

Effects of Nurse Visit Copayments: Does the Primary Care Use of the Poor Respond More?

Tapio Haaga, Petri Böckerman, Mika Kortelainen, and Janne Tukiainen*

August 2022

Abstract

We analyze the effects of a staggered adoption and a later nationwide abolition of a nurse visit copayment of approximately 10 euros on curative primary care use in Finnish adult population. We use difference-in-difference and event study methods and focus on the potential heterogeneity of the effects by income level. We find that the adoption reduces nurse visits by 9% to 12% during a one-year follow-up. There is notable heterogeneity in absolute terms: the number of visits decreases more at the lower end of the income distribution. However, such heterogeneity is less clear in relative terms. The estimates on the number of general practitioner (GP) visits are negative but close to zero and often insignificant (-2% to -5%). Regarding the nationwide abolition, we do not obtain clear causal conclusions since imposing a parallel trends assumption does not seem to be reasonable based on pre-trend patterns.

Keywords: Cost sharing, copayments, out-of-pocket costs, healthcare use, primary care, general practitioner, difference-in-differences

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

***Haaga:** University of Turku, and Finnish Institute for Health and Welfare (THL) (email: tapio.haaga@utu.fi). **Böckerman:** University of Jyväskylä, Labour Institute for Economic Research LABORE, and IZA Institute of Labor Economics (email: petri.boeckerman@labore.fi). **Kortelainen:** University of Turku, InFLAMES Research Flagship Center, VATT Institute for Economic Research, and Helsinki Graduate School of Economics (email: mika.kortelainen@utu.fi). **Tukiainen:** University of Turku, and VATT Institute for Economic Research (email: janne.tukiainen@utu.fi). **Acknowledgements:** We thank Mikko Peltola for support, and Lauri Sääksvuori and Jussi Tervola for comments and suggestions. We also thank all seminar participants who have provided comments to this study and our other related projects. This work is supported by THL, Finnish Ministry of Social Affairs and Health, and Yrjö Jahnsson Foundation (research grant No. 20197209). **Replication codes:** <https://github.com/tapiohaa/ASMA3>. **Pre-analysis plan:** <https://osf.io/skuv9/>.

Contents

1	Introduction	1
2	Institutional Background	5
3	Data	9
4	Results: Staggered Adoption	12
5	Results: Later Simultaneous Abolition	22
6	Discussion	27
A	Online Appendix	A1

1 Introduction

Primary care aims to provide comprehensive services from health promotion and disease prevention to treatment and rehabilitation as early and as close to the patient as possible. Well-functioning primary care improves health system efficiency, equitable access to healthcare, and public health (WHO, 2018). Aging populations, retiring healthcare professionals, and smaller working age cohorts in Europe and elsewhere make it difficult to meet the increased demand for primary care services. For example, the United States is projected to face a shortage of 18,000 to 48,000 primary care physicians by 2034 (AAMC, 2021). To increase supply and lower the costs, nurses are increasingly substituting doctors. Nurse practitioners (NP) have full practice authority in 31 states in the United States by 2022, allowing them to examine, diagnose, treat and prescribe medications to the full extent of their training and experience without the supervision of physicians (McMichael and Markowitz, 2022). Randomized trials have found that the effects of this substitution are mainly positive on patient outcomes and utilization (Laurant et al., 2018).

In addition to supply-side reforms, the growing demand for primary care can potentially be tackled by patient cost sharing, rationing by triage, and waiting times. A large literature has shown that out-of-pocket costs reduce the demand for health care (Einav and Finkelstein, 2018). Ideally, cost sharing can reduce healthcare spending if rational and well-informed patients do not seek care that they regard as low-value care. The situation is much more complicated in reality. Chandra et al. (2021) find that as-if random increases in out-of-pocket price reduce the consumption of drugs, including medically valuable statins and antihypertensives, and increase mortality among affected patients. Goldin et al. (2020) report that randomly assigned reminder letters about a tax penalty for not having health insurance (no changes in financial incentives *per se*) increase insurance coverage and reduce mortality.

Primary care is the first point of contact in the healthcare system. However, any needs-based prioritization and gatekeeping by healthcare professionals is conditional

on patients having contacted the system in the first place. Potentially, out-of-pocket costs at the entry are a barrier that affects population groups differently. They may contribute to inequality through some population groups having *de facto* a more limited access to primary care than others. Low-income individuals are arguably more responsive to out-of-pocket costs than the population on average. At the same time, the poor have worse health status. Consequently, the potential downside of increased cost sharing is that it can reduce the use of primary care among those people for whom the marginal benefit of health care services is greatest. Cost sharing in the form of small copayments can increase price transparency and reduce the risk of financial problems for the patients, but heterogenous utilization effects may still exist.

We analyze whether a small copayment of approximately 10 euros for curative primary care nurse visits affects primary care use in the Finnish adult population. The focus is on the potentially heterogeneous effects by income. To this end, we use the staggered adoption of the copayment for curative primary care nurse visits in municipalities in 2014-2019 and its nationwide abolition in July 2021 and difference-in-differences (DD) and event study methods, as well as comprehensive administrative data. Most primary care areas (or municipalities) adopted the copayment at some point between 2014 and 2019 to collect more revenue. In July 2021, the law on out-of-pocket costs was reformed, and nurse visits were set to be provided free of charge, abolishing the copayment nationally.

In Finnish primary care, nurses carry out triage and book appointments to nurses and general practitioners (GP). Unlike some states in the US with full practice authority, nurses have less independence in Finland and work closely with GPs even if they provide independent appointments. They consult GPs or book GP appointments for their patients if needed. GPs write the vast majority of prescriptions and authorize access to specialized healthcare. Nurses treat acute cases and patients with chronic conditions, such as diabetes, asthma or dementia.

We find that the copayment adoption reduces curative nurse visits by 9% to 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 at the bottom 40% of the income distribution than at the top 40%. The effect size is consistently larger the lower the income decile is. However, such heterogeneity is less clear in relative terms (percentage changes). We also examine whether copayments for curative nurse visits are causally linked to the changes in the number of GP visits to which access is rationed using gatekeeping and which proxy a professional-assessed need for diagnosis and treatment. We estimate a reduction of 2% to 5% in GP visits. Our preferred estimates are at the lower end of that interval and often insignificant.¹ Regarding the copayment abolition, we do not obtain any causal conclusions, because the parallel trends assumption is not plausible given the observed pre-trend patterns.

There is a large literature that studies the effects of task shifting from primary care GPs to nurses on healthcare delivery outcomes (Laurant et al., 2018; McMichael and Markowitz, 2022; Yang et al., 2021). However, we are not aware of previous studies that examine the effects of nurse visit copayments on primary care use. Most studies in the cost sharing literature examine completely different types of exposures, such as access to health insurance (Card et al., 2008; Kondo and Shigeoka, 2013) or changes in coinsurance rates (Shigeoka, 2014; Fukushima et al., 2016), and do not focus on primary care nurse visits. Previous studies based on moderate copayments in primary care have focused on copayments for doctor visits (Nilsson and Paul, 2018; Johansson et al., 2019; Ma and Nolan, 2017). We also contribute to the literature by examining the spillovers of primary care nurse visit copayments on GP use.

Results from our pre-registered heterogeneity analyses by income are our second contribution. The potential heterogeneity by income level is of interest for three reasons. The copayment is small, but it constitutes a larger fraction of the disposable income for low-income individuals. Primary care supply is effectively fixed in the short term, and

¹For GP visit estimates, we prefer specifications that allow for a linear pre-trend difference or that require parallel trends only from the last pre-treatment period on. See Section 4 for details.

we expect aggregate effects to be moderate. Observing a reduction in aggregate primary care utilization *per se* reveals only little about whether valuable care is missed. However, if the effects are heterogeneous by income, the copayment affects access to services in an unequal way. We estimate all DD analyses - both main results and complementary analyses - separately at the bottom 40% and the top 40% of the distribution of equivalised family disposable income and also use the triple difference estimator to conduct tests on this heterogeneity. Stratifying by income has not been common in the recent literature, potentially due to the lack of data on personal or family income. Of the studies that do so, Nilsson and Paul (2018) and Johansson et al. (2019) examine the effects of GP visit copayments for children and adolescents. Han et al. (2020) have several stratifying dimensions and study the effects on regular outpatient care at clinics and hospitals for small children.

Third, we estimate the effects of copayments on service use for the total adult population and use a staggered DD design with irreversible treatment using methods that are robust to the combination of staggered timing and heterogeneity of treatment effects. Age-based RD designs are commonly used in the cost sharing literature: of the abovementioned studies, only Kondo and Shigeoka (2013) and Ma and Nolan (2017) use some other research design. RD designs have their own strengths, but the continuity-based RD estimand is only informative at a specific cutoff where the policy changes, not for the total population. The RD estimates in Nilsson and Paul (2018) and Johansson et al. (2019) using Swedish data (institutionally the closest country to Finland) are estimated for children and adolescents at age cutoffs 7 and 20. Two other studies use a single-event DD design (Chandra et al., 2010, 2014). Previous studies based on staggered DD designs (e.g., Iizuka and Shigeoka, 2021) use conventional two-way fixed effects (TWFE) estimators while recent advances in DD and event-study methodologies (Goodman-Bacon, 2021; Sun and Abraham, 2021; Baker et al., 2022) show that those estimators may be significantly biased in the presence of treatment effect heterogeneity.

We use a registered pre-analysis plan (PAP; available at <https://osf.io/skuv9/>), accompanied by computer codes for replication, that specify in detail how we planned to clean the data, construct our analysis data, perform analyses, and report results *before* estimating any of the results. Only placebo results were estimated before the registration by using either placebos-in-time or randomized placebo policy assignment. All changes to the PAP are documented in Section A.2.

PAPs can improve research transparency and credibility in two important ways. First, the writing of statistical programs is separated from the effect estimation in order to avoid receiving feedback from the actual outcome data. Second, a clear distinction between confirmatory and exploratory analyses can be made. Thus, PAPs can reduce the possibility of data mining (Olken, 2015). Banerjee et al. (2020) view that analyses following a pre-specified PAP should be treated as exhibiting a lower risk of a false positive at a given confidence level than studies without a PAP. Brodeur et al. (2022) find that pre-registered RCTs with a PAP are less p-hacked than pre-registered RCTs without one.² Although p-hacking is more general in non-experimental causal inference in economics than in RCTs (Vivalt, 2019; Brodeur et al., 2020), there are only few non-experimental studies using a PAP (Christensen and Miguel, 2018). Following the examples of Neumark (2001), Neumark and Yen (2021), and Clemens and Strain (2021), our study offers one application of how PAPs can be used to improve transparency of non-experimental work.

2 Institutional Background

Primary care services are provided for the Finnish adult population by three sectors: publicly-funded primary care, occupational healthcare, and private clinics. These sectors cover different population groups and differ with respect to gatekeeping, out-of-pocket costs, and waiting times. Publicly-funded primary care is the main provider of primary care

²They define PAP as some form of a write-up document. In contrast, our PAP is much more detailed, with statistical programs and a placebo report included.

services for pensioners, the unemployed, and low-income individuals. Nurses do triage on the phone or at health stations and book appointments to both nurses and GPs. A referral is needed to consult a specialist. Copayments are moderate: at maximum 21 euros per GP visit and approximately 10 euros per curative nurse visit (if the area charges it). Waiting times vary and may be long for non-urgent care.

Employed workers are entitled under Finnish law to preventive occupational healthcare services, but many employers also provide and pay for additional curative services. Usually, a professional must be contacted before being allowed to book an appointment. These services are free of charge at the point of use, and waiting times are typically short. Private clinics do not use gatekeeping: one can book an appointment directly to a specialist. The state offers a small reimbursement for these services, but the out-of-pocket costs are still many times higher than in publicly-funded primary care. Private health insurance, which is common especially in families with children, can be purchased to cover the costs. Waiting times are short in these services.

Publicly-funded primary care in Finland is organized by municipalities that form primary care areas (officially, health centers) on their own or in cooperation with other municipalities. Every citizen has their designated health station, determined by the location of residence. In some areas, all health stations may be available on a visit-by-visit basis. Since 2014, all people have been able to choose their health station once a year, but active choices have been relatively rare. Municipal services are financed through transfers from the state, municipal taxes, copayments, and municipal bonds. The state guides municipal policies on copayments by setting which groups or services are exempted from copayments (in Act on Social and Health Care Client Fees) and maximum copayment levels (in the corresponding Government Decree). Within these constraints, primary care areas set their own copayment policies.

In July 2010, Finland adopted restricted prescription rights for nurses. Related to that, the law on copayments was changed to allow primary care areas to charge a

copayment for curative nurse visits. As a result, the law no longer specifies the professions (e.g., physicians) whose appointments can be subjected to copayments. However, the decree continued to explicitly mention only doctor appointments, potentially explaining why no areas immediately adopted the nurse visit copayment. Based on our data collection, the copayment was first introduced in January 2014. Many other areas also adopted it once they became aware of the possibility to collect more revenue: the staggered adoption is illustrated in terms of the number of municipalities and population in the top row in Figure 1. At the end of 2019, half of the population lived in a municipality that charged a copayment for curative nurse visits, and the vast majority of municipalities had adopted the copayment. Figure A1 shows the staggered adoption graphically, plotting the policies by municipality at the end of each year. Based on Figure A2, 80% of the municipalities with the copayment charged it only for the first three visits annually in Summer 2021, and by far the most common level for the per-visit copayment was approximately 11 euros, the population-weighted mean being 12 euros.

In July 2021, the government conducted a broader reform to the act to reduce barriers to access and health inequality. The key change was to exempt nurse visits from copayments. Consequently, more than 200 municipalities and almost three million people were affected by the policy change (see the bottom row of Figure 1). Transfers to municipalities were increased to compensate for reduced copayment revenue. Figure A3 shows the pre-abolition policies by municipality, also presenting the population size by using bubble size. The treatment group consists of many municipalities that are on average small and rural, while the comparison group contains the largest six cities by population size.

The state increased the maximum GP visit copayment from 14.70 euros in 2014 to 16.10 euros in 2015 and to 20.90 euros in 2016. All municipalities except Helsinki with zero copayments made the first increase in 1/2015. Municipalities responded differently to the latter increase: many made it instantly in 1/2016, some made it later, and some areas have not made the increase by 2022. In Figure A4, we examine whether the timing of these

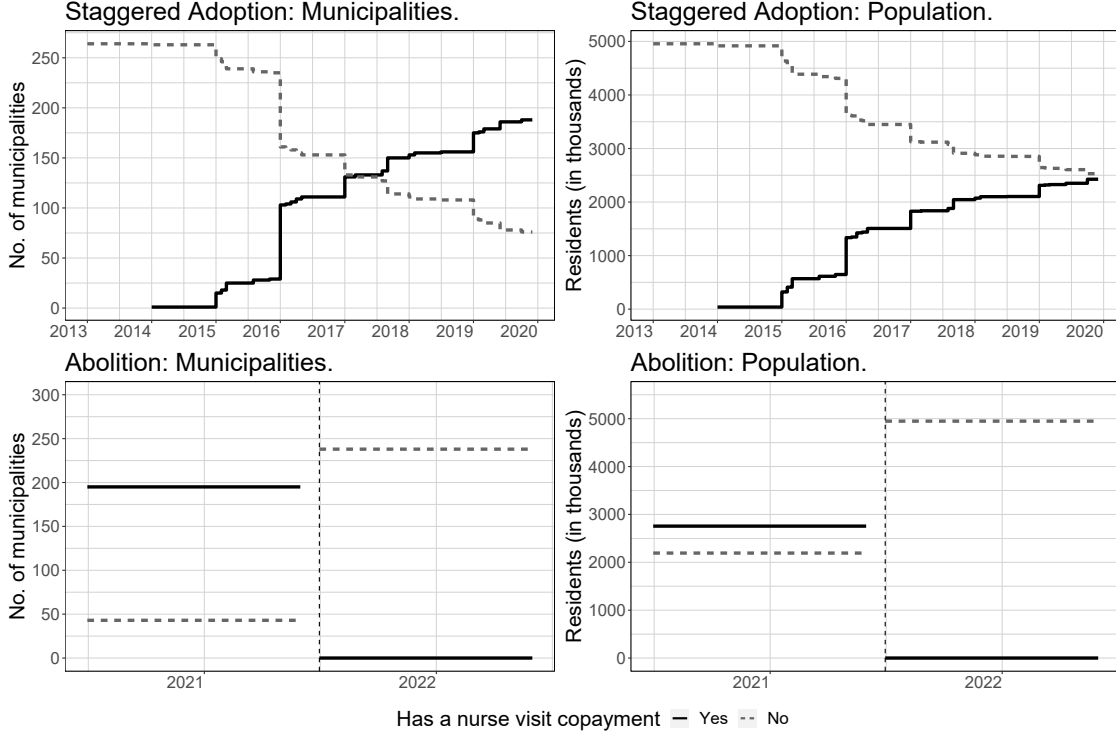


Figure 1: Staggered Adoption and Simultaneous Abolition of the Nurse Visit Copayment.

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 on the staggered adoption contains those municipalities whose policies on copayments for curative nurse visits we observe in our data collection. Regarding the abolition, we also require the adoption to have occurred at least 12 months before the start of the study window (7/2020). We assume that the effects have fully accumulated within 12 months after the adoption.

increases is correlated with the adoption of copayment for nurse visits. There is a 1 euro increase in GP visit copayments in the treated municipalities relative to the comparison municipalities at the time of the nurse visit copayment adoption using stacked data and balanced event-specific datasets (see sections 3 and 4 for methodology and how we construct the data). Our exposure is thus the sum of the nurse visit copayment adoption (about ten euros per visit) and a small increase in GP visit copayments (about one euro per visit). In contrast, there were no major changes to GP visit copayments at the time of the nurse visit copayment abolition.

The public primary care system is characterized by limited supply and excess demand for labor. Cohorts in medical schools are fixed in size, and the public sector

and the private sector compete for workforce. Primary care areas face challenges in hiring nurses at the prevailing wage level determined by collective bargaining between public sector employers and the nurses' union. Resource decisions are done in a context where central and local governments have been running fiscal deficits since the financial crisis of 2008, which is expected to continue for years due to the aging of the population. Consequently, our hypothesis is that the aggregate effects on service use are small but there may be heterogeneity by income level.

Several institutions protect financially vulnerable patients from healthcare costs. For example, there are separate annual out-of-pocket caps for public healthcare services and prescription drugs of 692 euros and 592 euros in 2022. Patients with low enough income and little wealth are entitled to apply for a last-resort benefit called social assistance to receive support for basic living costs such as out-of-pocket costs for public healthcare services and prescription drugs. The law on out-of-pocket costs in healthcare requires that, for some public services, patients can apply for an exemption of a lowered copayment if they are financially vulnerable. This right does not apply to nurse visit copayments, but some areas may still exempt individuals after application based on their criteria, e.g., the patient having received social assistance. Few primary care areas also provide more general exemptions to specific low-income groups, such as patients with the lowest pension or unemployment benefit.

3 Data

We combine several Finnish national administrative registers by using unique person IDs. Specifically, we use data on contacts in publicly-funded primary care, prescriptions, social assistance recipients, and socioeconomic characteristics of all individuals.³ The

³Primary care contacts are extracted from the Register of Primary Health Care Visits, administered by the Finnish Institute for Health and Welfare (THL). The data on social assistance recipients from the Register of Social Assistance are administered by THL. Prescriptions are obtained from the Kanta Prescription Center, administered by the Social Insurance Institution. The socioeconomic data comprise Statistics Finland's

socioeconomic data contain all individuals who have a permanent residence in Finland at the end of a given year. We observe age and the municipality of residence, which are used to link visits to copayment policies. We can also construct a variable for equivalised family disposable income and compute population sizes. Findata and Statistics Finland authorize data permissions for the datasets, and empirical work is conducted using Statistics Finland’s remote access system.

Additionally, we use publicly available data on each municipality’s primary care area in 2021 collected by the Association of Finnish Municipalities. Two THL’s publicly available registers listing social and healthcare provider organizations are linked to primary care contacts. We also create three tables mapping areas to copayment policies. The first contains information at the municipality level on whether the municipality had adopted a copayment for curative nurse visits by the end of 2019 and on the possible adoption date. Information was collected by observing municipal documents, websites, and news in local media. The search was based on the publicly available dataset on nurse and GP visits copayments (THL, 2019), which we also use to observe the GP visit copayments in 2013-2018. The third table is about pre-abolition policies and was collected by observing the websites of primary care areas in Summer 2021 before the abolition.⁴

In empirical analyses, we use data from 1/2013 to 12/2019 for the staggered copayment adoption and from 7/2020 to 5/2022 for the later simultaneous abolition. Separating these two study windows limits the adoption analyses to pre-pandemic (Covid-19) times and the abolition analyses to pandemic times, excluding the early 2020 when the supply and demand shocks of the pandemic were largest. For years 2013 to 2020, we use socioeconomic data from the end of the given year, but for years 2021 to 2022 we use data from the end of 2020 due to a data release lag. We only include those individual-by-year observations in which the individual is 25 years or older. The aim is to

FOLK modules "basic", "family", and "income".

⁴We thank Katja Ilmarinen who had done the same independently. We compared our table to Ilmarinen’s and, thus, verified its contents.

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. The set of secondary outcomes depend on whether we analyze the staggered adoption or the later simultaneous abolition.⁵ Regarding the adoption, we evaluate the effects on the share of individuals receiving social assistance and the annual sum of received basic social assistance.⁶ Regarding the abolition, we examine the effects on the annualized number of prescriptions per capita written by public-sector organizations. We refer to the abovementioned outcomes when using the following expressions: nurse and GP visits, share of social assistance recipients, sum of basic social assistance, and prescriptions. Our pre-registered choice is to estimate the effects separately for the bottom 40% and the top 40% of the equivalized disposable income distribution in all analyses.⁷ However, we also show the main estimates on nurse visits by income decile.

We discuss in detail how we clean and construct our analysis data in Section A.1 in the Online Appendix, also motivating our choices for data construction. For the analyses, we have an unbalanced panel at the municipality-by-time-period-by-income-decile-by-outcome level. Time period is month for all the other outcomes except for the sum of social assistance which is measured only annually. The panel is unbalanced because we have to exclude some observations due to quality problems, mainly for primary care outcomes. When the national primary care data collection started in 2011, not all areas were able to transfer data from

⁵We have access to the prescription data from 2018 to 2022 and the social assistance data from 2012 to 2019. For this reason, we use social assistance outcomes only in the adoption analyses and prescription outcomes in abolition analyses.

⁶Social assistance is a means-tested last-resort benefit for households. Using information on family relations, we define that an individual received social assistance if the person lived in a family in which someone is observed to receive social assistance. Similarly, we sum the amount of social assistance at the family level.

⁷For testing heterogeneity by income level, we use two groups instead of more for parsimony. Focusing on smaller groups than the bottom 40% and the top 40% has two disadvantages: larger variation in municipal outcomes due to smaller samples, and the fact that the share of social assistance recipients is larger at the bottom end of the income distribution, potentially attenuating estimates as social assistance can be applied to cover healthcare costs.

their electronic health record systems (EHR) to the national register. Later changes in the providers of 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 such observations that have data quality issues are provided in the Online Appendix in Section A.1.

4 Results: Staggered Adoption

We start our analysis by focusing on the adoption of the copayment using a staggered difference-in-differences (DD) design with an irreversible treatment. As a baseline, we use a stacked event-by-event design (Gormley and Matsa, 2011; Cengiz et al., 2019) that is robust to biases in conventional two-way fixed effects (TWFE) regression models caused by the combination of staggered treatment timing and treatment effect heterogeneity. The theoretical results show that both the static (Goodman-Bacon, 2021) and event study specifications suffer from these biases (Sun and Abraham, 2021). Baker et al. (2022) find that the biases can be relevant in real-world applications and that stacking is a robust alternative in Monte Carlo simulations.⁸ Moreover, we provide extensive robustness checks by using the Callaway and Sant’Anna (2021) (CS) estimator.

In stacking, 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 treatment window of 24 months before and 12 months after the copayment adoption.⁹ 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 primary care data quality issues (see Section A.1). In the main analysis, we require a balanced panel in event (or relative) time. These event-specific datasets are then stacked (or pooled) together for estimation which we conduct using TWFE regression models but now with dataset-specific

⁸Stacking also accommodates both static and dynamic specifications and triple difference models.

⁹Consequently, we avoid using earlier-treated units as a comparison for later treated units, which is the key problem in the conventional static TWFE models (Goodman-Bacon, 2021).

unit and time fixed effects. Due to substantial heterogeneity in municipality size, we weight by population size in all regressions.¹⁰ Standard errors are clustered by municipality.

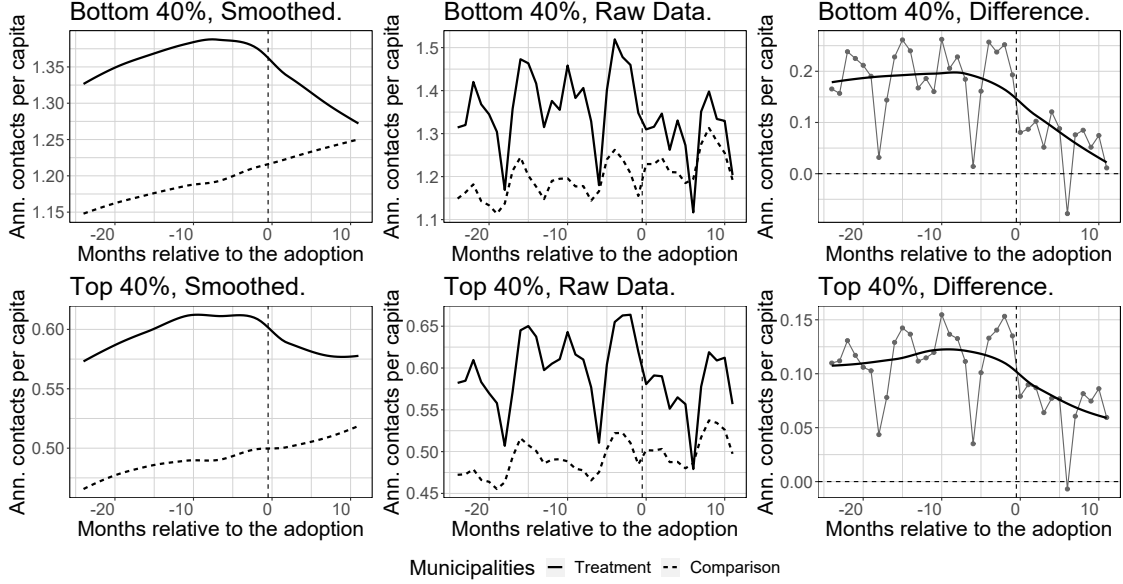


Figure 2: Adoption: Evolution in Nurse Visits.

Notes: The dataset is stacked, and event-specific datasets balanced. The outcome is the number of annualized curative nurse visits per capita. 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 equivalised family disposable income.

Pre-trend plots. Before presenting the regression results, we first plot in Figure 2 the trends in curative nurse visits for the bottom 40% and the top 40% of the distribution of equivalised family disposable income in treatment and comparison municipalities based on the stacked dataset. The same graphs, but for curative GP visits and social assistance outcomes, are shown in Figure A9 and Figure A10. According to these graphs, nurse use decreased in the treated municipalities after the adoption of the copayment compared to the comparison municipalities. The decrease was 0.10-0.15 annualized nurse visits at the bottom 40% of the income distribution and approximately 0.05 visits for the top 40%. There was an

¹⁰We discuss the choice of weighting in Section A.1. In robustness checks, we examine how sensitive the main results are to uniformly weighting municipalities.

increasing trend in nurse use in both municipality groups before the copayment adoption. After the adoption, the trend continued in the comparison municipalities, but the nurse use decreased in the treated municipalities. The effects on GP visits, in contrast, are small or zero. There may be a small decreasing pre-trend in GP visits in the treated areas relative to the comparisons.¹¹ Neither do we observe any clear effects on receiving social assistance.

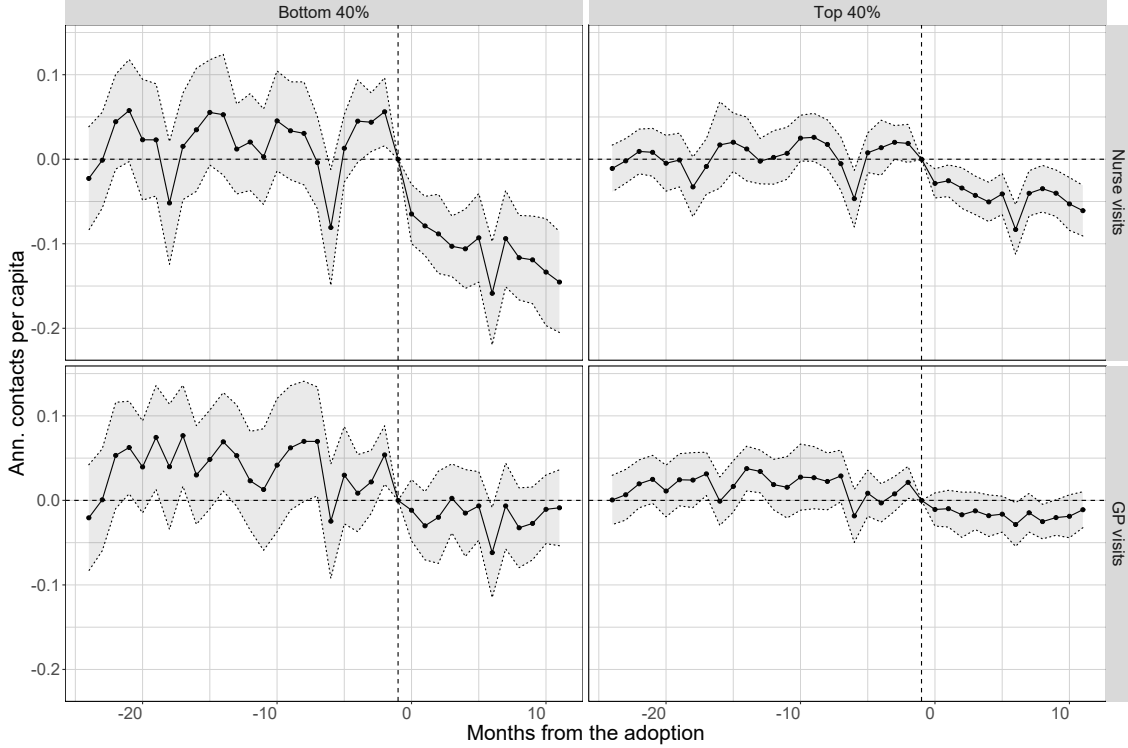


Figure 3: Adoption: Event-Study Plot on Primary Care Use with Stacked Data.

Notes: 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. Our TWFE regression model includes a full set of treatment indicators for 24 and 12 months before and after the treatment and event-specific municipality and time fixed effects, comparing the evolution of annualized contacts per capita between treated and unexposed municipalities. The last pre-treatment period, namely $t = -1$, is omitted as a reference. 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 equivalised family disposable income.

We continue assessing possible pre-trend differences by estimating a dynamic

¹¹For this reason, we prefer specifications that allow for a linear pre-trend difference in GP visits or assume parallel trends only from the last pre-treatment period on. The estimates from these specifications should be closer to zero than the estimates from the specification that is based on the assumption of parallel trends in every period.

regression model on the stacked data, comparing the evolution of annualized contacts per capita between the treated and unexposed municipalities. Our TWFE regression model includes a full set of treatment indicators for 24 and 12 months before and after the treatment,¹² and event-specific municipality and time fixed effects. Standard errors are clustered by municipality. The event study plots on nurse and GP visits are reported in Figure 3 and the plots for our social assistance outcomes are shown in Figure A11. Consistent with the pre-trend plots, the event-study plots show that nurse use 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, and no clear effects on receiving social assistance are observed. The reduction in nurse use in Figure 3 is larger at the bottom 40% of the income distribution in absolute terms. However, such a pattern is not evident in relative terms using the logarithm of the outcome (Figure A12). The observed reductions in service use (especially in nurse visits) in the treated municipalities occurring every twelve months are 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.

Main results. We produce aggregated treatment effect estimates using the static DD framework and the stacked data. Our TWFE regression model includes an indicator for post-treatment periods in the treated municipalities and event-specific municipality and time fixed effects. Standard errors are clustered by municipality. The assumption of parallel trends is plausible for nurse visits based on pre-trend and event-study plots. The results on annualized primary care contacts per capita are reported in Table 1. The results on social assistance outcomes are in Table A1 in the Online Appendix.

Both nurse visits and GP visits decrease in 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%)

¹²The last pre-treatment period, namely $t = -1$, is omitted as the reference.

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%. There is notable heterogeneity in the effects in absolute terms (i.e., the lower end of the income distribution is more sensitive to copayments). However, such a pattern is not apparent 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. The estimate on the annual sum of the family's received social assistance per capita is positive and marginally insignificant. However, the inclusion of a linear pre-trend difference attenuates both estimates towards zero.

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	0.0001	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. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by event-specific linear pre-trend differences in relative time. The mean of the estimated dynamic effects is reported. 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 equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

Robustness checks. In Table 1 and Table A1, we also provide the results using a

modified specification that allows for pre-trend differences. Based on pre-trend plots, this is mostly relevant for GP visits, as there is likely a small pre-trend difference. Specifically, we replace the static $\text{treat} \times \text{post}$ indicator by lags of every post-treatment period for the treated municipalities and by event-specific linear pre-trend differences in relative time between the treated and comparison municipalities. The average of the lags is reported in the tables. The estimates for nurse visits are insensitive. However, the estimates for GP visits are somewhat closer to zero once we allow for a linear pre-trend difference.¹³

In Panel A, Table A2 reports the stacked regression estimates on annualized primary care contacts per capita based on unbalanced event-specific datasets that have more municipalities and municipality-month observations than the balanced datasets in the main analysis. In Panel B, we present the estimates using the logarithm of annualized contacts per capita as the outcome. In both cases, the main findings are robust.

As an alternative to stacking, we also use the CS estimator (Callaway and Sant’Anna, 2021). The estimator is robust in settings with staggered treatment timing and treatment effect heterogeneity. The key building block in Callaway and Sant’Anna (2021) is a group-time average treatment effect¹⁴, allowing for treatment effect heterogeneity over groups and time. The authors propose several two-step plug-in estimators for these group-time average treatment effects: first estimate nuisance functions and then plug the fitted values of the nuisance functions into the sample analogue of the group-time ATT. In our application, we use outcome regression, weight by population, and cluster standard errors by municipality. As with stacking, we include events with at least a 12-month follow-up. However, this time we use a balanced dataset in calendar time, not in relative time.¹⁵ Using a balanced dataset in calendar time implies excluding those municipalities with poor primary care data quality. Our baseline is to exclude the years 2013 and 2019 when analyzing primary care outcomes with the CS estimator to increase the number of

¹³Considering all study results, the number of general practitioner (GP) visits is reduced by 2% to 5%. The estimates accounting for a linear pre-trend difference are closer to zero in that interval.

¹⁴Group is defined by the time period when units first receive the treatment.

¹⁵The CS algorithm in the R package *did* is slow with unbalanced datasets.

sample municipalities given that we restrict to a balanced panel.¹⁶ Regarding the social assistance outcomes, we use all data from 2013 to 2019. The data are aggregated to the municipality-by-time-period level for estimation.

The group-time ATTs can be aggregated to construct measures of overall treatment effects such as event study estimates or a single overall treatment effect estimate as in 2x2 DD designs. First, we provide event study estimates on the primary care outcomes in Figure A13 (logarithmized outcome: Figure A14) and on the social assistance outcomes in Figure A15. The findings on nurse visits are similar in both stacked event study plots and the CS event study plots. Regarding GP visits, the CS event study plots show virtually no effect, while the stacked event study plots suggest a small decrease. The CS event study plots reveal a decrease in the share of individuals receiving social assistance, but this finding is not supported by the stacked event study plots.

Next, we report a single estimate that is the weighted average of all group-time ATEs, weighted by group size.¹⁷ Regarding the primary care outcomes, we estimate the effects on annualized visits per capita (in Figure A16) and on the logarithm of annualized visits per capita (in Figure A17) in eight cases. We use either never-treated or not-yet-treated municipalities as the comparison group and exclude data from either 2013 or 2019 or both. The results on the social assistance outcomes are in Figure A18.

The CS estimates differ from the stacked estimates in two ways. First, the effects on nurse visits are farther from zero. Annualized nurse visits decrease by -0.18 to -0.23 visits (-13% to -17%) at the bottom 40% and by -0.08 to -0.10 visits (-13% to -16%) at the top 40%. This can plausibly be explained by the different estimand: the stacked

¹⁶The exclusion of year 2013 means that we trade off one event and 12 months of data for a greater number of municipalities. In many cases, the primary care visit data quality issues reported in Section A.1 occurred early in the panel. When 2013 is excluded, more municipalities have a balanced panel in the study period. However, the exclusion of 2019 allows us to keep one large, never treated municipality that changed its electronic health records system in Spring 2019 and consequently had data quality issues.

¹⁷Note that there are two differences to the estimand of the main analysis: the weighting procedure and also the length of the follow-up. In the main analysis, the effects were estimated using a 12-month follow-up. Here, follow-up varies by treatment group and is restricted only by the time each group participates in treatment. This implies putting more weight on the earlier-treated cohorts.

estimate computes the effect based on a one-year follow-up, while the CS estimate takes the average over all group-time ATEs and thus uses for most units a longer follow-up period. If the effects accumulate over some months and are not immediate, the CS estimate should mechanically be farther from zero. Second, the estimates on GP visits are all negative but close to zero and insignificant.

Effect heterogeneity by income level. Finally, we employ a triple difference (DDD) design using the stacked dataset to provide tests of treatment effect heterogeneity by the income level. Our hypothesis is that low-income individuals respond more strongly to copayment changes than high-income individuals. Based on this, we compare the evolution of the outcomes in the bottom 40% of the income distribution to that in the top 40% both in the treatment and comparison areas. The specification assumes parallel trends in ratios (Olden and Møen, 2022). If we use the number of annualized contacts as the outcome (or its logarithm), the DDD estimates are unbiased if the pre-trend differences are similar in levels (in percentage terms) in the bottom 40% and in the top 40%. We estimate the effects using the following specification:

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

Here, α is an intercept, m , g , t , and e denote municipality, socioeconomic group, month, and event-specific dataset. $Treat$ is a dummy for treated municipalities, $Affected$ is a dummy for the more affected group (the bottom 40%), $Post$ is a dummy for post-treatment periods, γ is the coefficient of interest, and ε is the error term. Note that the other coefficients except for the causal parameter of interest are event-specific. We weight by population size and cluster standard errors by municipality.

The results on annualized contacts per capita and logarithmized annualized contacts per capita are in Table 2, using both balanced and unbalanced event-specific datasets. In

absolute terms, the nurse visits decrease by -0.07 to -0.08 visits (-5.3% to -5.6%) in the bottom 40% relative to the top 40%, and the estimates are significant. GP visits appear to decrease more in the bottom 40%, but the estimates are insignificant. However, all the estimates are insignificant in relative terms (using the logarithmized outcome).

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	0.00001	0.058	0.00001	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 1. 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.

We also provide a more flexible analysis of treatment effect heterogeneity by income by estimating and plotting the effects on annualized nurse visits per capita and logarithmized annualized nurse visits per capita by income decile in Figure A19. We use stacking with balanced event-specific datasets to obtain the results. The pattern is clear in absolute terms:

the estimate is closer to zero the higher the income decile. The bottom 10% is, however, an exception. There appears to be heterogeneity also in relative terms: the lower end of the income distribution is somewhat more sensitive than the top end.

Estimates for all individuals. The focus of this study is on the potential treatment effect heterogeneity by income. However, aggregate estimates are also relevant to policymakers. We provide the results for the entire sample population in the Online Appendix: pre-trend plots (Figure A20), dynamic regression plots using stacking (Figure A21) and the CS estimator (Figure A22), and aggregated stacking (Table A3) and CS (Figure A23) estimates. Annualized nurse visits decrease by -0.09 to -0.10 visits (-9% to -10%) using stacking and by -0.13 to -0.16 visits (-13% to -17%) using the CS estimator. Note that the CS estimator has a longer follow-up than the one-year follow-up in our stacked regressions. Stacked estimates on GP visits show a decrease (-3% to -5%), but the corresponding CS estimates are closer to zero and insignificant.

Robustness to weighting municipalities uniformly. We repeated the above analyses, but instead of population weighting, we uniformly weighted municipalities when using the CS estimator and municipality-by-income-decile observations when using the TWFE regression. The pre-treatment healthcare use is now higher as small municipalities have a larger weight. Qualitatively, the main findings are robust to the weighting scheme: the copayment adoption is associated with a reduction in nurse use in the treated areas. The estimates on GP visits are negative but small, and the estimates on social assistance outcomes are inconclusive. The uniformly-weighted stacking results on annualized primary care contacts per capita are reported in Table A4: nurse visits decrease by -0.17 visits (-11.2%) in the bottom 40% and by -0.07 visits (-9.4%) in the top 40%. In contrast, the population-weighted estimates in Table 1 show smaller decreases: -0.13 visits (-9.3%) in the bottom 40% and -0.05 visits (-8.0%) in the top 40%. The rest of the figures and tables are provided in the replication codes folder.

5 Results: Later Simultaneous Abolition

We analyze the effects of the nationwide nurse visit copayment abolition of July 2021 using 12 pre-treatment and 11 post-treatment months, requiring a balanced panel. We assume that the effects accumulate fully within one year. For this reason, we exclude those municipalities that adopted the copayment less than 12 months before the start of the study window.

Background: the Covid-19 pandemic. Primary care utilization in 2020-2022 is to a large extent affected by the global Covid-19 pandemic. At the onset of the pandemic, major supply and demand shocks reduced the utilization of a broad range of services all around the world Moynihan et al. (2021). Interestingly, the number of curative nurse visits has not recovered to pre-pandemic levels by May 2022 in Finland (Figure A24). This finding is likely explained by supply-side factors. Public discussion about the shortage of nurses in publicly funded Finnish healthcare has intensified during the Covid-19 pandemic. The fact that nurses have been allocated to testing, tracing, and vaccination tasks has reduced the capacity of the system to treat non-Covid patients. Covid patients have also been a substantial additional burden on already strained public healthcare.

The observed reduction in nurse visits affects our ability to make valid inferences from the copayment abolition. First, it is likely that primary care utilization is even more supply-driven than in normal times. Assuming a shortage of appointment slots, the copayment abolition may not increase aggregate nurse visits even if demand for them increases. Patients who get appointments may also have more urgent conditions on average than before the pandemic. Second, the parallel trends assumption (PTA) required for causal inference may be less plausible in a context where the epidemiological situation and thus Covid-related burden to the system have varied during the pandemic. Moreover, the pandemic may have affected urban and rural areas differently. Urban areas are overrepresented in our comparison group for the abolition analyses.

Pre-trend plots. Figure 4 plots the trends in curative nurse and GP visits for all sample individuals. The differences in levels between the treatment areas are rather stable

in the nine months prior to the abolition. After the abolition, nurse use appears to increase in the treated areas compared to our comparison municipalities. This trend break is not observable in GP visits.

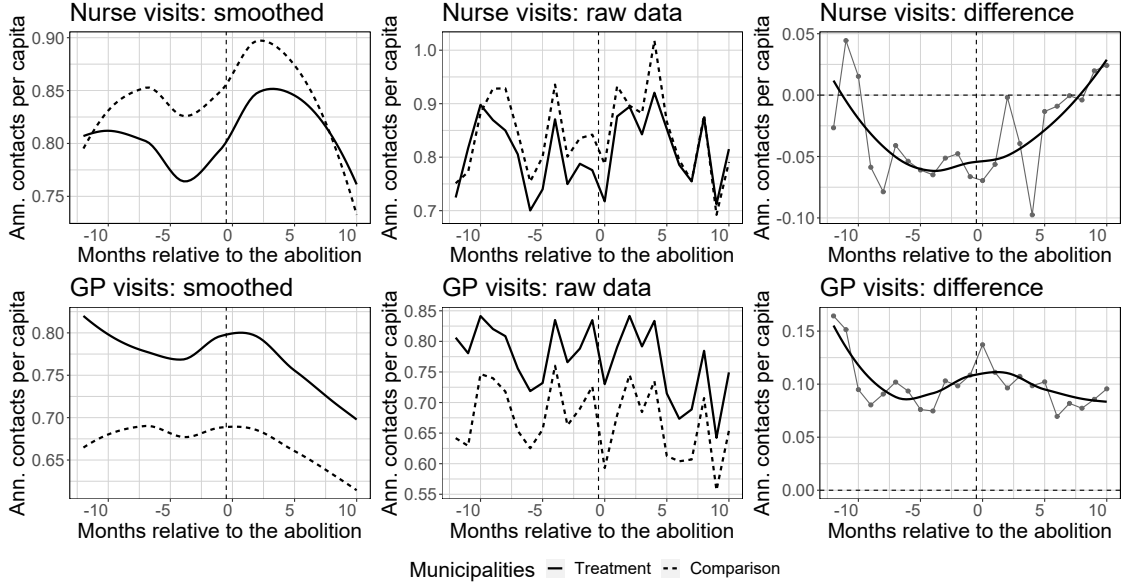


Figure 4: Abolition: Evolution in Outcomes, All Individuals.

Notes: This figure was not pre-registered. The outcomes are the number of annualized curative nurse visits and GP visits per capita. 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.

It seems possible that the abolition increased nurse use. However, we are not willing to make a much stronger conclusion that the observed increase in nurse visits is the causal effect of the copayment abolition. First, even if the pre-treatment differences in outcomes are stable for the nine months before the abolition, the same cannot be stated of the whole lead-up period. The nurse use was 0.04 to 0.08 annualized visits lower in the treated areas in 10/2020-6/2021 (months -9 to -1). In contrast, the nurse use was even higher in the treated areas than in our comparison areas in 7/2020-9/2020 (months -12 to -10). The quantitative magnitude of this relative decrease in pre-treatment periods is approximately similar to the observed relative increase in post-treatment periods. The observed relative changes in pre-treatment periods make it more plausible that there could

be changes of similar magnitude in the post-treatment periods even in the absence of the policy change. Second, the relative increase in nurse visits in the treated areas essentially occurs in 12/2021-5/2022 (months +5 to +10), six months after the policy change. Nurse use does not increase notably in the treated areas relative to the comparison areas in the first five post-treatment months.

In the Online Appendix, Figure A25 plots the trends in curative nurse visits for the bottom 40% and the top 40% of the distribution of equivalised family disposable income in treatment and comparison municipalities. The same graphs, but for curative GP visits and prescriptions, are in Figure A26 and Figure A27. Figure A28 plots the pre-trends in nurse and GP visits for all individuals but this time weighting municipality-by-income-decile observations uniformly instead of population weights.

Regression estimates. The above analysis illustrates visually that the conventional PTA is not plausible in this application. Consequently, few causal conclusions can be made. We still provide highly suggestive point estimates for the magnitude of the observed changes in nurse use, either caused causally by the copayment abolition or driven by unobserved confounders correlated with the treatment assignment. To highlight the suggestive nature of the estimates, we do not report standard errors in the table. At least it seems possible that the abolition increased nurse visits.

We use the following three methods that differ by the PTA required, weighted by population size. We view none of the three assumptions superior to the rest. First, the TWFE DD specification without a pre-trend difference contains a static indicator for post-treatment periods for treated municipalities, and municipality and time fixed effects. It assumes parallel trends throughout the study window. Second, the TWFE DD specification with a pre-trend difference replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by linear pre-trend differences in relative time between the treated and comparison areas. The average of the lags is reported. It assumes parallel trends in deviations of the outcome from a linear time trend. Third, the

CS estimator (see Section 4) assumes parallel trends but only from the last pre-treatment period on.

Table 3: Abolition: DD Comparisons, All Individuals.

Metric	Nurse Visits	GP Visits	Prescriptions
A. Annualized contacts per capita			
Level	0.799	0.791	4.301
Municipalities	230	230	238
Estimate (w/o trends)	0.018	−0.007	−0.006
Change (%)	2.298	−0.828	−0.141
Estimate (with trends)	0.088	0.038	−0.075
Change (%)	10.970	4.837	−1.740
Estimate (CS)	0.044	−0.012	−0.035
Change (%)	5.492	−1.506	−0.814
B. Logarithmized annualized contacts per capita			
Municipalities	229	230	238
Estimate (w/o trends)	3.816	0.435	0.145
Estimate (trends)	8.582	8.009	1.802
Estimate (CS)	5.489	−2.120	−0.386

Notes: This table was not pre-registered. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by a linear pre-trend difference in relative time. The mean of the estimated dynamic effects is reported. 3) The CS estimator with outcome regression (Callaway and Sant’Anna, 2021). Due to heterogeneity in municipality size, we weight by population size.

The estimates for all sample individuals are reported in Table 3. With respect to nurse visits, extrapolating a linear pre-trend difference (TWFE with trends) leads to the largest point estimate while the TWFE model without trends produces the smallest estimates. The abolition is associated with an increase of 2.3 to 11.0 percent in the treated

municipalities. The corresponding estimates for GP visits are in every case smaller and closer to zero, varying from -2.1% to $+8.0\%$. The estimates on prescriptions vary from -1.7% to $+1.8\%$.

Table A5 contains the corresponding estimates but after weighting municipalities uniformly instead of using population weights. The TWFE model without trends and the CS estimator produce estimates on nurse visits that vary between $+0.2\%$ and $+9.8\%$. The TWFE model extrapolating a linear pre-trend difference, in contrast, produces a large 22 to 23 percent increase. The estimates on GP visits vary between -4.8% and $+6.0\%$ and the estimates on prescriptions between -1.4% and $+0.7\%$.

The estimates for the bottom 40% and the top 40% of the income distribution are reported in Table A6. Regardless of the method, the estimates on nurse visits are always higher in relative terms for the top 40%. In two cases, they are also larger in absolute terms. Thus, there are no clear signs that the correlation between the copayment abolition and nurse use is larger at the lower end of the income distribution.

Time-placebo analyses. Our pre-analysis plan provides placebo estimates from experiments in which we fix the treatment and comparison municipalities and the treatment date (July 1st) but proceed as if the treatment occurred in 2018 or 2019. We exclude municipalities that adopted the copayment in the study period or less than 12 months before the start of the window. The implicit assumption is that the effects of copayment adoption have accumulated fully in a year, so that earlier adoptions do not confound our placebo analysis. These placebo results show that the estimates can be sensitive to the specific version of the PTA. However, in these time-placebo runs it is much easier to select a preferred PTA out of the three alternative than it is with the data from 2020-2022.

6 Discussion

We analyze the effects of a staggered adoption of a nurse visit copayment (approximately 10 euros) and its later simultaneous abolition on public primary care use in Finland. We find that the copayment adoption reduces curative nurse visits by 9% to 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 at the bottom 40% of the income distribution than at the top 40%. However, such heterogeneity is less clear in relative terms (percentage changes). The estimates also show small albeit sometimes insignificant reductions of 2% to 5% in GP visits. Regarding the abolition, we do not obtain causal conclusions since the parallel trends assumption is not plausible based on pre-trend patterns.

Our staggered DD design for the copayment adoption analysis has useful characteristics for strong internal validity. It allows us to construct a reasonable comparison group for the treated areas. The timing of the adoptions seem rather arbitrary, although the adopters are on average more rural and less populous than the country on average. The pooling of several event-times - 17 in the main analysis - reduces the risk that the estimates are driven by unobserved confounders that are not causally related to the copayment adoption. Our graphs show that the pre-treatment trends in nurse visits were parallel.¹⁸ The estimates are driven by a clear reduction in nurse use in the treated municipalities as soon as the copayment was adopted. In the comparison municipalities, however, nurse use evolves rather similarly to before treatment.

A key issue is whether there are unobserved confounders that are causally related to the copayment adoption. The leading candidate is data quality. Once the copayment is adopted, preventive and curative nurse visits must be distinguished so that the copayment can be charged for curative visits. Similarly, areas that charge the copayment for the first three visits annually need to count the number of contacts for invoicing. If these requirements

¹⁸This may not be the case for GP visits, however.

affect the data that are recorded to the EHR system and to the national register, the number of recorded nurse contacts may change even if the underlying utilization does not. However, we checked that the number of preventive outpatient nurse visits does not increase in the treated areas after the copayment adoption based on the stacked dataset. Thus, there is no evidence of preventive contacts crowding out curative contacts. Second, we find a small but negative effect on GP visits, suggesting that there may be a behavioral effect.

To make it easier to compare the effects, we convert our adoption 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 the change in quantity, 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 the "price" of a nurse visit. Following Nilsson and Paul (2018), we define the price as the share of the out-of-pocket costs of the total cost of the visit. The elasticity is sensitive to the chosen parameters, so we provide two estimates: a baseline and an upper bound in absolute value. The parameter values are listed in Section A.3. The baseline estimate is -0.41 and the upper bound is -1.24 . Nilsson and Paul (2018) report a semi-arc elasticity of -0.88 at the 20th birthday and -0.55 at the 7th birthday when previously exempted individuals face copayments of 10 to 15 euros for outpatient doctor visits in Sweden, Finland's neighbor and perhaps the closest possible comparison in terms of institutions. Also in Sweden, the estimates of the effect of a 10 euro copayment for GP visits at the 20th birthday map to an elasticity of -1.11 (Johansson et al., 2019).¹⁹ All these elasticity estimates are lower than the estimates of -2.11 and -2.26 that Brot-Goldberg et al. (2017) compute for the RAND Health Insurance Experiment.

Regarding the fiscal significance of the copayment adoption, we compare estimated copayment revenue to estimated savings from lower primary care use based on straightforward back-of-the-envelope calculations. Supposing 4 million sample individuals,

¹⁹We pick 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.

a 10-euro copayment, that only 70% of the copayment revenue are actually collected²⁰, a pre-treatment mean of 1.000 annualized nurse visits, an effect of -0.089 annualized visits²¹, and an average cost of 35 euros per nurse visit (Mäklin and Kokko, 2020), the copayment revenue would be 26 million in a year compared to savings of 12 million euros due to lower nurse use. However, if the estimated behavioral effect is larger, say -15% (Figure A23), then the collected copayment revenue would be 24 million and the savings 21 million.²² The main conclusion from these calculations is that if both the collection costs and production costs were higher than the values used above, then the fiscal importance of the short-term savings is notable relative to the collected revenue. This conclusion does not account for the potential health effects. If the reduced primary care use had health effects, these should be taken into account in the calculation.

We focus on heterogeneity analysis by income. After the copayment adoption, the reduction in nurse use in terms of the number of visits was largest at the bottom of the income distribution. This finding is consistent with Johansson et al. (2019) and Nilsson and Paul (2018). In contrast to these two Swedish studies, we observe heterogeneity only in levels but not in relative terms (effects compared to baseline utilization). The observed differences in sensitivity to copayments illustrate that public health policies can have unequal effects. Income-related heterogeneity can be especially important as low-income individuals have on average worse health status. Consequently, the marginal benefit of service use is arguably higher for them than for the rest.

Our study illustrates based on an application from health economics how a detailed pre-analysis plan (PAP) can be utilized. It is debated how strictly researchers should adhere to pre-registered plans *ex post* (see, e.g., Banerjee et al., 2020). In the adoption analyses,

²⁰This should account for collection costs, increased social assistance use, and the fact that the most common policy was to charge the copayment for the first three visits annually.

²¹These are an average over two specifications of Table A3.

²²According to the Finnish Ministry of Social Affairs and Health, the estimated collected copayment revenue for nurse visits was 14.7 million euros in 2017, not accounting for administrative costs. Our figures are computed as if our whole sample (those aged 25 or more) faced a 10-euro copayment. If only 39% did, as was approximately the case in 2017 (Figure 1), then the 24 and 26 million map to 9 and 10 million euros.

we were able to closely follow the pre-registered choices. In the abolition analyses, however, there was a need to change which analyses (that is, figures and tables) are reported and which are not. All changes to the PAP are documented in Section A.2. We generally recommend others to experiment with PAPs also in quasi-experimental work. Our PAP contains not just a general-level description of the project, but also statistical programs written without observing the real outcome data, and a corresponding placebo report. The potential extra credibility and transparency that detailed PAPs offer is especially desirable in policy evaluation, in which research questions are often straightforward to conceptualize (does the program "work" or not), studied phenomena can be politically controversial, and the results are potentially used to guide future policies.

CRedit author statement: **Haaga:** Conceptualization, Formal analysis, Writing - Original Draft, Writing - Review & Editing. **Böckerman:** Conceptualization, Writing - Review & Editing, Supervision. **Kortelainen:** Conceptualization, Writing - Review & Editing, Supervision. **Tukiainen:** Conceptualization, Writing - Review & Editing, Supervision.

References

- AAMC (2021). The Complexities of Physician Supply and Demand: Projections From 2019 to 2034. Prepared for the AAMC by IHS Markit Ltd.
- Baker, A. C., Larcker, D. F., and Wang, C. C. Y. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.
- Banerjee, A., Duflo, E., Finkelstein, A., Katz, L., Olken, B., and 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. Technical report.
- Brodeur, A., Cook, N., Hartley, J., and Heyes, A. (2022). Do Pre-Registration and Pre-Analysis Plans Reduce p-Hacking and Publication Bias?
- Brodeur, A., Cook, N., and Heyes, A. (2020). Methods Matter: p-Hacking and Publication Bias in Causal Analysis in Economics. *American Economic Review*, 110(11):3634–3660.
- Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., and 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(3):1261–1318.
- Callaway, B. and Sant’Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.
- Card, D., Dobkin, C., and Maestas, N. (2008). The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *The American Economic Review*, 98(5):2242–2258.
- Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.
- Chandra, A., Flack, E., and Obermeyer, Z. (2021). The Health Costs of Cost-Sharing. NBER Working Paper No. 28439-.

- Chandra, A., Gruber, J., and McKnight, R. (2010). Patient Cost-Sharing and Hospitalization Offsets in the Elderly. *American Economic Review*, 100(1):193–213.
- Chandra, A., Gruber, J., and McKnight, R. (2014). The impact of patient cost-sharing on low-income populations: Evidence from Massachusetts. *Journal of Health Economics*, 33:57–66.
- Christensen, G. and Miguel, E. (2018). Transparency, Reproducibility, and the Credibility of Economics Research. *Journal of Economic Literature*, 56(3):920–980.
- Clemens, J. and 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. Technical report.
- Einav, L. and Finkelstein, A. (2018). Moral Hazard in Health Insurance: What We Know and How We Know It. *Journal of the European Economic Association*, 16(4):957–982.
- Fukushima, K., Mizuoka, S., Yamamoto, S., and Iizuka, T. (2016). Patient cost sharing and medical expenditures for the Elderly. *Journal of Health Economics*, 45:115–130.
- Goldin, J., Lurie, I. Z., and McCubbin, J. (2020). Health Insurance and Mortality: Experimental Evidence from Taxpayer Outreach. *The Quarterly Journal of Economics*, 136(1):1–49.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.
- Gormley, T. A. and Matsa, D. A. (2011). Growing out of trouble? Corporate responses to liability risk. *The Review of Financial Studies*, 24(8):2781–2821.
- Han, H.-W., Lien, H.-M., and 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(3):238–278.

- Iizuka, T. and Shigeoka, H. (2021). Is Zero a Special Price? Evidence from Child Healthcare. *American Economic Journal: Applied Economics* (forthcoming).
- Johansson, N., Jakobsson, N., and 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(8):1271–1280.
- Kondo, A. and Shigeoka, H. (2013). Effects of universal health insurance on health care utilization, and supply-side responses: Evidence from Japan. *Journal of Public Economics*, 99:1–23.
- Laurant, M., van der Biezen, M., Wijers, N., Watananirun, K., Kontopantelis, E., and Van Vught, A. (2018). Nurses as substitutes for doctors in primary care. *Cochrane Database of Systematic Reviews*, (7).
- Ma, Y. and Nolan, A. (2017). Public Healthcare Entitlements and Healthcare Utilisation among the Older Population in Ireland. *Health Economics*, 26(11):1412–1428.
- Mäklin, S. and Kokko, P. (2020). Terveyden- ja sosiaalihuollon yksikkökustannukset Suomessa vuonna 2017.
- McMichael, B. J. and Markowitz, S. (2022). Toward a Uniform Classification of Nurse Practitioner Scope of Practice Laws. NBER Working Paper No. 28192.
- Moynihan, R., Sanders, S., Michaleff, Z. A., Scott, A. M., Clark, J., To, E. J., Jones, M., Kitchener, E., Fox, M., Johansson, M., Lang, E., Duggan, A., Scott, I., and Albarqouni, L. (2021). Impact of COVID-19 pandemic on utilisation of healthcare services: a systematic review. *BMJ Open*, 11(3):e045343.
- Neumark, D. (2001). The Employment Effects of Minimum Wages: Evidence from a Prespecified Research Design The Employment Effects of MinimumWages. *Industrial Relations: A Journal of Economy and Society*, 40:121–144.

- Neumark, D. and Yen, M. (2021). Effects of Recent Minimum Wage Policies in California and Nationwide: Results from a Pre-specified Analysis Plan. NBER Working Paper No. 28555. Technical report.
- Nilsson, A. and Paul, A. (2018). Patient cost-sharing, socioeconomic status, and children’s health care utilization. *Journal of Health Economics*, 59:109–124.
- Olden, A. and Møen, J. (2022). The triple difference estimator. *The Econometrics Journal*.
- Olken, B. A. (2015). Promises and Perils of Pre-analysis Plans. *Journal of Economic Perspectives*, 29(3):61–80.
- Shigeoka, H. (2014). The Effect of Patient Cost Sharing on Utilization, Health, and Risk Protection. *American Economic Review*, 104(7):2152–2184.
- Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.
- THL (2019). Data on Copayments in Primary Care.
- Vivalt, E. (2019). Specification Searching and Significance Inflation Across Time, Methods and Disciplines. *Oxford Bulletin of Economics and Statistics*, 81(4):797–816.
- WHO (2018). Building the economic case for primary health care: a scoping review. Technical report.
- Yang, B. K., Johantgen, M. E., Trinkoff, A. M., Idzik, S. R., Wince, J., and Tomlinson, C. (2021). State Nurse Practitioner Practice Regulations and U.S. Health Care Delivery Outcomes: A Systematic Review. *Medical Care Research and Review*, 78(3):183–196.

A Online Appendix

A.1 Constructing our Analysis Data

Copayment policies: In analyses, we do not use every municipality in mainland Finland. The policy is not observed for some municipalities. Some municipalities are excluded because they participated in such municipal mergers where some of the municipalities had a different copayment policy than others before the merger. Regarding the abolition, we exclude those areas that introduced the copayment between 8/2019 and 6/2020, less than 12 months before the start of the study window in 7/2020. Basically, we assume that the effects of the copayment adoption have accumulated fully within one year so that the areas are unaffected by the earlier adoption in the study period for the analysis of the abolition. Two municipalities are excluded because they abolished the copayment already some months before the national reform. Two municipalities were excluded because their nurse visit copayment covered only a very small set of nurse visits.

After the above restrictions, we have 264 out of 293 municipalities in mainland Finland for a study on staggered adoption. In total, they had 5.0 million residents compared to the Finnish population of 5.5 million. Regarding the simultaneous abolition, we have 249 municipalities with 5.0 million residents. Figure 1 illustrates the staggered adoption and the abolition in terms of treatment areas and treated population. Figure A1 (the staggered adoption) and Figure A3 (the abolition) show the municipal policies graphically. Figure A2 shows how the copayment level varied in Summer 2021.

Regarding the data on GP visit copayments using the 2013 municipal borders, we take into account municipal mergers and have to make choices about uncertainty in some municipality-month observations due to the observed documents not always being explicit and clear. These choices are observable in the replication codes.

Socioeconomic data: We exclude those ID-year observations where equivalised family disposable income is exactly zero (less than 1% of the rows) and only include those

observations where an individual is aged 25 years or more as we want to exclude minors, who are exempted from the copayment, and students, who have access to student healthcare. Without any other restrictions on the data, this leaves us with approximately four million individuals out of the population of 5.5 million. With the population remaining after the above two restrictions, we compute the distribution of the equivalised family disposable income and sort individuals into income deciles.

Primary care contacts: We extract curative primary care outpatient visits fulfilling the following conditions: 1) person ID and visit date are observed, 2) variables related to cancellations are missing, and 3) the healthcare professional was either a nurse or a doctor. The distinction between curative and preventive contacts is important, as the nurse visit copayment was charged for curative visits. In contrast, preventive nurse visits appear to include vaccinations as well (e.g., seasonal flu, or Covid-19) for which no copayments are charged. Since 2013, the coding rate with respect to the curative/preventive indicator has been close to one, but in 2012 the information was missing for approximately 7% of the rows. An exemption are the four municipalities who adopted the Apotti electronic health record system (EHR) system: Vantaa in Spring 2019, and Helsinki, Kerava, and Kauniainen in Spring 2021. The share of missing values has been large after the adoption of Apotti in these municipalities. Consequently, we start our analysis from 1/2013 and exclude the four Apotti municipalities from the abolition analyses and 2019 for Vantaa from the adoption analyses. The coding rate for profession containing both nurses and doctors has been rather steady, varying between 4 to 7 percent in 1/2013-12/2019 and between 5 and 10 percent in 1/2020-5/2022.

Weekend visits are excluded from the analysis to reduce the potential bias resulting from changes in the way emergency department visits are coded in the registers. During the study period, some primary care areas and hospitals have formed joint emergency departments, and these contacts may be coded either to the primary care register or to the specialized healthcare register. Duplicate contacts are dropped. That is, an individual

cannot have more than one curative visit on the same date and time with the same profession. Visits are linked to municipalities and, thus, to copayment policies via clients' municipality of residence.

Since 2019, the register also contains outpatient contacts in private clinics that we want to exclude. We do this by linking each visit in 2019-2022 to TOPI and SOTE organization registers that contain information on the visit provider. Both registers are continuously updated. We have an annual cross-sectional dataset on TOPI and two cross-sections on SOTE from early 2020 and early 2022. We use SOTE from early 2020 for years 2019 and 2020 and SOTE from early 2022 for years 2021 and 2022. In 2019-2022, the linking of TOPI does not work for 4% of the extracted visits while the same figure for SOTE is 1%. Then, we include those visits whose provider a) has a TOPI service area code that refers to health centers (120, 121, or 122) or b) is a public sector organization in SOTE.²³

After having aggregated primary care contacts to municipality-month observations, we exclude several observations due to quality issues. Not all areas were able to transfer data from their EHR to the national register when the national primary care data collection started in 2011. Changes in EHR systems can also be seen in the data as a sudden drop to a near-zero value in aggregate contacts. Even if these issues, mostly missing contacts, were unrelated to copayment policies, we want to mitigate the potential bias from missing visits by excluding the corresponding observations.

To identify suspiciously low or high values of service use, we first sum up curative nurse and GP visits at the municipality-month level and compute a distribution of means by permutationally excluding every combination of four consecutive months. The largest mean is our reference value with which we define an observation as suspiciously low if its value is less than X% of the largest mean. The threshold X depends on the outcome. July is not

²³There are also private sector organizations providing publicly-funded primary care if the primary care area has outsourced services. In 2020, there were 21 such contracts between primary care areas (mostly small municipalities) and private sector organizations according to the Association of Finnish Municipalities. However, all primary care contacts in health centers should have a TOPI service area code that refers to health centers.

considered because many people, both professionals and patients, are on vacation. Next, we again compute a distribution of means but this time after excluding the suspiciously low observations. We define an observation to be suspiciously high if its value is greater than $120 + X\%$ of the largest mean. We mark all municipality-year pairs to be excluded if they contain suspicious months. We conduct the algorithm separately before (from 1/2013 to 12/2019) and after (from 7/2020 to 5/2022) the onset of the COVID-19 pandemic.

90 municipalities (out of 293) have suspiciously low or high values of primary care use in the pre-pandemic study period using a threshold of $X = 0.4$. The evolution of the sum of curative nurse and GP visits in these municipalities is illustrated in Figure A5, Figure A6, and Figure A7. Gray segments highlight municipality-year pairs with suspiciously low values. Based on visual inspection, the algorithm appears to be good at detecting irregularities. Regarding the abolition analyses, we set a higher threshold of 0.55 to make the algorithm more sensitive. 8 municipalities have suspicious values (Figure A8). In the replication codes folder, we illustrate primary care use in the remaining municipalities for which we find no suspicious observations.

Social assistance recipients: The raw data contain ID-year observations for social assistance recipients, including monthly indicators for having received social assistance and annual sums of social assistance. One person is an applicant in the data even if the benefit is applied at the household level. Using data on family relations, we construct for each individual two variables: a monthly indicator for belonging to a family where someone received social assistance in a given month, and an annual sum of received basic social assistance. When aggregating outcomes to municipality-by-time-period level, the share of recipients is in percentages, and the sum of euros is per capita. We look for potentially missing values with the same algorithm as for primary care contacts with a threshold of 0.4. The algorithm comes up with 24 municipalities with susceptible municipality-year observations, but after a closer inspection we decide to exclude only one municipality-year observation, namely municipality no. 892 in 2016.

Prescriptions: We extract prescriptions written at public-sector units, containing both initializations of drug therapies and renewed prescriptions. Ideally, we would have wanted to consider only drug therapy initializations as patients can apply for renewals online or at pharmacies without a GP visit. However, we had some doubts about the quality of the variable and decided to include all prescriptions. Excluding private sector prescriptions may be a problem in a small set of mostly small municipalities that have outsourced their primary care services to a private provider. We search for potentially missing values with the same algorithm as with primary care contacts. The algorithm does not find any irregularities even if we use a higher threshold of 0.6, making it more sensitive to outliers.

Weighting by population size. In the main analysis, we weight all regressions by population size to increase the precision of our estimates. Finnish municipalities are heterogeneous by population size. In 2020, the smallest municipality in mainland Finland had 700 residents, while the largest had 657,000. The median population size was 6,000. Our effective municipal sample sizes are smaller, as we restrict to individuals aged 25 or older and focus on the bottom 40% or the top 40% of the income distribution. Therefore, it is obvious that the outcomes of small municipalities can be much noisier than the outcomes of large municipalities in our data.

The plausibly increased precision due to population weighting comes at a cost. First, the ATT estimates may not generalize to the whole country if the variation essentially comes from the largest municipalities. The ATTs may not need to be homogeneous by municipality size. Second, an institutional change (say, a change in the EHR system) in one large municipality can more easily bias the estimates than in a case where the variation comes from a large pool of municipalities, each municipality receiving the same weight.

As a robustness check, we estimate the results also without population weighting. This means that we uniformly weight municipalities when using the CS estimator and municipality-by-income-decile observations when estimating the results with the TWFE regression (stacked or not). We made the following commitments in the PAP. If the results are

reasonably insensitive to the form of weighting, we only report the main population-weighted estimates in the report and its appendices, and the uniformly-weighted results are provided in the replication folder in addition to the replication codes. If, however, the results are sensitive to weighting, we include some of these uniformly-weighted result tables to the report as well to provide balance.

A.2 Changes to the Pre-Analysis Plan

Data: detecting quality issues in the data. In the PAP, the algorithm which we use to detect quality issues in the primary care data was designed to find periods with abnormally low health care use. Here, we modify it so that it can detect periods with abnormally high values as well, which we think is a reasonable change. The implications can be seen for municipality no. 10 in Figure A5: now the spike in late 2019 is detected as an abnormally high value. We also use a higher threshold of 0.55 for the abolition analyses, making the algorithm more sensitive to outliers. Using a higher threshold of 0.6 for prescriptions does not change the number of excluded observations.

Data: the Apotti EHR system. We anticipated in the PAP that the adoption of the Apotti EHR system in three municipalities in Spring 2021 might cause data quality issues in the register data, which it did. The algorithm described above is able to detect these three municipalities, but we also manually exclude Vantaa from the abolition analyses. Vantaa adopted the Apotti EHR already in 2019. After that, the share of missing values in the curative/preventive categorical has been large.

Data: socioeconomic data. The latest statistical year available is currently 2020. We expect that the socioeconomic data from the end of 2021 will be available by June 2023. In the final research report, we plan use values from the end of 2021 for year 2022.

Data: social assistance recipients. We added the year 2019 to the analysis. These data were not yet extracted to us at the time of writing the PAP.

Data: does data quality improve when the copayment is adopted? Once the copayment is adopted, an area needs to distinguish between curative and preventive nurse visits and be able to count the number of curative nurse visits. This may affect data quality in a way in which the number of recorded curative nurse visits changes even if the underlying utilization does not. We checked that the number of preventive nurse visits does not increase in the treated municipalities after the copayment adoption, even if the number of curative nurse visits decreases considerably. The corresponding figures are available in the

replication folder.

Results: estimates on all individuals. In the PAP, we separately estimated the effects in the bottom 40% and the top 40% of the income distribution. None of the PAP results were estimated using the whole sample. This was motivated in two ways. First, we expected that the aggregate effects would be small due to supply rigidity in the public primary care system. If copayments reduce the use of primary care services among the poor, the waiting time may be reduced, potentially attracting more patients who previously used occupational or private healthcare. Second, we have a major focus on the potential heterogeneity of effects. Still, aggregate estimates are relevant to the policymaker as well. Thus, we do now provide the main results (pre-trend plots, and stacking and CS estimates) using all sample individuals.

Results: estimates by income decile. We committed in the PAP to show the main estimates for primary care outcomes by income quintile and decile” to allow for a more flexible analysis on treatment effect heterogeneity by income. Specifically, we now estimate and show the adoption results on the number of nurse visits by income decile using stacking with balanced event-specific datasets as in main analysis. However, we do not estimate the abolition results by income deciles, as the abolition results are highly sensitive to a specific version of the parallel trends assumption.

Results: weighting municipalities uniformly. We re-estimated all the results weighting our municipality-by-income-decile observations uniformly instead of population-weights as a robustness check. Regarding the adoption analyses, we include Table A4 in the appendix. Of the abolition analyses, Figure 4 and Table A5 are presented. The rest of the result tables and figures are available in the replication codes folder.

Adoption: potentially diverging trends. The PAP discussed our plans to adjust the analysis of the copayment adoption in a hypothetical case where pre-trend plots and event study plots hint of diverging pre-trends. We observe no such pre-trend differences in nurse visits, but there may be a small trend difference in GP visits. Overall, we view the

planned analyses as appropriate and sufficient. For GP visits, we recommend to put more weight to the specifications that allow for a linear pre-trend difference or assume parallel trends only from the last pre-treatment period on.

Abolition: follow-up length. The pre-registered follow-up for the abolition analyses was 12 months. In this version, we use an 11-month follow-up since we currently have data only until the end of 5/2022. This change is not related to the nurses' strike in April which we listed in the PAP as a potential reason to shorten the follow-up time.

Abolition: referrals as an outcome. Our initial plan was to use referrals to specialist care written by public primary care as an outcome that proxies GP-assessed need for diagnosis. However, we noticed already writing the PAP that there were too many referrals missing in the early 2010s so that referrals could not be used as an outcome for the adoption analyses. Consequently, the PAP used referrals as an outcome for the abolition analyses only. In this version of the study, we do not use referrals as an outcome at all. In contrast to our plans, we only observe referrals that have led to a specialist visit or a procedure by the time the data were extracted (June 2022). As waiting times can be long for nonurgent care and our follow-up ends in May 2022, we observe only a subset of referrals we would want to observe.

Abolition: results presented in the paper. Causal inference is based on a parallel trends assumption in the DD framework. We noticed already when writing the PAP that the abolition (time-placebo) estimates are sensitive to a specific version of the PTA. We thus wrote down that we may need to adjust our analysis based on pre-trend plots when analyzing the final data. The pre-trend differences do not behave as well as expected when writing the PAP. Consequently, we are not willing to impose a specific version of the parallel trends assumption. Thus, few causal conclusions can be drawn regarding the abolition. These findings affect the set of result tables and figures we present to the reader. In presentation, we deviate in many ways from the PAP. Interested readers can compare this study and the PAP to observe the differences.

Fixing bugs. The raw data on social assistance contain more than one row for some person-year pairs. Our codes did not previously account for that.

A.3 Parameter Values for Semi-Arc Elasticities

To obtain the price of care after the copayment adoption, we multiply a per-visit copayment of 10 euros by 0.8 that is selected to present the probability that the patient actually pays the copayment. The copayment is not always charged. As Figure A2 shows, by far the most common policy is to charge it for the first three visits annually. Finally, we divide the multiplication by the total cost of the visit. At baseline, we use the estimated average cost of 35 euros in 2017 in public healthcare (Mäklin and Kokko, 2020). To get the upper bound, we use a total cost of 61 euros. This is the total cost of a 20 minute nurse appointment in a major private provider in July 2022 (66 euros), adjusted to the 2017 prices.

Regarding the quantities, we use Table A3 as the baseline. These estimates present the stacked results for all individuals in a one-year follow-up. Specifically, we average over the two specifications in Panel A (both balanced and unbalanced event-specific datasets): the pre-treatment mean is thus 1.000 annualized nurse visits and the effect estimate is -0.089 . To get the upper bound, we fix the pre-treatment mean of 1.000 annualized visits, but use a -15% reduction as the effect estimate, which is based on Figure A23 showing the CS results for all individuals. In this case, the follow-up time is on average longer and depends on the municipality.

These parameter values lead us to the following semi-arc elasticities: the baseline is -0.41 and the upper bound is -1.24 .

A.4 Additional Figures and Tables

Table A1: Adoption: Social Assistance Use.

Metric	Share receiving	Euros received
Level	2.657	105.674
Estimate	−0.040	6.279
Std. error	0.039	3.279
P-value	0.311	0.057
Change (%)	−1.487	5.942
Estimate (trends)	−0.004	3.771
Change (%) (trends)	−0.133	3.569
Events	19	6
Treated areas	174	131
All areas	264	264

Notes: The dataset is stacked and balanced. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by event-specific linear pre-trend differences in relative time. The mean of the estimated dynamic effects is reported. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. Outcomes are the share of individuals in a family receiving social assistance (in percentages) and the annual sum of received basic social assistance (in euros). With the latter outcome, we only include events that occurred on January 1st.

Table A2: Adoption: DD Comparisons, Primary Care Use, Robustness Checks.

Metric	Nurse visits		GP visits	
	Bottom 40%	Top 40%	Bottom 40%	Top 40%
A. Unbalanced data				
Level	1.386	0.605	1.444	0.698
Estimate	−0.133	−0.050	−0.053	−0.031
Std. error	0.030	0.017	0.020	0.009
P-value	0.00001	0.004	0.008	0.001
Change (%)	−9.627	−8.219	−3.649	−4.386
Estimate (trends)	−0.129	−0.054	−0.032	−0.019
Change (%) (trends)	−9.316	−9.000	−2.228	−2.794
Events	19	19	19	19
Treated areas	175	175	175	175
All areas	264	264	264	264
B. Balanced data, logarithmized outcome				
Estimate	−10.198	−8.215	−5.113	−6.059
Std. error	2.464	3.326	2.001	2.098
P-value	0.0001	0.014	0.011	0.004
Estimate (trends)	−10.280	−8.593	−4.425	−4.555
Events	17	17	17	17
Treated areas	126	126	135	135
All areas	209	209	225	225

Notes: The dataset is stacked. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by event-specific linear pre-trend differences in relative time. The mean of the estimated dynamic effects is reported. 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 equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

Table A3: Adoption: DD Comparisons, Primary Care Use, All Individuals.

Metric	Contacts per capita		Log. contacts per capita	
	Balanced	Unbalanced	Balanced	Unbalanced
A. Nurse visits				
Level	0.995	1.004		
Estimate	−0.087	−0.091	−9.180	−9.669
Std. error	0.025	0.023	2.853	2.623
P-value	0.001	0.0001	0.002	0.0003
Change (%)	−8.737	−9.114		
Estimate (trends)	−0.098	−0.096	−9.756	−9.424
Change (%) (trends)	−9.797	−9.587		
Events	17	19	17	19
Treated areas	152	175	126	175
All areas	245	264	209	264
B. GP visits				
Level	1.091	1.085		
Estimate	−0.044	−0.041	−5.355	−5.172
Std. error	0.015	0.014	1.991	1.851
P-value	0.003	0.003	0.008	0.006
Change (%)	−4.048	−3.818		
Estimate (trends)	−0.039	−0.028	−4.589	−3.847
Change (%) (trends)	−3.607	−2.610		
Events	17	19	17	19
Treated areas	152	175	135	175
All areas	245	264	225	264

Notes: This table was not pre-registered. The dataset is stacked. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by event-specific linear pre-trend differences in relative time. The mean of the estimated dynamic effects is reported. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. Outcomes are the annualized number of curative nurse and GP visits (or their logarithm). Depending on the column, event-specific datasets are either balanced or unbalanced.

Table A4: Adoption: DD Comparisons, Primary Care Use, Uniform Weighting.

Metric	Nurse visits		GP visits	
	Bottom 40%	Top 40%	Bottom 40%	Top 40%
Level	1.512	0.754	1.558	0.857
Estimate	−0.170	−0.071	−0.046	−0.036
Std. error	0.024	0.014	0.017	0.010
P-value	0	0.00000	0.006	0.001
Change (%)	−11.224	−9.425	−2.934	−4.203
Estimate (trends)	−0.144	−0.093	−0.073	−0.054
Change (%) (trends)	−9.519	−12.392	−4.698	−6.314
Events	17	17	17	17
Treated areas	152	152	152	152
All areas	245	245	245	245

Notes: The dataset is stacked and balanced. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by event-specific linear pre-trend differences in relative time. The mean of the estimated dynamic effects is reported. We weight our municipality-by-income-decile observations uniformly. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

Table A5: Abolition: DD Comparisons, All Individuals, Uniform Weighting.

Metric	Nurse Visits	GP Visits	Prescriptions
A. Annualized contacts per capita			
Level	0.957	0.946	4.567
Municipalities	230	230	238
Estimate (w/o trends)	0.002	−0.041	−0.059
Change (%)	0.160	−4.293	−1.298
Estimate (with trends)	0.217	0.041	−0.031
Change (%)	22.675	4.367	−0.681
Estimate (CS)	0.097	0.059	0.025
Change (%)	9.697	6.002	0.528
B. Logarithmized annualized contacts per capita			
Municipalities	229	230	238
Estimate (w/o trends)	3.280	−4.808	−1.423
Estimate (with trends)	21.916	2.466	0.132
Estimate (CS)	9.750	2.157	0.706

Notes: This table was not pre-registered. The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by a linear pre-trend difference in relative time. The mean of the estimated dynamic effects is reported. 3) The CS estimator with outcome regression (Callaway and Sant’Anna, 2021). Due to heterogeneity in municipality size, we weight by population size.

Table A6: Abolition: Estimates by Income.

Metric	Nurse Visits	GP Visits	Prescriptions
A. Bottom 40%			
Level	1.064	1.052	5.923
Municipalities	230	230	238
Estimate (w/o trends)	−0.016	−0.013	−0.027
Change (%)	−1.459	−1.272	−0.462
Estimate (with trends)	0.084	0.053	−0.099
Change (%)	7.933	5.034	−1.668
Estimate (CS)	0.030	−0.015	−0.073
Change (%)	2.865	−1.433	−1.240
B. Top 40%			
Level	0.520	0.509	2.705
Municipalities	230	230	238
Estimate (w/o trends)	0.040	0.004	0.020
Change (%)	7.739	0.721	0.722
Estimate (with trends)	0.066	0.026	−0.042
Change (%)	12.638	5.024	−1.536
Estimate (CS)	0.039	−0.003	0.004
Change (%)	7.594	−0.612	0.138

Notes: The following methods are used: 1) a TWFE DID model without a pre-trend difference that includes an indicator for post-treatment periods in treated municipalities and municipality and time fixed effects, 2) a TWFE DID model with a pre-trend difference that replaces the static treat x post indicator by lags of every post-treatment period for treated municipalities and by a linear pre-trend difference in relative time. The mean of the estimated dynamic effects is reported. 3) The CS estimator with outcome regression (Callaway and Sant’Anna, 2021). Due to heterogeneity in municipality size, we weight by population size. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number contacts per capita.

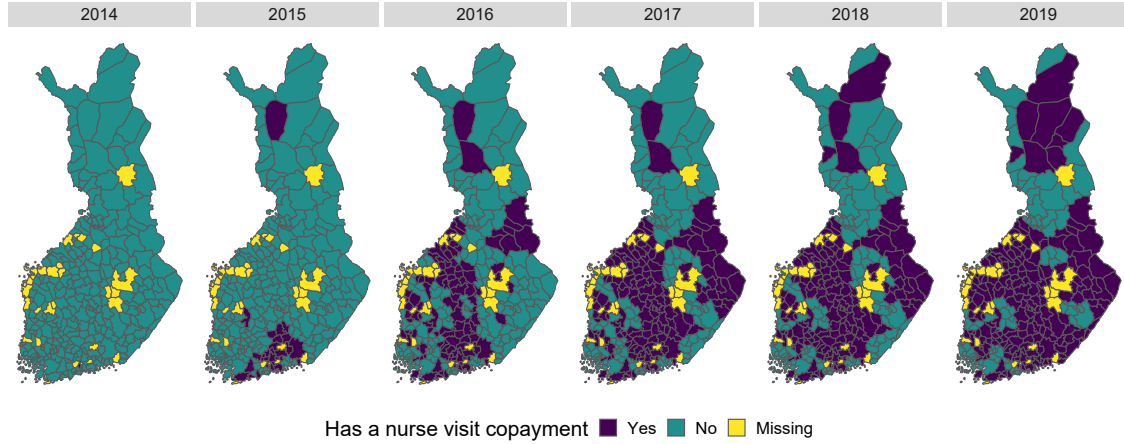


Figure A1: Staggered Adoption on Map.

Notes: The plot shows copayment policies by municipality at the end of a given year using the 2021 municipal boundaries (294 municipalities in mainland Finland). The sample contains those municipalities whose policies on copayments for curative nurse visits we observe in our data collection.

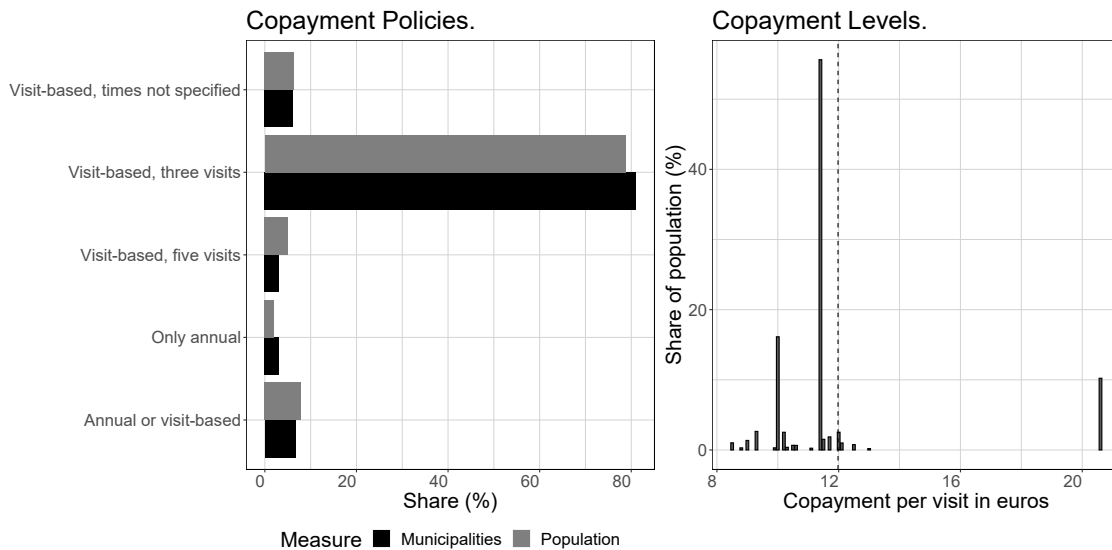


Figure A2: Copayment Levels and Policies in Summer 2021.

Notes: The plot shows how common different copayment options and levels were in Summer 2021 among the copayment municipalities of Section 5. That is, municipalities who adopted the copayment less than 12 months before the start of the study window (7/2020) are excluded - see Section A.1 for details. The population sizes are from the end of 2019. The annual copayment is often twice the amount of the per-visit copayment. Some municipalities allow the patient to choose between the annual and the per-visit copayment. In most municipalities, a per-visits copayment was charged for the first three visits annually.

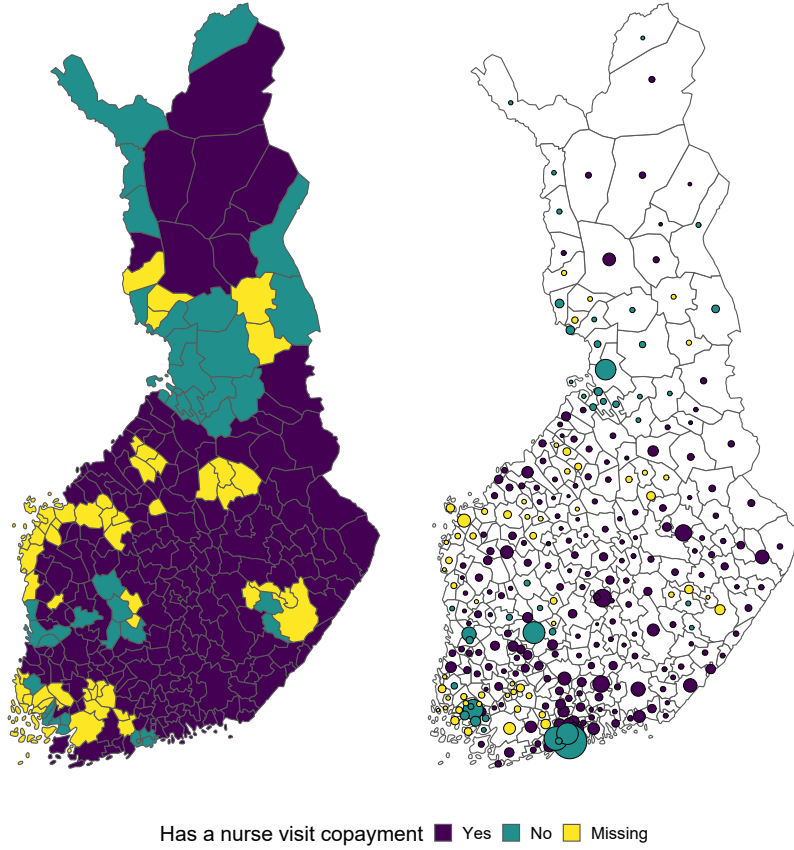


Figure A3: The Abolition on Map.

Notes: The plot shows copayment policies by municipality before the law change that abolished the nurse visit copayment in 7/2021, using the 2021 municipal boundaries (294 municipalities in mainland Finland). The group of missing municipalities contain municipalities for whom the policy is unobserved and municipalities who adopted the copayment less than 12 months before the start of the study window (7/2020), see Section A.1 for details. We assume that the effects have fully accumulated within 12 months after the adoption, which motivates the latter restriction. In the bubble plot, the size of the bubble is proportional to the 2018 population size.

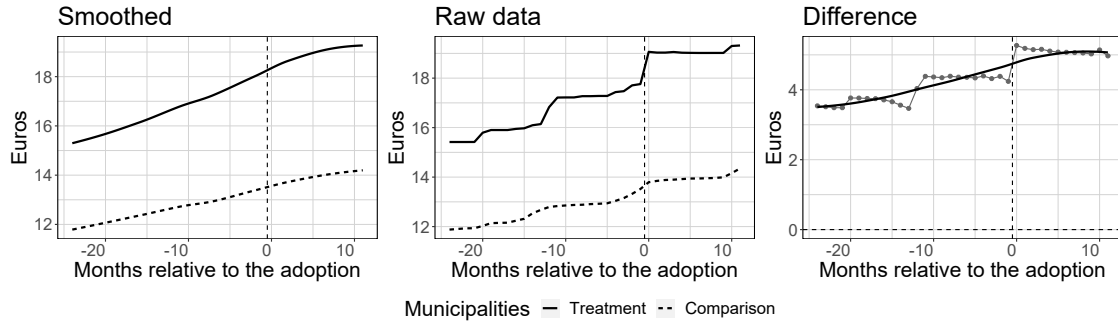


Figure A4: Adoption: Evolution in GP Visit Copayments.

Notes: The dataset is stacked and balanced. The outcome is the GP visit copayment, paid for the first three visits annually. 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.

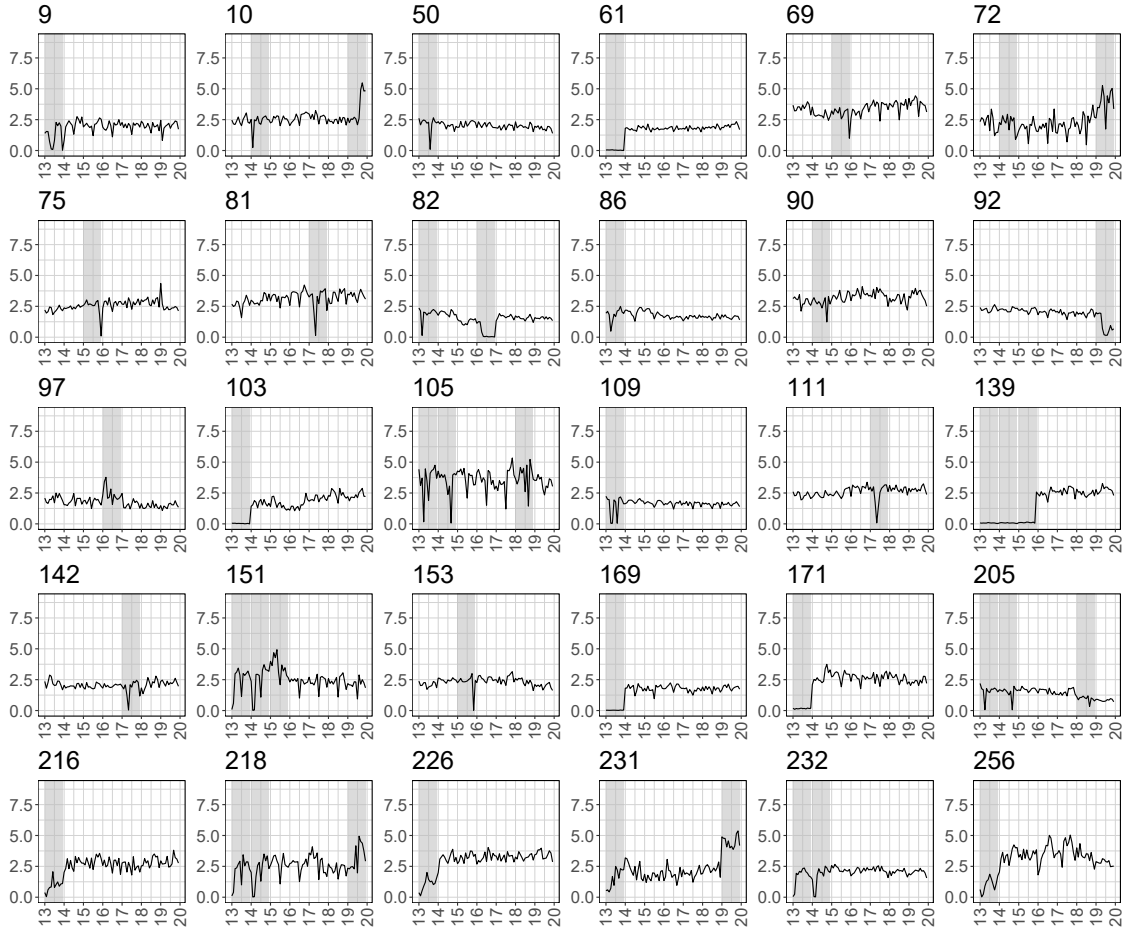


Figure A5: Adoption: Municipalities with Issues in the Primary Care Data, 1.

Notes: Using data from 1/2013 to 12/2019, we show the evolution in the annualized number of curative primary care visits (both nurse and GP visits; y axis) over time. Municipality-year observations highlighted by gray show values that are detected by our algorithm (see Section A.1) and that we view as suspiciously low or high and exclude from analyses.

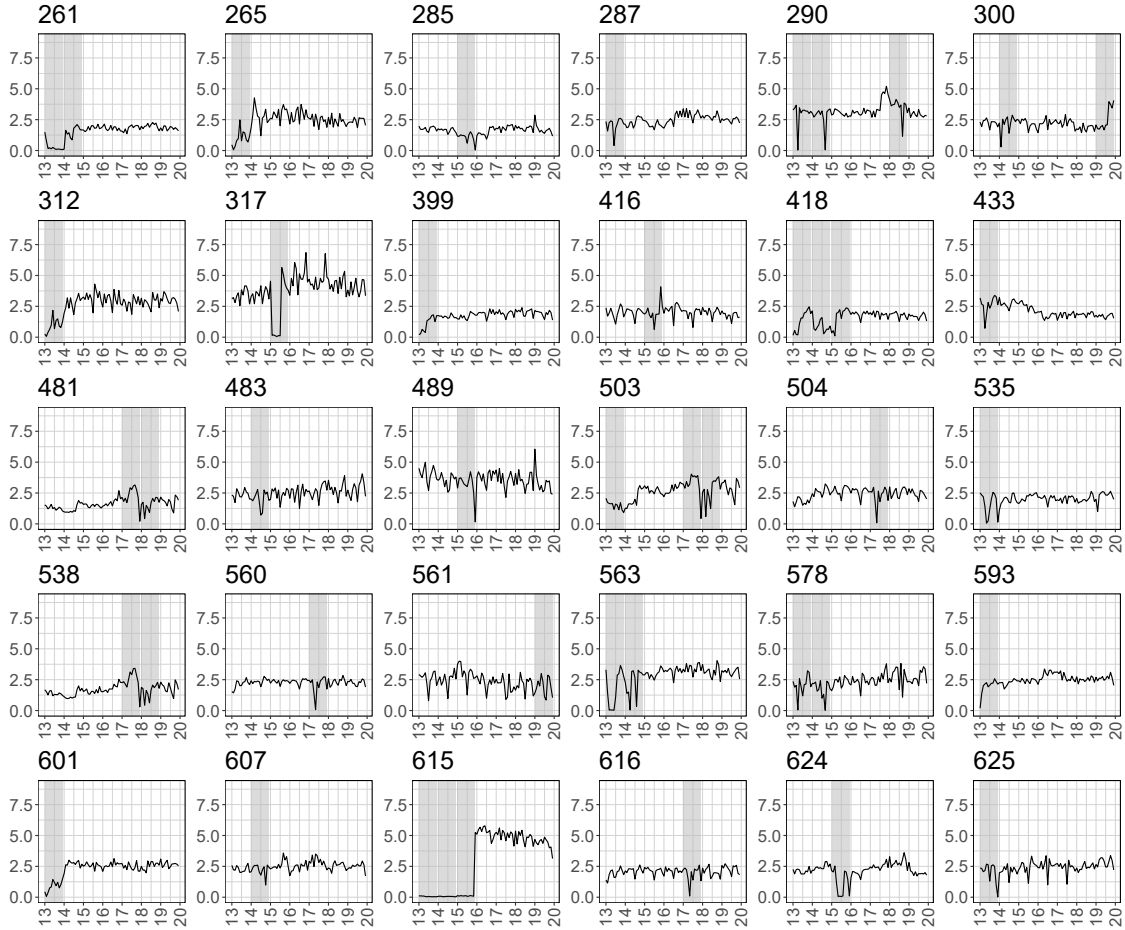


Figure A6: Adoption: Municipalities with Issues in the Primary Care Data, 2.

Notes: Using data from 1/2013 to 12/2019, we show the evolution in the annualized number of curative primary care visits (both nurse and GP visits; y axis) over time. Municipality-year observations highlighted by gray show values that are detected by our algorithm (see Section A.1) and that we view as suspiciously low or high and exclude from analyses.

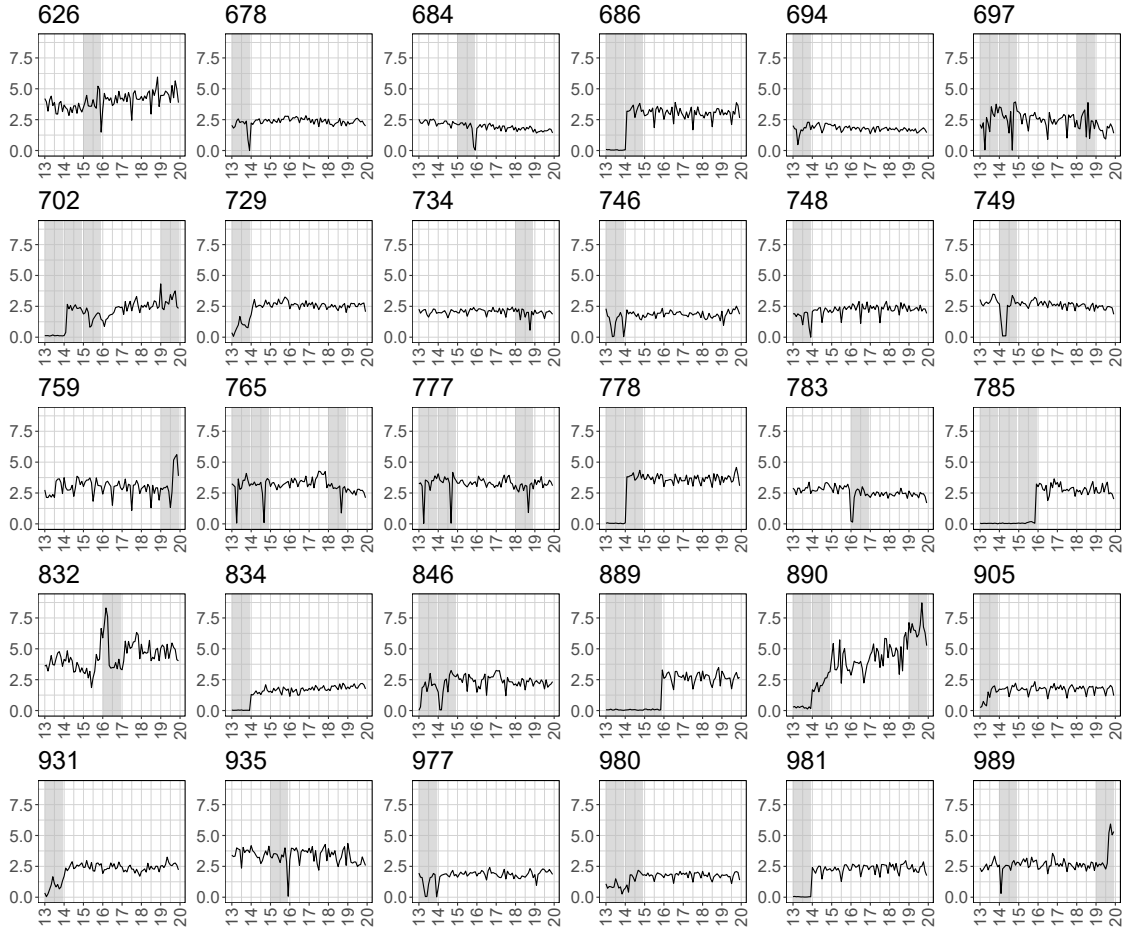


Figure A7: Adoption: Municipalities with Issues in the Primary Care Data, 3.

Notes: Using data from 1/2013 to 12/2019, we show the evolution in the annualized number of curative primary care visits (both nurse and GP visits; y axis) over time. Municipality-year observations highlighted by gray show values that are detected by our algorithm (see Section A.1) and that we view as suspiciously low or high and exclude from analyses.

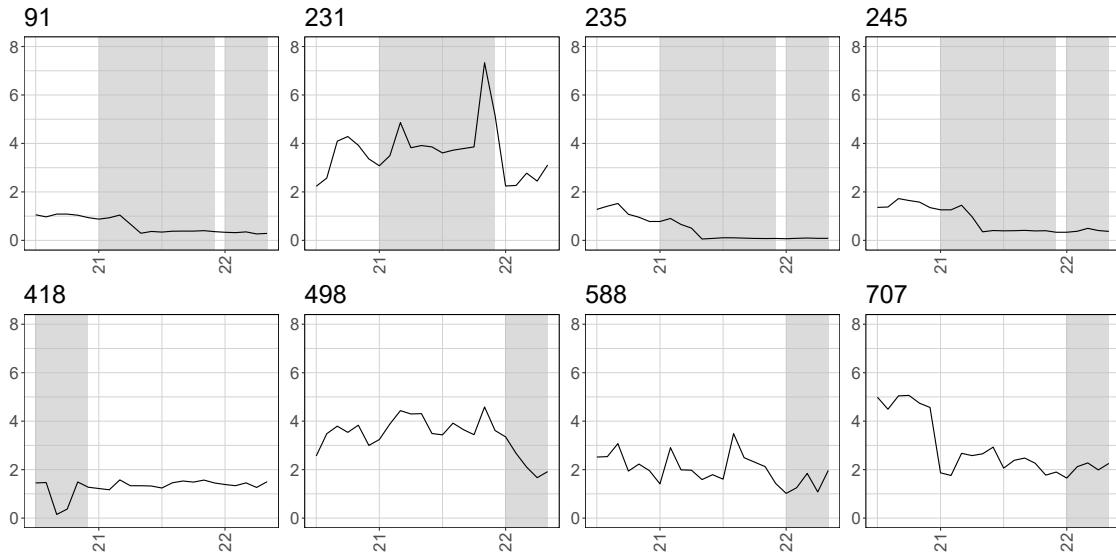


Figure A8: Abolition: Municipalities with Issues in the Primary Care Data.

Notes: Using data from 7/2020 to 5/2022, we show the evolution in the annualized number of curative primary care visits (both nurse and GP visits; y axis) over time. Municipality-year observations highlighted by gray show values that are detected by our algorithm (see Section A.1) and that we view as suspiciously low or high and exclude from analyses.

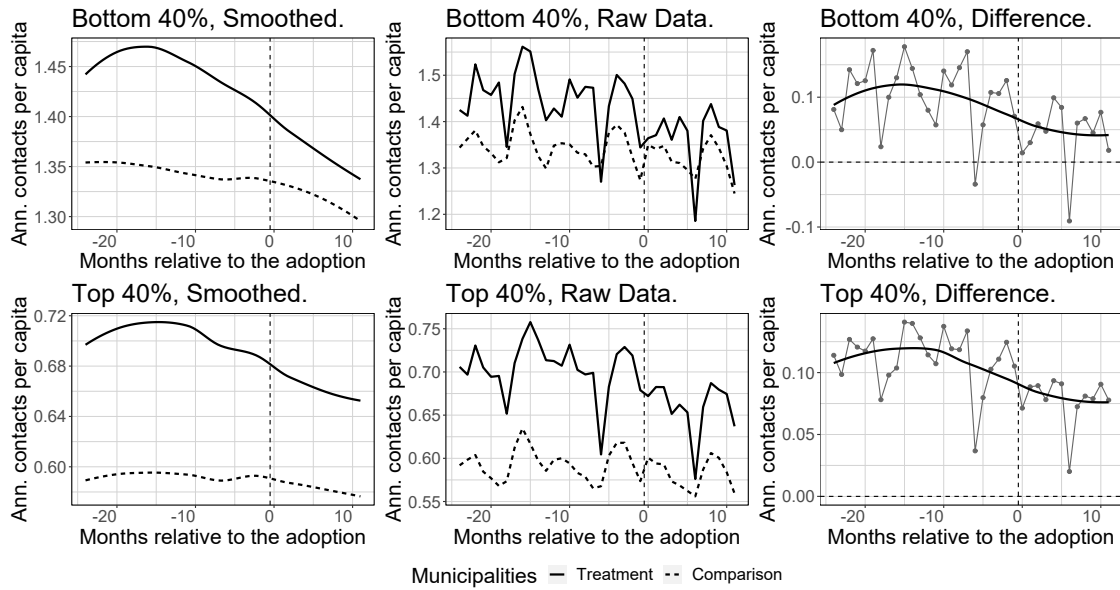


Figure A9: Adoption: Evolution in GP Visits.

Notes: The dataset is stacked and balanced. The outcome is the number of annualized curative GP visits per capita. 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 equivalised family disposable income.

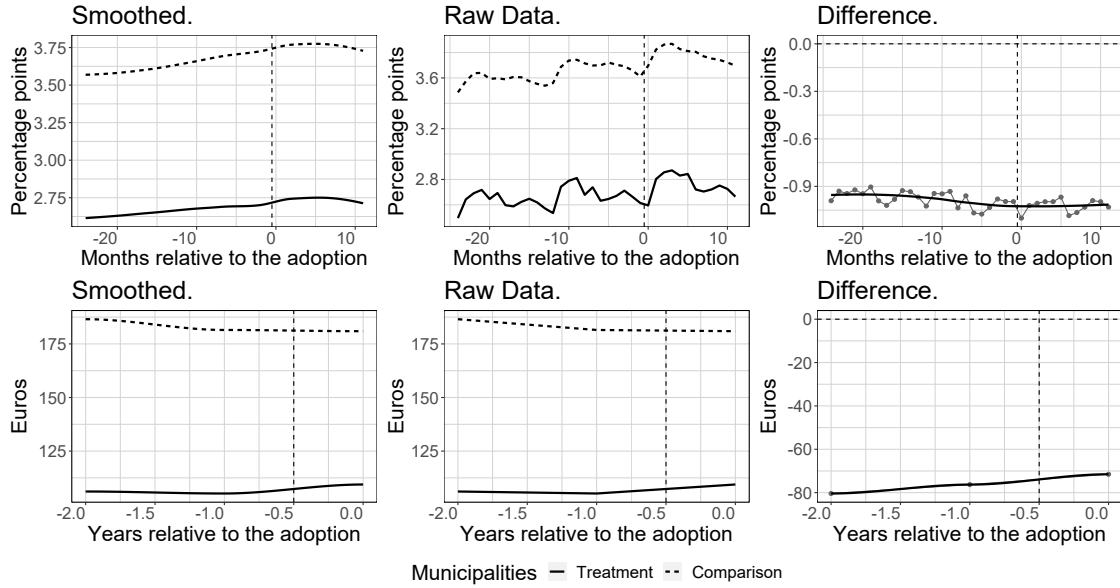


Figure A10: Adoption: Evolution in Social Assistance Recipients and Sums.

Notes: The dataset is stacked and balanced. In the top row, the outcome is the share of individuals in a family receiving social assistance (in percentages). In the bottom row, the outcome is the annual amount of social assistance received. As the latter is measured only annually, we include only those events that occurred on January 1st (and not in the middle of a year). 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.

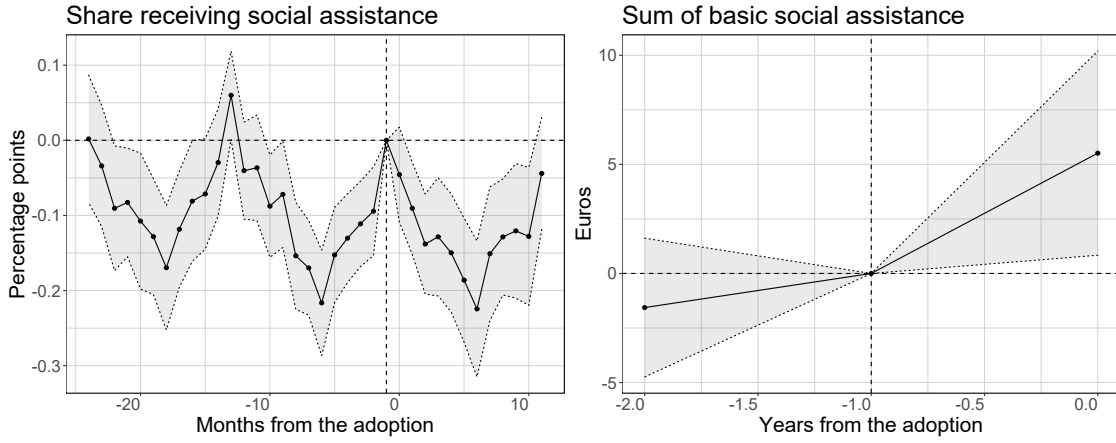


Figure A11: Adoption: Event-Study Plots on Social Assistance Use with Stacked Data.

Notes: 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. Our TWFE regression model includes a full set of treatment indicators for 24 and 12 months before and after the treatment (or two and one years before and after the policy change for the sum of received social assistance) and event-specific municipality and time fixed effects, comparing the evolution of annualized contacts per capita between treated and unexposed municipalities. The last pre-treatment period, namely $t = -1$, is omitted as a reference. With respect to the annual data on the sum of received social assistance, we include only events that occurred on January 1st. Due to heterogeneity in municipality size, we weight by population size. The standard errors are clustered by municipality.

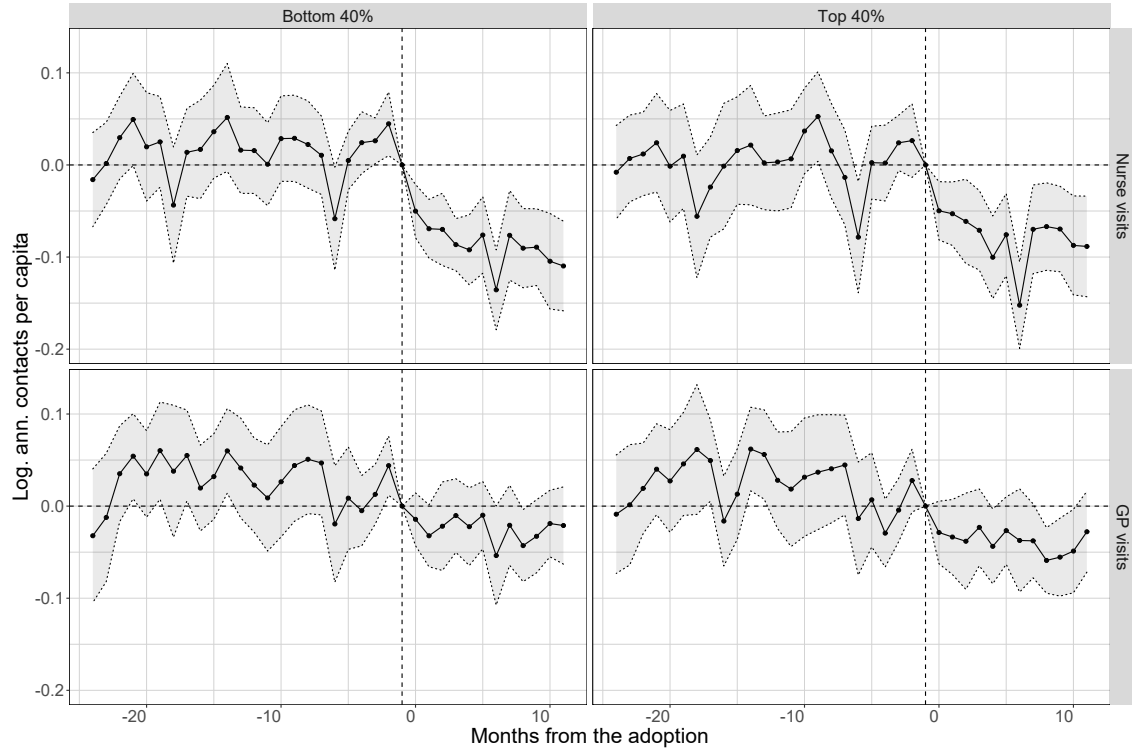


Figure A12: Adoption: Event-Study Plot on Logarithmized Primary Care Use with Stacked Data.

Notes: The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. The dataset is stacked, outcome logarithmized, and event-specific datasets balanced. Our TWFE regression model includes a full set of treatment indicators for 24 and 12 months before and after the treatment and event-specific municipality and time fixed effects, comparing the evolution of annualized contacts per capita between treated and unexposed municipalities. The last pre-treatment period, namely $t = -1$, is omitted as a reference. 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 equivalised family disposable income.

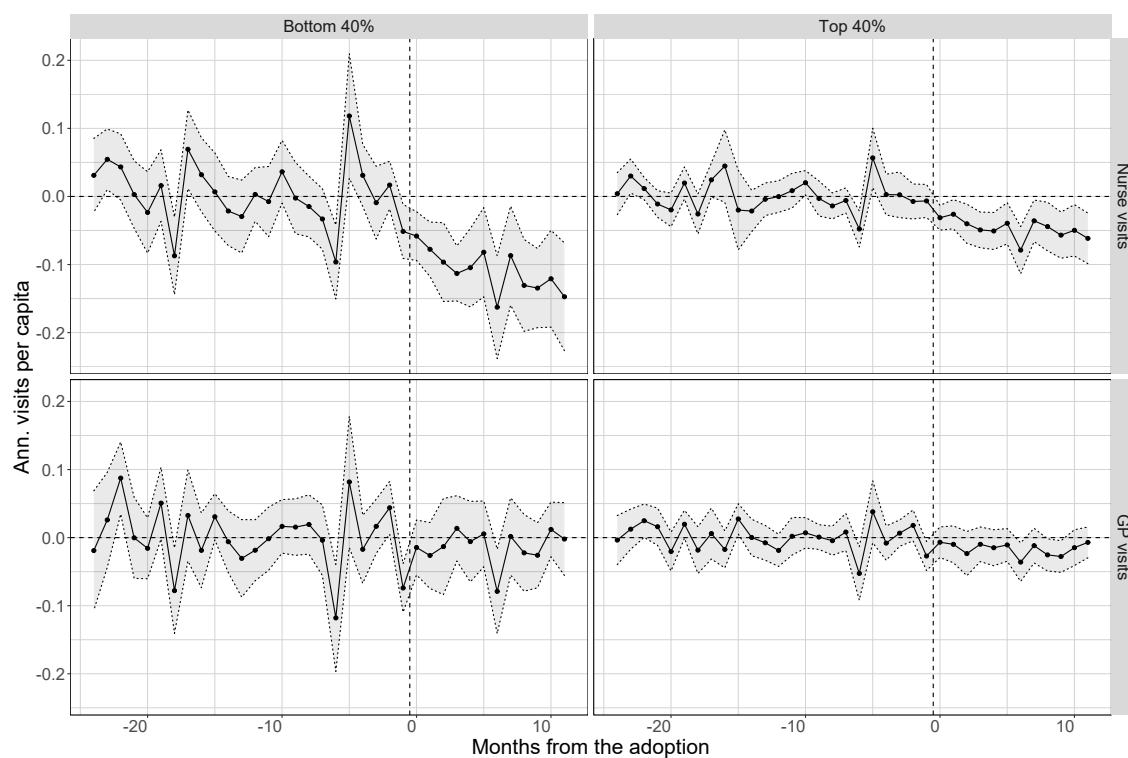


Figure A13: Adoption: Event-Study Plots Using the CS Estimator, Primary Care Use.

Notes: The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. Units that are not yet treated are used as a comparison. The dataset from 2014 to 2018 is balanced. The estimates compare the evolution of annualized contacts per capita between treated and unexposed municipalities. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

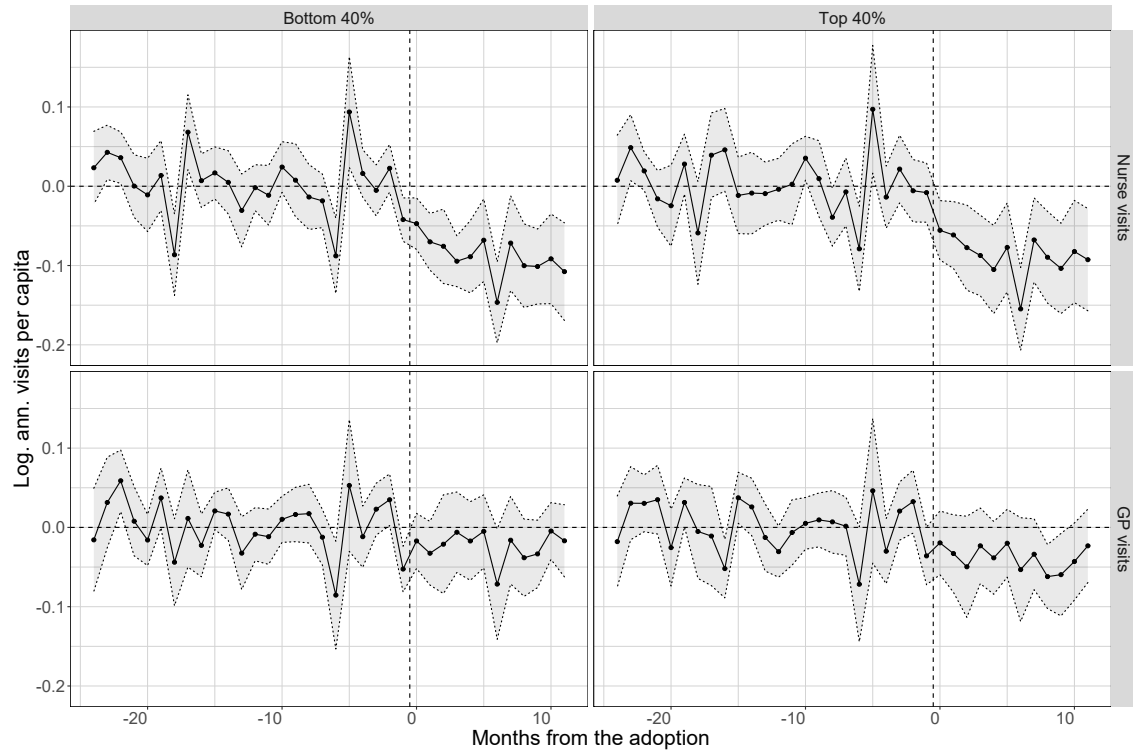


Figure A14: Adoption: Event-Study Plots Using the CS Estimator, Logarithmized Primary Care Use.

Notes: The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. Units that are not yet treated are used as a comparison. The dataset from 2014 to 2018 is balanced. The estimates compare the evolution of logarithmized annualized contacts per capita between treated and unexposed municipalities. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively.

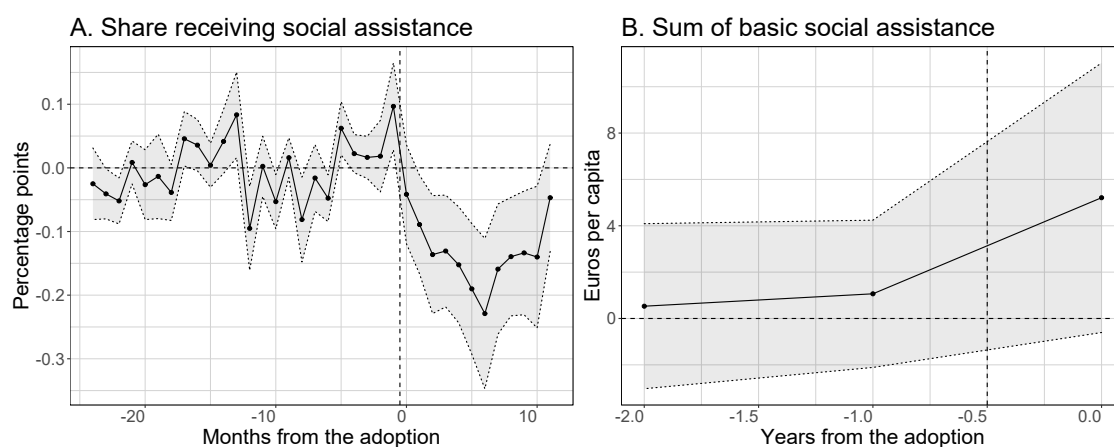


Figure A15: Adoption: Event-Study Plots Using the CS Estimator, Social Assistance Use.

Notes: The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. Units that are not yet treated are used as a comparison. The dataset is balanced. The estimates compare the evolution of outcomes between treated and unexposed municipalities. With respect to the annual data on the sum of received social assistance, we only include events that occurred on January 1st.

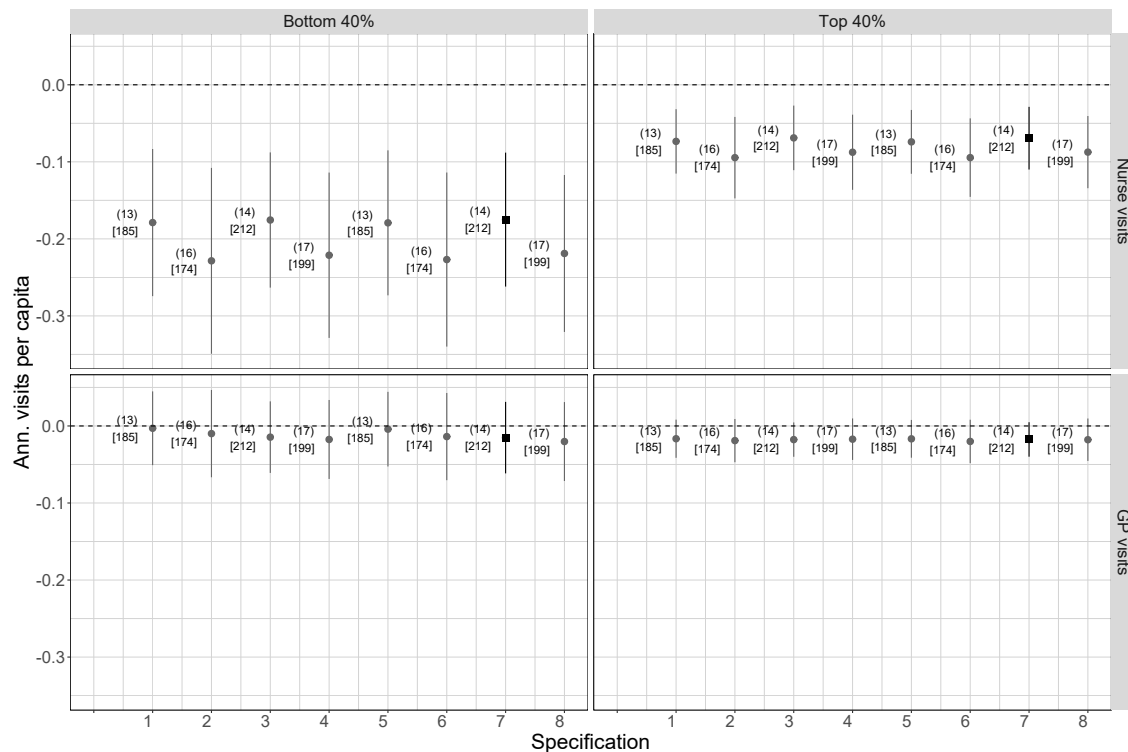


Figure A16: Adoption: the CS Estimator, Primary Care Use.

Notes: The point estimates represent static effect estimates for the treatment group. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. The estimates compare the evolution of annualized contacts per capita between treated and unexposed municipalities. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively. Comparison group consists of the never-treated areas in specifications 1-4 and the not-yet-treated areas in specifications 5-8. Study period start year is 2013 in specifications 1, 2, 5, and 6, and 2014 in specifications 3, 4, 7, and 8. Study period end year is 2018 in specifications 1, 3, 5, and 7, and 2019 in specifications 2, 4, 6, and 8. The baseline is highlighted by black.

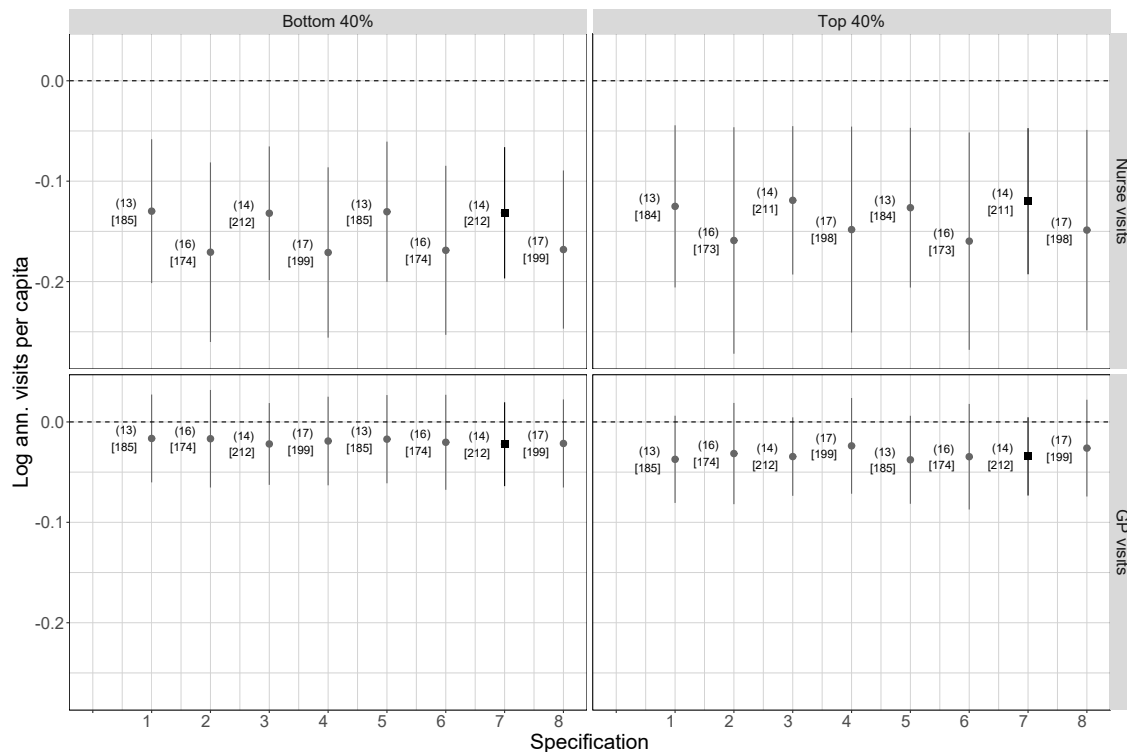


Figure A17: Adoption: the CS Estimator, Logarithmized Primary Care Use.

Notes: The point estimates represent static effect estimates for the treatment group. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. The estimates compare the evolution of logarithm of annualized contacts per capita between treated and unexposed municipalities. Bottom 40% and top 40% refer to the distribution of equivalised family disposable income. Outcomes are the annualized number of curative nurse and GP visits, respectively. Comparison group consists of the never-treated areas in specifications 1-4 and the not-yet-treated areas in specifications 5-8. Study period start year is 2013 in specifications 1, 2, 5, and 6, and 2014 in specifications 3, 4, 7, and 8. Study period end year is 2018 in specifications 1, 3, 5, and 7, and 2019 in specifications 2, 4, 6, and 8. The baseline is highlighted by black.

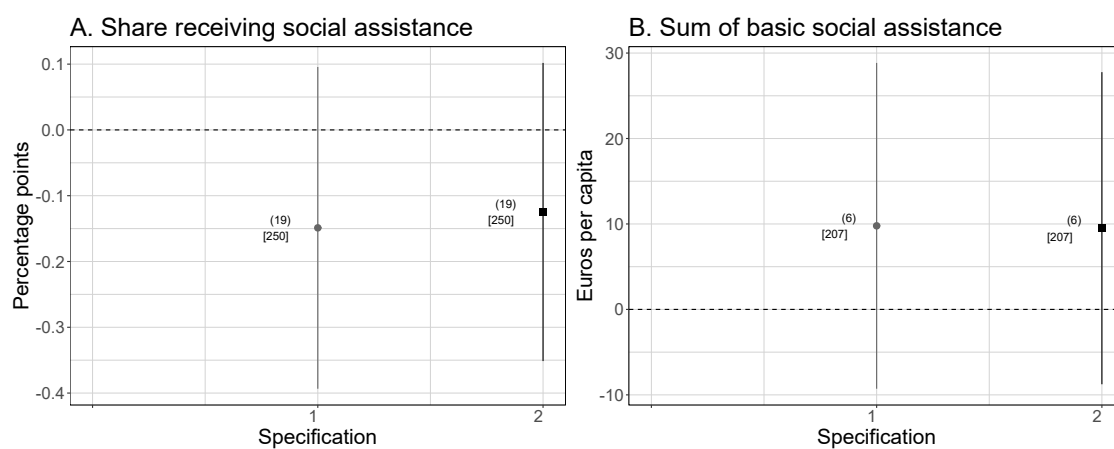


Figure A18: Adoption: the CS Estimator, Social Assistance Use.

Notes: The point estimates represent static effect estimates for the treatment group. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. The estimates compare the evolution of outcomes between treated and unexposed municipalities. Outcomes are the share of individuals in a family receiving social assistance (in percentages) and the annual sum of received social assistance. Specifications (comparison units): 1) the never-treated, and 2) the not-yet-treated. The baseline is highlighted by black.

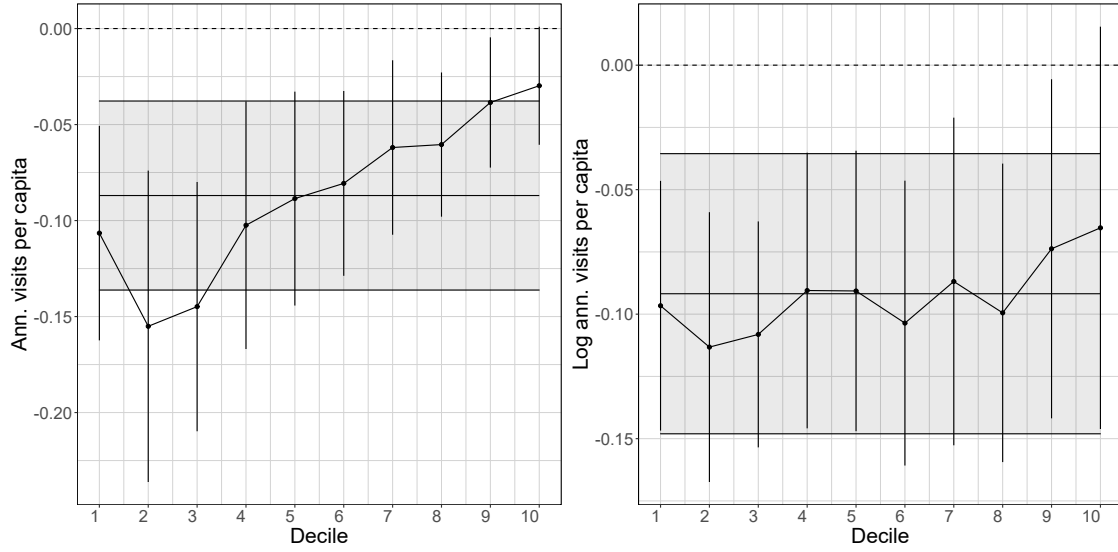


Figure A19: Adoption: Estimates on Nurse Visits by Income Decile.

Notes: This figure was not pre-registered. The dataset is stacked and balanced. We use a TWFE DID model that includes an indicator for post-treatment periods in treated municipalities and event-specific municipality and time fixed effects. Due to heterogeneity in municipality size, we weight by population size. Standard errors are clustered by municipality. We use the distribution of equivalised 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.

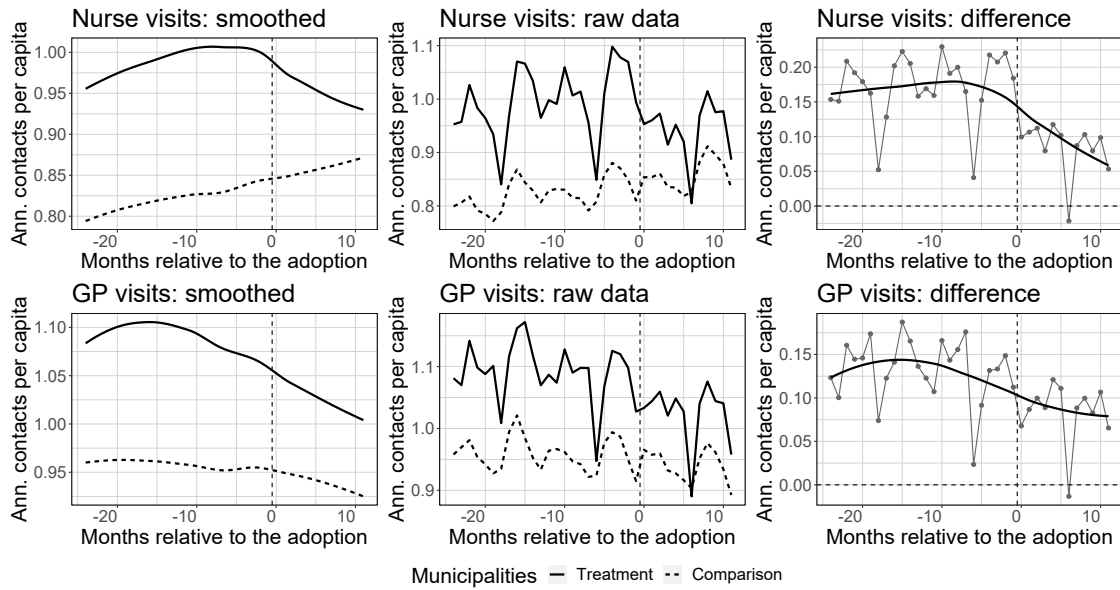


Figure A20: Adoption: Evolution in Outcomes, All Individuals.

Notes: This figure was not pre-registered. The dataset is stacked, and event-specific datasets balanced. The outcomes are the number of annualized curative nurse visits and GP visits per capita. 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.

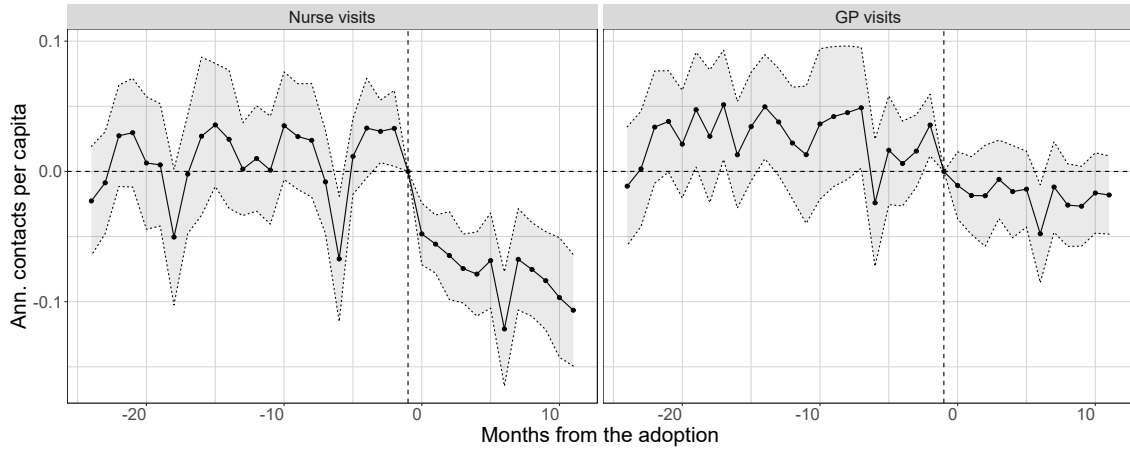


Figure A21: Adoption: Stacked Event-Study Plot on Primary Care Use, All Individuals.

Notes: This figure was not pre-registered. 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. Our TWFE regression model includes a full set of treatment indicators for 24 and 12 months before and after the treatment and event-specific municipality and time fixed effects, comparing the evolution of annualized contacts per capita between treated and unexposed municipalities. The last pre-treatment period, namely $t = -1$, is omitted as a reference. Due to heterogeneity in municipality size, we weight by population size. The standard errors are clustered by municipality.

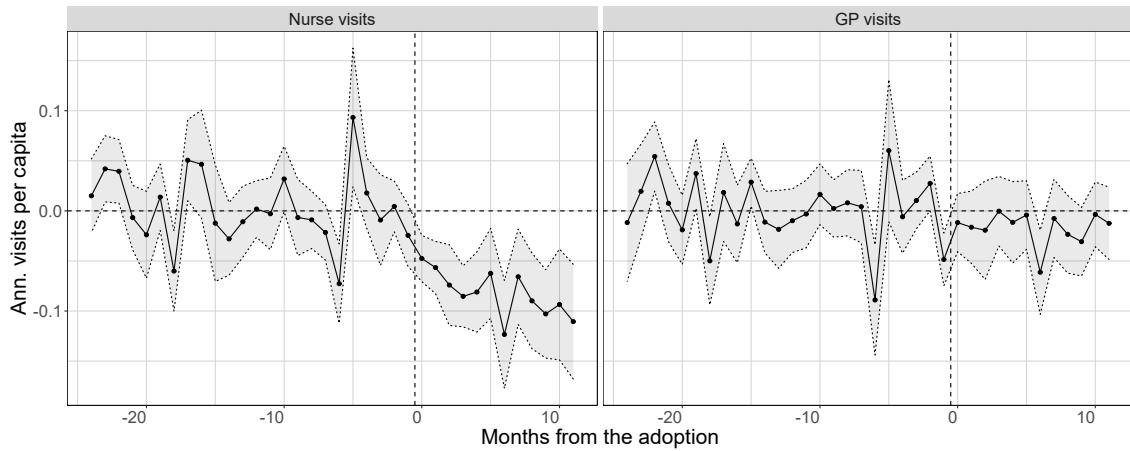


Figure A22: Adoption: the CS Event-Study Plots, Primary Care Use, All Individuals.

Notes: This figure was not pre-registered. The point estimates represent effect estimates for the treatment group as a function of time relative to the copayment adoption. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. Units that are not yet treated are used as a comparison. The dataset from 2014 to 2018 is balanced. The estimates compare the evolution of annualized contacts per capita between treated and unexposed municipalities. Outcomes are the annualized number of curative nurse and GP visits.

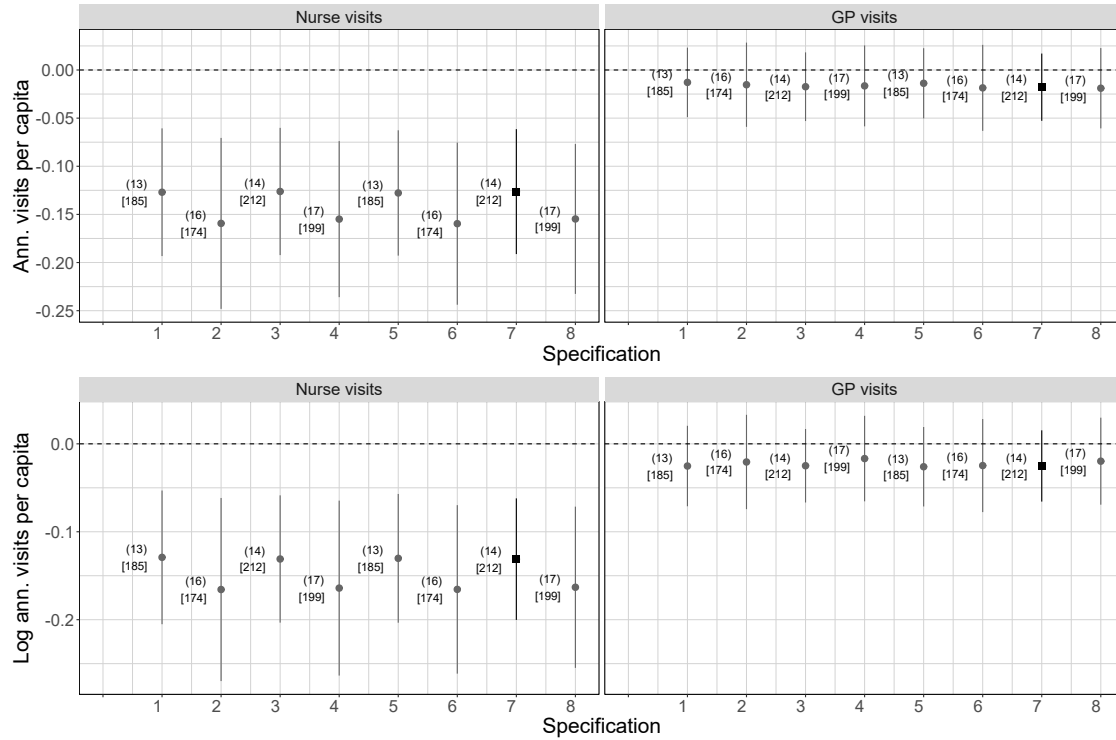


Figure A23: Adoption: the CS Estimator, Primary Care Use, All Individuals.

Notes: This figure was not pre-registered. The point estimates represent static effect estimates for the treatment group. We use the CS estimator (Callaway and Sant'Anna, 2021) with outcome regression, weight by population size, and cluster standard errors by municipality. The estimates compare the evolution of annualized contacts per capita (or its logarithm) between treated and unexposed municipalities. Outcomes are the annualized number of curative nurse and GP visits, respectively. Comparison group consists of the never-treated areas in specifications 1-4 and the not-yet-treated areas in specifications 5-8. Study period start year is 2013 in specifications 1, 2, 5, and 6, and 2014 in specifications 3, 4, 7, and 8. Study period end year is 2018 in specifications 1, 3, 5, and 7, and 2019 in specifications 2, 4, 6, and 8. The baseline is highlighted by black.

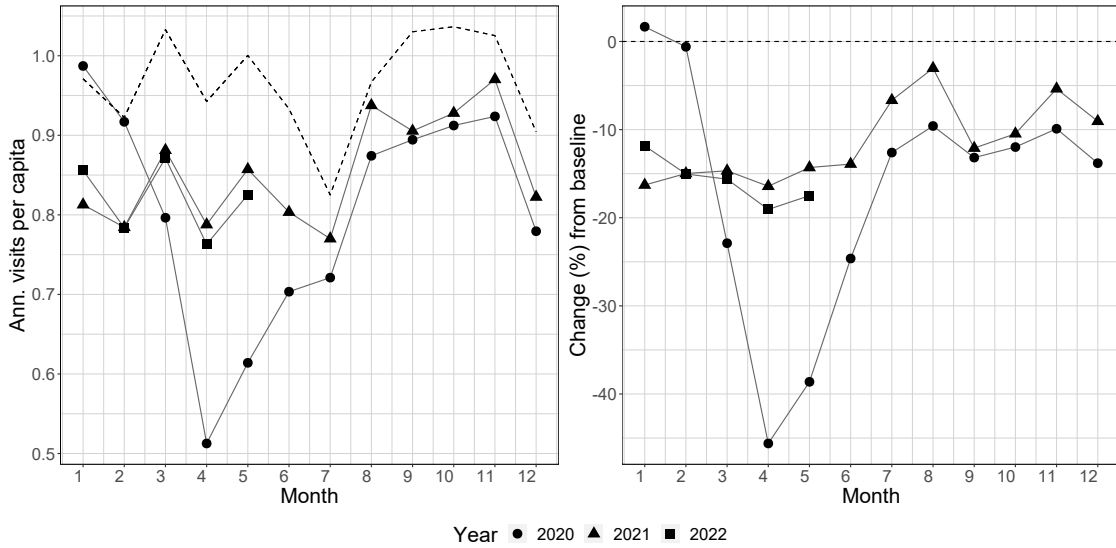


Figure A24: The Evolution of Nurse Visits during the Covid-19 Pandemic.

Notes: The figure shows the evolution of curative nurse visits in primary care relative to a baseline (monthly means from 2018-2019). On the left, the baseline is depicted by the dashed line. On the right, we show a change relative to the baseline. The data are from individuals aged 25 or more residing in Mainland Finland. Four municipalities that adopted the Apotti EHR system either in 2019 or 2021 are excluded due to missing values in the curative/preventive categorical. To filter out supply-side noise due to the changing number of workdays in a given month, we divide our measure for healthcare use by the number of workdays in a given month and then multiply it by the mean number of monthly workdays over the years.

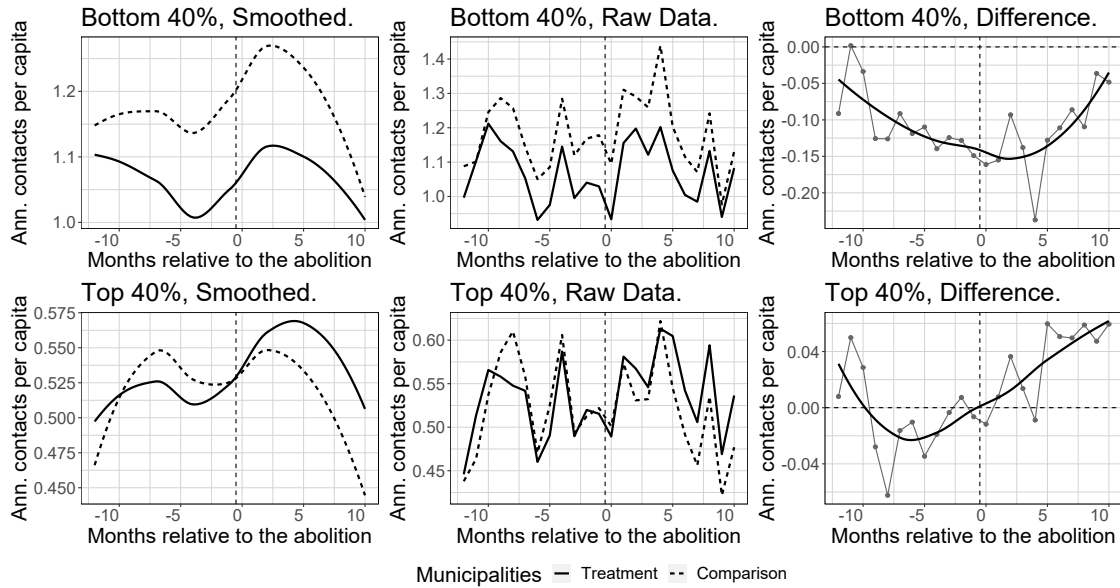


Figure A25: Abolition: Evolution in Nurse Visits.

Notes: The outcome is the number of annualized curative nurse visits per capita. 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 equalised family disposable income.

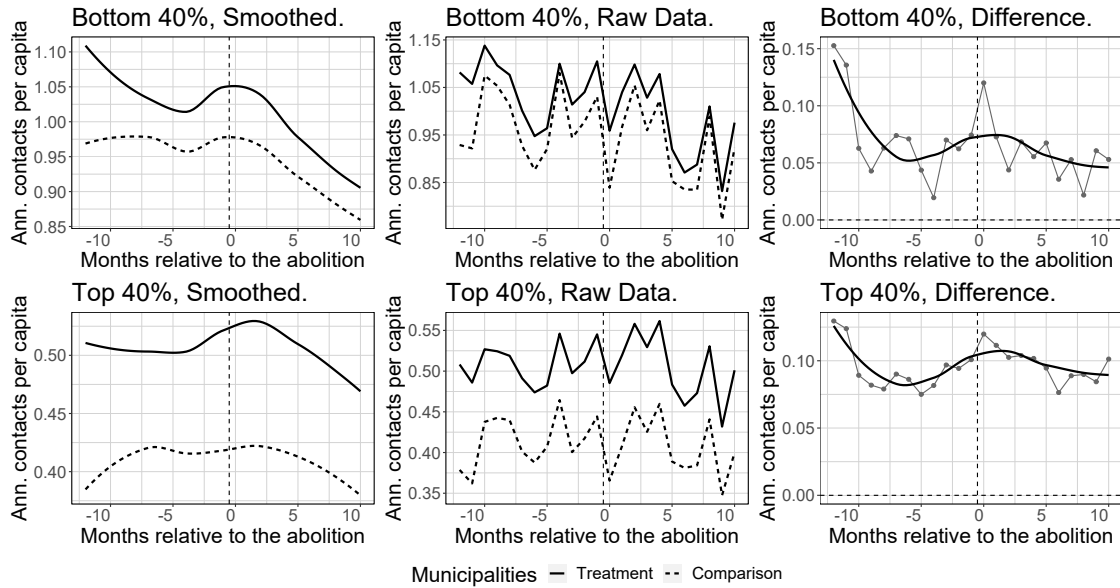


Figure A26: Abolition: Evolution in GP Visits.

Notes: The outcome is the number of annualized curative GP visits per capita. 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 equalised family disposable income.

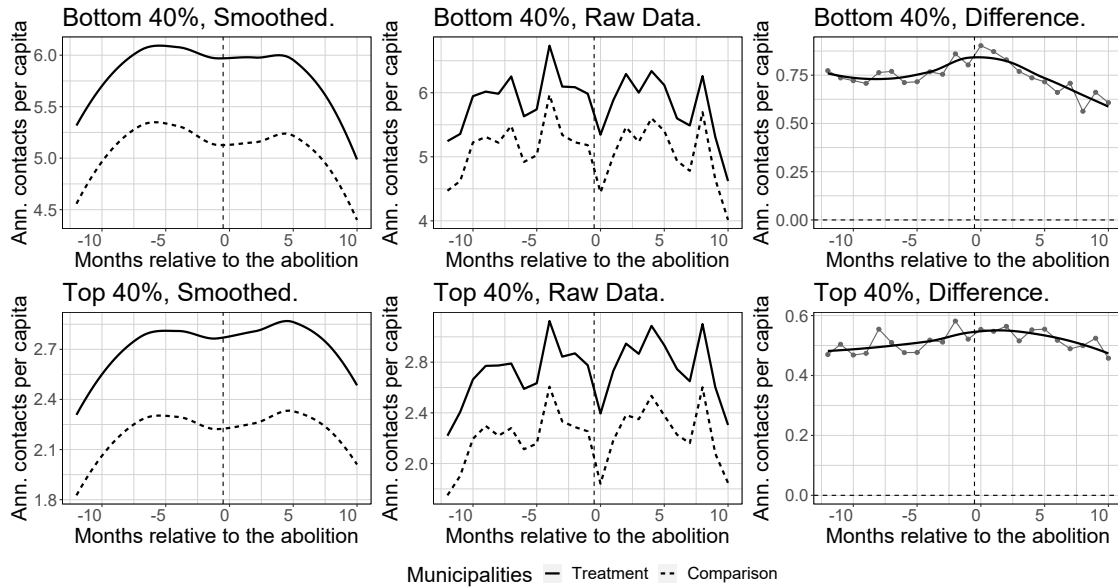


Figure A27: Abolition: Evolution Prescriptions.

Notes: The outcome is the number of annualized prescriptions. 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 equalised family disposable income.

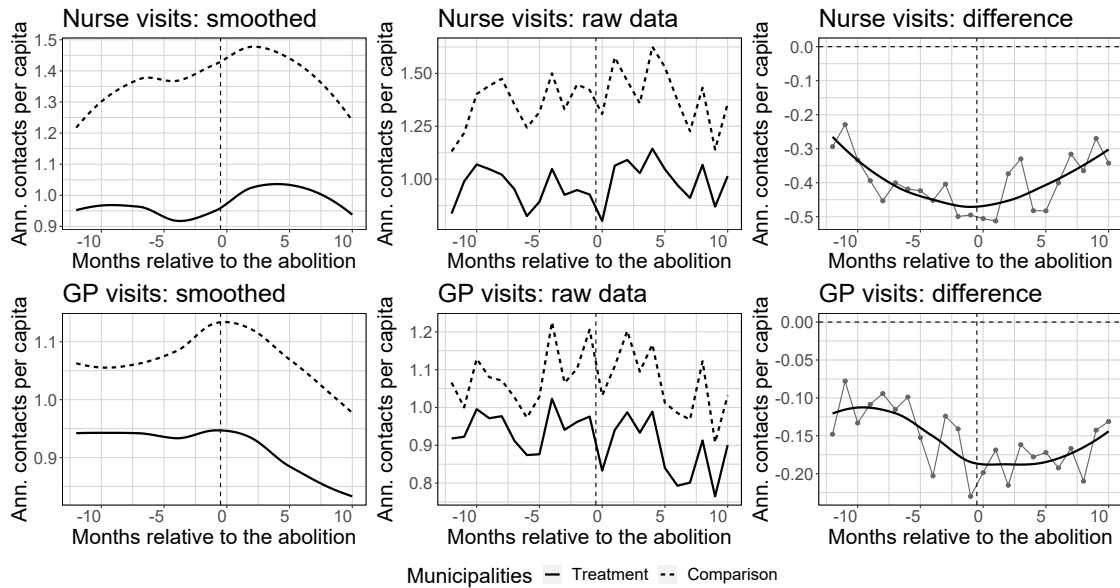


Figure A28: Abolition: Evolution in Outcomes, All Individuals, Uniform Weighting.

Notes: This figure was not pre-registered. Municipalities are weighted uniformly instead of population weights. The outcomes are the number of annualized curative nurse visits and GP visits per capita. 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.