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

November 2023[1*]

Abstract

Insurance coverage increases health care consumption, but less is known whether moderate copayments affect adults' primary care utilization in a system characterized by gatekeeping. We analyze whether abolishing a 14-euro copayment for visits to general practitioners (GP) in Helsinki, the capital of Finland, increased the number of GP visits among adults and especially among low-income individuals. Using a difference-in-differences (DD) design and combining several administrative registers from 2011 to 2014, we find that the abolition is associated with only a small increase in GP visits (+0.04 visits annually, or +4.4%, for all adults). The increase is driven by low-income adults (+0.06 visits, or +4.5%, at the bottom 40%). Although the point estimates are rather robustly positive, the conclusions regarding the statistical significance are sensitive to how we account for clustering in a setting characterized by only one treated cluster and a finite number of comparison clusters.

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

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

[1*] **Acknowledgements:** We thank Mikko Peltola, Heikki Kauppi, and THL for support and Liisa T. Laine, Tuomas Markkula, Mikko Nurminen, Jukka Pirttilä, Lauri Sääksvuori, Jussi Tervola, and Maria Vaalavuo 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 Yrjö Jahnsson Foundation (research grant No. 20197209) and by the Finnish Ministry of Social Affairs and Health. **Replication codes:** https://github.com/tapiohaa/ASMA2. **Working paper versions:** https://osf.io/8q5b2/.

1. Introduction

Out-of-pocket costs reduce health care utilization (Einav and Finkelstein, 2018). Most of the literature is based on variation in insurance coverage, but studies have also exploited variation in cost-sharing schemes, such as coinsurance rates, deductibles, and copayments. Copayments have potentially useful features as a policy instrument. They are transparent and understandable for patients, easy to bill, and they can yield fiscal revenue. Their level is usually low, which mitigates financial risks to patients but generates less revenue. Copayments are widely utilized in tax-funded public healthcare systems, including the Nordic countries.[2] A central policy concern is that fixed copayments may create a greater barrier to access for low-income patients, who are on average sicker and need more services.

We examine whether low copayments affect primary care general practitioner (GP) use in a system where patients are triaged by nurses at the entry (gatekeeping) and waiting times vary for non-urgent visits. Our interest is in the potential heterogeneity of the effects by income level. The capital Helsinki, Finland's most populous city, abolished its GP visit copayment of 14 euros in January 2013 to reduce health inequality. Using a difference-in-differences (DD) design and the fact that other municipalities continued to charge the copayment, we examine the effects of the abolition on public primary care GP use (our primary outcome), emergency department (ED) visits, specialist consultations, and social assistance use (a last-resort benefit that also covers healthcare costs) based on administrative register data from 2011 to 2014. The synthetic control method complements the DD results.

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

[2] Denmark is currently the only Nordic country that does not charge a copayment for primary care general practitioner (GP) visits. In Sweden, some regions do not charge the GP visit copayment.

quantitative magnitude of the estimates is modest or small. Consistent with the small effects for our primary outcome, we do not find significant effects for our secondary outcomes.[3]

Statistical inference is challenging in our setting due to the availability of only one treated cluster and a finite number of comparison clusters. Our inference results are inconclusive without strict and seemingly implausible assumptions. Although the point estimates are rather robustly positive, the conclusions on statistical significance are sensitive to how we account for clustering. For instance, the p-value for the estimate for all individuals (+0.04 visits annually or +4.4%) is 0.00, 0.01, 0.07, or 0.22, depending on the method (Table 1).

Previously, five studies have examined the impacts of GP visit copayments of 10 to 18 euros on GP use in the Nordic countries (Beck Olsen and Melberg, 2018; Haaga et al., 2023a; Johansson et al., 2019; Magnussen Landsem and Magnussen, 2018; Nilsson and Paul, 2018). They all focus on children or adolescents and exploit the fact that adolescents under a given age are exempted from copayments. Four of the studies employ an age-based regression discontinuity (RD) design, while Beck Olsen and Melberg (2018) construct a synthetic control for individuals aged 12 to 15 from other age cohorts. In the RD studies, the number of GP visits decreases by 4–12% after the copayment is charged. Beck Olsen and Melberg (2018) report large estimates for free care: +22% for women and +14% for men.

Our study relates to these studies, but we examine the effects for the whole adult population and not only for those aged 12 to 15 or for adolescents at a specific birthday. Moreover, we study the impacts of a copayment abolition (i.e., a policy change) instead of individuals aging out of an exemption. As age-based policy rules are foreseeable, such schemes may be more sensitive to the strategic behavior of individuals who decide when to contact primary care. Besides the above-mentioned studies focusing on adolescents, Jakobsson and Svensson (2016) report that a 33% increase (circa 5 euros) in the GP visit copayment did not affect the total number of GP visits in an 8-month follow-up in Sweden. While their exposure is a change in the intensity of the copayment, we examine a copayment abolition. The distinction between an adoption and an abolition is also relevant, as the effects of increased and decreased out-of-pocket costs may not be symmetric (Hayen et al., 2021; Iizuka and Shigeoka, 2021; Remmerswaal et al., 2019).

[3] However, we observe an unexpected reduction in dentist visits, a potential placebo outcome, of similar magnitude to the increase in GP visits.

Furthermore, we contribute to the understanding of whether low-income individuals respond more strongly to changes in cost-sharing. Previously, Nilsson and Paul (2018) and Johansson et al. (2019) have found that patients at the lower end of the income distribution are more sensitive to copayments than high-income individuals in both absolute (the number of visits) and relative (compared to baseline utilization) terms, while Haaga et al. (2023b) report heterogeneity but only in absolute terms. In Haaga et al. (2023a), the evidence for heterogeneity by income is weaker.

Section 2 introduces the context and Section 3 the data. Section 4 presents our empirical approach and Section 5 the results. Section 6 concludes. Our Appendix contains additional figures and tables.

2. Institutional Background

Primary care for Finnish adults is provided by three sectors, targeting different subpopulations and differing by the level of out-of-pocket costs, waiting times and the strictness of gatekeeping. Public primary care charges copayments of about 14 euros for GP visits. Waiting times for non-urgent conditions can be long. There is also gatekeeping at the point of entry and in accessing specialists. Patients are triaged by nurses when they call or visit their health station, determined by their address. Pensioners, low-income individuals, and the unemployed disproportionately rely on public primary care. In contrast, employed people often prefer occupational healthcare, which is free of charge, or private clinics over public primary care due to faster access and less or no gatekeeping.

Municipalities organize public primary care on their own or in cooperation with other municipalities. The services are financed by state transfers, municipal taxes, copayments (a minor share), and borrowing. The state sets the upper limits for copayments and determines which services and groups are nationally exempted. Helsinki charged a copayment of 13.80 euros for the first three GP visits annually before abolishing it in January 2013. Its health care services committee assumed (6 March 2012) that the copayment abolition would reduce health inequality.[4] The committee also noted that many patients of public primary care are pensioners or unemployed and thus not entitled to occupational healthcare that is free of charge. Minors, war veterans and war invalids were exempted already before the abolition, and copayments were not charged for

[4] The mechanism was not stated explicitly, but we assume the aim was to increase primary care use at the lower end of the income distribution.

preventive services, such as health checks. Our comparison municipalities had a similar per-visit copayment or an annual copayment which was twice the amount of the per-visit copayment. Copayments continued to be charged in other services, such as dentist visits in public primary care and ED visits and specialist consultations at hospitals.

3. Data

We combine Finnish national-level administrative registers to construct five outcomes[5]: GP visits in public primary care (the primary outcome), ED visits and specialist consultations at hospitals, an indicator of belonging to a family where someone received social assistance[6] (the secondary outcomes), and dentist visits in public primary care (a plausible placebo). We also use publicly available data on municipal copayment policies (Finnish Institute for Health and Welfare) and municipal characteristics (Statistics Finland, Sotkanet, and Social Insurance Institution of Finland).

Our analysis dataset is a person-month panel from 2011–2014. Individuals and their visits are linked to policies using the municipality of residence.[7] We first exclude small municipalities with less than 30,000 residents in 2012, with 36 municipalities remaining.[8] We exclude municipalities with missing copayment policy or which changed from a per-visit copayment to an annual copayment (or *vice versa*) in 2013–2014. One municipality (Espoo) is excluded, because it adopted exemptions for several low-income groups in 8/2011. These restrictions lead to a sample of 28 municipalities. Our sample individuals are those who were 25 years or older at the end of 2011 and who were observed to reside in the same sample municipality throughout 2011–2014. This leaves us with 380,000 people in Helsinki and 1.35 million people in the 27 comparison areas.

An individual may have had multiple contacts on a given day, but we treat these as one visit. We only include GP and dentist visits from Monday to Friday to reduce the share of acute visits outside of normal office hours, which have a different copayment. Specialist consultations do not include repeated visits to treat the same health problem.

---

[5] Public primary care contacts are from the Register of Primary Health Care Visits, specialized healthcare contacts from the Care Register for Health Care, and social assistance data from the Register of Social Assistance. All three registers are administered by the Finnish Institute for Health and Welfare (THL). Socioeconomic data are from Statistics Finland's FOLK modules (basic, family, and income).

[6] Social assistance is a last-resort benefit for those with low income and little wealth to cover basic living expenses.

[7] Only since 2014 have individuals had an option to choose another public primary care provider than the one determined by municipality of residence. However, these changes have been rare.

[8] Helsinki is by far the largest municipality in Finland with 600,000 residents in 2012. We exclude small and rural municipalities where primary care demand and supply may differ considerably from Helsinki.

The final sample sizes vary across outcomes as some municipalities are excluded for data quality reasons. In a DD design, missing visits correlated with the treatment would bias the results. Two types of data quality concerns are noteworthy here. First, the register on primary care contacts started in its current form in 2011. Not all areas transferred high-quality data to the register at the beginning.[9] Second, changes in municipal IT systems may have led to drastic but brief reductions in observed contacts.

We detect and exclude municipalities with data quality concerns as follows: 1) compute a distribution of mean contacts by permutationally excluding every combination of four consecutive months, 2) mark an observation as invalid if its value is less than X% of the largest observed mean (July is not considered due to holidays), and 3) exclude municipalities with invalid observations. The threshold X varies by outcome based on what we think works well in detecting outliers. We use 50% for GP visits (19 comparison municipalities remain), 30% for ED visits (23), 40% for specialist consultations (24), 55% for dentist visits (17), and 40% for the social assistance indicator (27). We show the evolution of outcomes and highlight the detected municipality-year observations in Figure A1 for GP visits, Figure A2 for ED visits, Figure A3 for specialist consultations, Figure A4 for dentist visits, and Figure A5 for social assistance use. The sample municipalities for the main outcomes are illustrated in Figure A6. The background statistics for Helsinki, the 19 comparisons (for the GP visits outcome), and the remaining municipalities are in Table A1.

4. Empirical Approach

We use a difference-in-differences (DD) design, comparing individuals in Helsinki to individuals living in comparison municipalities that continued to charge the GP visit copayment. The key identifying assumption is the parallel trends assumption (PTA): the outcomes of individuals in the treated and comparison municipalities would have evolved similarly in the absence of treatment. Figure 1 shows an increasing pre-trend in GP visits in Helsinki relative to the comparison units. The same pattern also exists separately at the bottom 40% and the top 40% of the income distribution (Figure A7). Helsinki is by far the largest and most urban municipality and in many ways an outlier (Table A1), so the trend difference is not a complete surprise.

We make a modified PTA to account for diverging pre-trends: we assume that the PTA holds after subtracting a linear pre-trend difference from the data (detrending). That is, there should

[9] For instance, the number of GP visits was suspiciously low in Rovaniemi (698) in 2011 in Figure A1.

be no time-variant confounders once a linear pre-trend difference is eliminated. Specifically, we fit a linear trend difference in time with OLS between Helsinki and the comparisons using only pre-treatment data. The estimated trend difference is then subtracted from the outcomes to construct a transformed outcome variable.[10] Our baseline assumption is that the trend difference does not change in post-treatment periods. Still, we examine the sensitivity of the results to the trend difference slowing down or accelerating by changing the slope of the estimated pre-trend difference for the post-treatment periods.

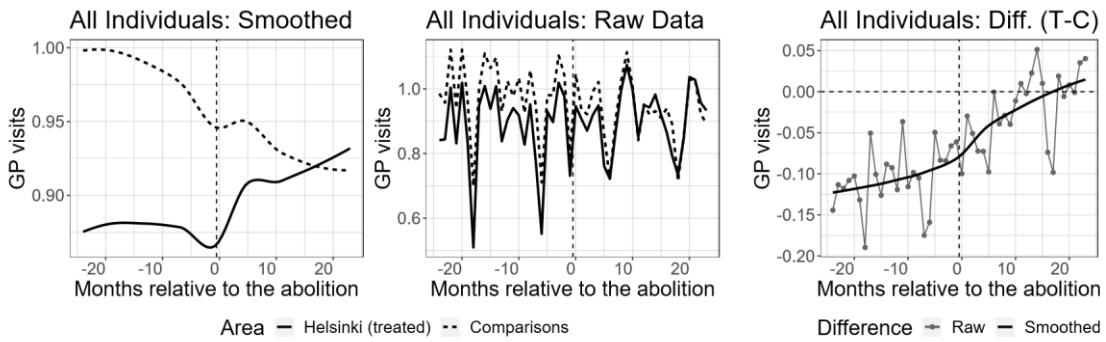

Figure 1: Trends in GP Visits.

Notes: The outcome is the number of annualized GP visits per capita. We show 1) smoothed conditional means fitted with local linear regression, 2) the raw data, and 3) the difference in outcomes between Helsinki and the comparison areas. The sample is described in Section 3.

Figure 2 shows that the difference in GP visits between Helsinki and the comparisons varied around zero in pre-treatment periods after removing the estimated linear trend difference.[11] GP visits increased modestly in Helsinki after the copayment abolition relative to the comparison municipalities. This increase is driven by the lower end of the income distribution. We fit the following regression model using the detrended data:

$$y_{imt} = \alpha + \beta_1 Post_t + \beta_2 Treat_m \times Post_t + \gamma_m + \varepsilon_{imt} \qquad (1)$$

[10] The same idea has been earlier used by Bhuller et al. (2013) and Goodman-Bacon (2021). An alternative is to control for a linear pre-trend difference in one regression (Bilinski and Hatfield, 2020).

[11] GP use appears to be low in Helsinki in June or July relative to the comparisons, explained by holidays.

where y is the outcome for individual i in municipality m at time t, α is an intercept, $Post$ is an indicator for post-abolition periods, $Treat$ is an indicator for Helsinki (the treated area), γ_m denote municipality fixed effects, and β_2 is the coefficient of interest.

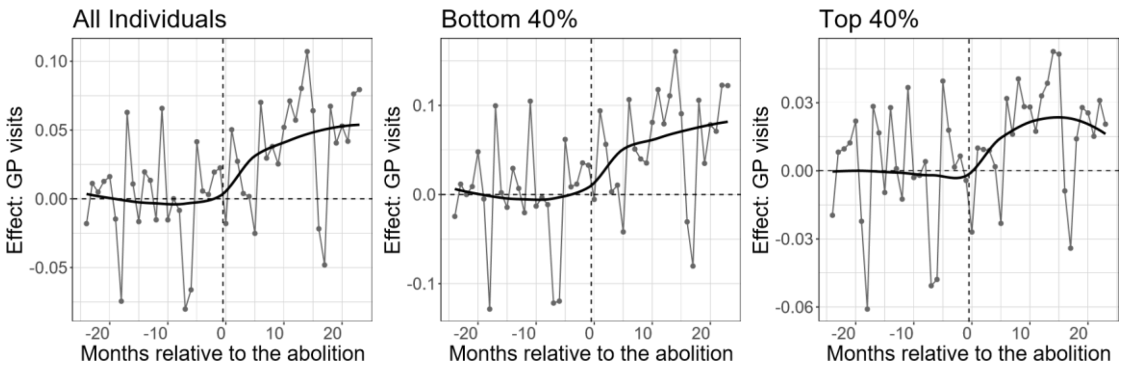

Figure 2: Trends in GP Visits after Removing a Linear Pre-Trend Difference.

Notes: We show the difference in outcomes between Helsinki and the comparison areas after subtracting a linear pre-trend difference from the outcomes, estimated with OLS using only pre-abolition data. The plot shows the raw difference and its smoothed conditional mean, fitted with local linear regression. We use the distribution of equivalized family disposable income to extract the bottom 40% and the top 40%.

Inference. The setting is challenging inference-wise; see Roth et al. (2022) for a discussion on advances in econometrics for the DD context with a small number of clusters. In principle, it is advisable to cluster standard errors at the municipality level, given that treatment assignment occurs at that specific level (Abadie et al., 2023). However, we have only one treated cluster and a finite number of smaller comparison clusters. Thus, the conventional asymptotic methods for estimating clustered standard errors are not applicable. Hagemann (2020) provides a rearrangement procedure to conduct inference in a setting like ours, but the approach requires no cluster-specific heterogeneity in trends in untreated potential outcomes so that any single untreated cluster could be used as a counterfactual. This strict assumption does not seem to be valid in our application even after removing a linear pre-trend difference from each sample municipality (Table A2). Alternatively, we could increase the number of clusters by using postal code areas, but that is not the level of the treatment assignment, and the postal code area is often missing.[12]

[12] The postal code area is obtained by 1) reading all public primary care contacts from 2011–2014, 2) including only those person-by-postal-code rows which can be linked to the real 2015 postal codes, and 3) excluding those person-

With no ideal choice available, we use several methods for inference. We cluster using analytical formulas by 1) postal code area and 2) municipality. We also provide confidence intervals based on the 3) unrestricted (WCU) and 4) restricted (WCR) versions of the wild cluster bootstrap (Roodman et al., 2019), clustering at the municipality level. For the main results, we also show IID and robust (HC1) standard errors after ignoring the time series information by aggregating the data at the municipality-by-post-treatment-indicator level (Bertrand et al., 2004).[13]

Effect heterogeneity by income level. We estimate a triple difference (DDD) model to test whether low-income individuals respond more strongly to the copayment abolition, comparing the outcomes of the bottom 40% of the income distribution to that at the top 40% in both Helsinki and comparison areas. As the baseline, we first estimate the linear pre-trend differences separately for the bottom 40% and the top 40% of the income distribution and subtract the estimated trends from the outcome data. The PTA after detrending is now assumed in ratios concerning the relative outcomes of the income groups (Olden and Møen, 2022). We estimate the following specification:

$$\begin{aligned} y_{igmt} = \alpha + \beta_1 Helsinki_m + \beta_2 Affected_g + \beta_3 Post_t + \beta_4 Helsinki_m \times Affected_g \\ + \beta_5 Helsinki_m \times Post_t + \beta_6 Affected_g \times Post_t \\ + \gamma Helsinki_m \times Affected_g \times Post_t + \varepsilon_{igmt} \end{aligned} \qquad (2)$$

where y is the outcome for individual i in income group g in municipality m at time t, α is an intercept, $Post$ is an indicator for post-abolition periods, $Helsinki$ is an indicator for Helsinki (the treated area), $Affected$ is an indicator for the bottom 40% of the income distribution, and γ is the coefficient of interest.

Complementary synthetic control (SC) analysis. We use the demeaned SC estimator (Ferman and Pinto, 2021) by subtracting the pre-treatment outcome mean from each municipality before computing the weights for available comparison units that optimize pre-treatment fit (MSE) in the matching variables between the treated area and the SC, assuming no time-varying confounders. The matching variables include all pre-treatment outcome values (Ferman et al., 2020). The weights are restricted to be positive and sum up to one to avoid extrapolation. To reduce

by-postal-code rows where the person has multiple observed postal codes (18% of the individuals). The postal code area information is missing for 34% of the population in Helsinki and for 22% in the comparison areas. For these individuals, we use the municipality of residence as the cluster.

[13] Note that neither of these methods accounts for the uncertainty induced by estimating the linear pre-trend difference between Helsinki and the comparisons, which is removed from the data (detrending).

the risk of overfitting, we only include donors with at least 40,000 sample individuals, resulting in 9 donors for the GP visits outcome.

5. Main Results

The DD results on GP visits using Specification 1 with the detrended data are reported in Table 1. The copayment abolition is associated with an increase in annualized GP visits: +0.04 (+4.4%) for the whole sample, +0.06 (+4.5%) for the bottom 40%, and +0.02 (+3.3%) for the top 40% of the income distribution. Significance conclusions are sensitive to the inference method. The estimate for all individuals is significant in four cases out of six at the 10% level and in three cases at the 5% level. The estimate for the bottom 40% is significant in five cases at the 10% level and in four cases at the 5% level. The estimate for the top 40% is insignificant at the 5% level in all cases.

Table 1: DD Estimates: GP Visits.

	All	Bottom 40%	Top 40%
Mean	0.868	1.306	0.513
Estimate	0.038	0.059	0.017
Change (%)	4.43%	4.51%	3.33%
SE (IID)	0.020 (p=0.072)	0.026 (p=0.033)	0.017 (p=0.323)
SE (HC1)	0.012 (p=0.005)	0.014 (p=0.001)	0.010 (p=0.099)
SE (CL: postal code)	0.032 (p=0.224)	0.032 (p=0.064)	0.036 (p=0.635)
SE (CL: municipality)	0.012 (p=0.004)	0.014 (p=0.000)	0.010 (p=0.090)
CI WCU	[0.012; 0.065]	[0.027; 0.091]	[-0.005; 0.039]
CI WCR	[-0.048; 0.124]	[-0.050; 0.177]	[-0.042; 0.077]
Individuals	1,365,486	541,431	555,529

Notes: We estimate Specification 1. The pre-abolition mean is computed in Helsinki for 2012, and the change in percentage terms compares the estimate to this mean. Several methods are used for statistical significance testing. First, we show IID and robust (HC1) standard errors and corresponding p-values after ignoring the time series information by aggregating the data at the municipality-by-post-treatment-indicator level (Bertrand et al., 2004). Second, we use analytical formulas and cluster by postal code area and by municipality. We also provide confidence intervals from the unrestricted (WCU) and restricted (WCR) wild cluster bootstrap (Roodman et al., 2019), clustering by municipality. Before estimation, we remove a linear pre-trend difference from the data: we compute outcome means over time by policy group and calculate their difference using only pre-treatment data, then fit a linear trend difference with ordinary least squares (OLS), and finally subtract the estimated linear pre-trend difference from the outcome data. The observed pre-trend difference is assumed to extrapolate to the post-abolition periods. Bottom 40% and top 40% are based on the equivalized family disposable income distribution.

Sensitivity of the estimates. The effect estimates grow (attenuate) if we assume that the trend difference would have slowed down (accelerated) after the abolition (Figure 8). Here, we multiply the estimated pre-treatment trend slope with different values and use the transformed slope for post-treatment periods. Our estimates are positive for all sensible multipliers and remain so even if the slope of the pre-trend difference doubled. Assuming no underlying trend difference in the post-abolition periods produces estimates of +0.06 (+6.9%) for the whole sample, +0.09 (+6.7%) for the bottom 40%, and +0.03 (+5.9%) for the top 40% of the income distribution (Table A3). Similarly, assuming that the slope of the trend-difference accelerates by 50% in post-treatment periods yields smaller estimates: +0.03 (+3.2%) for the whole sample, +0.05 (+3.4%) for the bottom 40%, and +0.01 (+2.1%) for the top 40% of the income distribution.

We also report bounds-based confidence sets as proposed by Rambachan and Roth (2022), varying how much the slope of the trend difference is allowed to deviate from linearity between consecutive periods (Figure A9). We do not reject the null of no effects at the 10% level even if exact linearity is assumed: the confidence interval for all individuals is from –1% to +10%. Once the exact linearity is relaxed, the confidence intervals widen considerably and contain larger negative values.

Figure A10 shows the robustness of the results to small changes in the comparison group. Based on the leave-one-out results, the estimates grow noticeably if either Vantaa (92) or Kouvola (286) are excluded and decrease if either Turku (853) or Joensuu (167) are excluded.[14] The estimates for all individuals vary between +0.02 (+2.6%) and +0.05 (+6.2%). The leave-two-out estimates for all individuals remain positive, while a couple of leave-three-out combinations out of 969 produce negative point estimates. Table A4 presents the results from using a monthly indicator of having any GP visits as the outcome and from replacing municipality fixed effects with postal code area fixed effects. The main findings are robust.

Effect heterogeneity by income level. Figure 3 shows the effects by income decile: they are larger in absolute terms for low-income individuals, but such a pattern is not observable in relative terms (% changes). The DDD estimates are reported in Table 2. GP use increased by +0.04 (+3.2%) annualized visits at the bottom 40% relative to the top 40% when the pre-trend difference

[14] We consider Vantaa to be an important comparison area as it belongs to the Helsinki metropolitan area and is large and urban. Turku as a large city is similarly an important comparison.

is extrapolated to the post treatment periods (1.0 x slope; our baseline). Alternatively, assuming that there was no underlying trend difference in the post-abolition periods (0 x slope) yields a somewhat larger estimate: +0.06 (+4.4%) visits. Similarly, assuming that the trend difference accelerated in the post-abolition periods (1.5 x slope) leads to a smaller estimate: +0.03 (+2.6%) visits. Detrending attenuates the estimates: the DDD estimate is +0.07 (+5.7%) on the raw data. The estimates are robust to changing the outcome from the number of GP visits to a monthly indicator of having any GP visits. Clustering analytically and the WCU bootstrap rejects at the 10% level, but the WCR bootstrap does not reject in any case after detrending at the 5% level.

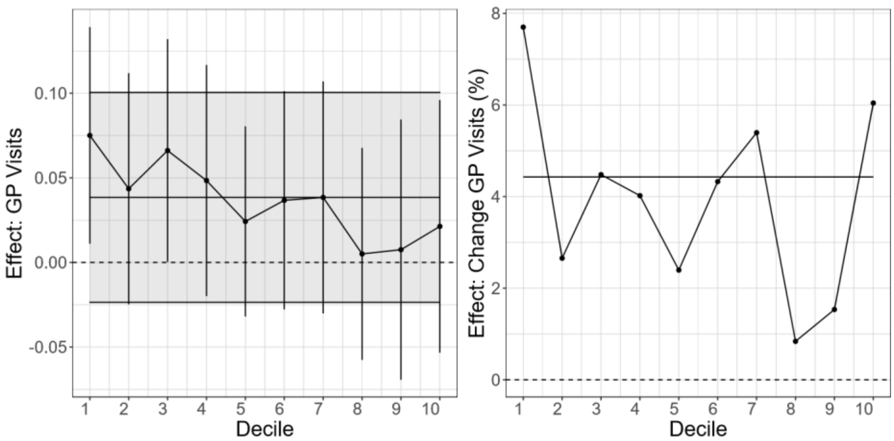

Figure 3: DD Estimates by Income Decile.

Notes: We estimate Specification 1 and cluster standard errors by postal code area. The effects are shown by income deciles (equivalized family disposable income). On the left, the gray block centered at the black horizontal line shows the estimate and its confidence interval for the whole sample. On the right, we map the point estimates to percentage changes by dividing the estimate by mean GP use in Helsinki in 2012 and multiply by 100. Before estimation, we remove a linear pre-trend difference from the data: we compute outcome means over time by policy group and calculate their difference using only pre-treatment data, then fit a linear trend difference with ordinary least squares (OLS), and finally subtract the estimated linear pre-trend difference from the outcome data. The observed pre-trend difference is assumed to extrapolate to the post-abolition periods.

Table 2: DDD Estimates: GP Visits.

A. Outcome: the number of GP visits

	No detrending	0 x slope	1.0 x slope	1.5 x slope
Mean	1.306	1.306	1.306	1.306
Estimate	0.074	0.057	0.042	0.034
Change (%)	5.68%	4.39%	3.20%	2.61%
SE (postal code)	0.018 (p=0.000)	0.018 (p=0.002)	0.018 (p=0.024)	0.018 (p=0.066)
SE (municipality)	0.009 (p=0.000)	0.009 (p=0.000)	0.009 (p=0.000)	0.009 (p=0.001)
CI WCU	[0.054; 0.094]	[0.038; 0.077]	[0.022; 0.062]	[0.014; 0.054]
CI WCR	[0.015; 0.134]	[-0.002; 0.117]	[-0.018; 0.101]	[-0.025; 0.094]

B. Outcome: the indicator of having any GP visits

	No detrending	0 x slope	1.0 x slope	1.5 x slope
Mean	9.243	9.243	9.243	9.243
Estimate	0.443	0.369	0.301	0.267
Change (%)	4.79%	3.99%	3.26%	2.89%
SE (postal code)	0.128 (p=0.001)	0.128 (p=0.004)	0.128 (p=0.019)	0.128 (p=0.037)
SE (municipality)	0.061 (p=0.000)	0.061 (p=0.000)	0.061 (p=0.000)	0.061 (p=0.000)
CI WCU	[0.302; 0.583]	[0.229; 0.509]	[0.161; 0.441]	[0.127; 0.407]
CI WCR	[0.025; 0.866]	[-0.049; 0.793]	[-0.117; 0.725]	[-0.151; 0.691]

Notes: We estimate Specification 2. The pre-abolition mean is computed at the bottom 40% of the income distribution in Helsinki for 2012, and the change in percentage terms compares the estimate to this mean. For statistical significance, we report standard errors and corresponding p-values using analytical formulas and cluster by postal code area and by municipality. We also provide confidence intervals from the unrestricted (WCU) and restricted (WCR) wild cluster bootstrap (Roodman et al., 2019), clustering by municipality. In the first column, we use raw data without detrending. Otherwise, we remove a linear pre-trend difference from the data before estimation: we compute outcome means over time by policy group and calculate their difference using only pre-treatment data, then fit a linear trend difference with ordinary least squares (OLS), and finally subtract the estimated linear pre-trend difference from the outcome data. The multiplier of the slope of the linear trend difference is varied for the post-abolition periods in columns (0, the baseline 1.0, and 1.5). If the multiplier is larger (smaller) than 1, the trend difference is expected to accelerate (slow down) in post-abolition periods. Sample size is 1,096,960 individuals.

6. Supplementary Analyses

Synthetic control results. Figure 4 shows that there may be a small increasing pre-trend in GP visits for Helsinki compared to our synthetic control. We report both raw and detrended results, preferring the latter. The detrended results are computed by subtracting a linear pre-trend difference, estimated with OLS using the pre-treatment data, from the raw gaps. The detrended SC estimate for all individuals is +0.037 annualized visits (+4.3%), essentially the same as the corresponding DD estimate of Table 1. The estimate without detrending is larger: +0.070 (+8.0%). Figure 4 also shows the SC results on the difference and the ratio of GP visits between the bottom 40% and the top 40% of the income distribution. GP use increased in absolute terms by +0.054 annualized visits (+6.8%) in the bottom 40% compared to the top 40% based on detrending (without detrending: +0.079 visits, or +10.0%). In contrast, the point estimates are close to zero in relative terms. The SC results are robust to a leave-two-out analysis in which we report the average results after permutatively excluding each two-donor pair from the donor pool (Figure A11).

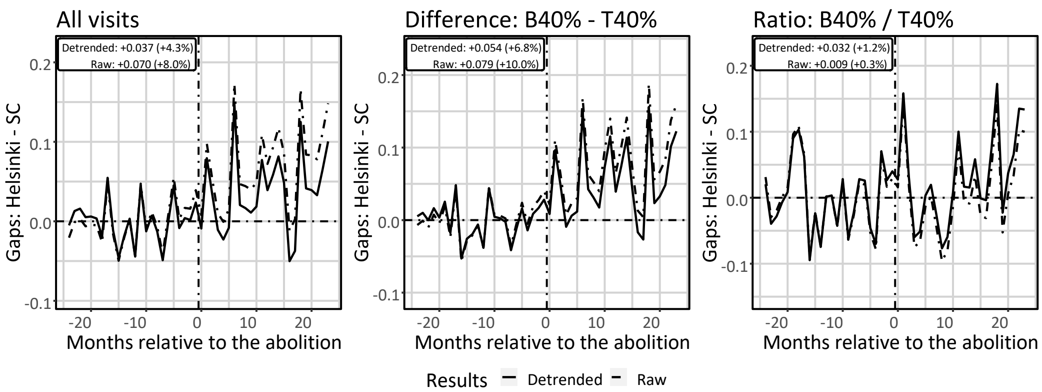

Figure 4: Synthetic Control Estimates: GP Visits.

Notes: The plots show the difference in outcomes between Helsinki and its synthetic control (gaps). The donor pool contains municipalities with more than 40,000 sample individuals. Pre-treatment lags are used as matching variables. The SC weights are reported in Table A5. We subtract from each municipality its pre-treatment outcome mean (demeaning) before estimation. B40% and T40% refer to the bottom 40% and the top 40% of the equivalized disposable income distribution. The detrended results show the gaps after subtracting a linear pre-trend difference. In the top left corner, we show aggregated treatment effect estimates from averaging all post-treatment gaps. The pre-abolition mean is computed in Helsinki for 2012, and the change in percentage terms compares the estimate to this mean.

Secondary outcomes. Given the small effects on GP visits, we expect null or small effects on our secondary outcomes. Figure A12 shows the trends in raw outcomes in Helsinki and the comparison areas. For social assistance use, there is a clear increasing trend in Helsinki. The slope of the pre-trend difference was larger before the abolition, but we think that few conclusions can be made regarding the use of social assistance. Figure A13 shows the detrended difference in ED visits and specialist consultations between Helsinki and the comparison areas. Nothing striking seems to happen after the abolition. In absolute terms, the changes are small compared to the observed increase in GP visits. The corresponding regression results are in Table A6. The increase in ED visits at the bottom 40% of the income distribution translates to a +3.0% increase, but it is insignificant. The other estimates are close to zero and insignificant.

Placebo outcome: dentist visits. Interestingly, dentist visits decreased in Helsinki after the GP visit copayment abolition relative to the comparison areas (Figure A14 and Figure A15). This results from a reduction in dentist use in Helsinki, which has partially recovered since April 2014 relative to the comparison areas. The DD estimates on dentist visits in Table A7 are of similar magnitude in absolute terms than the estimates on GP visits, but with a different sign. The DDD estimates in Table A8 show that it is the lower end of the income distribution whose dentist use decreased more in Helsinki after the GP visit copayment abolition. We searched for possible explanations for the observed reduction in dentist use but did not find convincing candidates.[15] As dentist visits were *ex ante* a plausible placebo, detecting effects (and not a precise null) creates some doubts abouts our main results on GP visits.

Time placebo. We estimate the effects of a placebo intervention using only pre-abolition data from 2011–2012 and proceed as if Helsinki abolished the copayment in January 2012 (Table A9). All other aspects of data processing and analysis remain fixed. The placebo estimates are negative (not positive) and closer to zero than the policy estimates of Table 1.

[15] A partial explanation is that Helsinki reduced the supply of vouchers for private dentist visits from July 2013 to the end of the year for budgetary reasons (Helsinki's social and healthcare services committee, September 17th, 2013). There were 7,600 voucher visits in the first half of 2013, converting to 0.025 annualized visits per capita. However, the estimated reduction in 7/2013–3/2014 was much larger.

7. Conclusion

Overall, our finding that smaller out-of-pocket costs increase primary care use is in accordance with the earlier literature, but the magnitude of our estimates is modest. Our main estimate of the effect of the copayment abolition on the number of annualized GP visits for all individuals (+0.038 visits, or +4.4%) maps to a semi-arc elasticity of 0.26.[16] For comparison, Nilsson and Paul (2018) report semi-arc elasticities of 0.88 at the 20th birthday and 0.55 at the 7th birthday for doctor visit copayments in Sweden. Moreover, Johansson et al. (2019) report a semi-arc elasticity of 1.11 at the 20th birthday in Sweden for GP visit copayments.[17] Regarding the heterogeneity of the effects by income level, we find some evidence to support the hypothesis of low-income individuals responding more strongly to copayments. However, this heterogeneity is only present in absolute terms (the number of visits) but not in relative terms (compared to the baseline). Our standard errors are sensitive to how we account for clustering. Depending on the inference method, confidence intervals can be rather narrow, or they can include large point estimates as well. We advise against strong conclusions on statistical significance.

The small effect sizes may be explained by several factors: gatekeeping at the entry, supply that is relatively insensitive to changes in demand, and by the fact that the copayment was charged only for the first three visits annually before the abolition. Our analysis does not account for some visits consequently having been free of charge. Future studies should formally test whether Finnish adults respond to the spot price change after the third GP or nurse visit of the calendar year. More specifically, such studies could test for discontinuities in the frequency of visiting primary care after the third visit of the year, also exploiting the staggered adoption of the nurse visit copayment in the 2010s and the abolition of the GP visit copayment in Helsinki in 2013. As another potential explanation for small effects, the effects of increased and decreased cost-sharing may not be symmetric. Indeed, some recent studies have concluded that framing the change as a loss may have larger effects than framing it as a gain (Hayen et al., 2021; Iizuka and Shigeoka, 2021; Remmerswaal et al., 2019).

---

[16] The semi-arc elasticity captures the change in GP visits, normalized by the baseline, divided by the price change (Brot-Goldberg et al., 2017): $(q1\text{-}q0)(q1+q0)(p1\text{-}p0)2=(0.868+0.038\text{-}0.868)(0.868+0.038+0.868)(0\text{-}13.8/83)2$. As in Nilsson and Paul (2018), our price is the share of the out-of-pocket costs of the total cost of the visit. The average production cost of a GP visit was 83 euros in 2017 (Mäklin and Kokko, 2020).

[17] We computed the elasticity from the estimates for all individuals in Table 1 in Johansson et al. (2019), using a copayment of SEK 100 and the total cost of SEK 1500 per visit, a figure appearing in the study.

There are several limitations. Helsinki is in many ways unique in Finland, which can explain the observed pre-trend differences. Our point estimation relies on the detrended PTA, and statistical inference is complicated in the presence of only one treated cluster and a finite number of comparison clusters. Conclusions regarding the statistical significance are sensitive to how we account for clustering. Interestingly, the number of dentist visits (a potential placebo) decreased in Helsinki after the GP visit copayment abolition by a similar magnitude as GP use increased.

Our findings suggest that the abolition of a 14-euro copayment did not lead to a large increase in GP use. Unfortunately, we do not have data on waiting times. The copayment abolition was a clear improvement in terms of reduced costs for low-income individuals, including the unemployed and pensioners, who disproportionately rely on public primary care. In this sense, the policy reduced inequality in barriers to access. However, the policy most likely did not greatly reduce health inequalities, because the first-order effects on service use were so moderate.

References

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's Tobacco control program. *Journal of the American Statistical Association*, 105(490):493–505.

Abadie, A., Athey, S., Imbens, G., Wooldridge, J. (2023). When should you adjust standard errors for clustering? *The Quarterly Journal of Economics*, 138(1):1–35

Beck Olsen, C. and Melberg, H. O. (2018). Did adolescents in Norway respond to the elimination of copayments for general practitioner services? *Health Economics*, 27(7):1120–1130.

Bertrand, M., Duflo, E., and Mullainathan, S. (2004). How Much Should We Trust Differences-In-Differences Estimates? *The Quarterly Journal of Economics*, 119(1):249–275.

Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). Broadband Internet: An Information Superhighway to Sex Crime? *The Review of Economic Studies*, 80(4):1237–1266.

Bilinski, A. and Hatfield, L. A. (2020). Nothing to see here? Non-inferiority approaches to parallel trends and other model assumptions.

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.

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.

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.

Farbmacher, H. and Winter, J. (2013). Per-period co-payments and the demand for health care: Evidence from survey and claims data. *Health Economics*, 22:1111–1123.

Ferman, B. and Pinto, C. (2021). Synthetic controls with imperfect pretreatment fit. *Quantitative Economics*, 12:1197–1221.

Ferman, B., Pinto, C., and Possebom, V. (2020). Cherry Picking with Synthetic Controls. *Journal of Policy Analysis and Management*, 39(2):510–532.

Giacomo, M. D., Piacenza, M., Siciliani, L., and Turati, G. (2022). The effect of co-payments on the take-up of prenatal tests. *Journal of Health Economics*, 81:102553.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*.

Haaga, T., Böckerman, P., Kortelainen, M., and Tukiainen, J. (2023a). Do adolescents from low-income families respond more to cost-sharing in primary care?

Haaga, T., Böckerman, P., Kortelainen, M., and Tukiainen, J. (2023b). Effects of nurse visit copayment on primary care use: Do low-income households pay the price?

Hagemann, A. (2020). Inference with a single treated cluster.

Hayen, A. P., Klein, T. J., and Salm, M. (2021). Does the framing of patient cost-sharing incentives matter? The effects of deductibles vs. no-claim refunds. *Journal of Health Economics*, 80:102520.

Iizuka, T. and Shigeoka, H. (2021). Asymmetric Demand Response When Prices Increase and Decrease: The Case of Child Healthcare. *The Review of Economics and Statistics (accepted)*.

Jakobsson, N. and Svensson, M. (2016). The effect of copayments on primary care utilization: results from a quasi-experiment. *Applied Economics*, 48(39):3752–3762.

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.

Ma, Y. and Nolan, A. (2017). Public healthcare entitlements and healthcare utilisation among the older population in Ireland. *Health Economics*, 26:1412–1428.

Magnussen Landsem, M. and Magnussen, J. (2018). The effect of copayments on the utilization of the GP service in Norway. *Social Science & Medicine*, 205:99–106.

Mäklin, S. and Kokko, P. (2020). Terveyden- ja sosiaalihuollon yksikkökustannukset Suomessa vuonna 2017.

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*.

Rambachan, A. and Roth, J. (2022). A More Credible Approach to Parallel Trends.

Remmerswaal, M., Boone, J., Bijlsma, M., and Douven, R. (2019). Cost-sharing design matters: A comparison of the rebate and deductible in healthcare. *Journal of Public Economics*, 170:83–97.

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., and Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal*, 19(1):4–60.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., and Poe, J. (2022). What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature.