean relative effect size being 66% [37]. In another study, the findings on the impacts of behavioral RCT nudge interventions are compared between academic journals and the records of two governmental nudge units [38]. Studies in academic journals are plausibly selected, while selection is not an issue in the nudge unit records that contain all trials the unit has ever conducted. The mean impacts are much smaller in the nudge unit records and only 16% of the mean effect size in academic journals. Of this difference, 60–70% is estimated to be explained by publication bias in academic journals, while different nudge characteristics explain most of the remaining difference [38].

Inflated p-values. The concept of the "garden of forking paths" by [39]

[5] The validity of the findings may not be the only objective for publishing research [32]. For instance, policymakers with limited time for consuming research may benefit if the published studies are selected as being surprising. However, that selection invalidates the validity of statistical inference [32].

[6] I am not aware of studies examining the extent of publication bias in the cost-sharing literature.

[7] Publication bias is tackled by registered reports that are accepted in journals based on the research question and study design *before* the estimates are constructed. Their tests do not depend on findings.

illustrates why interpreting the *estimates* in social sciences is often challenging even in the absence of publication bias. This is a challenge especially in nonexperimental literature where pre-registration and pre-analysis plans are rare. By definition, frequentist statistical tests should not depend on the observed realizations of the data. Thus, if researchers use p-values or report confidence intervals, as most do, they should have chosen exactly the *same* way to process and analyze the data under *every* possible realization of the data [39].

In practice, economics papers typically test a large number of hypotheses that are often contingent on other, previous tests in the same paper [40]. My understanding is that for most (nonexperimental) economics projects the workflow is more or less exploratory: the analysis is adjusted and somewhat-formed hypotheses refined along the process, according to what seems reasonable based on the observed data. For instance, data or outcomes are pooled to increase sample size if the observed data appear noisy and the confidence intervals "too wide". A seminar audience may encourage you to focus more on a new, "novel" outcome. A referee may suggest you use an alternative estimator that has attractive efficiency properties when your estimates are not that precise in some subgroup analyses.

All the changes listed above seem reasonable and innocent, and many, including myself, have happily considered and adopted them. However, these changes would make the analysis and the tests contingent on the observed data. If these changes are adopted, the published p-values cannot be taken at face value anymore. The garden of forking paths is essentially a multiple comparisons problem where a massive number of *potential* comparisons arises when the details of data analysis become contingent on the observed data [39]. Thus, it is a fundamental issue for confirmatory inference [39; 40; 33].

The problem with the current norms is that if one reads a typical well-published nonexperimental article, one often does not know how much the analysis is contingent on the observed data. Ideally, researchers should communicate these contingencies to the reader by either using a pre-analysis plan or discussing the research workflow and the forking paths in detail. Without this transparency, it may be hard to take the *estimates* at face value, even if the study was otherwise inspiring and important.

Advances in econometrics. A recent econometric literature has shown that the conventional two-way fixed effects regression models, that have been commonly used to analyze staggered DID settings, can be seriously biased in the presence of staggered treatments and treatment effect heterogeneity (see [41] for a review). These new econometric results have thus put in question the credibility of the earlier *estimates* from several studies and contexts where the assumption of no treatment effect heterogeneity seems strict. As the methods are constantly developing, so should the way we weigh earlier estimates. Table 1 shows that many earlier Nordic studies use an RD design. The credibility of their findings is thus related to the future

advances in RD methods.

The way forward? The arguments above illustrate the fundamental challenges in interpreting the *estimates* of earlier studies. Nonetheless, these challenges do not make the estimates uninformative. First, future meta-analyses can and should carefully estimate the size of the publication bias (see [34]). Such an analysis should also be done in the cost-sharing literature. Second, confirmatory hypothesis testing is not the only lens to look at data — generating hypotheses can be equally important. Indeed, some journals, such as the ones of the American Economic Association, have begun to deemphasize statistical significance, e.g., by not using asterisks to denote significance. The phrase 'correlation is not causation' is already well-known. We should also distinguish between confirmatory claims, such as "we have statistical evidence that D causes Y", and exploratory claims, such as "it is a plausible assumption based on data that D causes Y". It could be useful to consider (many of) the previously published findings as plausible hypotheses about the world rather than definitive confirmatory assertions.

1.4 Advancing the Use of Pre-Analysis Plans

In 1983, long before the design-based "credibility revolution" in applied econometrics [42] took place, Edward Leamer saw its state rather negatively:

"This is a sad and decidedly unscientific state of affairs we find ourselves in. Hardly anyone takes data analyses seriously. Or perhaps more accurately, hardly anyone takes anyone else's data analyses seriously." [43]

Although large individual-level administrative datasets and better designs have arguably significantly improved the quality of observational causal inference since Leamer's critique, better data and design alone are not sufficient for a credibility revolution. To make confirmatory data analysis credible, we need both good designs and good research workflows, namely being more transparent in how we wander in the garden of forking paths. One proposed solution is to use pre-analysis plans (PAPs) [40; 44; 45]. This proposal does not eliminate publication bias if journals still select studies based on the point estimates *ex post*, but it is otherwise attractive.

In their ideal form, PAPs specify a full mapping from data to which statistics will be reported [33]. This full mapping contains not just the details of each step of the data cleaning, construction, and analysis, but also a plan on how the estimates are presented in tables, figures, and text. Ideally, there is also a fraud-proof firewall between the design phase of the PAP and observing the relevant data, e.g., by pre-specifying the study before the intervention is administered. In practice, PAPs often do not reach these ideals, especially if the study is observational and based on a historical exposure. In these cases, PAPs constrain the analysis and reporting of the estimates to some extent and limit the garden of forking paths.

A complementary solution proposal to PAPs is open data and codes[8] for replication studies. Reproducing an analysis that explored the garden of forking paths also results in inflated p-values. The (nonexperimental) replication should thus be independent and blinded from the original data construction and analysis choices and essentially be pre-registered. However, such blinded replications are costly because all the empirical tasks need to be conducted from scratch by a new team.[9]

We used PAPs in two thesis essays in order to increase the transparency of our confirmatory policy evaluation. For one essay [20], general-level pre-registration to the AEA RCT Registry was done before starting the trial. A more detailed statistical plan was registered while the trial was going on but before getting access to any post-treatment outcome data, which can be verified. The detailed statistical plan contained all planned statistical programs for data construction and analyses and a corresponding placebo report, which shows the planned way to present the analyses in a paper. The placebo report was written using a time placebo as if the intervention was conducted two years earlier.[10] After analyzing the post-treatment data, we put out a populated PAP showing the estimates of the planned analyses, and rather strictly followed the PAP. Finally, we wrote a summary of the key (null) estimates in a letter format as a distinct object from the populated PAP (the fourth essay of this thesis). The populated PAP illustrates what we *planned* to learn, while the final paper reports what we think we *learned*, the distinction proposed by [44]. The final statistical codes for the populated PAP and the thesis essay are also publicly available.

In another essay [17], we used the PAP in a nonexperimental context and analyzed a historical reform using (mostly) pre-existing data to which we already had access. Definitive blindness was thus not possible. At the time of writing the PAP, we had already worked on two other thesis essays although on different exposures [18; 19]. Thus, this kind of PAP is not as "pure" as the way described above and definitely not fraud-proof. However, this application is arguably very interesting, because the use of PAPs in nonexperimental work is currently rare [45], meaning that there is much room for improvement.

The rationale to use a PAP in this context comes from not having previously linked municipal nurse use and municipal policies on nurse visits copayments. Thus, it was possible to blind the relation between policy changes and outcomes. Specifically, we randomly assigned municipalities into "fake" placebo policies at the start using the real observed adoption dates, making the placebo treatment indicator independent of outcomes (see [40] for the benefits). The aim was to shield

[8] All the codes of the thesis essays are publicly available.

[9] Blind replications are important even if costly. Recent many-analyst studies in economics and finance have found that the variation in data construction and analysis can be large across replications and lead to notable differences in estimates [46; 47]. This uncertainty is not captured by standard errors.

[10] An alternative would have been to additionally blind the treatment assignment [40].

ourselves from learning about the causal relation of interest when fixing choices on data cleaning, data construction, analysis, and reporting. For credibility of statistical inference, this kind of pre-registration is superior to having no pre-registration, and it provides opportunities to assess credibility and potential biasing influences, at least subjectively [48].

Moreover, some analyses of the second essay ([17]) were pre-specified in a more conventional manner, based on the fact that we did not have access to post-treatment outcomes at the time of registering the PAP. For the subset of analyses, the PAP was written as if the treatment occurred years earlier (placebo-in-time).

When writing the research paper after registering the PAP and unblinding the data, we were flexible to reasonable *ex post* changes and additions in analysis, but reported and discussed these changes in a hopefully clear and transparent way. Thus, the final research paper is a mix of pre-registered confirmatory primary analyses and non-pre-registered exploratory and supplementary analyses. The main body of the paper separates tables and figures that were not registered by label "post-blind" while also mentioning this in the table and figure notes. There is also an appendix section that lists and discusses the changes made *ex post*.

My thoughts for future projects are the following. First, using a PAP should be a default choice in both experimental and nonexperimental *confirmatory* policy evaluation. Deviating from this principle is possible if the research is *exploratory* or for other good reasons. An alternative is to be more transparent on how the analysis is contingent on data. Second, a pre-analysis plan is just a plan, and analyses can be refined *ex post* if the reasons are transparently communicated to the reader, a sentiment echoed by [44]. Third, a good research paper probably has both confirmatory and exploratory elements in it. The set of pre-specified confirmatory analyses should be narrow enough to leave room for exploratory analyses that may be contingent on data, an idea also expressed by [40]. Fourth, it should be clear which results were pre-specified and which were not, even when the reader is skimming.

2 Summary of the Essays

This section summarizes the thesis essays. The summaries are available in Finnish in our policy report [6], which also summarizes the earlier Nordic literature, lists over 260 relevant references from all around the world, and provides an accessible introduction to the causal inference methods used in the thesis essays.

2.1 Effects of Nurse Visit Copayment on Primary Care Use: Do Low-Income Households Pay the Price?

The first essay [17] examines the impacts of a staggered adoption and a later simultaneous abolition of a copayment for curative nurse visits in PHC on the number of nurse and GP visits among those aged 25 or above. The staggered adoption occurred in 2014–2021 in PHC areas and was induced by fiscal deficits, the example of other primary care areas, and an increased awareness of the legal possibility to charge the copayment. A clear majority of municipalities charged the copayment by Summer 2021, and the most common policy was to charge it for the first three visits annually, the mean being 12 euros.

Using a DID design, we compare the evolution of nurse visits in municipalities that adopted the copayment, before and after the adoption (first difference), to the same evolution in our comparison municipalities (second difference) that did not adopt the copayment (never-treated) or that adopted it later (later-treated). The group of municipalities that adopted the copayment is not representative, but the timing of adoption conditional on the adoption seems more arbitrary. Causal inference in DID designs relies on a parallel trends assumption (PTA): the outcomes for the treated cohort and for the not-yet-treated comparisons would have followed parallel trends in the absence of treatment. The assumption seems plausible, as nurse use evolved similarly in both policy groups for two years preceding the copayment adoption.

The results show that the number of nurse visits decreased by 9–12% (0.09–0.10 visits per capita per year) in a one-year follow-up after the copayment adoption. Before the adoption, nurse use was increasing in both policy groups. After the adoption, nurse use decreased in the treated cohorts, while the increasing trend continued in the comparison areas. Nurse visits decreased in both the bottom 40% and the top 40 % of the income distribution. The estimated decrease in the number of visits is more than two times larger at the bottom 40% of the income distribution than

at the top 40%, and the difference is statistically significant. However, heterogeneity by income level is less clear-cut and statistically insignificant in relative terms. The relevant dimension of heterogeneity depends on the context.

If the results accumulate gradually, the baseline results based on a one-year follow-up underestimate long-term effects. An alternative method using a longer follow-up estimates reductions of 13% to 17%. The nurse visit copayment may also reduce GP visits if contacts that would otherwise lead to GP appointments either directly or through a nurse visit are missed. We estimate a 2–5% reduction in GP visits, but our preferred estimates are closer to zero and often insignificant.

Using a simple DID design, we also analyze the effects of a law reform that abolished the nurse visit copayment nationally in July 2021, comparing the evolution in nurse visits between municipalities that removed the copayment and municipalities that had never adopted the copayment. However, the PTA is not credible in a pre-specified 12-month time window before the policy change, plausibly related to the COVID-19 pandemic. As we are not willing to impose the PTA, we do not reach any causal conclusions.

We are among the first to focus on the impacts of cost-sharing for *nurse* visits in PHC. We do so for the adult population by analyzing direct effects on nurse visits and, importantly, also indirect effects on GP use. Moreover, we use a staggered DID design, which we analyze with methods that are robust to the staggered design and heterogeneity of treatment effects. We also demonstrate the use of pre-analysis plans (PAPs) in non-experimental economics by illustrating how blind analysis and a corresponding detailed PAP can be used to limit the concerns for the "garden of forking paths" when analyzing historical events. We are probably the first to use a detailed PAP in an observational study in the cost-sharing literature and to conduct pre-specified heterogeneity analysis by income level.

2.2 Do Adolescents from Low-Income Families Respond More to Cost-Sharing in Primary Care?

The second essay [18] examines whether GP visits decrease at the 18th birthday when previously exempted adolescents start paying copayments of 14 to 21 euros for GP visits in many municipalities, using an age-based regression discontinuity (RD) design. We also exploit municipal-level variation in whether there was a discontinuity in policy to construct treatment and comparison areas. The comparison municipalities either charged no copayment at all or exempted students. The key identification assumption in RD designs is that the expected potential outcomes are continuous functions of the running variable at the cutoff. We account for other factors that potentially cause discontinuities at the 18th birthday by estimating models that subtract from the discontinuities in the copayment areas the discontinuities in the comparison areas (RD-DID). The results are estimated for

women only because call-up health checks concentrate for men close to the 18th birthday.

The results show that GP visits decrease by 4–5% in the copayment areas relative to the comparison areas at the 18th birthday. These RD-DID estimates are driven by statistically significant reductions in the copayment areas, while the estimates in the comparison areas are close to zero and insignificant. For instance, the RD estimates for all individuals vary between -0.04 and -0.05 annualized visits (-4.8% to -5.8%) in the copayment areas compared to estimates of -0.00 to -0.01 visits (-0.3% to -1.0%) in the exemption areas.

We do find that the largest reductions in the RD-DID estimates are for the bottom 20% of the income distribution. Their GP use decreases by 0.08–0.10 annualized visits (7–10%). Our income measure is the equivalized family disposable income in the year when an individual turns 17. The estimates attenuate for the bottom 40%. However, the estimates are also larger than average for the top 50%, showing reductions of 0.05–0.06 visits (6–8%). Thus, our results do not support the hypothesis that the effects are overwhelmingly concentrated at the lower end of the income distribution while the top end does not respond noticeably.

Our study contributes data-wise and methodologically to the earlier Nordic literature that estimates the effects of copayments for adolescents at a specific birthday using RD designs. First, we combine recent data from 2011–2019 with a design that also includes comparison areas. Second, the methodological RD literature has evolved, and we use data-driven bandwidth selection leading to mean-squared-error (MSE) optimal point estimation and bias-corrected robust inference, and also account for the discreteness of our running variable. Applying state-of-the-art methods is possible because our data on birth dates and visits are more granular than in the earlier studies.

2.3 Does Abolishing a Copayment Increase Doctor Visits? A Comparative Case Study

The third essay [19] examines whether GP use increased among those aged 25 or above in Helsinki, the capital, after it abolished its 14-euro copayment for GP visits in PHC in 2013 in an effort to reduce health inequality. Using a DID design and the synthetic control method (SC) as a complement, we compare the evolution in GP visits between Helsinki and a comparison group of large and middle-sized cities with no change in copayments.

Pre-trends are not parallel in our raw outcomes as GP use increased in Helsinki relative to comparison municipalities already before the copayment abolition. Causal inference in the DID framework thus relies on a modified PTA: we assume that the PTA holds after subtracting a linear pre-trend difference from the data (detrending). Specifically, we fit a linear trend difference in time with OLS between Helsinki and

the comparisons using only pre-treatment data. The estimated trend difference is then subtracted from the outcomes to construct a transformed outcome variable.

The results show that post-abolition GP use per capita increased moderately ($+0.04$ visits annually, or $+4.4\%$, for all sample adults) in Helsinki after subtracting an increasing linear pre-trend difference. The overall estimates are driven by low-income individuals. The results show an increase of $+0.06$ visits ($+4.5\%$) at the bottom 40% of the income distribution and $+0.02$ visits ($+3.3\%$) at the top 40%. The effect size is larger in absolute terms for low-income groups, but such heterogeneity is less clear or unobservable in relative terms. The effect sizes increase (decrease) if the increasing pre-trend difference is assumed to slow down (accelerate) in the post-treatment periods.

Although the point estimates are rather robustly positive, conclusions regarding the statistical significance are sensitive to how we account for clustering. The challenges are caused by the setting with only one treated cluster and a small number of comparison clusters.

Our study differs from most of the earlier Nordic literature by estimating the effects for the whole adult population and not just for adolescents at a specific birthday. The intervention is somewhat different: we study the impacts of a copayment abolition (i.e., a policy change) instead of individuals aging out of an exemption. Our analysis also illustrates some practical examples, such as the detrending, on conducting program evaluation in a setting where pre-trends are not exactly parallel and where there is only one treated cluster and a small number of comparison clusters.

2.4 The Impact of an Informational Campaign on Primary Care Utilization among Older Citizens: Evidence from a Randomized Field Experiment

The fourth essay [20] examines the impacts of an informational campaign on PHC use. We sent informational letters to randomly selected households to remind citizens aged 55 and above about the importance of early detection and treatment of health conditions and informed that the law reform of July 2021 abolished the copayment for nurse visits in PHC. The campaign took place in late 2021, and we hypothesized that unmet health needs had accumulated during the COVID-19 pandemic.

We had two aims. First, we aimed to analyze how effective PHC services are for marginal patients based on exogenous variation in PHC use. Second, we wanted to use a credible design to test whether information about a copayment abolition affects PHC use. Most of the program evaluation literature in economics, also in the cost-sharing literature, focuses on intent-to-treat effects and bypasses the role of communication, implicitly assuming that people are aware of the reforms. Instead of randomizing different cost-sharing schemes, we exploited exogenous variation in

information about a change in cost-sharing. As our trial occurred shortly after the policy change, many patients were plausibly not aware of it.

Regarding the intent-to-treat impacts of any reminder or information about the copayment abolition, the effect estimates on annualized nurse and GP visits are close to zero and statistically insignificant in a 6-month follow-up, the largest relative change being 0.7% in absolute value. When receiving any reminder is the intervention, our confidence intervals do not include effects larger than $+3.8\%$ for nurse visits and $+2.5\%$ for GP visits. Overall, we find no heterogeneity in the null average effects. When the copayment information is the intervention (smaller sample size), the confidence intervals are wider and include increases of up to $+6.9\%$ for nurse visits and $+3.6\%$ for GP visits. The effect estimates on the indicator of having any nurse or GP visits are also insignificant and mostly close to zero.

To potentially explain the lack of (large) observable effects, access to curative PHC visits is limited by gatekeeping and waiting times. Moreover, the tightness of supply was exacerbated by the COVID-19 pandemic. Additionally, if people feel well at the time of receiving a reminder, they probably do not get an appointment through the triage. There is a recent discussion on whether nudge interventions can have large impacts on average or whether there is sizeable publication bias. For us, our study was a healthy reminder that nudges do not always produce large impacts. It appears that our intervention did not induce a strong first stage for our first study aim. Furthermore, the lack of precision and data on first contacts (calls, etc.) make interpreting results challenging for our second study aim.

List of References

[1] Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen. Effects of nurse visit copayment on primary care use: Do low-income households pay the price? *Journal of Health Economics*, 102866, 2024. URL https://doi.org/10.1016/j.jhealeco.2024.102866.

[2] Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen. Does abolishing a copayment increase doctor visits? A comparative case study. *The B.E. Journal of Economic Analysis & Policy*, 24(1):187–204, 2024. URL https://doi.org/10.1515/bejeap-2023-0056.

[3] Liran Einav and Amy Finkelstein. Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, 16:957–982, 2018. URL https://doi.org/10.1093/JEEA/JVY017.

[4] J. P. Newhouse and the Insurance Experiment Group. Free for all? Lessons from the RAND Health Insurance Experiment. 1993. URL https://doi.org/10.7249/CB199.

[5] Toshiaki Iizuka and Hitoshi Shigeoka. Is zero a special price? Evidence from child health care. *American Economic Journal: Applied Economics*, 14:381–410, 2022. URL https://doi.org/10.1257/app.20210184.

[6] Tapio Haaga, Lauri Sääksvuori, and Jussi Tervola. Asiakasmaksujen ja lääkkeiden omavastuiden vaikutukset terveydenhuollon käyttöön. Hankkeen väliraportti. *THL Työpaperi 57/2022*, 57, 2022. URL https://urn.fi/URN:ISBN:978-952-408-003-3.

[7] Astrid Kiil and Kurt Houlberg. How does copayment for health care services affect demand, health and redistribution? A systematic review of the empirical evidence from 1990 to 2011. *European Journal of Health Economics*, 15:813–828, 2014. URL https://doi.org/10.1007/s10198-013-0526-8.

[8] Niklas Jakobsson and Mikael Svensson. The effect of copayments on primary care utilization: results from a quasi-experiment. *Applied Economics*, 48:3752–3762, 8 2016. URL https://doi.org/10.1080/00036846.2016.1145346.

[9] Anton Nilsson and Alexander Paul. Patient cost-sharing, socioeconomic status, and children's health care utilization. *Journal of Health Economics*, 59:109–124, 2018. URL https://doi.org/10.1016/j.jhealeco.2018.03.006.

[10] Naimi Johansson, Niklas Jakobsson, and Mikael Svensson. Effects of primary care cost-sharing among young adults: varying impact across income groups and gender. *The European Journal of Health Economics*, 20(8):1271–1280, 2019. URL https://doi.org/10.1007/s10198-019-01095-6.

[11] Lina Maria Ellegård, Gustav Kjellsson, and Linn Mattisson. An app call a day keeps the patient away? Substitution of online and in-person doctor consultations among young adults. 2022. URL http://hdl.handle.net/2077/68545.

[12] Mari Magnussen Landsem and Jon Magnussen. The effect of copayments on the utilization of the GP service in Norway. *Social Science & Medicine*, 205:99–106, 2018. URL https://doi.org/10.1016/j.socscimed.2018.03.034.

[13] Camilla Beck Olsen and Hans Olav Melberg. Did adolescents in Norway respond to the elimination of copayments for general practitioner services? *Health Economics*, 27(7):1120–1130, 2018. URL https://doi.org/10.1002/hec.3660.

[14] Naimi Johansson, Sonja C de New, Johannes S Kunz, Dennis Petrie, and Mikael Svensson. Reductions in out-of-pocket prices and forward-looking moral hazard in health care demand. *Journal of Health Economics*, 87, 2023. URL https://doi.org/10.1016/j.jhealeco.2022.102710.

[15] Benjamin Ly Serena. Revisiting offsets of psychotherapy coverage. 2021. URL https://www.econ.ku.dk/cebi/publikationer/working-papers/CEBI_WP_05-21.pdf.

[16] Marie Kruse, Kim Rose Olsen, and Christian Volmar Skovsgaard. Co-payment and adolescents' use of psychologist treatment: Spill over effects on mental health care and on suicide attempts. *Health Economics*, 31:92–114, 2022. URL https://doi.org/10.1002/hec.4582.

[17] Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen. Effects of nurse visit copayment on primary care use: Do low-income households pay the price? Version 4. 2023. URL https://osf.io/skuv9/.

[18] Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen. Do adolescents from low-income families respond more to cost-sharing in primary care? Version 2. 2023. URL https://osf.io/r83sg.

[19] Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen. Does abolishing a copayment increase doctor visits? A comparative case study. Version 2. 2023. URL https://osf.io/v7b5s.

[20] Tapio Haaga, Lauri Sääksvuori, and Jussi Tervola. The impact of an informational campaign on primary care utilization among older citizens: Evidence from a randomized field experiment. Version 2. 2023. URL https://osf.io/g72b5.

[21] Toshiaki Iizuka and Hitoshi Shigeoka. Asymmetric demand response when prices increase and decrease: The case of child healthcare. *The Review of Economics and Statistics*, 2021. URL https://doi.org/10.1162/rest_a_01110.

[22] Minke Remmerswaal, Jan Boone, Michiel Bijlsma, and Rudy Douven. Cost-sharing design matters: A comparison of the rebate and deductible in healthcare. *Journal of Public Economics*, 170:83–97, 2019. URL https://doi.org/10.1016/j.jpubeco.2019.01.008.

[23] Arthur P Hayen, Tobias J Klein, and Martin Salm. Does the framing of patient cost-sharing incentives matter? The effects of deductibles vs. no-claim refunds. *Journal of Health Economics*, 80, 2021. URL https://doi.org/10.1016/j.jhealeco.2021.102520.

[24] Aviva Aron-Dine, Liran Einav, Amy Finkelstein, and Mark Cullen. Moral hazard in health insurance: Do dynamic incentives matter? *The Review of Economics and Statistics*, 97:725–741, 2015. URL https://doi.org/10.1162/REST_a_00518.

[25] Liran Einav, Amy Finkelstein, and Paul Schrimpf. The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D. *The Quarterly Journal of Economics*, 130: 841–899, 5 2015. URL https://doi.org/10.1093/qje/qjv005.

[26] Marika Cabral. Claim timing and ex post adverse selection. *The Review of Economic Studies*, 84: 1–44, 2017. URL https://doi.org/10.1093/restud/rdw022.

[27] Marianne Simonsen, Lars Skipper, Niels Skipper, and Anne Illemann Christensen. Spot price biases in non-linear health insurance contracts. *Journal of Public Economics*, 203, 2021. URL https://doi.org/10.1016/j.jpubeco.2021.104508.

[28] Zarek C Brot-Goldberg, Amitabh Chandra, Benjamin R Handel, and Jonathan T Kolstad. What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *The Quarterly Journal of Economics*, 132:1261–1318, 2017. URL https://doi.org/10.1093/qje/qjx013.

[29] Haizhen Lin and Daniel W Sacks. Intertemporal substitution in health care demand: Evidence from the RAND Health Insurance Experiment. *Journal of Public Economics*, 175:29–43, 2019. URL https://doi.org/10.1016/j.jpubeco.2019.04.006.

[30] Audrey Guo and Jonathan Zhang. What to expect when you are expecting: Are health care consumers forward-looking? *Journal of Health Economics*, 67, 2019. URL https://doi.org/10.1016/j.jhealeco.2019.06.003.

[31] Aviva Aron-Dine, Liran Einav, and Amy Finkelstein. The RAND Health Insurance Experiment, three decades later. *Journal of Economic Perspectives*, 27:197–222, 2013. URL https://doi.org/10.1257/jep.27.1.197.

[32] Alexander Frankel and Maximilian Kasy. Which findings should be published? *American Economic Journal: Microeconomics*, 14:1–38, 2022. URL https://doi.org/10.1257/mic.20190133.

[33] Maximilian Kasy. Of forking paths and tied hands: Selective publication of findings, and what economists should do about it. *Journal of Economic Perspectives*, 35:175–192, 2021. URL https://doi.org/10.1257/jep.35.3.175.

[34] Isaiah Andrews and Maximilian Kasy. Identification of and correction for publication bias. *American Economic Review*, 109:2766–2794, 2019. URL https://doi.org/10.1257/aer.20180310.

[35] Abel Brodeur, Nikolai Cook, and Anthony Heyes. Methods matter: p-hacking and publication bias in causal analysis in economics. *American Economic Review*, 110:3634–3660, 2020. URL https://doi.org/10.1257/aer.20190687.

[36] Eva Vivalt. Specification searching and significance inflation across time, methods and disciplines. *Oxford Bulletin of Economics and Statistics*, 81:797–816, 2019. URL https://doi.org/10.1111/obes.12289.

[37] Colin F Camerer, Anna Dreber, Eskil Forsell, Teck-Hua Ho, Jürgen Huber, Magnus Johannesson, Michael Kirchler, Johan Almenberg, Adam Altmejd, Taizan Chan, Emma Heikensten, Felix Holzmeister, Taisuke Imai, Siri Isaksson, Gideon Nave, Thomas Pfeiffer, Michael Razen, and Hang Wu. Evaluating replicability of laboratory experiments in economics. *Science*, 351: 1433–1436, 2016. URL https://doi.org/10.1126/science.aaf0918.

[38] Stefano DellaVigna and Elizabeth Linos. RCTs to scale: Comprehensive evidence from two nudge units. *Econometrica*, 90:81–116, 2022. URL https://doi.org/10.3982/ECTA18709.

[39] Andrew Gelman and Eric Loken. The statistical crisis in science. *American Scientist*, 102: 460–465, 2014. URL https://doi.org/10.1511/2014.111.460.

[40] Benjamin A Olken. Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, 29:61–80, 2015. URL https://doi.org/10.1257/jep.29.3.61.

[41] Andrew C. Baker, David F. Larcker, and Charles C. Y. Wang. How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144:370–395, 2022. URL https://doi.org/10.1016/j.jfineco.2022.01.004.

[42] Joshua Angrist and Jörn-Steffen Pischke. The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, 24:3–30, 2010. URL https://doi.org/10.1257/jep.24.2.3.

[43] Edward E Leamer. Let's take the con out of econometrics. *The American Economic Review*, 73: 31–43, 1983. URL http://www.jstor.org/stable/1803924.

[44] Abhijit Banerjee, Esther Duflo, Amy Finkelstein, Lawrence Katz, Benjamin Olken, and Anja Sautmann. In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. NBER Working Paper No. 26993. 2020. URL https://doi.org/10.3386/w26993.

[45] Jeffrey Clemens and Michael Strain. The heterogeneous effects of large and small minimum wage changes: Evidence over the short and medium run using a pre-analysis plan. NBER Working Paper No. 29264. 2021. URL https://doi.org/10.3386/w29264.

[46] Nick Huntington-Klein, Andreu Arenas, Emily Beam, Marco Bertoni, Jeffrey R Bloem, Pralhad Burli, Naibin Chen, Paul Grieco, Godwin Ekpe, Todd Pugatch, Martin Saavedra, and Yaniv Stopnitzky. The influence of hidden researcher decisions in applied microeconomics. *Economic Inquiry*, 59:944–960, 2021. URL https://doi.org/10.1111/ecin.12992.

[47] Albert. J. Menkveld, Anna Dreber, Felix Holzmeister, Juergen Huber, Magnus Johannesson, Michael Kirchler, Sebastian Neusuess, Michael Razen, Utz Weitzel, and et al. Non-standard errors. *Journal of Finance (forthcoming)*, 2023. URL https://doi.org/10.2139/ssrn.3961574.

[48] Brian A Nosek, Charles R Ebersole, Alexander C DeHaven, and David T Mellor. The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115:2600–2606, 2018. URL https://doi.org/10.1073/pnas.1708274114.

TURUN
YLIOPISTO
UNIVERSITY
OF TURKU

Painosalama, Turku, Finland 2024

ISBN 978-951-29-9668-1 (PRINT)
ISBN 978-951-29-9669-8 (PDF)
ISSN 2343-3159 (Painettu/Print)
ISSN 2343-3167 (Verkkojulkaisu/Online)