

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

Tapio Haaga\* <sup>† a,b</sup>, Petri Böckerman<sup>c,d</sup>, Mika Kortelainen<sup>a,b</sup>, and Janne Tukiainen<sup>a</sup>

<sup>a</sup>Turku School of Economics, FI-20014 University of Turku

<sup>b</sup>Finnish Institute for Health and Welfare (THL), P.O. Box 30, FI-00271 Helsinki

<sup>c</sup>Jyväskylä University School of Business and Economics, P.O. Box 35, FI-40014 University of Jyväskylä

<sup>d</sup>Labour Institute for Economic Research LABORE, Arkadiankatu 7, FI-00100 Helsinki

August 2023

## Abstract

Nurses are increasingly treating primary care patients, yet nurse visits are understudied in the cost-sharing literature. We employ a staggered difference-in-differences design to examine the effects of adopting a 10-euro copayment for nurse visits on the primary care use of Finnish adults. We find that the copayment reduced nurse visits by 9–12% during a one-year follow-up. There is heterogeneity by income in absolute terms, but not in relative terms. The effects on general practitioner (GP) use are negative but small, with varying statistical significance. We also analyze the subsequent nationwide abolishment of the copayment. However, we refrain from drawing any causal conclusions from this due to the lack of credibility in the parallel trends assumption. Overall, our analysis suggests that moderate copayments can create a greater barrier to accessing care for low-income individuals. Additionally, it provides an example of using a pre-analysis plan for retrospective observational data.

**Keywords:** Cost-sharing, copayment, primary care, difference-in-differences, pre-analysis plan, blind analysis

**JEL codes:** I18, I14, I13, H42, I11

---

\*Corresponding author. E-mail address: tapio.haaga@utu.fi (T. Haaga)

<sup>†</sup>We thank Mikko Peltola for support, and editor Owen O'Donnell, two anonymous reviewers, Heather Royer, Henri Salokangas, Markku Siikanen, Lauri Sääksvuori, Jussi Tervola, and several seminar participants for comments and suggestions.

# 1 Introduction

An aging population puts pressure on primary care systems that provide comprehensive and accessible services as early and as close to the patient as possible. One solution to cope with this pressure is cost-sharing policies that aim to curb wasteful health care utilization. Therefore, to limit demand (and collect revenue), universal healthcare systems often charge moderate-sized copayments. While smaller copayments reduce the financial risks for patients, it is notable that even they can have significant effects on service utilization (Iizuka & Shigeoka, 2022). As needs-based prioritization by primary care professionals is conditional on patients having initiated the contact, cost-sharing may put more screening responsibility on the patients themselves. However, the patients' consumption choices can be far from optimal (Chandra et al., 2023).

A key concern is that copayments constitute a barrier to care that disproportionately affects low-income patients, potentially leading to a reduction in valuable care and contributing to inequality. Given the potential for large utilization effects, policymakers should set copayment levels, whether zero or nonzero, to achieve specific objectives (Iizuka & Shigeoka, 2022). Some countries, such as the UK and Germany, do not charge copayments for primary care GP visits. Several others, such as Ireland, Sweden, Denmark, and Finland, or regions within them, have recently abolished copayments of specific entry-level health services for specific vulnerable groups, such as low-income individuals, minors, and the elderly, for better accessibility. These services include visits to primary care nurses and general practitioners (GP) and to psychologists. Despite the vital role nurses play in many primary care systems, nurse visits and their copayments are understudied in the cost-sharing literature. Internationally, nurses are increasingly substituted for physicians in primary care settings to address physician shortages and improve efficiency (Maier & Aiken, 2016). This includes tasks such as examination, diagnosis, and the treatment of patients.

We analyze whether adopting a 10-euro copayment for curative primary outpatient care nurse visits affects the primary care use of Finnish adults. Our focus is on the heterogeneous effects by income level, as specified in our pre-analysis plan (Haaga et al., 2022). To this end, we employ

a difference-in-differences (DD) design and exploit the staggered adoption of the copayment in primary care areas in 2014–2019, as well as comprehensive administrative data. Most Finnish municipalities adopted the copayment at some point in 2014–2019 to collect more revenue, and the exact timing of the adoption in the treated areas is rather arbitrary. In Finland, nurses triage primary care patients and book appointments if needed. Many patients with acute infectious diseases or chronic conditions are directed to nurse visits rather than GP visits. Nurses can consult GPs, who write the vast majority of prescriptions, or book GP appointments for their patients if needed.

We find that the copayment adoption reduced the number of curative nurse visits by 9–12% during a one-year follow-up. There is statistically significant heterogeneity by income in absolute terms: the estimated decrease in the number of visits is more than two times larger for individuals at the bottom 40% of the income distribution, who also have a higher baseline service use, than at the top 40%. The effect size increases as income decreases. However, heterogeneity by income level is much weaker and statistically insignificant in relative terms (percentage changes). Moreover, we examine whether nurse visit copayments have spillover effects on GP use. The triaged patients do not have perfect control on which professional, if any, they get to visit. Thus, the number of GP visits may decrease if the nurse visit copayment reduced first contacts to health stations and a fraction of the missed contacts would have led to GP visits. We estimate a 2–5% reduction in GP visits, but our preferred estimates are closer to zero and often insignificant.

The relevant dimension of heterogeneity by income is context-specific. Our findings suggest that copayments can create a greater barrier to care for low-income individuals in terms of visits. This can be important for politicians who prioritize equality and the well-being of the most disadvantaged and consider cost-sharing as a means to influence the distribution of public resources. On the other hand, equal elasticities, defined in relative terms, may be appropriate for both low-income and high-income individuals when authorities make simple predictions about the impacts of cost-sharing changes. A relevant question is whether the copayments increase inequality in health. With this in mind, we also conducted some post-blind exploratory analyses. First, those with a prescription for diabetes or hypertension, whose baseline nurse use is much

higher, responded more strongly in absolute terms to the copayment than those without, but we find no evidence for differences in relative terms. Second, we observe statistically insignificant and imprecisely estimated increases in emergency department visits and unplanned hospitalizations for ambulatory care sensitive conditions. In short, these results do not offer conclusive evidence supporting either the existence or nonexistence of health effects.

The nurse visit copayment was abolished nationwide in July 2021. We also examine the impacts of this reform using a DD design, as stated in the pre-analysis plan. However, we refrain from drawing any causal conclusions from the abolishment policy due to the lack of credibility in the parallel trends assumption based on pre-trend patterns. Using the approach of Rambachan and Roth (2023) (post-blind), we demonstrate that even small post-treatment violations of parallel trends, compared to the observed pre-treatment violations of parallel trends, would make the confidence intervals very wide. In essence, the abolishment analyses are imprecise and provide evidence supporting neither the existence nor the nonexistence of effects.

Our analysis relates most closely to studies that analyze the impacts of moderate copayments on primary care use in public health insurance systems covering all citizens. Such studies have evaluated the effects of copayments for GP visits in the Nordic countries among children and adolescents (Johansson et al., 2019; Landsem & Magnussen, 2018; Nilsson & Paul, 2018; Olsen & Melberg, 2018) and the elderly (Johansson et al., 2023), and among Irish children (Nolan & Layte, 2017) and the elderly (Ma & Nolan, 2017). In Germany, Farbmacher and Winter (2013) study the impacts of doctor visit copayments, paid once in each calendar quarter, and focus on adolescents – see also the model-based approaches of Farbmacher et al. (2017) and Kunz and Winkelmann (2017) to account for the nonlinear nature of the copayment. Regarding other services, Han et al. (2020) and Iizuka and Shigeoka (2022) examine the impacts of copayments among small children and adolescents in Taiwan and Japan, while Kruse et al. (2022) study the effects of abolishing a copayment for psychologist treatment among Danish adolescents.

Despite the vital role of nurses in many primary care systems, we are among the first to focus on nurse visits and their copayments. We expect that nurse visits are more responsive than GP

visits to copayments of equal size: i) Nurse visits tend to be more accessible, with less gatekeeping and shorter waiting times. This implies lower indirect costs of seeking care and potentially a larger role for the copayment in affecting demand. ii) The equal-sized copayment represents a larger proportion of the production costs for nurse visits, i.e., a higher coinsurance rate. Patients may value GP visits more as GPs have more extensive education, the authority to write prescriptions, and they can make referrals to specialists.<sup>1</sup>

We study a large adult population and both the direct effects on nurse visits and also the indirect effects on GP use. Moreover, we employ a staggered DD design with methods that are robust to heterogeneous effects. These features have several important benefits: i) the availability of several events reduces the risk of external shocks systematically biasing the estimates, ii) the short lag between the policy decision and its implementation limits anticipation effects, and iii) we provide estimates for the larger adult population in a one-year follow-up instead of individual-level effects local to a specific birthday. To our understanding, we are among the first to use a pre-analysis plan in a nonexperimental study on cost-sharing and to conduct pre-specified heterogeneity analysis by income level. Several earlier studies have examined treatment effect heterogeneity by income but by focusing on narrow age groups, mainly children or adolescents (Han et al., 2020; Johansson et al., 2019; Kruse et al., 2022; Nilsson & Paul, 2018), or the elderly (Johansson et al., 2023).

The rest of the paper is structured as follows. Section 1.1 discusses pre-analysis plans (PAP) and outlines our PAP-based workflow. Section 2 presents the institutional background and Section 3 the data. Section 4 describes our empirical approach for the staggered adoption, and Section 5 reports the results. Section 6 summarizes the analyses for the copayment abolishment. Section 7 concludes. The Online Appendix contains additional figures and tables, post-blind supplementary analyses for the copayment adoption, regression analyses for the copayment abolishment, a description of the data construction, and documentation of deviations from the PAP.

---

<sup>1</sup>Copayment increases for prescription drugs reportedly reduced doctor visits in Germany (Winkelmann, 2004).

## 1.1 Pre-analysis plans, and research transparency

One proposal to improve the distinctiveness between hypothesis generation and hypothesis testing is to commit to a pre-analysis plan (PAP) *ex ante* (Nosek et al., 2018; Olken, 2015). In its purest form, a PAP specifies a complete mapping from data to what statistics and estimates will be reported and how (Kasy, 2021). In practice, there is considerable heterogeneity about what PAPs should and do entail (Banerjee et al., 2020; Ofosu & Posner, 2023). PAPs have thus far been uncommon in observational economics, in contrast to experimental economics (Banerjee et al., 2020). Early observational contributions include Bohm and Lind (1993) and Neumark (2001), but many other studies experimenting with PAPs (in varying ways) are now published or ongoing.<sup>2</sup>

Observational research is often retrospective and with no way to prove that the PAP truly preceded the analysis. Debates about PAPs for observational studies have emphasized scenarios with a verifiable firewall between design and analysis, for example, by registering the PAP before the intervention is implemented (Burlig, 2018; Miguel, 2021; Ofosu & Posner, 2023). However, a strict requirement of credible firewalls for the benefit of the reader can be counterproductive if it hinders the adoption of PAPs in common situations where such firewalls are impossible. We view PAPs as a self-control tool to improve the quality and transparency of our research, for example, by avoiding multiple-testing or hindsight bias. Nosek et al. (2018) also argue that PAPs accompanied by transparent reporting of what was and was not known in advance maximize the diagnosticity of statistical inference even with pre-existing data, relative to having no pre-registration.

Our study uses a PAP and analysis blinding with retrospective observational data in the absence of a fraud-proof firewall (Haaga et al., 2022). Besides being a self-control tool for us, the workflow provides additional transparency to readers. For the analysis of the staggered adoption of the copayment, we had all data available when designing the analyses. Our outcomes and the policies on nurse visit copayments had not been linked previously, which allowed us to blind the causal relation of interest for the PAP. Specifically, we randomly assigned municipalities into

---

<sup>2</sup>Recent examples of PAP-based observational studies in economics include Neumark and Yen (2022) and Clemens and Strain (2021) in labor economics, Altındağ and O’Connell (2023) in development economics, and Cesarini et al. (2016), Lindqvist et al. (2020), and Clemens et al. (2020) in health economics.

placebo policies, using the real adoption dates but randomly assigning municipalities into them, and specified our statistical approach and demonstrated the reporting of the results in the PAP without observing the real results.<sup>3</sup> For the copayment abolishment, we did not have access to microdata on outcomes from 2021–2022 (and thus, no post-treatment outcomes) when writing our PAP, verifiable by Statistics Finland. We wrote the statistical programs as if the abolishment occurred years earlier on July 1st, 2018. Some aggregated post-treatment outcome data were publicly available at the municipality level; however, these data did not influence our work.

In practice, our PAP included a placebo report based on blinded data, mimicking the structure of the research paper, and the corresponding statistical codes (Haaga et al., 2022). The benefits of using blinded data for PAPs have previously been discussed by Olken (2015) and Nosek et al. (2018). PAPs should not prevent being flexible to reasonable changes *ex post*. Deviations from plans are common, and PAPs can still provide substantial benefit if researchers transparently report all changes and their reasons (Nosek et al., 2018). We document and discuss our changes in Online Appendix Section A.4. This article is a combination of pre-registered confirmatory analyses and post-blind exploratory and supplementary analyses. We separate tables and figures that were not pre-registered by using the label “post-blind” in the text and in the table and figure notes.

One deviation from our PAP is worth highlighting here. Our plan was to examine and report both the staggered adoption and the later abolishment of the copayment with equal weight. The final paper focuses on the staggered adoption because its design turned out to be much more credible, and we moved parts of the abolishment analyses to the Online Appendix. For the abolishment, we could not credibly assume parallel trends, the essential identification assumption, based on pre-treatment outcomes observed *ex post*. According to Banerjee et al. (2020), the final paper should be written and judged as a distinct object from “the results of the PAP”, and there are cases where it is reasonable to exclude pre-registered analyses – if infeasible or irrelevant – from the final paper and report them elsewhere, as long as all changes are documented transparently.

---

<sup>3</sup>We had previously examined the impacts of GP visit copayments using similar data from the 2010s by 1) focusing on exempted minors (Haaga et al., 2023a), 2) exploiting the abolishment of the GP visit copayment in one large municipality in 2013 (Haaga et al., 2023b), and 3) conducting exploratory analyses on the impacts of GP visit copayment increases.

## 2 Institutional Background

**Publicly funded primary outpatient care in Finland.** Primary outpatient care is provided for adults by three sectors: public primary care, occupational healthcare, and private clinics. These sectors target different patient populations and differ with respect to gatekeeping, out-of-pocket costs, and waiting times. Publicly funded primary care is the main provider for those not entitled to occupational care or who cannot afford the fees for private physicians. These groups include pensioners, the unemployed, and low-income individuals. The employed and more affluent adults usually prefer private clinics or occupational healthcare due to faster access and much less gatekeeping. Occupational healthcare is also free of charge at the point of use. Public primary care charges copayments, 21 euros per GP visit at maximum and approximately 10 euros per curative nurse visit, and has waiting times that may be long for nonurgent care.

Municipalities form publicly funded primary care areas (health centers) on their own or in cooperation with others. Every citizen has their designated health station determined by where they live. In some primary care areas, all health stations may be available on a visit-by-visit basis. Since 2014, citizens have been able to choose their health station once a year, but active choices have not been common. Municipal services are financed through state transfers, municipal taxes, copayments, and borrowing. The state guides copayment policies by setting which groups or services are exempted (based on the Social and Health Care Client Fees Act) and maximum copayment levels (in the corresponding government decree). Within these constraints, primary care areas set their own policies.

The public system is characterized by limited supply and labor shortages. Cohort sizes in medical schools are fixed, and the public and the private sector compete for doctors. Primary care areas face challenges in hiring nurses at the prevailing wage level determined by sector-based collective bargaining. The challenges are reinforced by the fact that central and local governments have been running fiscal deficits for years, which is expected to continue.

**The vital role of nurses in public primary outpatient care in Finland.** Patients can contact public health stations via phone (the primary channel) or visit in person. Upon contact,



patients undergo triage, a process primarily handled by nurses or public health nurses who then schedule appointments with professionals if necessary. As a result of task-shifting from GPs to nurses, an increasing number of patients are directed toward nurse visits instead of GP visits. Nurses offer a range of services including urgent care for conditions such as flu and stomach flu and monitoring visits for patients with chronic conditions, such as diabetes, hypertension, COPD, asthma, and dementia. Nurses can also provide wound care, administer drug injections and infusions, perform ear irrigation, and direct patients to GPs. Specialist visits in public healthcare require a doctor's referral, and most prescriptions are written by doctors.<sup>4</sup> A recent trend in work arrangements is that the triaging nurses not only guide patients to the appropriate professional but try to resolve the cases immediately. This often involves real-time consultation with a GP based on nurse–GP pairs or larger multidisciplinary teams, without the need for the patient to see the GP directly.

In a cross-country comparison that includes Europe, the USA, Canada, Australia, and New Zealand, Finland has been classified as part of a group of countries where there has been “extensive” task shifting from GPs to nurses in primary care (Maier & Aiken, 2016). This group also includes the USA, the UK, Canada, Australia, Ireland, the Netherlands, and New Zealand. Moreover, as of 2019, Finland had the second highest ratio of nurses to doctors among OECD countries (OECD, 2021), with all the countries identified by Maier and Aiken (2016) as having “extensive” task shifting surpassing the OECD average. Regarding curative visits at the municipal level in Finland, utilization of GP visits shows a strong correlation with nurse use, while there is noticeable variation in the ratio of nurse visits to GP visits (Figure A1, post-blind).

**Copayments for nurse visits.** Finland adopted restricted prescription rights for nurses in 2010. Related to that, the law on copayments was changed to allow primary care areas to charge a copayment for curative nurse visits, no longer specifying professions (e.g., physicians) whose visits can be subject to copayments. However, the decree continued to mention explicitly only doctor visits. This confusion likely explains why no areas immediately adopted the nurse visit

---

<sup>4</sup>Less than 700 nurses had restricted prescription rights in 2023, and they can describe drugs to treat, for example, urinary tract infections or tonsillitis.

copayment, first introduced in 2014. Many other areas adopted it to collect more revenue once they became aware of the possibility.

The staggered adoption is illustrated in Figure 1. At the end of 2019, half of the population lived in areas charging the copayment, with the vast majority of municipalities charging it. The Online Appendix includes a complementary map (Figure A2). In 2021, 80% of the municipalities with the copayment charged it for three visits annually, and by far the most common per-visit copayment was 11 euros, the population-weighted mean being 12 euros (Figure A3). No significant area-specific changes have been made to the levels after the adoption, except for minor inflation adjustments.

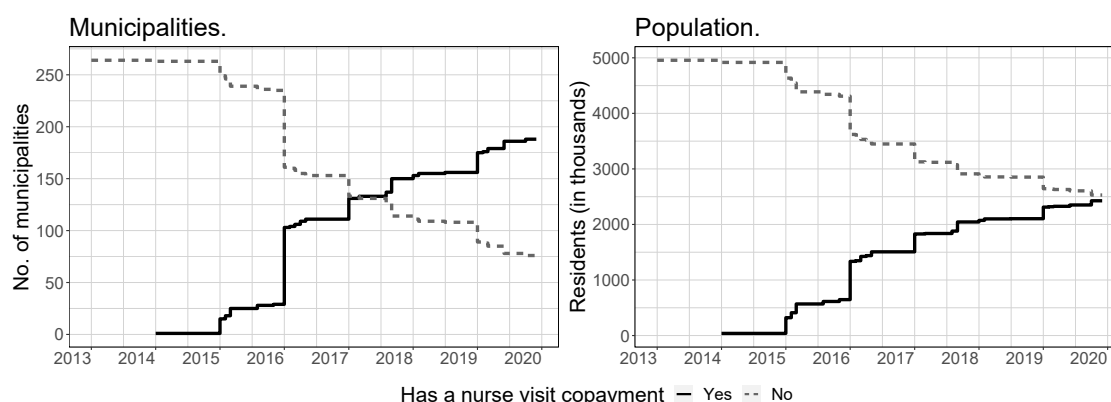


Figure 1: Staggered Adoption of the Nurse Visit Copayment.

*Notes:* We take municipalities in mainland Finland (293 in 2022) and use the 2022 municipal boundaries and population sizes from the end of 2019. The sample contains those municipalities whose policies on copayments for curative nurse visits we observe in our data collection.

However, the nurse visit copayment was later abolished nationally in July 2021 when the national government conducted a reform to the act on copayments to reduce barriers to access and health inequality. The reform included no major changes to GP visit copayments. More than 200 municipalities, on average disproportionately small and rural, and almost three million people were affected by the nurse visit copayment abolishment (Figure A4; Figure A5; Table A1).

Several policies protect financially vulnerable low-income patients from healthcare costs. There are annual out-of-pocket caps for public healthcare services and prescription drugs of 692

euros and 592 euros, respectively (in 2022). Households with the lowest incomes and only little wealth can apply for social assistance, a means-tested last-resort benefit, which can also cover out-of-pocket costs for public health care and prescription drugs. The law on copayments requires that, for some public services, financially vulnerable patients can apply for an exemption or a lowered copayment. This right does not apply to nurse visit copayments, but some areas may still exempt individuals based on applications. A few primary care areas provide general exemptions to specific low-income groups, such as those with the lowest national pension or those receiving social assistance.

### 3 Data

We combine several Finnish national administrative registers using person IDs. The data contain contacts in publicly funded primary care and hospitals, social assistance recipients, drug prescriptions since 2018, and socioeconomic characteristics of all individuals who have a permanent residence at year's end.<sup>5</sup> We observe age and the municipality of residence, which are used to link visits to copayment policies. We construct a variable for equivalized family disposable income for each individual and calculate population sizes for each municipality.

We also use publicly available data on each municipality's primary care area in 2021 (from the Association of Finnish Municipalities). Two publicly available registers listing social and healthcare organizations are linked to primary care contacts (THL). We create three tables mapping areas to copayment policies. The first reports whether a given municipality had adopted the nurse visit copayment by the end of 2019 and the possible adoption date. These data were collected from municipal documents, websites, and news in local media. The search was based on the publicly available dataset on nurse and GP visits copayments (from THL), which we also use

---

<sup>5</sup>The data on primary care use (Register of Primary Health Care Visits), specialized healthcare use (Care Register for Health Care), and social assistance recipients (Register of Social Assistance) are all administered by the Finnish Institute for Health and Welfare (THL). The socioeconomic data comprise Statistics Finland's FOLK modules "basic", "family", and "income". Prescriptions (Kanta Prescription Center) are administered by the Social Insurance Institution of Finland.

for GP visit copayments in 2013–2018 (the second table). The third table reports the copayments in the summer of 2021, collected from the websites of primary care areas.<sup>6</sup>

The analysis for the staggered adoption of the copayment uses data from 2013–2019, restricting to pre-pandemic (COVID-19) years. We include those individual-by-year observations in which the person is 25 years or older. The aim is to exclude minors, who are exempted from the copayment, and students, who have access to student healthcare. The primary outcomes are the annualized number of curative nurse and GP visits per capita in publicly funded primary care, constructed by multiplying monthly visits per capita by 12. The secondary outcomes include the share of individuals receiving social assistance and the annual sum of received basic social assistance, both defined at the family level.<sup>7</sup> Our pre-specified choice is to estimate the effects separately for individuals at the bottom 40% and the top 40% of the equivalized family disposable income distribution in all analyses.<sup>8</sup>

We discuss in detail how we clean and construct our analysis data in Online Appendix Section A.3. Ultimately, we have an unbalanced panel at the municipality-by-time-period-by-income-decile-by-outcome level. Time period is a month except for the annually measured sum of social assistance. The panel is unbalanced because we exclude some observations due to quality problems, mainly for primary care outcomes. When the national data collection started in 2011, not all areas were able to transfer primary care data from their electronic health record (EHR) systems to the national register. Later changes in the EHR systems may also be visible in the data as a sudden but short drop to a near zero value in aggregate contacts. The details of how we detect and exclude observations with data quality concerns are provided in Online Appendix Section A.3.

The analysis of the nationwide abolishment of the nurse visit copayment in July 2021 is based on 12 pre-treatment and 11 post-treatment months (from 7/2020 to 5/2022), requiring a balanced panel. Socioeconomic data are from 2020. The observation window excludes early 2020

---

<sup>6</sup>We thank Katja Ilmarinen, who had gathered the same information independently, for allowing us to cross-validate our information.

<sup>7</sup>Social assistance is a means-tested last-resort benefit for households.

<sup>8</sup>We focus on two groups for parsimony. Using smaller groups than the bottom 40% and the top 40% has two disadvantages: smaller samples and larger variation, and the fact that the share of social assistance recipients is larger at the bottom of the income distribution, potentially attenuating estimates as the benefit can cover copayments.

when the supply and demand shocks caused by the COVID-19 pandemic were largest. For data construction, we assume that the effects of the earlier adoption of the copayment fully accumulate within one year, excluding municipalities that adopted the copayment less than 12 months before the start of the observation window.

## 4 Empirical Approach for the Staggered Adoption

**Research design.** We use a staggered difference-in-differences (DD) design with an irreversible treatment. For each event, we have both never-treated and later-treated municipalities as controls.<sup>9</sup> The never-treated areas include the six largest cities and differ from the treated areas. We do not view the decision to adopt the nurse visit copayment as quasi-random. However, the decision on when to adopt conditional on adopting seems much more arbitrary. Our interpretation of municipal decision-making is that there is potential randomness in the timing of when public servants became aware of the possibility to charge the copayment and, consequently, in the treatment timing.<sup>10</sup>

In staggered settings, two alternative identification assumptions are relevant: a model-based parallel trends assumption (PTA) or a stricter design-based assumption of (quasi-)random treatment timing. The causal interpretation of our main analysis relies on the validity of the PTA, i.e., we assume that the outcomes for the treated cohort and for the comparisons (not-yet-treated or never-treated) would have followed parallel trends in the absence of treatment. The PTA may hold in both levels and logs, or in levels but not logs, or *vice versa* (Roth & Sant’Anna, 2023). As stated in our PAP, we report the main results on primary care use in both levels and logs.<sup>11</sup> Both models seem plausible based on pre-treatment trends in event-study plots reported in Section 5.<sup>12</sup> However, similar pre-trends are not sufficient for the PTA to be valid,

---

<sup>9</sup>To be specific, copayment policies are set at the primary care area level. However, we analyze the data at the municipality level for practical purposes.

<sup>10</sup>The earlier-treated and later-treated municipalities, defined by median event date, are rather similar in some key features (Table A1, post-blind).

<sup>11</sup>The results in levels are shown first while most of the results in logs are documented in the Online Appendix.

<sup>12</sup>Complementarily, we will estimate models that allow for smooth (linear) pre-trend differences and models that do not use the never-treated municipalities.

and the PTA is inherently untestable. Another caveat is that in DD designs it is not possible to rule out with complete certainty the influence of other exposures concurrent with the intervention of interest. If they exist, we can only identify the net effect. However, the staggered design (several events) reduces the risk of external shocks systematically biasing our results.

Still, two potential threats to identification should be noted. First, the fiscal challenges that led many municipalities to adopt the nurse visit copayment may also have led to other concurrent policy changes affecting health care use. Most importantly, GP visit copayments could have been increased simultaneously. The central government increased the maximum GP visit copayment from 16.10 euros in 2015 to 20.90 euros in 2016. Municipalities responded differently: many made the increase instantly in 1/2016, some made it later, and some not. However, we find that the nurse visit copayment adoption was correlated with only a 1-euro increase in GP visit copayments (small in both absolute terms and percentages) based on the methods of our main analysis (Figure A6, post-blind). We chose to model the setting as a single-treatment design in line with our PAP, not controlling for GP visit copayments.<sup>13</sup> We are not aware of anecdotal evidence for supply reductions motivated by cost savings, such as health station closures, shorter opening times, or staff reductions, concurrent to the copayment adoption. If they existed, we would expect to detect rather similar reductions in both nurse and GP visits in the main analysis, which is not the case.

Second, once the copayment is adopted, preventive and curative visits must be distinguished for charging purposes. If this affects how contacts are recorded, the number of recorded visits may change even if the underlying use does not. However, we find no evidence of preventive-labeled visits crowding out curative-labeled visits: adopting the copayment for curative nurse visits was not correlated with the number of preventive nurse visits based on the methods of our main analysis (Figure A7, post-blind).

We do not have to account for utilization spillovers across areas due to Finland's publicly funded primary care system where nonurgent care is provided by a designated health station

---

<sup>13</sup>With this choice, we likely trade off some bias for lower variance and more external validity compared to using the estimator proposed by de Chaisemartin and D'Haultfœuille (2023).

determined by the location of residence. Neither do we expect noticeable anticipation effects. The implementation time from the political decision is often a month or less, and few citizens likely pay much attention to the minutes of the municipal committees.

**Econometric methods.** We use stacked regressions (Cengiz et al., 2019; Gormley & Matsa, 2011) as our baseline. For robustness checks, we use the Callaway and Sant’Anna (2021) (CS) estimator. Both estimators are robust to biases in conventional two-way fixed effects (TWFE) regression models caused by staggered treatments and treatment effect heterogeneity (Baker et al., 2022) and thus are attractive *ex ante* when designing the PAP. If misspecified, the conventional models project heterogeneous treatment effects onto group and time fixed effects, rather than treatment status (Gardner, 2021). Both the static (Goodman-Bacon, 2021) and event-study (Sun & Abraham, 2021) TWFE specifications suffer from these biases.

In practice, stacking ensures that earlier-treated units are not used as controls for later-treated units. It transforms the staggered setting into event-specific datasets that are ultimately stacked (or pooled) together before conventional TWFE regressions are fitted. We first create a separate dataset for each event, including the treatment cohort and all clean controls that are unexposed (not-yet-treated) in the window of 24 months before and 12 months after the copayment adoption. We only use data from the 36-month window and include events with at least 12 post-treatment months. Depending on the outcome, we exclude several municipality–year observations due to data quality concerns (see Online Appendix Section A.3). We require balanced panels in event (or relative) time as the baseline. These event-specific datasets are stacked for estimation. Our static TWFE specification includes event-specific unit and time fixed effects:

$$y_{mte} = \alpha_{me} + \gamma_e + \delta^{DD} D_{mt} + \varepsilon_{mte}. \quad (1)$$

Here, subscripts  $m$ ,  $t$ , and  $e$  denote municipality, month, and event-specific datasets, respectively, and  $\alpha_{me}$  and  $\gamma_e$  represent event-specific municipality and calendar month fixed effects.  $D_{mt}$  is a dummy for post-treatment periods in the treated municipalities. We weight by population due to

heterogeneity in municipality size.<sup>14</sup> Standard errors are clustered by municipality.<sup>15</sup> We estimate the model separately for individuals at the bottom 40% and the top 40% of the income distribution (equivalized family disposable income).

We complementarily use two other stacked specifications modified from Model 1. Model 1.1, a dynamic event-study version of Model 1, is used for event-study plots:

$$y_{mte} = \alpha_{me} + \gamma_{te} + \sum_{l=-24, l \neq -1}^{11} \mu_l D_{mte}^l + \varepsilon_{mte}. \quad (1.1)$$

The coefficients of interest,  $\mu_l$ , represent all leads and lags ( $l = -1$  omitted as a reference).  $D_{mte}^l$  is a dummy for the treated areas for observations  $l$  months from the copayment adoption in the event-specific dataset  $e$ . Both relative time  $l$  and the treated are defined by  $e$ .

Next, Model 1.2 allows for a linear pre-trend difference between the treated and the comparisons in contrast to Model 1. Its PTA concerns deviations from a linear pre-trend difference: Model 1.2 is preferable to Model 1 if there exists a linear pre-trend difference that would have continued in the absence of treatment.

$$y_{mte} = \alpha_{me} + \gamma_{te} + \theta d_{me} t_e + \sum_{l=0}^{11} \mu_l D_{mte}^l + \varepsilon_{mte}. \quad (1.2)$$

The coefficients of interest,  $\mu_l$ , represent all lags.  $D_{mte}^l$  is a dummy for the treated areas for observations  $l$  months from the copayment adoption in the event-specific dataset  $e$ , as in Model 1.1. Regarding the linear pre-trends,  $d_{me}$  is a dummy for the municipality being treated in the event-specific dataset  $e$ , and  $t_e$  denotes time relative to that event. We report the mean of  $\mu_l$  over  $l$  as our point estimate.

The stacking estimator is efficient as it uses OLS to derive weights on the event-specific DD estimates, trading off bias for efficiency. However, the use of variance weighting may lead to

---

<sup>14</sup>We discuss weighting in Online Appendix Section A.3 and present uniformly weighted robustness checks in Online Appendix Section A.1.

<sup>15</sup>Our stacked data contain multiple copies of the same observation, as a municipality can belong to several event-specific datasets. Thus, clustering at the municipality-by-event level is not appropriate.



inconsistency for the *sample-average* ATT (average treatment effect on the treated) (Baker et al., 2022). In fact, Gardner (2021) shows that the estimator identifies an average of event-specific ATTs, weighted by event-specific treatment variance and sample size. For the PAP, we valued the simplicity and implementability of stacking and its ability to accommodate triple difference models for testing treatment effect heterogeneity.

As an alternative to stacking, we use the CS estimator (Callaway & Sant’Anna, 2021). The aim is to identify a group-time average treatment effect, allowing for treatment effect heterogeneity over cohorts and time. The group-time ATTs can be aggregated to construct measures of overall treatment effects. We provide both event-study-type estimates and a static estimate that is the average of all group-time ATTs, weighted by group size.<sup>16</sup> The authors propose several two-step plug-in estimators for group-time ATTs: first estimate nuisance functions and then plug their fitted values into the sample analogue of the group-time ATT. When the never-treated units are used for comparison, the PTA is assumed only from the last pre-treatment period on.<sup>17</sup> We use outcome regression, weight by population, and cluster standard errors by municipality. Events with at least 12 follow-up months are included. The dataset is balanced in calendar time, excluding municipalities with data quality concerns in the study window. Our baseline is to exclude the years 2013 and 2019 when analyzing primary care use to increase the number of sample municipalities.<sup>18</sup> Regarding social assistance use, we use all data from 2013–2019. The data are aggregated to the municipality-by-time-period level for estimation.

**Assessing treatment effect heterogeneity.** We focus on whether the copayment had heterogeneous utilization effects by income. The analysis compares the bottom 40% of the income distribution (equivalized family disposable income) to the top 40% throughout the study, as defined in our PAP. First, we estimate the effects separately for individuals at the bottom 40% and the top

---

<sup>16</sup>Our stacking analyses use a 12-month follow-up. Here, follow-up varies by treatment group. This also implies putting more weight on the earlier-treated cohorts.

<sup>17</sup>Thus, the assumption does not restrict pre-treatment trends. However, the PTA is different and restricts pre-trends when the not-yet-treated are used as comparisons (Callaway & Sant’Anna, 2021).

<sup>18</sup>The exclusion of 2013 trades off one event and 12 months of data for a larger number of municipalities. In many cases, the primary care data quality concerns reported in Online Appendix Section A.3 occurred early in the panel. The exclusion of 2019 leads us to keep one large never-treated municipality that changed its EHR system in Spring 2019.

40% of the income distribution by using the DD framework described above to a) visually illustrate the design and results and b) to estimate the magnitude of the effects separately for income groups.

Second, we use a triple difference (DDD) model with the stacked data to formally test whether there is heterogeneity in effects by income. In the DDD model, we compare the evolution of outcomes at the bottom 40% of the income distribution to that at the top 40% in both treatment and comparison areas. The PTA is now assumed in ratios (Olden & Møen, 2022), concerning the relative outcomes of the income groups. We use the following specification:

$$\begin{aligned}
y_{mgte} = & \alpha + \beta_{1e}Treat_{me} + \beta_{2e}Bottom40_{ge} + \beta_{3e}Post_{te} + \beta_{4e}Treat_{me} \times Bottom40_{ge} \\
& + \beta_{5e}Treat_{me} \times Post_{te} + \beta_{6e}Bottom40_{ge} \times Post_{te} \\
& + \gamma Treat_{me} \times Bottom40_{ge} \times Post_{te} + \epsilon_{mgte}.
\end{aligned} \tag{2}$$

Here, subscripts  $m$ ,  $g$ ,  $t$ , and  $e$  denote municipality, income group, time (month), and event-specific datasets, respectively.  $Treat_{me}$  and  $Post_{te}$  are two sets of indicators.  $Treat_{me}$  indicates whether municipality  $m$  is used in the treatment group in the dataset  $e$ . Consequently, the never-treated municipalities always have  $Treat_{me} = 0$ , but the later-treated can get either 1 or 0 for  $Treat_{me}$  depending on whether they are used in the treatment group or in the comparison group for the earlier treated in the dataset  $e$ .  $Post_{te}$  indicates the calendar months following the treatment month in the dataset  $e$ .  $Bottom40_{ge}$  is a dummy for the bottom 40% of the income distribution (0 for those in the top 40%), and  $\gamma$  is the coefficient of interest. Other coefficients (from  $\beta_{1e}$  to  $\beta_{5e}$ ) are event-specific. We again weight by population size and cluster standard errors by municipality.

## 5 Results: Staggered Adoption

**Pre-trend plots.** Figure 2 plots the trends in curative nurse visits for individuals at the bottom 40% and the top 40% of the income distribution in treatment and comparison municipalities based on the stacked dataset. We find that nurse visits decreased in the treated municipalities after the adoption of the copayment compared to the comparison municipalities. The decrease was

0.10–0.15 annualized visits at the bottom 40% of the income distribution and approximately 0.05 visits at the top 40%. Nurse visits were increasing in both policy groups before the copayment adoption, and the PTA is arguably plausible. After the adoption, the growth continued in the comparison municipalities, but nurse visits decreased in the treated municipalities. The effects on GP visits, in contrast, are small or zero, and there may be a small decreasing pre-trend in GP visits in the treated areas relative to the comparisons (Figure A8).<sup>19</sup> Neither do we observe any clear effects on receiving social assistance (Figure A9).

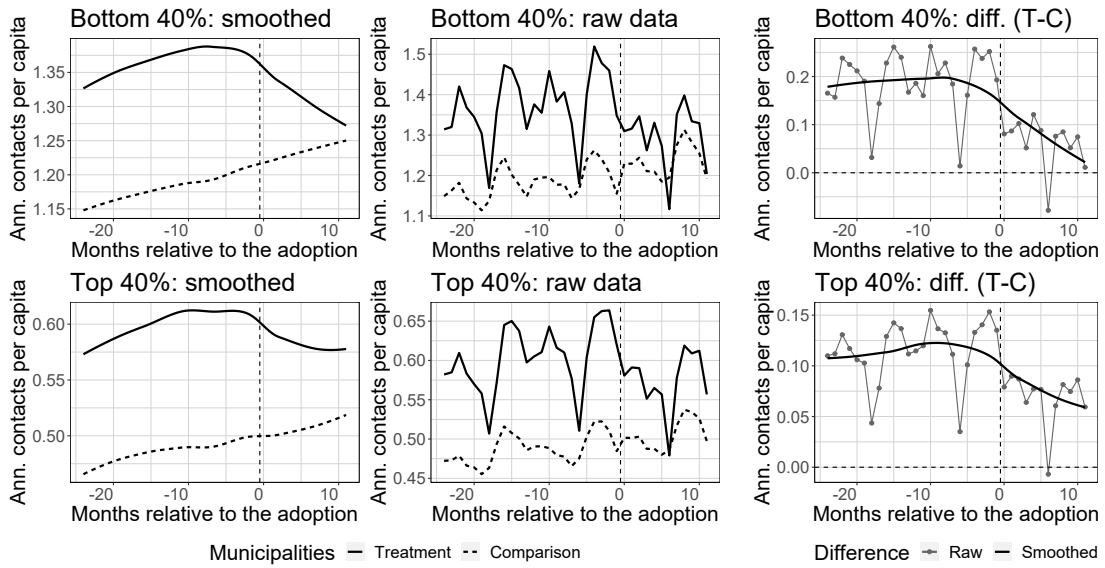


Figure 2: Adoption: Evolution in Nurse Visits.

*Notes:* The dataset is stacked, and event-specific datasets balanced. The outcome is the number of annualized curative nurse visits per capita. We weight by population due to heterogeneity in municipality size. Treatment municipalities adopted the nurse visits copayment at time 0 in relative time. The left column contains smoothed conditional means, fitted with local linear regression. The raw data is illustrated in the middle column, while the difference between treatment and comparison areas is depicted in the right column. Bottom 40% and top 40% refer to the distribution of equivalized family disposable income. The observed reductions in nurse use occurring every twelve months are likely explained by summer holidays and reduced supply in July, which appears to disproportionately affect the treated municipalities that are smaller on average. Figure 1 shows that January was a common adoption month.

We also estimate dynamic event-study regressions using the stacked data (Model 1.1),

<sup>19</sup>For GP visits, we consequently prefer the stacked specification that allows for a linear pre-trend difference or the CS estimator with the never-treated units as comparisons that assumes parallel trends only from the last pre-treatment period on. These estimates should be closer to zero than the estimates from our baseline stacked specification, assuming parallel trends in every period.

comparing the evolution of outcomes between the treated and unexposed municipalities. The event-study plots on nurse and GP visits are reported in Figure 3. Consistent with the pre-trend plots, nurse visits decreased in the treated municipalities after the copayment adoption compared to the comparison municipalities. The potential effects on GP visits are negative but close to zero. No clear effects on receiving social assistance are observed (Figure A10). The reduction in nurse visits is larger at the bottom 40% of the income distribution in absolute terms. However, we do not find such a pattern in relative terms using logarithmic outcomes (Figure A11, post-blind).

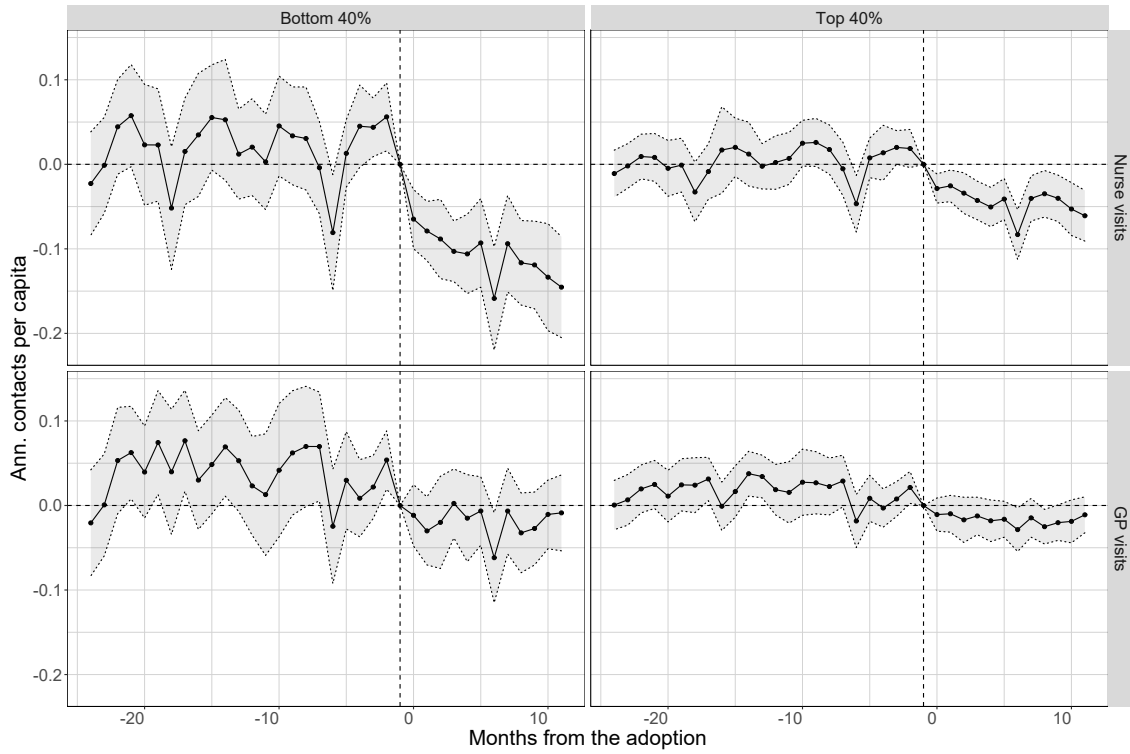


Figure 3: Adoption: Event-Study Plot on Primary Care Use with Stacked Data.

*Notes:* We use Model 1.1. The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. The dataset is stacked, and event-specific datasets balanced. We compare the evolution of annualized contacts per capita between treated and unexposed municipalities. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. Bottom 40% and top 40% refer to the distribution of equivalized family disposable income.

**Main results.** We construct static treatment effect estimates by fitting Model 1 to the stacked data. The results on annualized primary care contacts per capita are reported in Table 1.

Both nurse visits and GP visits decrease for individuals at both the bottom 40% and the top 40% of the income distribution, and these reductions are statistically significant. Annualized nurse visits decrease by  $-0.13$  visits ( $-9.3\%$ ) at the bottom 40% and by  $-0.05$  visits ( $-8.0\%$ ) at the top 40%. The reductions in GP visits are closer to zero:  $-0.06$  visits ( $-3.9\%$ ) at the bottom 40% and  $-0.03$  ( $-4.7\%$ ) at the top 40%. The estimates are clearly heterogeneous in absolute terms, and the lower end of the income distribution responds more. They also have much higher baseline utilization. However, such heterogeneity is not found in relative terms. The estimates on social assistance outcomes are insignificant: the estimate on the share of the population receiving social assistance is close to zero, while the estimate on the annual sum of received social assistance is positive and marginally insignificant (Table A2). However, the inclusion of a linear pre-trend difference attenuates both estimates.

Table 1: Adoption: DD Comparisons, Primary Care Use.

Metric	Nurse visits		GP visits	
	Bottom 40%	Top 40%	Bottom 40%	Top 40%
Level	1.373	0.603	1.450	0.705
Estimate	$-0.127$	$-0.048$	$-0.056$	$-0.033$
Std. error	0.032	0.018	0.021	0.010
P-value	0.000	0.008	0.007	0.001
Change (%)	$-9.252$	$-8.020$	$-3.879$	$-4.669$
Estimate (trends)	$-0.131$	$-0.057$	$-0.048$	$-0.026$
Change (%) (trends)	$-9.565$	$-9.490$	$-3.298$	$-3.706$
Events	17	17	17	17
Treated areas	152	152	152	152
All areas	245	245	245	245

*Notes:* The dataset is stacked and balanced. Our baseline model is Model 1, but we also use its extension (Model 1.2) allowing for differential linear pre-trend (“trends”). Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. Bottom 40% and top 40% refer to the distribution of equalized family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

We did not anticipate the lack of heterogeneity in relative terms. The earlier literature

has found, focusing on Swedish children and adolescents and GP visits, clear-cut heterogeneity by income in effects in both absolute and relative terms (Johansson et al., 2019; Nilsson & Paul, 2018). Our PAP did not explicitly discuss which of the two dimensions is more relevant. Accordingly, we highlight the difference in estimates. Our interpretation is that while in our analysis moderate copayments pose a greater barrier to accessing care for low-income individuals in terms of visits, the lack of clear-cut heterogeneity in relative terms nudges our posterior beliefs toward uncertainty about the strength of the income hypothesis. Recently, Johansson et al. (2023) found no evidence of heterogeneity by income at the 85th birthday in Sweden.

**Robustness checks.** The stacking estimates for nurse visits are not sensitive but attenuate for GP visits and social assistance outcomes when using Model 1.2, which allows for linear pre-trend differences (Table 1 and Table A2).<sup>20</sup> The stacking results on primary care use are robust to logarithmic outcomes and to unbalanced event-specific datasets that have more municipalities and observations than the balanced datasets in the main analysis (Table A3).

Next, we report the results constructed using the CS estimator (see Section 4). The findings on nurse visits are similar in both stacked and CS event-study plots. Regarding GP visits, the CS estimates show virtually no effect, while the stacked estimates suggest a small decrease (Figure A12 and Figure A13, post-blind). The CS event-study plots point to a decrease in the share of individuals receiving social assistance (Figure A14), but this finding is not supported by the stacked results.

Our static CS effect estimate averages over all group-time ATTs, weighted by group size. We estimate the effects on primary care use in eight cases, using either never-treated or not-yet-treated municipalities as the comparison group and exclude data from either 2013 or 2019 or both. The CS estimates differ from the stacked estimates in two ways (Figure A15 and Figure A16). First, the effects on nurse visits are larger. Annualized nurse use decreases by  $-0.18$  to  $-0.23$  visits ( $-13\%$  to  $-17\%$ ) for individuals at the bottom 40% and by  $-0.08$  to  $-0.10$  visits ( $-13\%$  to  $-16\%$ ) at the top 40%. Different estimands and accumulating effects plausibly explain

---

<sup>20</sup>Considering all our results, the number of general practitioner (GP) visits is reduced by 2–5%. The estimates accounting for a linear pre-trend difference are closer to zero in that interval.

this. The stacked results are based on a one-year follow-up, while the CS estimand averages over all group-time ATEs, having for most units a longer follow-up. Second, the estimates on GP visits are all negative but close to zero and insignificant. The results on social assistance outcomes are insignificant (Figure A17).

**Effect heterogeneity by income level.** We apply a triple difference (DDD) model (Model 2) using the stacked data to formally test treatment effect heterogeneity by income, comparing the evolution of outcomes at the bottom 40% of the income distribution to that at the top 40% both in the treatment and comparison areas. The results on primary care use are presented in Table 2. In absolute terms, nurse use decreases by  $-0.07$  to  $-0.08$  annualized visits ( $-5.3\%$  to  $-5.6\%$ ) for individuals at the bottom 40% relative to the top 40%, and the estimates are significant. GP visits appear to decrease more at the bottom 40%, but the estimates are insignificant. In relative terms (logs), all the estimates are insignificant.

Complementarily, we plot the effects on primary care use by income decile in Figure 4 (post-blind). We use stacking with balanced event-specific datasets and Model 1 to obtain the results. The pattern is clear in absolute terms: the estimate attenuates as income increases. The bottom 10% is an exception. A plausible partial explanation is that the share of social assistance recipients is decreasing in income. About 37% (80%) of social assistance recipients are at the bottom 10% (30%) of the income distribution (Figure A18, post-blind). Social assistance can cover copayments for these individuals, potentially attenuating the effects. Apart from the first decile, there may be heterogeneity also in relative terms: the lower end of the income distribution is somewhat more sensitive than the top.

**Estimates for all individuals.** We provide the results for the entire sample population in Online Appendix Section A.1 (post-blind). Annualized nurse use decreases by  $-0.09$  to  $-0.10$  visits ( $-9\%$  to  $-10\%$ ) using stacking (one-year follow-up) and by  $-0.13$  to  $-0.16$  visits ( $-13\%$  to  $-17\%$ ) using the CS estimator (a longer follow-up). Stacked estimates on GP visits show a decrease ( $-3\%$  to  $-5\%$ ), but the corresponding CS estimates are closer to zero and insignificant.

**Potential health effects.** We examine the heterogeneity of the effects with respect to

Table 2: Adoption: DDD Comparisons, Primary Care Use.

Metric	Balanced datasets		Unbalanced datasets	
	Nurse Visits	GP Visits	Nurse Visits	GP Visits
A. Annualized contacts per capita				
Level	1.373	1.450	1.386	1.444
Estimate	−0.073	−0.025	−0.078	−0.021
Std. error	0.017	0.013	0.017	0.013
P-value	0.000	0.058	0.000	0.115
Change (%)	−5.304	−1.718	−5.635	−1.462
Events	17	17	19	19
Treated areas	152	152	175	175
All areas	245	245	264	264
B. Logarithmized annualized contacts per capita				
Estimate	−1.544	0.711	−1.575	0.818
Std. error	1.559	0.903	1.523	0.925
P-value	0.322	0.431	0.301	0.377
Events	17	17	19	19
Treated areas	126	135	175	175
All areas	209	225	264	264

*Notes:* The dataset is stacked. We use Model 2. Estimates and standard errors are multiplied by 100 if the outcome is the logarithm of annualized contacts per capita. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. Estimates for logarithmized outcomes are multiplied by 100.



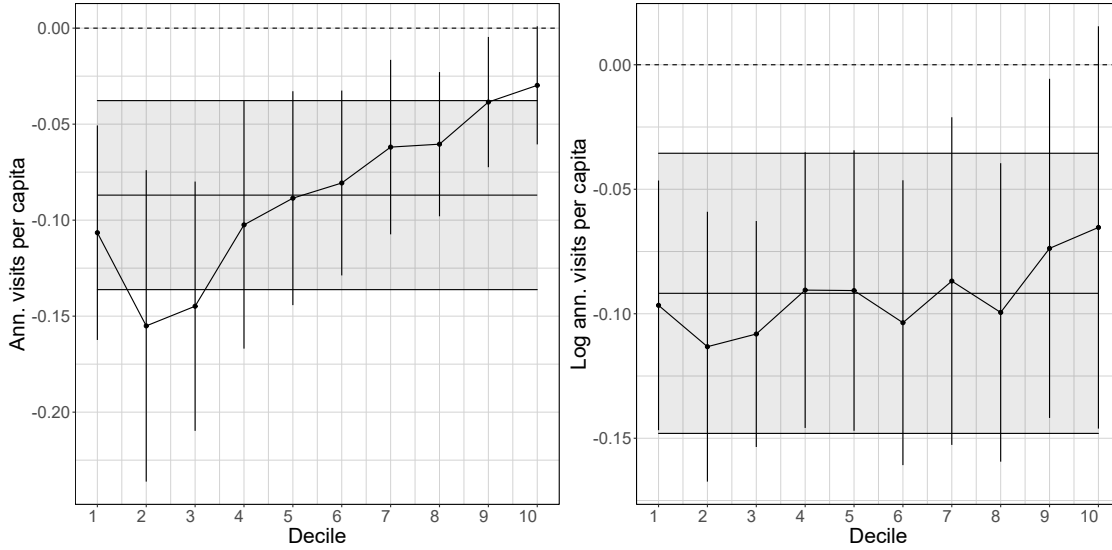


Figure 4: Adoption: Estimates on Nurse Visits by Income Decile.

*Notes:* This figure was not pre-registered and is post-blind. The dataset is stacked and balanced. We use Model 1. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. We use the distribution of equivalized family disposable income. Outcomes are the annualized number of curative nurse visits and its logarithm. The gray block, centered at the black horizontal line, shows the ATT estimate for the whole population and its confidence interval.

having received a drug prescription in 2018–2019 with an ATC (anatomical therapeutic chemical) code referring to diabetes or hypertension (A10, C02–C03, and C07–C09), proxying a diagnosis of these conditions. Those with a prescription for diabetes or hypertension responded more strongly in absolute terms to the copayment adoption, but we found no difference in relative terms (Table A4, post-blind). Online Appendix Section A.1 presents the results on emergency department (ED) visits and unplanned hospitalizations for ambulatory care sensitive conditions (ACSC), both outcomes being post-blind. The effects – more exploratory than confirmatory – are positive but imprecisely estimated and statistically insignificant for both outcomes. To summarize, these findings do not offer sufficient evidence either for the existence or nonexistence of health effects.

**Additional Supplementary Analyses.** Online Appendix Section A.1 adds and discusses more (post-blind) analyses. The effect estimates for nurse visits are larger when a longer follow-up is used, and the main findings are qualitatively robust to weighting municipalities uniformly. We also assume quasi-random treatment timing among the later-treated municipalities, excluding the

never-treated, and use the estimator proposed by Roth and Sant’Anna (2023). On average, the nurse visit copayment reduces nurse visits but has no effect on GP use, mostly in line with our main results. However, there is a caveat to our main heterogeneity findings: the differences in the effects in absolute terms are not that large or clear using the RS estimator. Thus, the relative effects appear even larger for individuals at the top 40%.

## 6 National Copayment Abolishment

Primary care utilization in 2020–2022 was affected by the COVID-19 pandemic. Curative nurse visits had not recovered to the pre-pandemic levels by May 2022 (Figure A19). This pattern is plausibly explained by the supply-side factors: nurses had been allocated to test, trace, and vaccinate. Consequently, primary care use was likely more supply-driven and gatekeeping stricter than during pre-pandemic times, attenuating the impact of the copayment. Second, the parallel trends assumption is less credible *ex ante* as i) the treatment group has an overrepresentation of rural areas, and ii) the epidemiological situation varied regionally throughout the pandemic, potentially correlated with the imbalanced baseline features. As shown below, the pre-treatment patterns in outcomes observed *ex post* do suggest that there is considerable uncertainty about the validity of the parallel trends assumption.

**Pre-trend plots.** Figure 5 plots the trends in curative nurse visits for the bottom 40% and the top 40% of the distribution of equivalized family disposable income in the treatment and comparison municipalities, relative to the copayment abolishment. There are large fluctuations in nurse use between the treated municipalities and their comparison group during the observation window. Nurse visits were increasing more in the treated areas for both income groups after the copayment abolishment, compared to the periods immediately prior to the reform, although there was not much difference (if any) between the income groups in absolute terms. However, we do not view these patterns as causal and emphasize caution in interpreting them. The difference in nurse visits fluctuated considerably before the copayment abolishment, which makes it plausible

that there could be fluctuations of similar magnitude also during the post-treatment periods even in the absence of the policy change. The graphs corresponding to Figure 5, but for curative GP visits and prescriptions, are in Figure A20 and Figure A21. We observe largely similar pre-treatment fluctuations in the difference of GP use between the policy groups as for nurse visits, but GP visits do not appear to increase after the nurse visit copayment abolishment.

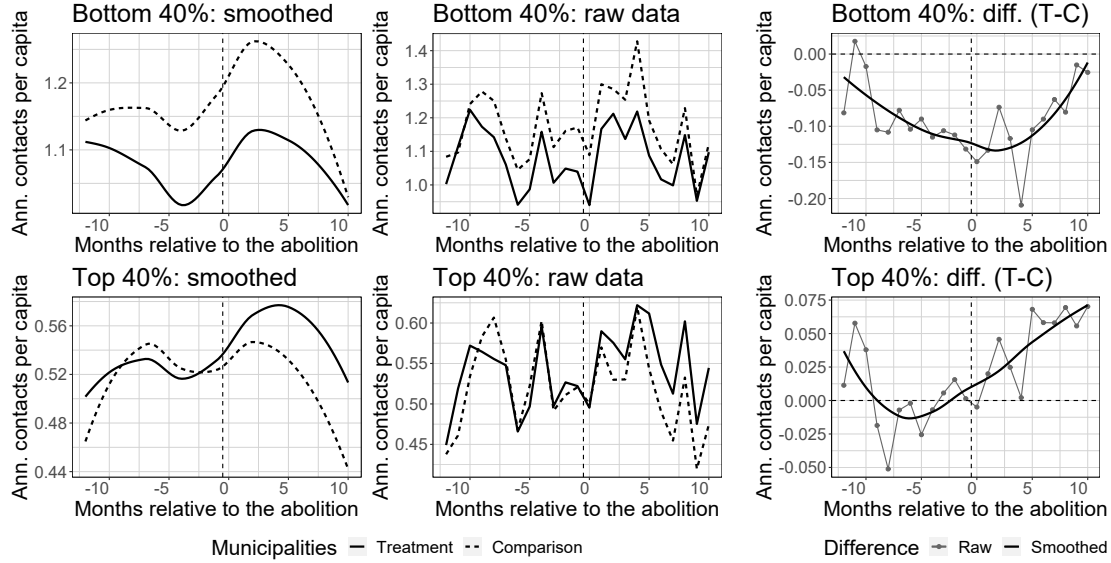


Figure 5: Abolishment: Evolution in Nurse Visits.

*Notes:* The outcome is the number of annualized curative nurse visits per capita. We weight by population due to heterogeneity in municipality size. The left column contains smoothed conditional means, fitted with local linear regression. The raw data are illustrated in the middle column, while the difference between treatment and comparison areas is depicted in the right column. Bottom 40% and top 40% refer to the distribution of equivalized family disposable income.

Figure 6 (post-blind) illustrates that even small violations of parallel trends, which we consider very likely, lead to major uncertainty about the effects of the abolishment. In essence, the abolishment analyses are imprecise and provide evidence neither for the existence nor the absence of effects. The figure reports confidence intervals (CIs) at the 5% level as proposed by Rambachan and Roth (2023), bounding the maximum post-treatment violation of parallel trends between consecutive periods by  $\bar{M}$  times the maximum pre-treatment violation of parallel trends. When naively assuming parallel trends, the estimates in the CIs are predominantly positive (they

do include zero) for both the bottom 40% and the top 40% of the income distribution. The CIs noticeably widen when allowing for small post-treatment violations of parallel trends. For instance, the CI for the bottom 40% includes estimates from  $-10\%$  to  $+20\%$  when post-treatment violations of parallel trends are at maximum (only) 0.15 times the maximum pre-treatment violation of parallel trends. In contrast to the PAP, we present the regression tables of the abolishment analyses in Online Appendix Section A.2 (Table A10 and Table A12 as key ones).

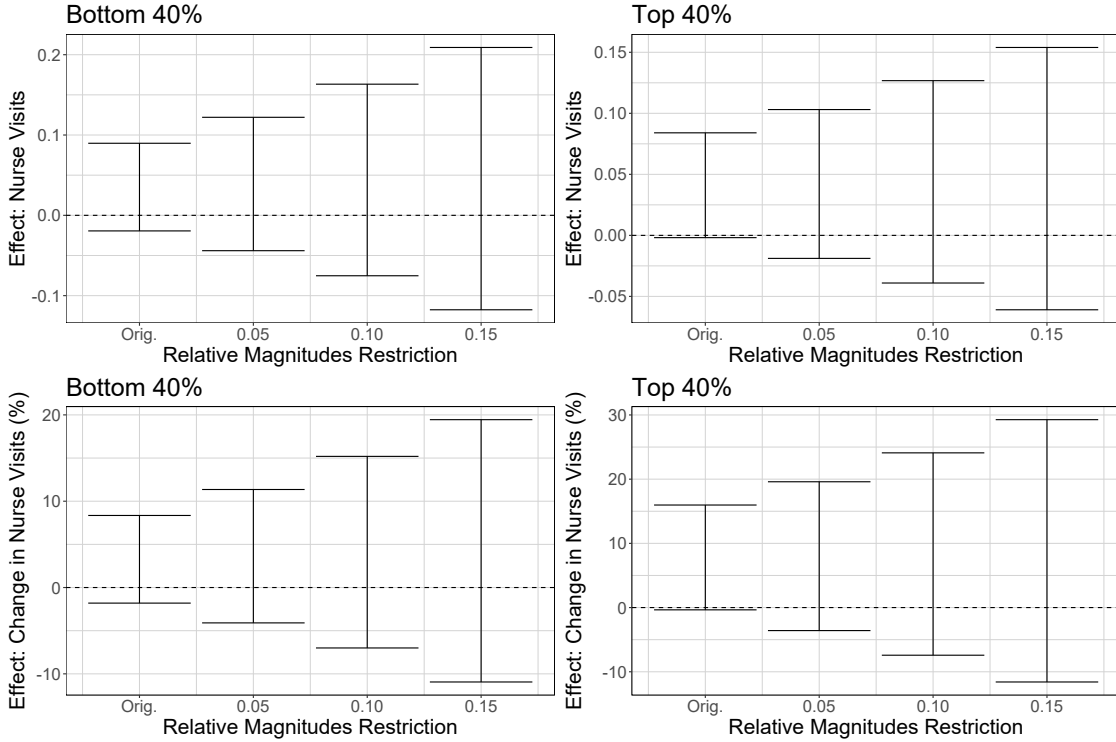


Figure 6: Abolishment and Nurse Visits: Bounding Violations of Parallel Trends.

*Notes:* This figure was not pre-registered and is post-blind. We apply the method proposed by Rambachan and Roth (2023) to construct confidence intervals at the 5% level by bounding violations of parallel trends. The effects represent the estimated change in the number of annualized nurse visits during an 11-month follow-up. The change in percentage terms compares the estimate to the pre-treatment mean in the treated municipalities. Bottom 40% and top 40% are based on the equivalized family disposable income distribution. First, we estimate a population-weighted event study specification that includes dynamic treatment indicators for the treated municipalities ( $D_{mt}^l$ ), normalized at time  $t = -1$ , and municipality ( $m$ ) and time ( $t$ ) fixed effects:  $y_{mt} = \alpha_m + \gamma_t + \sum_{l=-12, l \neq -1}^{10} \mu_l D_{mt}^l + \varepsilon_{mt}$ . Standard errors are clustered by municipality. Next, we use the “relative magnitudes” restriction,  $\Delta^{RM}(\bar{M})$ , and construct conditional-least favorable hybrid confidence sets (C-LF) for the average of the estimated post-treatment effects using the R package *HonestDiD*.  $\Delta^{RM}(\bar{M})$  bounds the maximum post-treatment violation of parallel trends between consecutive periods by  $\bar{M}$  times the maximum pre-treatment violation of parallel trends.

## 7 Conclusion

We analyze the effects of a staggered adoption of a nurse visit copayment (approximately 10 euros) on public primary care use of Finnish adults. Moreover, we provide an example of how a pre-analysis plan and analysis blinding can be used retrospectively in nonexperimental hypothesis testing. We find that the copayment adoption reduced curative nurse visits by 9–12% during a one-year follow-up. There is statistically significant heterogeneity by income in absolute terms: the decrease in the number of visits is more than two times larger for individuals at the bottom 40% of the income distribution than at the top 40%. However, such heterogeneity is much weaker and statistically insignificant in relative terms (percentage changes).<sup>21</sup> The estimates for GP visits are negative but close to zero and often insignificant (ranging from  $-2\%$  to  $-5\%$ ). Moreover, we analyze the later nationwide abolishment of the nurse visit copayment, but do not obtain causal conclusions from this reform because the parallel trend trends assumption required for these analyses, also pre-registered, was not credible and the corresponding uncertainty high.

The earlier studies on Swedish children and adolescents have found heterogeneity by income in both absolute and relative terms for GP visit copayments (Johansson et al., 2019; Nilsson & Paul, 2018). Based on their findings, we did not anticipate observing heterogeneity being so weak in relative terms in our data. While our analysis suggests that moderate copayments pose a greater barrier to accessing care for low-income individuals in terms of visits, the lack of clear-cut heterogeneity in relative terms nudges our posterior beliefs toward uncertainty about the strength of the income hypothesis. Recently, Johansson et al. (2023) found no evidence of heterogeneity by income at the 85th birthday in Sweden.

The relevant dimension of heterogeneity by income is context-specific, as argued earlier. The observed heterogeneity in absolute terms is important for many policy debates. Assuming that each visit corresponds to a genuine health concern, the decrease in visits indicates that these

---

<sup>21</sup>However, there is a caveat to these heterogeneity results. In a robustness check using the Roth and Sant’Anna (2023) estimator and later-treated areas as controls, added after registering our pre-analysis plan, the differences in effects in absolute terms are not clear-cut. Thus, the relative effects appear even larger for the top 40%. However, the results from other robustness checks are in line with the finding of treatment effect heterogeneity by income in absolute terms.

concerns might not be addressed. This situation could result in reduced treatment for chronic health problems and the emergence of untreated, potentially exacerbating new health concerns. Disproportionately, these consequences would affect low-income individuals, who already have poorer baseline health and, potentially, a higher marginal benefit of care. Consequently, ensuring that low-income individuals have good access to primary care might be considered essential, particularly for politicians who prioritize the well-being of the most disadvantaged and view health issues as a partial indicator of this disadvantage.

To compare the effect sizes, we convert our estimates for all individuals to the semi-arc elasticity of the number of nurse visits with respect to price, as in Brot-Goldberg et al. (2017). These elasticities represent changes in quantities, normalized by the baseline, divided by the price change:  $\frac{(q_1 - q_0)/(q_1 + q_0)}{(p_1 - p_0)/2}$ . Here,  $q$  denotes the number of nurse visits, and  $p$  is their “price”. Following Nilsson and Paul (2018), we define the price as the share of out-of-pocket costs of the total cost of the visit. The elasticity is sensitive to the selected parameters, so we provide two estimates: the baseline and a large estimate. The parameter values are listed in Online Appendix Section A.5. The baseline for nurse visits is  $-0.41$  and the large estimate  $-1.24$ .<sup>22</sup> For doctor visits, Nilsson and Paul (2018) report a semi-arc elasticity of  $-0.88$  at the 20th birthday and  $-0.55$  at the 7th birthday when individuals face copayments of 10–15 euros for outpatient doctor visits in Sweden. Also in Sweden, the effect of a 10-euro copayment for GP visits at the 20th birthday maps to an elasticity of  $-1.11$  (Johansson et al., 2019).<sup>23</sup> All these Nordic estimates are lower than the elasticities of  $-2.11$  and  $-2.26$  that Brot-Goldberg et al. (2017) calculate based on the RAND Health Insurance Experiment.

Regarding the fiscal impacts, our back-of-the-envelope calculations compare the estimated copayment revenue to the estimated savings from reduced primary care utilization. Assuming 4 million sample individuals, a 10-euro copayment, that only 70% of the copayment

---

<sup>22</sup>The corresponding elasticity of the nurse visit copayment on GP visits is close to zero as we estimate only small reductions in GP use.

<sup>23</sup>We use the estimates for all individuals from their Table 1 and use a copayment of SEK 100 and the total cost of SEK 1500 per visit.

revenue are collected,<sup>24</sup> a pre-treatment mean of 1.000 annualized nurse visits, an effect of  $-0.089$  annualized visits,<sup>25</sup> and an average cost of 35 euros per nurse visit (Mäklin & Kokko, 2020), annual copayment revenue would be 26 million compared to saving 12 million due to lower nurse use. If the utilization effect is larger, say  $-15\%$  (Figure A25), the copayment revenue would be 24 million and the savings 21 million. If both the collection costs and production costs were higher than the values used above, then the fiscal importance of the short-term savings from reduced service use is prominent relative to the collected revenue. However, this conclusion does not account for the potentially important health effects.

Gaining additional (confirmatory) insights into the health and broader welfare implications of cost-sharing in further studies is crucial, especially concerning low-income households. Neither does our analysis account for the fact that the copayment was charged only for the first three visits annually in many municipalities. A similar cap also frequently exists for GP visits. This feature creates a promising opportunity to study how dynamic incentives (i.e., the spot price is reduced to zero after the third visit) affect primary care utilization. The dynamic incentives in cost-sharing have received growing attention in the recent literature (see, e.g., Aron-Dine et al., 2015; Cabral, 2017; Farbmacher & Winter, 2013; Simonsen et al., 2021). Based on this research, our hypothesis is that individuals exhibit myopic behavior and react to spot prices. For now, we have not examined the relationship of the nurse or GP visit copayments and the distribution of nurse or GP visits over a calendar year or the frequency of visits relative to the third visits of the year (intentional choice). Consequently, we can still blind the causal relationship of interest if we design a PAP for the corresponding confirmatory tests.

---

<sup>24</sup>This share is an *ad hoc* choice to account for collection costs, increased social assistance use, and the fact that the copayment was often charged for the first three visits annually.

<sup>25</sup>These are averages based on two different specifications of Table A5.

### **CRedit authorship contribution statement**

**Tapio Haaga:** Conceptualization, Formal analysis, Writing–Original Draft, Writing–Review & Editing. **Petri Böckerman:** Conceptualization, Writing–Review & Editing, Supervision. **Mika Kortelainen:** Conceptualization, Writing–Review & Editing, Supervision. **Janne Tukiainen:** Conceptualization, Writing–Review & Editing, Supervision.

### **Funding**

This work is supported by the Finnish Institute for Health and Welfare, the Finnish Ministry of Social Affairs and Health, and Yrjö Jahnsson Foundation (research grant No. 20197209).

### **Declaration of Competing Interest**

None.

### **Supplementary Materials**

Pre-analysis plan and earlier working paper versions: <https://osf.io/skuv9/>. Replication codes: <https://github.com/tapiohaa/ASMA3>. The (online) appendix is attached to the submission.



## References

- Altındağ, O., & O'Connell, S. D. (2023). The short-lived effects of unconditional cash transfers to refugees. *Journal of Development Economics*, 160, 102942. <https://doi.org/10.1016/j.jdeveco.2022.102942>
- Aron-Dine, A., Einav, L., Finkelstein, A., & Cullen, M. (2015). Moral hazard in health insurance: Do dynamic incentives matter? *The Review of Economics and Statistics*, 97, 725–741. [https://doi.org/10.1162/REST\\_a\\_00518](https://doi.org/10.1162/REST_a_00518)
- Baker, A. C., Larcker, D. F., & Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144, 370–395. <https://doi.org/10.1016/j.jfineco.2022.01.004>
- Banerjee, A., Duflo, E., Finkelstein, A., Katz, L., Olken, B., & Sautmann, A. (2020). In praise of moderation: Suggestions for the scope and use of pre-analysis plans for RCTs in economics. NBER Working Paper No. 26993. <https://doi.org/10.3386/w26993>
- Bohm, P., & Lind, H. (1993). Policy evaluation quality: A quasi-experimental study of regional employment subsidies in Sweden. *Regional Science and Urban Economics*, 23, 51–65. [https://doi.org/10.1016/0166-0462\(93\)90028-D](https://doi.org/10.1016/0166-0462(93)90028-D)
- Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., & Kolstad, J. T. (2017). 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. <https://doi.org/10.1093/qje/qjx013>
- Burlig, F. (2018). Improving transparency in observational social science research: A pre-analysis plan approach. *Economics Letters*, 168, 56–60. <https://doi.org/10.1016/j.econlet.2018.03.036>
- Cabral, M. (2017). Claim timing and ex post adverse selection. *The Review of Economic Studies*, 84, 1–44. <https://doi.org/10.1093/restud/rdw022>
- Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225, 200–230. <https://doi.org/10.1016/j.jeconom.2020.12.001>

- Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134, 1405–1454. <https://doi.org/10.1093/qje/qjz014>
- Cesarini, D., Lindqvist, E., Östling, R., & Wallace, B. (2016). Wealth, health, and child development: Evidence from administrative data on Swedish lottery players. *The Quarterly Journal of Economics*, 131, 687–738. <https://doi.org/10.1093/qje/qjw001>
- Chandra, A., Flack, E., & Obermeyer, Z. (2023). The health costs of cost-sharing. NBER Working Paper No. 28439. <https://doi.org/10.3386/w28439>
- Clemens, J., McNichols, D., & Sabia, J. (2020). The long-run effects of the Affordable Care Act: A pre-committed research design over the COVID-19 recession and recovery. NBER Working Paper No. 27999. <https://doi.org/10.3386/w27999>
- Clemens, J., & Strain, M. (2021). 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. <https://doi.org/10.3386/w29264>
- de Chaisemartin, C., & D’Haultfœuille, X. (2023). Two-way fixed effects and differences-in-differences estimators with several treatments. *Journal of Econometrics*, 236, 105480. <https://doi.org/10.1016/j.jeconom.2023.105480>
- Farbmacher, H., Ihle, P., Schubert, I., Winter, J., & Wuppermann, A. (2017). Heterogeneous effects of a nonlinear price schedule for outpatient care. *Health Economics*, 26, 1234–1248. <https://doi.org/10.1002/hec.3395>
- Farbmacher, H., & Winter, J. (2013). Per-period co-payments and the demand for health care: Evidence from survey and claims data. *Health Economics*, 22, 1111–1123. <https://doi.org/10.1002/hec.2955>
- Gardner, J. (2021). Two-stage differences in differences.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*. <https://doi.org/10.1016/j.jeconom.2021.03.014>

- Gormley, T. A., & Matsa, D. A. (2011). Growing out of trouble? Corporate responses to liability risk. *The Review of Financial Studies*, 24, 2781–2821. <https://doi.org/10.1093/rfs/hhr011>
- Haaga, T., Böckerman, P., Kortelainen, M., & Tukiainen, J. (2022). Cost sharing and primary care use: Evidence from staggered copayment adoption and later abolition. A pre-analysis plan. <https://doi.org/10.17605/OSF.IO/FV2GA>
- Haaga, T., Böckerman, P., Kortelainen, M., & Tukiainen, J. (2023a). Do adolescents from low-income families respond more to cost-sharing in primary care? Version 2. <https://osf.io/vmuzf/>
- Haaga, T., Böckerman, P., Kortelainen, M., & Tukiainen, J. (2023b). Does abolishing a copayment increase doctor visits? A comparative case study. Version 2. <https://osf.io/v7b5s>
- Han, H.-W., Lien, H.-M., & Yang, T.-T. (2020). Patient cost-sharing and healthcare utilization in early childhood: Evidence from a regression discontinuity design. *American Economic Journal: Economic Policy*, 12, 238–278. <https://doi.org/10.1257/pol.20170009>
- Iizuka, T., & Shigeoka, H. (2022). Is zero a special price? Evidence from child health care. *American Economic Journal: Applied Economics*, 14, 381–410. <https://doi.org/10.1257/app.20210184>
- Johansson, N., de New, S. C., Kunz, J. S., Petrie, D., & Svensson, M. (2023). Reductions in out-of-pocket prices and forward-looking moral hazard in health care demand. *Journal of Health Economics*, 87, 102710. <https://doi.org/10.1016/j.jhealeco.2022.102710>
- Johansson, N., Jakobsson, N., & Svensson, M. (2019). Effects of primary care cost-sharing among young adults: Varying impact across income groups and gender. *The European Journal of Health Economics*, 20, 1271–1280. <https://doi.org/10.1007/s10198-019-01095-6>
- Kasy, M. (2021). Of forking paths and tied hands: Selective publication of findings, and what economists should do about it. *Journal of Economic Perspectives*, 35, 175–192. <https://doi.org/10.1257/jep.35.3.175>

- Kruse, M., Olsen, K. R., & Skovsgaard, C. V. (2022). Co-payment and adolescents' use of psychologist treatment: Spill over effects on mental health care and on suicide attempts. *Health Economics*, 31, 92–114. <https://doi.org/10.1002/hec.4582>
- Kunz, J. S., & Winkelmann, R. (2017). An econometric model of healthcare demand with nonlinear pricing. *Health Economics*, 26, 691–702. <https://doi.org/10.1002/hec.3343>
- Landsem, M. M., & Magnussen, J. (2018). The effect of copayments on the utilization of the GP service in Norway. *Social Science & Medicine*, 205, 99–106. <https://doi.org/10.1016/j.socscimed.2018.03.034>
- Lindqvist, E., Östling, R., & Cesarini, D. (2020). Long-run effects of lottery wealth on psychological well-being. *The Review of Economic Studies*, 87, 2703–2726. <https://doi.org/10.1093/restud/rdaa006>
- Ma, Y., & Nolan, A. (2017). Public healthcare entitlements and healthcare utilisation among the older population in Ireland. *Health Economics*, 26, 1412–1428. <https://doi.org/10.1002/hec.3429>
- Maier, C. B., & Aiken, L. H. (2016). Task shifting from physicians to nurses in primary care in 39 countries: A cross-country comparative study. *European Journal of Public Health*, 26, 927–934. <https://doi.org/10.1093/eurpub/ckw098>
- Mäklin, S., & Kokko, P. (2020). Terveystien- ja sosiaalihuollon yksikkökustannukset Suomessa vuonna 2017. <https://urn.fi/URN:ISBN:978-952-343-493-6>
- Miguel, E. (2021). Evidence on research transparency in economics. *Journal of Economic Perspectives*, 35, 193–214. <https://doi.org/10.1257/jep.35.3.193>
- Neumark, D. (2001). The employment effects of minimum wages: Evidence from a prespecified research design the employment effects of minimum wages. *Industrial Relations: A Journal of Economy and Society*, 40, 121–144. <https://doi.org/10.1111/0019-8676.00199>
- Neumark, D., & Yen, M. (2022). Effects of recent minimum wage policies in California and nationwide: Results from a pre-specified analysis plan. *Industrial Relations: A Journal of Economy and Society*, 61, 228–255. <https://doi.org/10.1111/irel.12297>

- Nilsson, A., & Paul, A. (2018). Patient cost-sharing, socioeconomic status, and children's health care utilization. *Journal of Health Economics*, 59, 109–124. <https://doi.org/10.1016/j.jhealeco.2018.03.006>
- Nolan, A., & Layte, R. (2017). The impact of transitions in insurance coverage on GP visiting among children in Ireland. *Social Science & Medicine*, 180, 94–100. <https://doi.org/10.1016/j.socscimed.2017.03.026>
- Nosek, B. A., Ebersole, C. R., DeHaven, A. C., & Mellor, D. T. (2018). The preregistration revolution. *Proceedings of the National Academy of Sciences*, 115, 2600–2606. <https://doi.org/10.1073/pnas.1708274114>
- OECD. (2021). OECD Health Statistics 2021, accessed 26/06/2023. <https://stat.link/m5nfxa>
- Ofori, G. K., & Posner, D. N. (2023). Pre-analysis plans: An early stocktaking (2021/03/31). *Perspectives on Politics*, 21, 174–190. <https://doi.org/10.1017/S1537592721000931>
- Olden, A., & Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*, 25, 531–553. <https://doi.org/10.1093/ectj/utac010>
- Olken, B. A. (2015). Promises and perils of pre-analysis plans. *Journal of Economic Perspectives*, 29, 61–80. <https://doi.org/10.1257/jep.29.3.61>
- Olsen, C. B., & Melberg, H. O. (2018). Did adolescents in Norway respond to the elimination of copayments for general practitioner services? *Health Economics*, 27, 1120–1130. <https://doi.org/10.1002/hec.3660>
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *The Review of Economic Studies*. <https://doi.org/10.1093/restud/rdad018>
- Roth, J., & Sant'Anna, P. H. C. (2023). Efficient estimation for staggered rollout designs. *Journal of Political Economy Microeconomics*. <https://doi.org/10.1086/726581>
- Roth, J., & Sant'Anna, P. H. (2023). When is parallel trends sensitive to functional form? *Econometrica*, 91, 737–747. <https://doi.org/10.3982/ECTA19402>

- Simonsen, M., Skipper, L., Skipper, N., & Christensen, A. I. (2021). Spot price biases in non-linear health insurance contracts. *Journal of Public Economics*, 203, 104508. <https://doi.org/10.1016/j.jpubeco.2021.104508>
- Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225, 175–199. <https://doi.org/10.1016/j.jeconom.2020.09.006>
- Winkelmann, R. (2004). Co-payments for prescription drugs and the demand for doctor visits – evidence from a natural experiment. *Health Economics*, 13, 1081–1089. <https://doi.org/10.1002/hec.868>