Gender Differences in Healthcare Utilisation - Evidence from Unexpected Adverse Health Shocks

Nadja van 't Hoff[1], Giovanni Mellace[2], Seetha Menon[3]

Abstract

This paper is the first to provide causal evidence of gender differences in healthcare utilisation to better understand the male-female health-survival paradox, where women live longer but experience worse health outcomes. Using rich Danish administrative healthcare data, we apply a staggered difference-in-differences approach that exploits the randomness in treatment timing to estimate the causal impact of adverse health shocks, such as first-time non-fatal heart attacks or strokes, on healthcare use. Our findings suggest that, in response to such a shock, men consistently use more healthcare than women, highlighting the underlying factors driving gender disparities in health outcomes. These insights contribute to the broader discourse on healthcare equity and inform policy interventions aimed at addressing these imbalances.

Keywords: Cardiovascular risk; healthcare utilisation; gender.

JEL Codes: I12, I14, J16.

[1]Assistant Professor, Quantitative Economics Section, University of Amsterdam

[2]Professor, Department of Economics, University of Southern Denmark

[3]Associate Professor, Department of Economics, University of Southern Denmark

Funding: Danish Independent Research Foundation Grant ID - 2033-00076B

1 Introduction

Women live longer than men by nearly half a decade (Blaakilde, 2012), yet they tend to report worse health than men of the same age, particularly in later life (Case and Paxson, 2005; Pedersen et al., 2014). This puzzling contradiction—known as the male-female health-survival paradox—has been well documented but remains only partially understood (Cutler and Glaeser, 2005).

This paper investigates whether differences in healthcare utilisation between men and women causally contribute to this paradox. Research consistently shows that men and women differ in their healthcare-seeking behaviours and treatment patterns (Galdas et al., 2005; Juel and Christensen, 2007; Nobili et al., 2011; Oksuzyan et al., 2011, 2008; van Loenen et al., 2015; Schulman et al., 1999). For instance, men are more likely to receive invasive treatments, while women often face delays or less aggressive care (Smith et al., 2009). These disparities stem from multiple factors, including inequalities in prevention, diagnosis, treatment, biological differences affecting health outcomes, and the historical exclusion of women from clinical trials, which limits understanding of treatment efficacy in women (Legato et al., 2016). Even gender concordance between patient and provider has been shown to affect medical evaluations for women but not for men, suggesting that provider discretion may reinforce gender-based disparities (Cabral and Dillender, 2024). However, much of this evidence remains descriptive and, the magnitude and underlying causes of these differences, particularly in response to acute health events, are not well understood.

Identifying gender differences in healthcare utilisation is challenging because these differences often coincide with variations in underlying health status and behaviours, making it difficult to disentangle gender from correlated health factors. Moreover, men and women differ in their health-seeking behaviours and reporting of symptoms, introducing selection and measurement biases.

To address this gap, we estimate sex-specific responses in healthcare utilisation following severe, unanticipated health shocks—specifically, first-time non-fatal heart attacks or strokes. These events are high-salience sudden shocks, the timing of which is plausibly exogenous for individuals without a prior cardiovascular diagnosis. As conceptualised in Grossman's model

of health demand (Grossman, 1972, 2000), such shocks represent a sudden depreciation of health capital, prompting individuals to invest in healthcare—through doctor visits, medication, and other services—to restore or maintain their health. Economic research has used such shocks to study subjective survival probabilities (Baji and Bíró, 2018), healthcare productivity spillovers (Chandra and Staiger, 2007), returns to local healthcare spending (Doyle, 2011), family labour supply responses (Fadlon and Nielsen, 2021), spouse's labour supply and mental health (Böckerman et al., 2025), allocative inefficiency and comparative advantage across hospitals (Chandra and Staiger, 2020), and changes in family health behaviours (Fadlon and Nielsen, 2019).

Building on this literature, we compare men and women before and after these shocks, to isolate the causal effect of gender on post-shock healthcare utilisation. Our empirical analysis leverages comprehensive Danish administrative data recording individual-level healthcare utilisation over time, including visits to general practitioners (GPs), prescription drug use, and days hospitalised. We complement these utilisation measures with healthcare expenditure outcomes—specifically, patient out-of-pocket payments for statins and provider fees for GP visits—to assess whether differences in use translate into differences in healthcare expenditure. There are a number of strengths of our research setting. First, the Danish context is particularly well suited for this research due to its universal, free healthcare system, which minimizes confounding from financial barriers or differential insurance coverage. Second, Denmark ranks among the most gender-equal countries globally (World Bank, 2019), limiting the influence of institutionalized gender bias in care delivery. These features enable us to focus on behavioural responses rather than structural constraints.

We find a persistent gender gap in post-shock healthcare engagement: men increase their use of healthcare services—particularly GP visits and statin uptake—significantly more than women in the three to five years following a health shock. This difference is robust to differences in survival and shock severity, challenging the assumption that women's longer lifespans and poorer self-reported health naturally translate to greater health-seeking behaviour. Healthcare expenditure patterns mirror these results: men spend more on statins than women in the five years post-shock, while GP visit fees show no consistent gender difference - suggesting

that higher utilisation by men does not mechanically translate into higher healthcare expenditure. These results provide new causal evidence that gender differences in healthcare use are dynamic and may have important implications for long-term health outcomes.

We test the robustness of our findings using matched samples, alternative estimators, and multiple definitions of the control group—all of which produce consistent results. To better understand the gender gap in post-shock healthcare use, we examine potential mechanisms and rule out explanations based on survival differences, household composition, or cohort effects. Our results point to both demand- and supply-side drivers: for instance, high-income men increase GP visits more than high-income women, and we find suggestive evidence that women are prescribed statins at lower rates than men.

The paper is organised as follows. Section 2 introduces the data and defines key variables. Section 3 outlines the staggered difference-in-differences identification strategy. Section 4 discusses the assumptions underlying the empirical approach. Section 5 presents the main findings. Section 6 explores robustness checks and Section 7 investigates potential mechanisms that may drive these results. Section 8 discusses the broader implications and concludes.

2 Data

Measuring underlying health status presents significant methodological challenges, particularly in survey data where measurement error is common. Inaccuracies in measuring healthcare utilization or the timing of health shocks can introduce substantial bias into our estimates. To address these limitations, we use reliable and objective administrative measures of health events and healthcare utilisation, characterised by high completeness and verified accuracy through multiple validation studies (Schmidt et al., 2015; Özcan et al., 2016). Our analysis draws on three comprehensive Danish administrative datasets. The Danish National Patient Register provides longitudinal, individual-level disease diagnosis data. We link this to the Danish Health Insurance database, which captures all visits to healthcare practitioners, and the Danish Prescription Drug database, which records all prescriptions filled at pharmacies nationwide.

2.1 Analytical Sample

To leverage the convergence of cardiovascular disease risk between men and women at older ages, we restrict our sample to Danish adults aged 50-70 who experienced their first health shock during 1995-2018 (Section 2.2 provides further details).

2.2 Measures of health shock

We use the first diagnosis of a heart attack (acute myocardial infraction) or a stroke (cerebral infraction) at the individual level to create our measure of an adverse health shock as done in recent literature (Fadlon and Nielsen, 2019). To ensure that the health shock is exogenous, we exclude individuals with a high pre-existing cardiovascular risk, evaluated by calculating CHA_2DS_2-VASC scores. A CHA_2DS_2-VASC score assigns points for congestive heart failure, hypertension, age > 75 (double), diabetes, previous stroke/transient ischaemic attack/thromboembolism (double), vascular disease, age: 65–74, sex (female). Specifically, we exclude men with a risk score of 2 or higher and women with a risk score of 3 or higher, as defined by the European Society of Cardiology and commonly applied by medical professionals.

Table 1: Distribution of health shocks between men and women.

	Man	Woman	p-value
Acute myocardial infarction	58.98%	50.05%	0.00
Cerebral infarction	41.02%	49.95%	0.00
	100%	100%	
Nr. unique individuals	30,103	15,140	

Table 1 shows the distribution of health shocks between men and women in the analytical sample. The last column shows the p-value from Chi-square tests of homogeneity which indicates that we reject the null of equal distributions. In particular, the distribution of health shocks differs by gender: women have roughly equal chances of experiencing either a stroke or heart attack, while men are significantly more likely to experience heart attacks than strokes. This gender difference in shock distribution, together with other potentially unobserved underlying gender differences, motivates our decision to analyse the data separately by gender. To ensure robustness, Section 6.1 presents results using a matched sample and Section 6.2 presents results

using a triple difference-in-differences estimation strategy that allows for direct comparison between men and women within a single estimation framework.

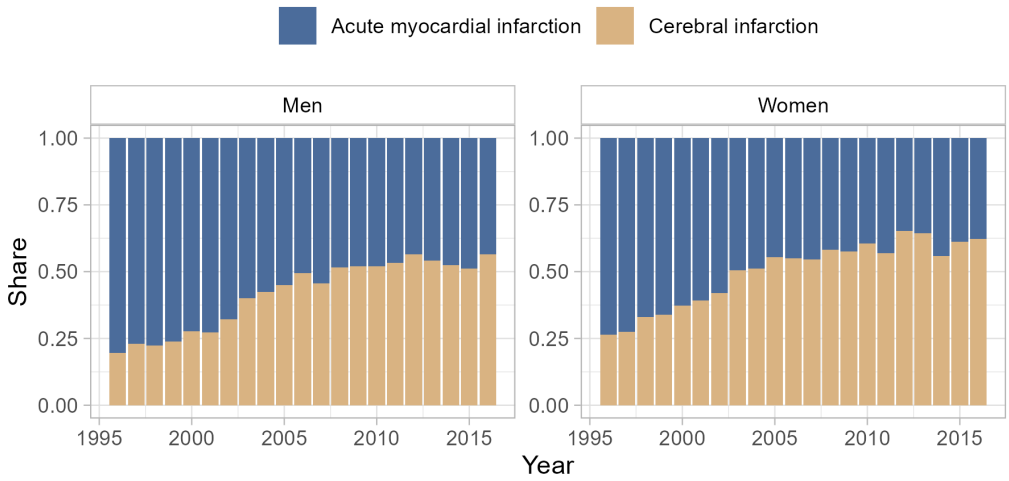


Figure 1: Share of observations by gender by health shock over time.

Figure 1 shows that while men experience health shocks more often compared to women across the years, the distribution of health shocks over time is relatively similar between both groups. Figure 2 shows the years until death after the health shock for men and women separately which shows similar rates of survival. We exclude individuals who die within the first year of the health shock as we cannot observe healthcare utilisation outcomes for these individuals post-shock. Furthermore, we ensure our results are not driven by differential rates of survival in Section 7.2.

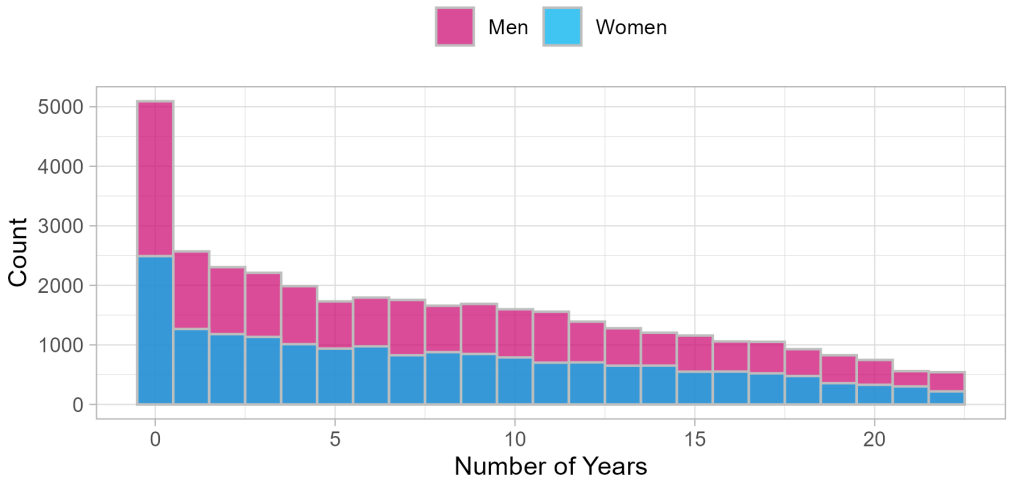


Figure 2: Comparison of death rates after shock for men and women.

2.3 Measures of healthcare utilisation

We examine gender differences in the effect of a health shock on three measures of healthcare utilisation: (1) defined daily dose (DDD) of statin consumption, (2) number of visits to the general practitioner and (3) number of days spent in the hospital as an in-patient or outpatient user.

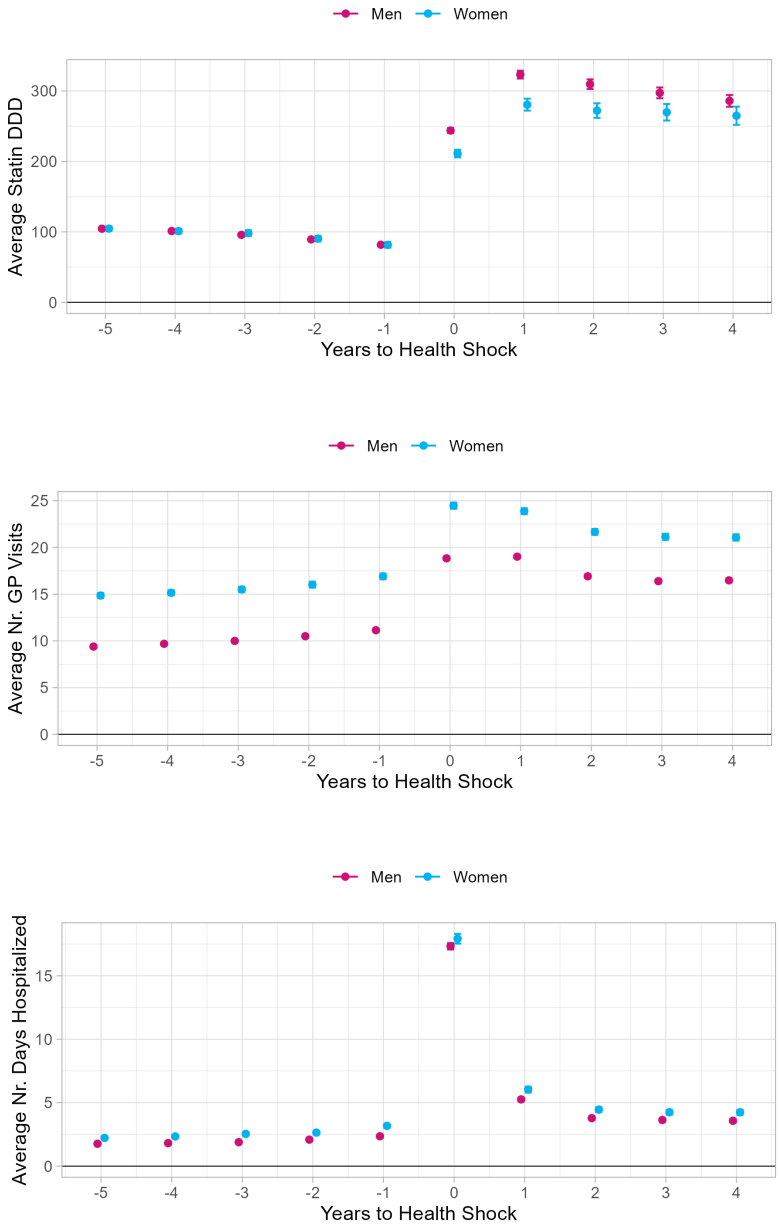


Figure 3: Plots of healthcare utilisation outcomes by gender over time

Figure 3 plots the outcomes over years by gender for each healthcare utilisation outcome where zero indicates the year of treatment. While the utilisation of all three outcomes examines increases post an adverse health shock, the level of utilisation is markedly different by gender

Table 2: Healthcare utilisation in pre- and post-treatment period

	5 years pre-treatment	5 years post-treatment	p-value
Men			
Statins DDD	180.83	1574.25	0.00
Nr. GP visits	42.54	84.45	0.00
Nr. Days hospitalized	8.29	32.90	0.00
Women			
Statins DDD	208.39	1467.11	0.00
Nr. GP visits	68.62	107.87	0.00
Nr. Days hospitalized	11.21	36.05	0.00

for statin consumption and average number of visits to the GP with men consuming more statins and women visiting the GP more.

Table 2 presents healthcare utilisation outcomes in the 5 years before health shock and in the 5 years after health shock for each gender. The last column shows the p-values of a test of equal distributions which is rejected in all outcomes indicating that healthcare utilisation significantly increases post a health shock for all outcomes and for both genders.

3 Identification, estimation and inference

Our goal is to estimate gender-specific causal effects of adverse health shocks (myocardial or cardiac infarction) on three healthcare utilisation outcomes: statin usage, doctor visit frequency, and hospitalization days. Utilising a quasi-experimental design, we exploit the random timing of the health shocks. The core idea is to compare individuals who have experienced a health shock with *not-yet-treated* individuals who will face the same shock later. We make these comparisons separately by gender: women who experienced a health shock at a given time serve as the treatment group, while women who experienced it later constitute the control group. Men are analysed using the same logic.

Given that our dataset spans 23 years, the plausibility of the random timing assumption for health shocks becomes questionable when comparing individuals who experienced shocks too far apart in time. For example, individuals who remained untreated by 2018 may differ systematically from those who were treated in 1996, potentially reflecting changes in medical practices, patient demographics, or underlying health trends over this extended period. To

address this issue, we implement a five-year rolling window that restricts comparisons to temporally proximate cohorts.[4] This approach prevents inappropriate temporal mismatches, such as comparing someone treated in 1996 with someone treated in 2018. The restriction allows us to estimate up to four pre-treatment and five post-treatment effects for each individual.

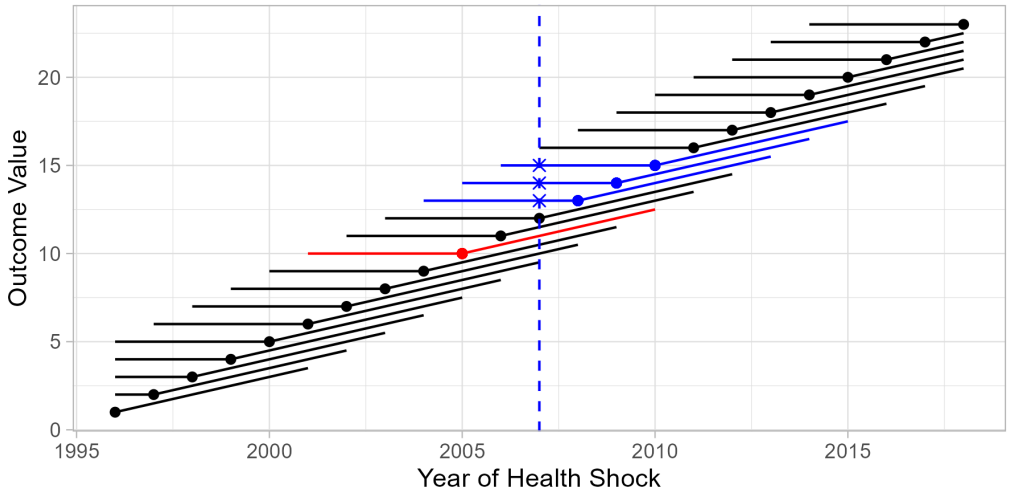


Figure 4: Explanation of five-year rolling window.

Figure 4 provides a graphical representation of how the five-year window determines control group composition for each treatment cohort. To illustrate with a concrete example: individuals who experience their health shock in 2005 (as indicated in red) can be compared against control groups consisting of those treated in 2008, 2009, and 2010 (as indicated in blue) when estimating effects two years post-treatment (in 2007). This framework ensures reasonable temporal comparability between treated and control units. The approach does create some limitations at the dataset boundaries—cohorts treated in 2015 or later have fewer available control groups, while cohorts treated before 2000 have fewer periods available for pre-treatment effect estimation.

Although the timing of receiving treatment can plausibly be considered randomly assigned, we strengthen our identification strategy by controlling for time-invariant individual characteristics using a Difference-in-Differences (DiD) methodology. Since men and women may exhibit different underlying trends unrelated to health shocks, we apply DiD separately by gender. This approach ensures that any observed differences in outcomes can be attributed to the health shock itself rather than to pre-existing gender-specific trends.

[4]Section 6.3 presents sensitivity analyses using alternative control group definitions.

Recent methodological advances have demonstrated that two-way fixed effects models perform poorly in settings with staggered treatment adoption and heterogeneous treatment effects (Borusyak et al., 2024; Goodman-Bacon, 2021; De Chaisemartin and d'Haultfoeuille, 2020)—both features present in our empirical context. We therefore implement the doubly robust estimator proposed by Callaway and Sant'Anna (2021), which adapts the Sant'Anna and Zhao (2020) methodology to accommodate multiple time periods. This estimator is particularly appropriate for our analysis because it enables the estimation of treatment effects for each cohort in each time period, controlling for observable covariates while preserving the ability to capture heterogeneous treatment effects across groups and time. In our analysis, we control for age and income. This is essential for accommodating covariate-specific trends. For example, controlling for covariates is crucial when examining the longitudinal consumption of statins, as these trends may be influenced by income and age.

Similar to Callaway and Sant'Anna (2021) our goal is to identify and estimate *group-time average treatment effects* defined as:

$$ATT(g,t) = E[Y_t(g) - Y_t(0)|G_g = 1],$$

where $Y_t(g)$ denotes the potential outcome in year t for those first treated in year g. Groups are defined by the year an individual first experiences a health shock. Let G denote the year when an individual experiences their first adverse health shock. We define a binary indicator G_g which equals one if an individual experiences the health shock in year g, and a binary variable $D_{i,t}$ which equals one if individual i has experienced the health shock either in or before year t. Under the standard difference-in-differences assumptions and abstracting from covariates for simplicity, the ATT for group g at time t is identified by

$$ATT(g,t) = E[Y_t - Y_{g-1}|G_g = 1] - E[Y_t - Y_{g-1}|D_t = 0].$$

For example, considering the ATT for individuals experiencing the health shock in 1996 ($g = 1996$), assessed two years post-treatment ($t = 1998$), using the not-yet-treated group as

control group (individuals treated in the years following 1998), the expression is as follows:

$$ATT(1996, 1998) = E[Y_{1998} - Y_{1995}|G_{1996} = 1] - E[Y_{1998} - Y_{1995}|D_{1998} = 0].$$

Including covariates, the group-time ATT are identified as follows:

$$ATT(g,t) = E\left[\left(\frac{G_g}{E(G_g)} - \frac{\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}}{E\left[\frac{p_{g,t}(X)(1-D_t)(1-G_g)}{1-p_{g,t}(X)}\right]}\right)(Y_t - Y_{g-1} - m_{g,t}(X))\right], \quad (1)$$

where $m_{g,t} = E[Y_t - Y_{g-1}|X, D_t = 0, G_g = 0]$ denotes the outcome regression for the not-yet-treated group and $p_{g,s}(X) = P(G_g = 1|X, G_g + (1 - D_s)(1 - G_g) = 1)$ the propensity score. The latter gives the probability of experiencing the health shock in year g, given the covariates X and being treated for the first time in year g or being not-yet-treated in year s.

Estimation proceeds in two steps. In the first step, the propensity score $p_{g,t}(X)$ and the outcome model $m_{g,t}(X)$ are estimated using logistic and linear regressions, respectively. In the second step, the fitted values from these models are substituted into the sample analogue of Equation (1). This doubly robust estimator attains the semiparametric efficiency bound when both nuisance models are correctly specified. An important advantage of the approach in Callaway and Sant'Anna (2021) is the availability of simultaneous confidence bands that account for multiple testing, obtained via a multiplier-type bootstrap procedure.[5]

Under the two assumptions of no treatment anticipation and random timing of health shocks, this procedure consistently estimates the group-specific ATTs. The first assumption requires that health shocks are unexpected. We ensure this by restricting our analytical sample to individuals with low CHA_2DS_2-VASC scores, confirming their good health prior to the shock, and including only those experiencing a single health shock within a five-year window around treatment. Even if the second assumption (random timing) is violated, we can still consistently estimate group ATTs under a conditional parallel trends assumption. This assumption requires that the not-yet-treated groups exhibit similar pre- and post-treatment trends. Specifically, the

[5]Specifically, the bootstrap reweights the influence function contributions using random Rademacher multipliers (± 1 draws at the individual level). The $(1 - \alpha)$-quantile of the maximum absolute studentized t-statistics across event times provides the uniform critical value for simultaneous confidence intervals.

trajectory of our outcome variable (e.g., statin use) for the group experiencing the health shock would have followed the same path as the not-yet-treated group over time, had the health shock not occurred. We further discuss the plausibility of our identification strategy in Section 4.

4 Plausibility of identifying assumptions

The two most crucial assumptions for our identification strategy are the no anticipation assumption and the assumption of random treatment timing within the five-year treatment timing. If the latter assumption fails, we still identify the effect under the standard parallel trends assumption. While both of these assumptions are essentially not testable, we can obtain some suggestive evidence on the plausibility of these assumptions. Figure 5 presents short DiD estimates in the pre-treatment period. The absence of significant pre-treatment effects provides reassuring evidence for our identifying assumptions. Apart from a negligible effect on GP visits in the immediate pre-treatment period, we detect no statistically significant anticipatory responses. This pattern strongly supports both the no treatment anticipation assumption and the validity of our random treatment timing framework.

For the parallel trends analysis presented in Appendix A, we construct a control group to address the fact that all sample individuals eventually experience treatment. Our control group comprises individuals who remain untreated at the time of a given cohort's treatment but will receive treatment within the subsequent five years. To facilitate comparison, we assign these control individuals a "placebo" treatment date coinciding with the actual treatment timing of the focal cohort. Data for control group members are then extracted over a five-year window surrounding their assigned placebo event. Our control group construction strategy closely resembles the methodology described by Fadlon and Nielsen (2019), who similarly leverage future untreated cohorts to create comparison groups. However, we implement a more flexible variant of their approach. Whereas Fadlon and Nielsen (2019) impose a strict constraint requiring control group members to be treated exactly five years after the focal treatment cohort, we adopt a more inclusive criterion that incorporates individuals treated at any point during the subsequent five-year window. This modification increases statistical power while preserv-

ing temporal comparability. The resulting parallel trends analysis yields reassuringly similar patterns for both men and women, supporting the validity of our identification strategy.

5 Results

We use the estimator of Callaway and Sant'Anna (2021) as our preferred specification, since it allows for uniform inference and does not exhibit the same issues as TWFE when treatment effects are heterogeneous.[6]

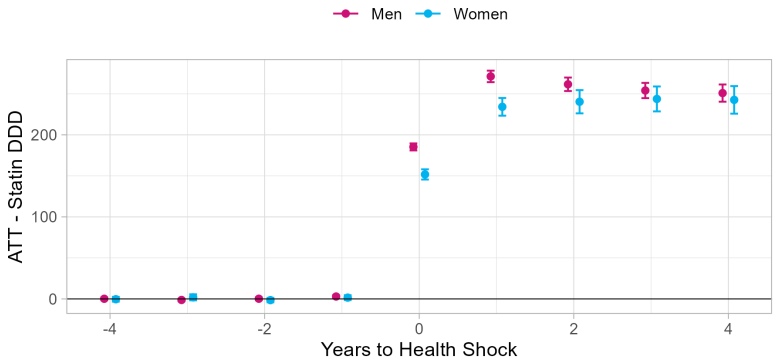
Figure 5 depicts evolution of treatment effects over time for the three different outcomes. Tables 5, 6, and 7 presents the ATT estimates for statins, GP visits and days hospitalised respectively. The group-time average treatment effects reported in Figure 5 indicate that statin consumption increases by approximately 251 to 271 DDD in the years following the health shock for men (see Panel B in Table 5), with an aggregate ATT of 259. For women, this increase ranges approximately between 234 and 244 DDD in the years following the health shock (see Panel A in Table 5), with an aggregate ATT of 240 DDD over the four post-treatment years. These findings indicate that men consume more statins than women following a health shock, with the difference being statistically significant in the first two years post-shock.[7]

It is worth noting that while we see the dynamic ATT estimate for the year of the health shock (year 0) is lower for women than for men, both are likely underestimated. These results includes individuals who experience the health shock late in the year (e.g., December), suggesting that the actual estimate for this period may be higher. In subsequent years, the effects of the health shock on statin consumption remain relatively stable over time for both men and women.

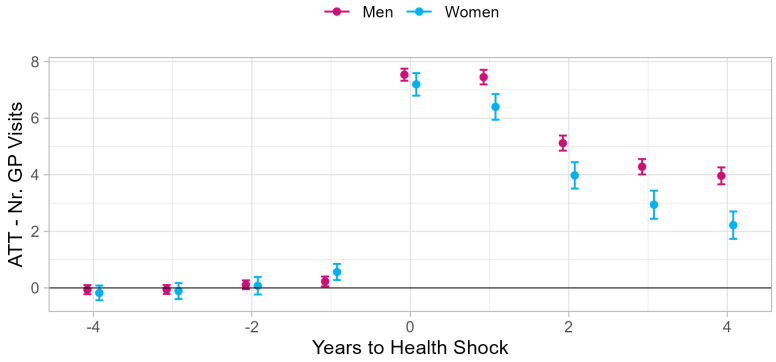
Turning to our second outcome, healthcare utilisation measured as the number of GP visits, Figure 5 shows both men and women increase their average number of visits by approximately

[6]Callaway and Sant'Anna (2021) is not the only method that addresses issues with the traditional TWFE model. Several alternatives to TWFE exist but are less suitable for our setting. For instance, De Chaisemartin and d'Haultfoeuille (2020) does not capture the evolution of treatment effects over time, Borusyak et al. (2024) requires more pre-treatment periods to predict counterfactual trends, and Sun and Abraham (2021) impose unconditional parallel trends assumptions.

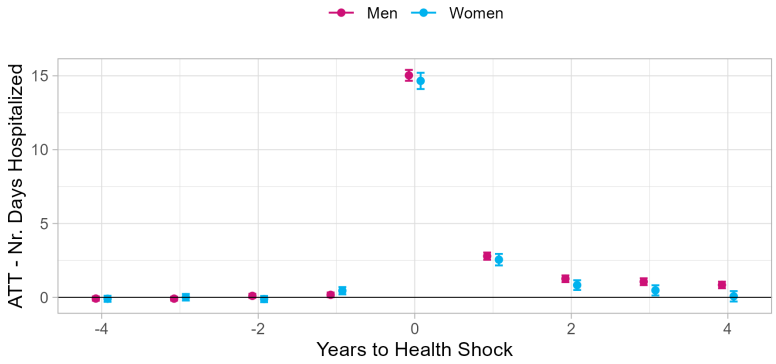
[7]While the effects for men and women can be interpreted as causal, this observed difference should not be interpreted as causal.



(a) Dynamic ATT for statin DDD.



(b) Dynamic ATT for number of GP visits.



(c) Dynamic ATT for days hospitalized.

Figure 5: Dynamic ATT plots for various healthcare measures. Each panel shows the ATT for a specific measure, illustrating the impact of interventions over time. Simultaneous 95% confidence bands obtained through bootstrapping are depicted.

6 to 8 per year during the year of the health shock and the year following it. These numbers decrease in the third and fourth year after the health shock as reported in Table 6. Over the four years after the treatment year, men visit a GP on average 1 to 2 times more per year than women. This difference is statistically significant and grows larger over time.

For the third outcome, days hospitalised, Figure 5 demonstrates that both men and women are hospitalised for approximately 15 days on average during the year of the health shock. In subsequent years, hospitalisation rates decline substantially a shown in Table 7, with no statistically significant differences observed between men and women.

Table 3 aggregates the post-shock group–time ATTs over the five years after the health shock and benchmarks them against cohort-weighted pre-shock means to aid scale interpretation. Three patterns stand out. (i) *Statins:* men exhibit a larger post-shock increase than women in both absolute and relative terms: 242.9 vs. 219.5 DDD (gap ≈ 23.4 DDD), corresponding to 1313.7% vs. 1147.3% relative to very low pre-shock use. This aligns with the event-study evidence where the male–female difference is most pronounced in the first two post-shock years. (ii) *GP visits:* men increase visits by 5.84 per year versus 4.77 for women (gap ≈ 1.07), and because men start from a lower baseline (8.12 vs. 13.82), the relative change is about twice as large for men (71.9% vs. 34.5%). (iii) *Days hospitalised:* post-shock days hospitalised are similar (women 4.22, men 4.67); the modest absolute gap (≈ 0.45 days) and the large percentages primarily reflect small pre-shock baselines rather than a persistent gender difference in admissions. Across outcomes, effects are precisely estimated (tight confidence intervals, and low p-values).

Our results show that men, on average, use slightly more healthcare than women following a health shock—particularly in statin use and GP visits—while no gender difference is observed in hospitalisations, which are less influenced by individual behaviour. This suggests the gap likely reflects differences in healthcare-seeking rather than underlying health needs. Because our study spans ages when many women transition through menopause, a natural question is whether this change could account for the patterns we observe. Although menopause represents a clinically significant transition for women, with a documented rise in physician consultations and hormone replacement therapy use (Conti et al., 2025), current clinical guidelines still rec-

Table 3: ATT aggregated over the five years following the health shock. The final two columns report (i) the cohort-weighted pre-shock mean of each outcome (measured over the five years before the shock) and (ii) the ATT expressed as a percentage of that baseline mean.

	ATT	CI lower	CI upper	Std. err.	Nr. obs.	p-value	Baseline mean	Rel. % change
Statin DDD								
Women	219.48	212.85	226.11	3.40	15140	0.00	19.13	1147.29%
Men	242.90	238.47	247.34	2.26	30103	0.00	18.49	1313.65%
No. of GP visits								
Women	4.77	4.51	5.03	0.13	15140	0.00	13.82	34.54%
Men	5.84	5.70	5.99	0.07	30103	0.00	8.12	71.90%
No. of days hospitalised								
Women	4.22	4.04	4.39	0.09	15140	0.00	2.09	201.95%
Men	4.67	4.56	4.78	0.06	30103	0.00	1.58	295.83%

ommend statins for postmenopausal women at elevated cardiovascular risk regardless of HRT use (Hirsh, 2019; Witting et al., 2022). Therefore, while menopause may affect baseline health and treatment patterns, it does not appear to explain the persistent gender gap in post-shock healthcare utilisation we observe.

6 Robustness Checks

6.1 Estimation with matched sample

As a robustness check, we match women and men based on their pre-treatment covariate values and then implement the estimator proposed by Callaway and Sant'Anna (2021) in Section C.1 of the Appendix. We use nearest neighbour matching on propensity score to balance covariate distributions between men and women, which ensures that the effect of the health shock is less affected by systematic differences between men and women. Propensity score matching ensures comparability between the two groups and that differences across groups might be attributed to gender-specific responses rather than confounding by pre-treatment covariates. Our results in the matched sample are consistent with our main findings.

6.2 Triple Difference-in-Difference

In addition to matching men women as done in Section 6.1, we run a triple difference-in-difference estimation using the Fadlon and Nielsen (2019) method. We continue to find higher consumption of statins and higher number of visits to the GP for men than for women as shown in Figure 6.

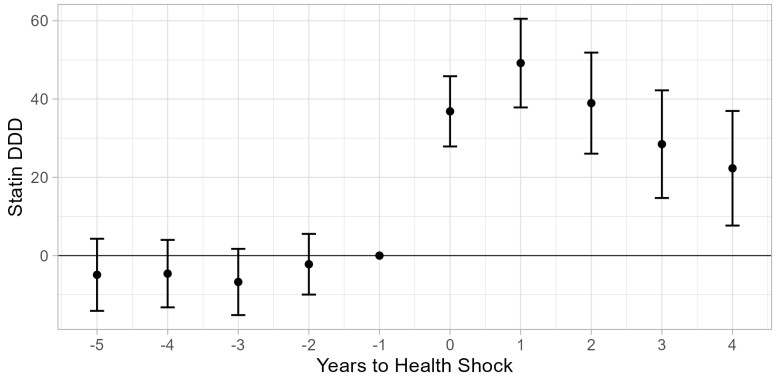
6.3 Two-Way Fixed Effects estimator

As a robustness check, we implement the Two-Way Fixed Effects (TWFE) estimator as in Fadlon and Nielsen (2019). While Callaway and Sant'Anna (2021) use the not-yet-treated as a control group, Fadlon and Nielsen (2019) use as a control group those individuals who are treated exactly five years later. Moreover, when treatment effects are heterogeneous, the causal parameters estimated by the two methods differ, with the one estimated by TWFE potentially having problematic (non-convex) weights and comparison groups ("forbidden controls").

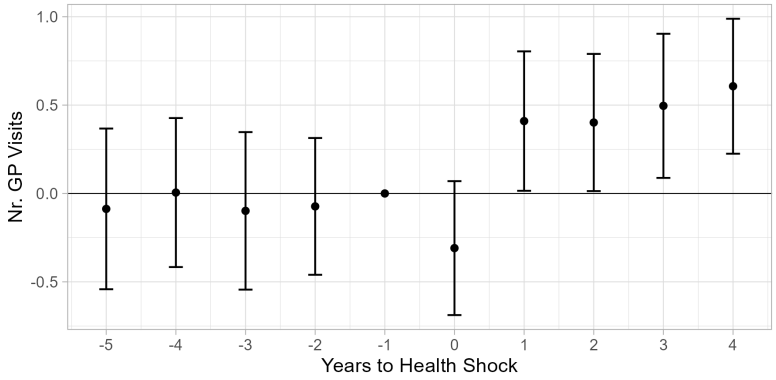
The results, presented in Figure 24 in Section C.3, are reassuringly consistent with those obtained using our main specification. Figures 24a and 24b reveal several significant non-zero effects in the pre-treatment period. Differences between TWFE pre-trend estimates and Callaway and Sant'Anna (2021) pre-trend estimates may arise because TWFE captures long-run trends while Callaway and Sant'Anna (2021) focus on short-run changes, and because the two methods target different causal estimands.

6.4 Alternative control groups by treatment year proximity

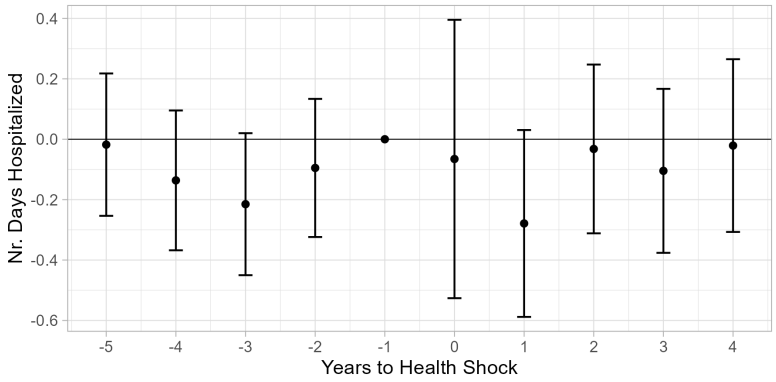
A potential concern is that effects may be driven by our choice of restricting the analysis such that the treatment year of the controls is at most five years apart from the treatment year of the treated. We ensure our results remain consistent when using alternative control groups by restricting the analysis such that treatment year of controls is at most three years and at most seven years apart from the treatment year of the treated. Using a three-year window, presented in Figure 25 in Appendix C.5, ensures greater comparability between treated and control units. Our results remain consistent with our main specification. Using an eight-year



(a) Dynamic ATT for statin DDD.



(b) Dynamic ATT for number of GP visits.



(c) Dynamic ATT for days hospitalized.

Figure 6: Triple DiD results using the Fadlon and Nielsen (2019) estimator. Estimates represent the interaction effect between treatment, gender, and time. Each panel shows the coefficient from this interaction for a different outcome.

window, presented in Figure 26 increases the number of observations, potentially improving the precision of our estimates. We find that the difference in consumption of statins between men and women disappears after year 2 while the difference in number of visits to the GP increases over time throughout the 7 years.

7 Mechanisms

Our results indicate higher healthcare utilisation for men than for women particularly with regards to statin consumption and number of visits to the GP for up to five years after the adverse health shock. Below we examine several alternative mechanisms that could be driving our results.

7.1 Severity of health shock

Our findings indicate that women have fewer visits to the general practitioner (GP) than men following a health shock. This result contrasts with previous epidemiological evidence, which suggests that men typically use primary healthcare services less frequently than women, both before and after hospitalization (Höhn et al., 2020). Our findings are in line with more recent evidence examining the effect of a stroke diagnosis in Australia where the authors find that women had fewer consultations with medical doctors and allied health practitioners (Sibbritt et al., 2024).

This finding of higher healthcare utilisation for men could be driven by higher intensity shocks for men. We leverage ICD-10 codes to separate myocardial infractions into ST-Elevation Myocardial Infarction (STEMI), which is generally considered more severe and non-ST-Elevation Myocardial Infarction (NSTEMI), which is generally considered less severe. The proportion of women with myocardial infarction experiencing high intensity was 25.29%, compared to 34.14% among men. We then examine if our estimated effects are different based on the severity of this health shock. We do not have a measure of severity for stroke and therefore only use the sample of individuals who had a heart attack. The results are shown in Figure 27 in Section D. While women with less severe shocks use less statins and visit the GP fewer

times than men with less severe shocks, the aggregate gender difference in healthcare utilisation appears to be unrelated to underlying health needs.

7.2 Gender differences in survival rates

We examine whether differential survival rates between men and women in our analytical sample influence our results by estimating the effects using a balanced panel in Section D.2 in the Appendix. Approximately 74% of men and 77% of women remain in the sample throughout the analysis period. The findings are similar to our main specification using the unbalanced panel. We also directly estimate the effect of the health shock on mortality in Figure 7 and Table 4 find no differential effects of survival between men and women. This suggests that the results are not driven by differential survival rates between men and women.

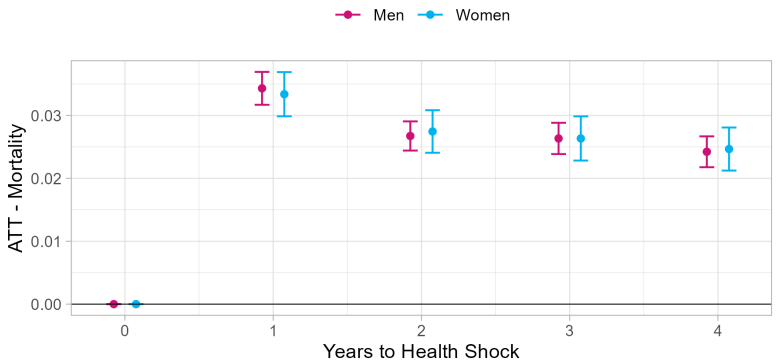


Figure 7: Dynamic ATT for mortality.

Table 4: Dynamic ATT for mortality

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
0	0.00	0.00	0.00	15140	
1	0.03	0.03	0.04	14363	0.00
2	0.03	0.02	0.03	13046	0.00
3	0.03	0.02	0.03	11934	0.00
4	0.02	0.02	0.03	10918	0.00
Panel B: Men					
0	0.00	0.00	0.00	30103	
1	0.03	0.03	0.04	28547	0.00
2	0.03	0.02	0.03	25945	0.00
3	0.03	0.02	0.03	23790	0.00
4	0.02	0.02	0.03	21944	0.00

7.3 Heterogeneous effect by health shock

To assess whether our main results differ based on the type of health shock, we estimate the treatment effects for myocardial infarction and cerebral infarction separately. Our findings, presented in Figure 32 in Section D.3, show that men with a heart attack use more significantly more healthcare than women with a heart attack.

7.4 Heterogeneous effect by income & household type

The context of our research is Denmark, a nation characterised by universal and free public healthcare access to all residents with patients generally incurring very low out-of-pocket costs. We would therefore not expect to see differences in healthcare driven by income differentials due to access. Nonetheless income may proxy for preferences or health awareness. To examine if our results differ by income we split the analytical sample by median income and estimate the results for low income and high income individuals separately in Figure 33 in Section D.4.We find weak evidence that men with a high income visit the GP more than women with high income but this difference is only statistically significant in the later periods.

Figure 34 in Appendix D.5 depicts effects for single households versus non-single households (where non-single means couples, other couples, multi-family households). We find weak evidence that single men visit the GP more than single women but this difference is only statistically significant in the later periods.

7.5 Statins: supply v. demand

Our findings are in line with recent evidence showing women were less likely to be prescribed high-intensity statins in the Netherlands (Kiss et al., 2024). Detailed research in the US has also shown that women were less likely to have been offered statins, more likely to discontinue statin use if offered, and more likely to decline statins when offered (Nanna et al., 2019). Thus gender differences in statin utilisation could be due to both differences in provision through lower prescription rates or differences in patient preferences.

Our results do not indicate differential trends by gender in statin consumption by gender

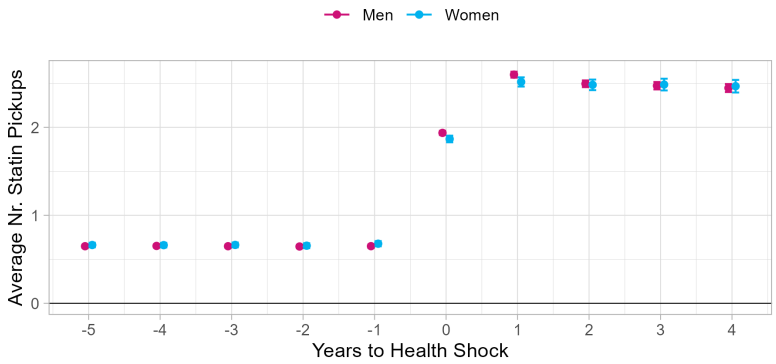


Figure 8: Raw trend of number of times statins picked up.

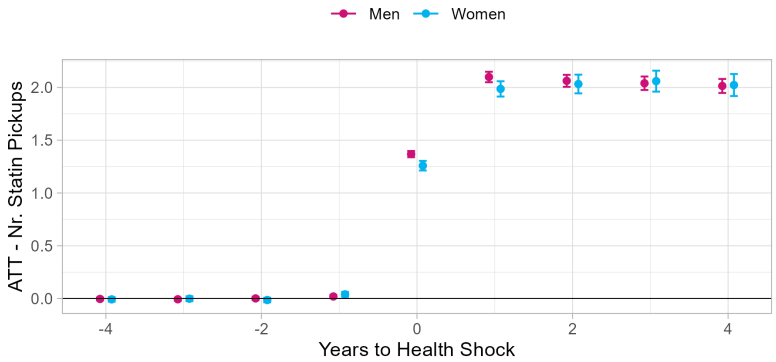


Figure 9: Dynamic ATT for number of times statins picked up.

post health shock but instead show differences in the level of consumption which suggests that these results are not driven by women discontinuing statin use after having taken it up. We also examine if women pick up less statins than men by using number of times statins was picked up from the pharmacy per year as an outcome. Figure 8 shows the raw trend for the average number of times statins were picked up by time to treatment while Figure 9 graphs the dynamic ATT estimates for men and women separately (corresponding table with estimates can be found in Table 10 in Section D.6). We do not see statistically significant differences between men and women in the number of times they pick statins in the 5 years post treatment. This suggest that women are not picking up less statins than men and that our findings are likely due to lower prescription rates for women.

7.6 Healthcare expenditure

While our main analysis focuses on healthcare utilisation as measured by service use and prescription drug uptake, examining the associated costs of these services offers additional insight into the drivers of gender differences in care. To this end, we analyse two financial outcomes:

patient out-of-pocket payments for statins in Danish kroner (DKK) and the fees received by general practitioners for patient visits. All prices are deflated to 2015 levels using the annual Consumer Price Index from Statistics Denmark[8].

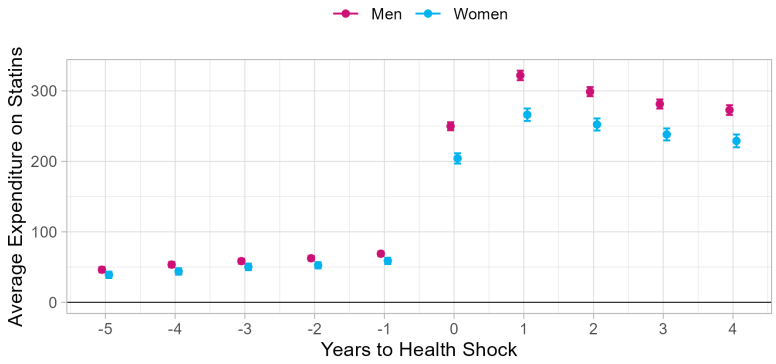


Figure 10: Raw trend of expenditure on statins.

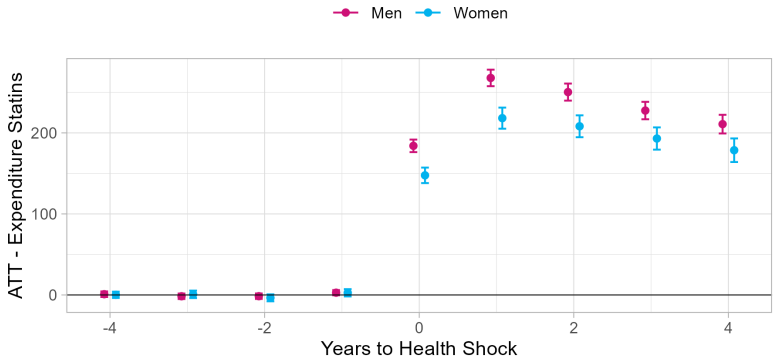


Figure 11: Dynamic ATT for expenditure on statins.

Figure 10 plots average statin expenditures over time, while Figure 11 shows the estimated causal effect of a health shock on statin spending. The results mirror our main findings: men spend significantly more on statins than women in the five years following a shock, with some convergence in later years. Table 9 in Section D.7) reports the corresponding ATT estimates, showing that average annual post-shock spending is DKK 288 for men and DKK 189 for women.

Figures 12 and 13 present analogous results for GP visit fees. Here, we find no consistent gender difference in spending, suggesting that although men increase GP visits more than women post-shock, this does not translate into higher per-patient fees. Table 10 in Section D.7 reports the corresponding estimates.

[8]Obtained from https://www.statbank.dk/20072.

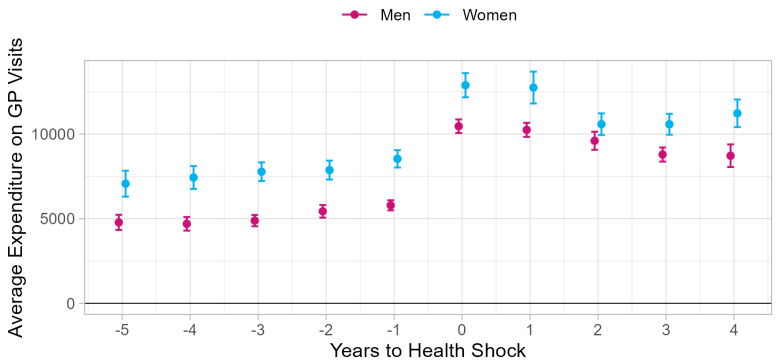


Figure 12: Raw trend of number of expenditure on GP visits.

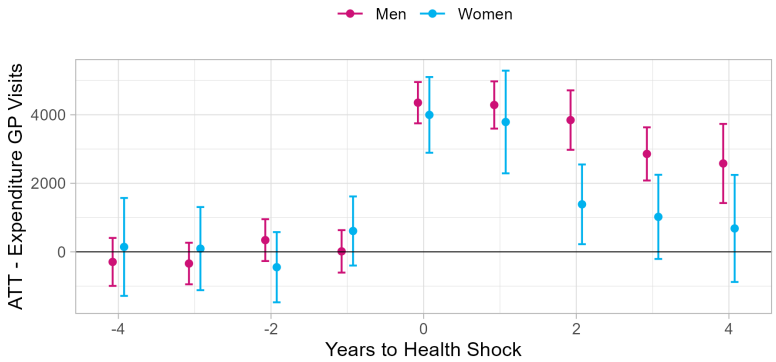


Figure 13: Dynamic ATT for expenditure on GP visits.

7.7 Group-specific effects

Lastly, we examine group-specific ATTs, which represent the aggregated ATTs for different treatment cohorts based on the year in which they were treated. This approach allows us to assess whether treatment effects vary across cohorts. We find that the differences of our main estimates are not driven by particular groups as shown in Figure 35 in Appendix D.8.

8 Conclusion

This paper provides the first causal evidence of gender differences in healthcare utilisation following severe, unanticipated health shocks. Using rich administrative data from Denmark and a staggered difference-in-differences design, we find that men consistently use more healthcare than women—particularly in the form of statin consumption and GP visits—in the years following a first-time, non-fatal heart attack or stroke. These effects persist for up to five years post-shock and are statistically significant. Importantly, the gender difference is not observed in hospitalization rates, which are less subject to individual behaviour, suggesting the disparity

is rooted in health-seeking behaviour or provider discretion rather than underlying health status or access constraints. Healthcare expenditure patterns reflect these utilisation gaps: men incur higher statin costs, while GP visit fees show no systematic gender difference.

Our findings challenge common assumptions embedded in the male-female health-survival paradox. While women live longer and report worse health, their lower responsiveness to acute health shocks calls into question the notion that they engage more actively with healthcare services. These results also stand in contrast to prior descriptive literature, which often documents higher female use of routine and preventive care.

We subject our results to a range of robustness checks. Estimates from matched samples, alternative estimators, and control group definitions remain consistent with our main findings. In addition, we explore several mechanisms that could drive the observed gender gap. The results are not explained by differential survival rates, household composition, or cohort effects. We find that high-income men visit general practitioners more frequently than high-income women, and that men with less severe health shocks increase their healthcare use more than similarly affected women. Notably, we find that women are likely prescribed statins at lower rates than men, rather than discontinuing treatment or failing to fill prescriptions—suggesting both demand- and supply-side influences.

Our findings, while robust within the Danish context, may not generalise to countries with different healthcare systems, gender norms, or institutional barriers. Further research is needed to disentangle the relative contributions of patient preferences versus provider decisions in shaping treatment uptake—particularly in prescription behaviour and follow-up care.

The policy implications of these findings are clear. Even in a universal healthcare system that ranks high on gender equity, men and women respond differently to major health events. These results suggest that even well-designed systems require tailored strategies to close gender gaps, as uniform policies may fail to accommodate meaningful variation in patient behaviour and provider response. More broadly, recognising and addressing behavioural, situational, and institutional sources of gender disparities in care is critical for improving health equity across the life course.

References

Baji, P. and A. Bíró (2018). Adaptation or Recovery After Health Shocks? Evidence Using Subjective and Objective Health Measures. *Health Economics 27*(5), 850–864.

Blaakilde, A. L. (2012). Døde Mænd og Syge Kvinder. Køn, Alder og Ulighed i Sundhed. *Kvinder, Køn & Forskning* (4).

Böckerman, P., M. Kortelainen, H. Salokangas, and M. Vaalavuo (2025). A Family Affair? Long-Term Economic and Mental Health Effects of Spousal Cancer. *Journal of Population Economics 38*(1), 19.

Borusyak, K., X. Jaravel, and J. Spiess (2024). Revisiting Event-Study Designs: Robust and Efficient Estimation. *Review of Economic Studies*, rdae007.

Cabral, M. and M. Dillender (2024). Gender Differences in Medical Evaluations: Evidence From Randomly Assigned Doctors. *American Economic Review 114*(2), 462–499.

Callaway, B. and P. H. Sant'Anna (2021). Difference-in-Differences with Multiple Time Periods. *Journal of Econometrics 225*(2), 200–230.

Case, A. and C. Paxson (2005). Sex Differences in Morbidity and Mortality. *Demography 42*(2), 189–214.

Chandra, A. and D. O. Staiger (2007). Productivity Spillovers in Health Care: Evidence From the Treatment of Heart Attacks. *Journal of Political Economy 115*(1), 103–140.

Chandra, A. and D. O. Staiger (2020). Identifying Sources of Inefficiency in Healthcare. *The Quarterly Journal of Economics 135*(2), 785–843.

Conti, G., R. Ginja, P. Persson, and B. Willage (2025). The Menopause" Penalty". Technical report, National Bureau of Economic Research.

Cutler, D. M. and E. Glaeser (2005). What Explains Differences in Smoking, Drinking, and Other Health-Related Behaviors? *American Economic Review 95*(2), 238–242.

De Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–2996.

Doyle, J. J. (2011). Returns to Local-Area Health Care Spending: Evidence From Health Shocks to Patients Far From Home. *American Economic Journal: Applied Economics 3*(3), 221–43.

Fadlon, I. and T. H. Nielsen (2019, September). Family Health Behaviors. *American Economic Review 109*(9), 3162–91.

Fadlon, I. and T. H. Nielsen (2021). Family Labor Supply Responses to Severe Health Shocks: Evidence From Danish Administrative Records. *American Economic Journal: Applied Economics 13*(3), 1–30.

Galdas, P. M., F. Cheater, and P. Marshall (2005). Men and Health Help-Seeking Behaviour: Literature Review. *Journal of Advanced Nursing 49*(6), 616–623.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics 225*(2), 254–277.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy 80*(2), 223–255.

Grossman, M. (2000). The Human Capital Model. In *Handbook of Health Economics*, Volume 1, pp. 347–408. Elsevier.

Hirsh, B. J. (2019). Guidelines for Lipid-Lowering in Women – What Have We Learned? Accessed: 2025-09-28.

Höhn, A., J. Gampe, R. Lindahl-Jacobsen, K. Christensen, and A. Oksuyzan (2020). Do Men Avoid Seeking Medical Advice? A Register-Based Analysis of Gender-Specific Changes in Primary Healthcare Use After First Hospitalisation at Ages 60+ in Denmark. *J Epidemiol Community Health 74*(7), 573–579.

Juel, K. and K. Christensen (2007). Are Men Seeking Medical Advice Too Late? Contacts to General Practitioners and Hospital Admissions in Denmark 2005. *Journal of Public Health 30*(1), 111–113.

Kiss, P. A., A. Uijl, A. R. de Boer, T. C. Duk, D. E. Grobbee, M. Hollander, E. Smits, M. C. Sturkenboom, and S. A. Peters (2024). Sex Differences in the Intensity of Statin Prescriptions at Initiation in a Primary Care Setting. *Heart 110*(15), 981–987.

Legato, M. J., P. A. Johnson, and J. E. Manson (2016). Consideration of Sex Differences in Medicine to Improve Health Care and Patient Outcomes. *Jama 316*(18), 1865–1866.

Nanna, M. G., T. Y. Wang, Q. Xiang, A. C. Goldberg, J. G. Robinson, V. L. Roger, S. S. Virani, P. W. Wilson, M. J. Louie, A. Koren, et al. (2019). Sex Differences in the Use of Statins in Community Practice: Patient and Provider Assessment of Lipid Management Registry. *Circulation: Cardiovascular Quality and Outcomes 12*(8), e005562.

Nobili, A., C. Franchi, L. Pasina, M. Tettamanti, M. Baviera, L. Monesi, C. Roncaglioni, E. Riva, U. Lucca, A. Bortolotti, et al. (2011). Drug Utilization and Polypharmacy in an Italian Elderly Population: The EPIFARM-Elderly Project. *Pharmacoepidemiology and Drug Safety 20*(5), 488–496.

Oksuzyan, A., R. Jacobsen, K. Glaser, C. Tomassini, J. W. Vaupel, and K. Christensen (2011). Sex Differences in Medication and Primary Healthcare Use Before and After Spousal Bereavement at Older Ages in Denmark: Nationwide Register Study of Over 6000 Bereavements. *Journal of Aging Research 2011*.

Oksuzyan, A., K. Juel, J. W. Vaupel, and K. Christensen (2008). Men: Good Health and High Mortality. Sex Differences in Health and Aging. *Aging Clinical and Experimental Research 20*(2), 91–102.

Özcan, C., K. Juel, J. F. Lassen, L. M. von Kappelgaard, P. E. Mortensen, and G. Gislason (2016). The Danish Heart Registry. *Clinical Epidemiology 8*, 503.

Pedersen, P. V., K. B. H. Johansen, O. Ekholm, and K. Juel (2014). Sundhed og Trivsel i et Kønsperspektiv. *Statens Institut for Folkesundhed, Det Sundhedsvidenskabelige Fakultet*.

Sant'Anna, P. H. and J. Zhao (2020). Doubly Robust Difference-in-Differences Estimators. *Journal of Econometrics 219*(1), 101–122.

Schmidt, M., S. A. J. Schmidt, J. L. Sandegaard, V. Ehrenstein, L. Pedersen, and H. T. Sørensen (2015). The Danish National Patient Registry: A Review of Content, Data Quality, and Research Potential. *Clinical Epidemiology 7*, 449.

Schulman, K. A., J. A. Berlin, W. Harless, J. F. Kerner, S. Sistrunk, B. J. Gersh, R. Dube, C. K. Taleghani, J. E. Burke, S. Williams, et al. (1999). The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization. *New England Journal of Medicine 340*(8), 618–626.

Sibbritt, D., J. Bayes, W. Peng, and J. Adams (2024). Demographic Factors Affect Stroke-Related Healthcare Utilisation Among Australian Stroke Survivors. *Scientific Reports 14*(1), 21241.

Smith, D. B., P. Murphy, P. Santos, M. Phillips, and M. Wilde (2009). Gender Differences in the Colorado Stroke Registry. *Stroke 40*(4), 1078–1081.

Sun, L. and S. Abraham (2021). Estimating Dynamic Treatment Effects in Event Studies With Heterogeneous Treatment Effects. *Journal of Econometrics 225*(2), 175–199.

van Loenen, T., M. J. van den Berg, M. J. Faber, and G. P. Westert (2015). Propensity to Seek Healthcare in Different Healthcare Systems: Analysis of Patient Data in 34 Countries. *BMC health services research 15*(1), 465.

Witting, C., A. Devareddy, and F. Rodriguez (2022). Review of Lipid-Lowering Therapy in Women from Reproductive to Postmenopausal Years. *Reviews in cardiovascular medicine 23*(5), 183.

World Bank (2019). Women, Business and the Law: A Decade of Reform. World Bank Group. Available at: https://openknowledge.worldbank.org/bitstream/handle/10986/31327/WBL2019.pdf.

Appendix

A Trend plots

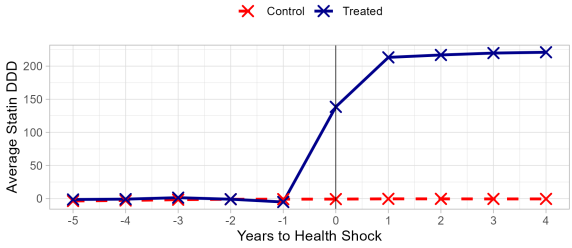


Figure 14: Statin DDD - Women

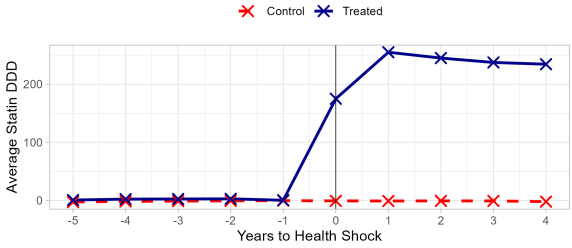


Figure 15: Statin DDD - Men

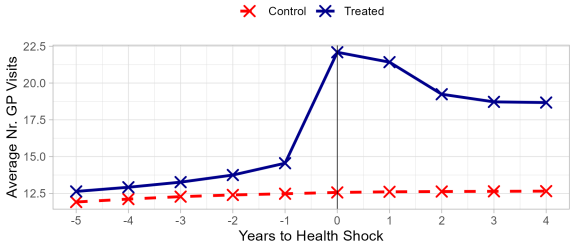


Figure 16: GP Visits - Women

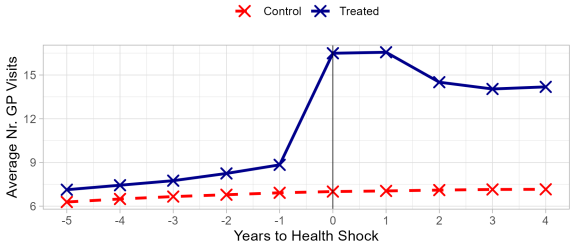


Figure 17: GP Visits - Men

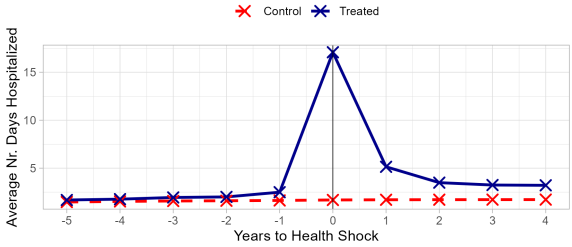


Figure 18: Days Treated - Women



Figure 19: Days Treated - Men

Figure 20: Parallel trends for control and treated group, where control group are the not-yet-treated with restricted time window (non-shifting), here added as a placebo group. De-trended with year fixed effects for men and women separately.

B Tables corresponding to main results/figures Callaway and Sant'Anna estimator

Table 5: Statin DDD Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	-0.37	-3.28	2.53	12927	0.75
-3	1.96	-1.72	5.65	13616	0.16
-2	-1.46	-4.13	1.20	14313	0.13
-1	1.59	-1.34	4.53	14342	0.15
0	151.76	145.55	157.96	15140	0.00
1	234.14	223.37	244.92	14363	0.00
2	240.36	226.21	254.51	13046	0.00
3	243.78	228.59	258.97	11934	0.00
4	242.62	225.82	259.41	10918	0.00
Panel B: Men					
-4	0.20	-1.32	1.71	24934	0.74
-3	-1.28	-2.71	0.16	26468	0.02
-2	0.24	-1.33	1.80	28110	0.70
-1	2.94	1.26	4.62	28307	0.00
0	185.29	181.01	189.56	30103	0.00
1	271.23	264.23	278.23	28547	0.00
2	261.61	253.39	269.84	25945	0.00
3	254.04	244.75	263.34	23790	0.00
4	250.89	240.41	261.38	21944	0.00

Table 6: GP Visits Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	-0.18	-0.44	0.08	12927	0.08
-3	-0.11	-0.39	0.17	13616	0.29
-2	0.08	-0.23	0.39	14313	0.51
-1	0.56	0.28	0.85	14342	0.00
0	7.20	6.80	7.59	15140	0.00
1	6.40	5.95	6.86	14363	0.00
2	3.98	3.51	4.45	13046	0.00
3	2.94	2.44	3.44	11934	0.00
4	2.22	1.73	2.70	10918	0.00
Panel B: Men					
-4	-0.06	-0.22	0.10	24934	0.31
-3	-0.06	-0.22	0.10	26468	0.34
-2	0.11	-0.05	0.27	28110	0.05
-1	0.23	0.06	0.40	28307	0.00
0	7.54	7.32	7.75	30103	0.00
1	7.45	7.19	7.71	28547	0.00
2	5.12	4.85	5.38	25945	0.00
3	4.28	4.01	4.56	23790	0.00
4	3.96	3.66	4.26	21944	0.00

Table 7: Days Hospitalised Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	-0.10	-0.30	0.11	12927	0.19
-3	0.01	-0.21	0.23	13616	0.91
-2	-0.12	-0.32	0.09	14313	0.11
-1	0.45	0.20	0.70	14342	0.00
0	14.65	14.11	15.21	15140	0.00
1	2.55	2.16	2.94	14363	0.00
2	0.83	0.50	1.16	13046	0.00
3	0.47	0.13	0.82	11934	0.00
4	0.07	-0.28	0.42	10918	0.57
Panel B: Men					
-4	-0.08	-0.22	0.06	24934	0.11
-3	-0.09	-0.22	0.04	26468	0.07
-2	0.10	-0.04	0.24	28110	0.05
-1	0.17	0.01	0.33	28307	0.00
0	15.03	14.66	15.40	30103	0.00
1	2.79	2.54	3.04	28547	0.00
2	1.26	1.03	1.49	25945	0.00
3	1.06	0.83	1.29	23790	0.00
4	0.84	0.62	1.06	21944	0.00

C Tables and Figures for Robustness Checks

C.1 Matched sample

Nearest neighbour matching on propensity score using logit model for estimation in the pre-treatment year.

	Means Men	Means Women	Std Mean Diff	Var. Ratio	eCDF Mean	eCDF Max	Std. Pair Dist.
Distance	0.37	0.37	0.01	1.02	0.00	0.01	0.01
Region Capital	0.28	0.29	-0.01		0.00	0.00	0.86
Region Central Jutland	0.21	0.21	0.01		0.00	0.00	0.81
Region North Jutland	0.13	0.12	0.01		0.00	0.00	0.64
Region South Denmark	0.22	0.22	-0.00		0.00	0.00	0.82
Region Zealand	0.16	0.16	-0.01		0.00	0.00	0.72
Married couple	0.53	0.52	0.02		0.01	0.01	0.88
Multi-family households	0.09	0.10	-0.00		0.00	0.00	0.58
Other couples	0.07	0.07	0.00		0.00	0.00	0.12
Single	0.31	0.31	-0.02		0.01	0.01	0.70
Income (DKK)	176338	179615	-0.03	0.98	0.01	0.04	0.59
Age	59.94	59.90	0.01	1.01	0.00	0.01	1.11
Education in months	138.11	137.75	0.01	0.97	0.04	0.14	0.70

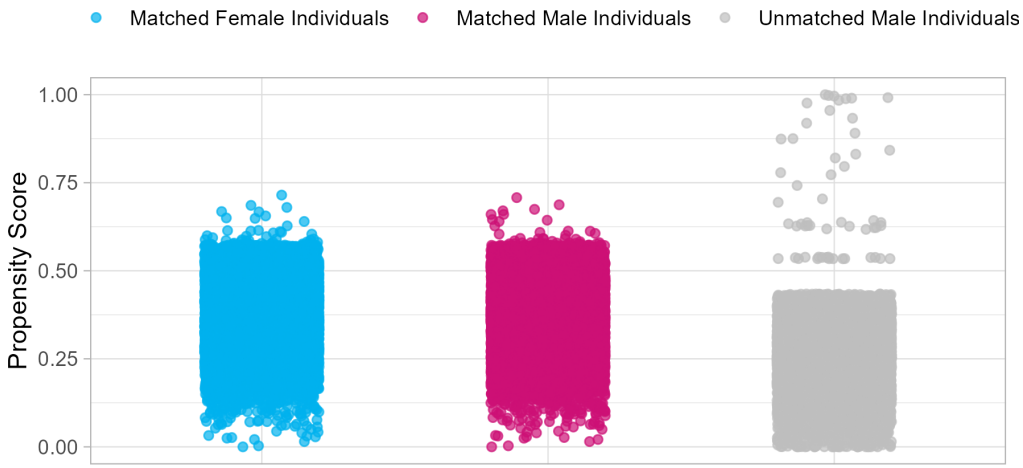


Figure 21: Propensity score distribution before and after matching to illustrate overlap between men and women.

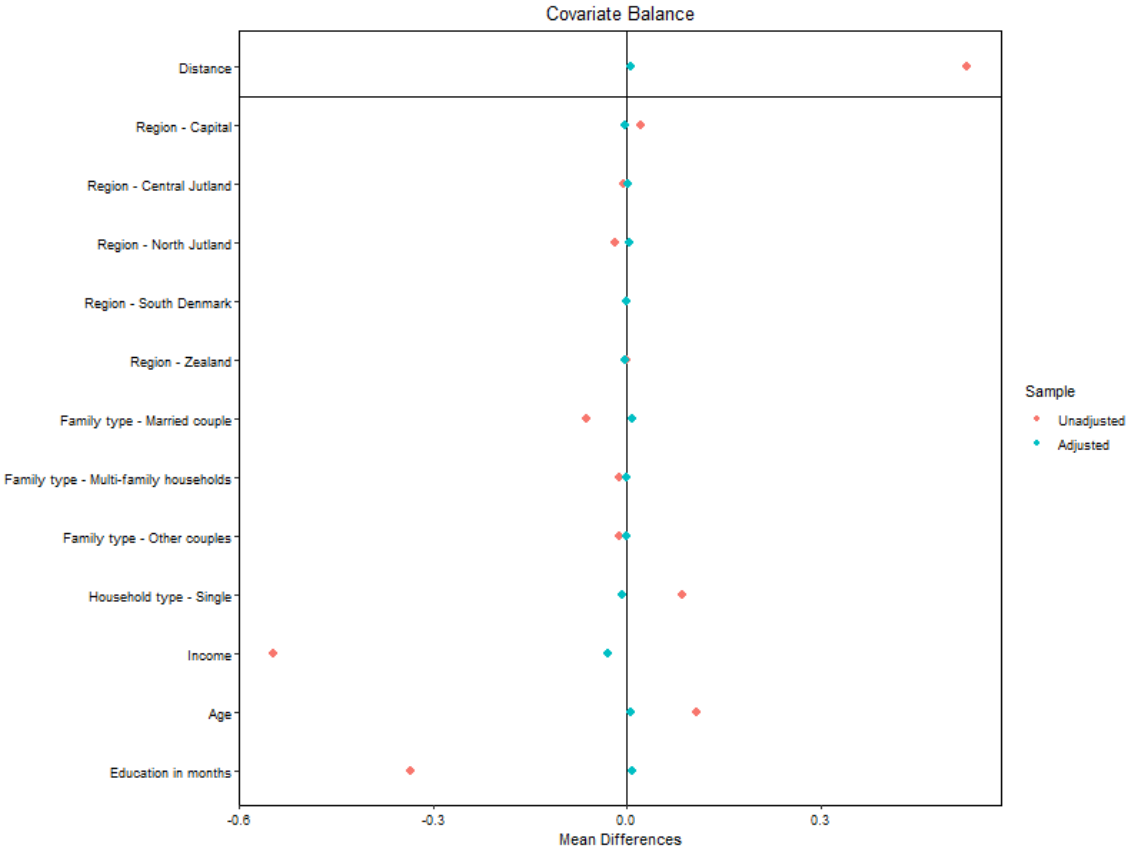
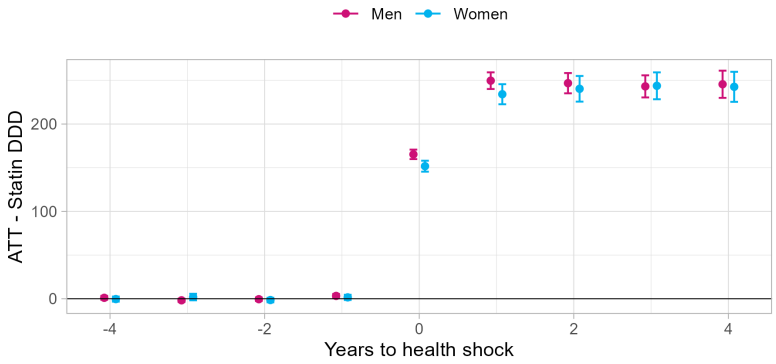
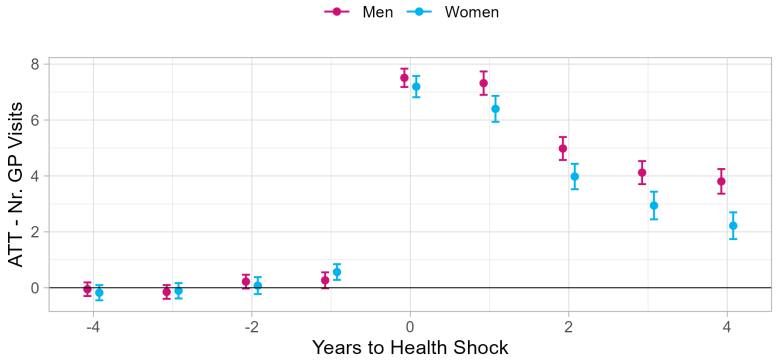


Figure 22: Standardized mean differences of baseline covariates before and after propensity score matching to asses covariate balance between men and women.

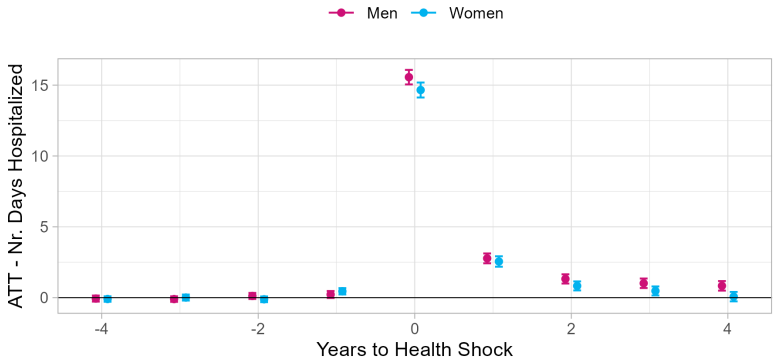
C.2 Results for matched sample



(a) Dynamic ATT for statin DDD in matched sample.



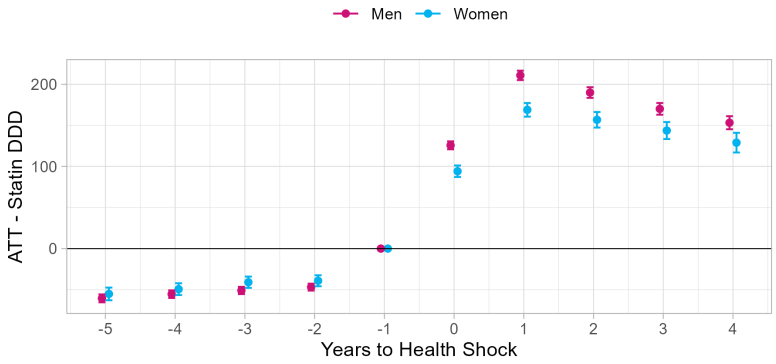
(b) Dynamic ATT for number of GP visits in matched sample.



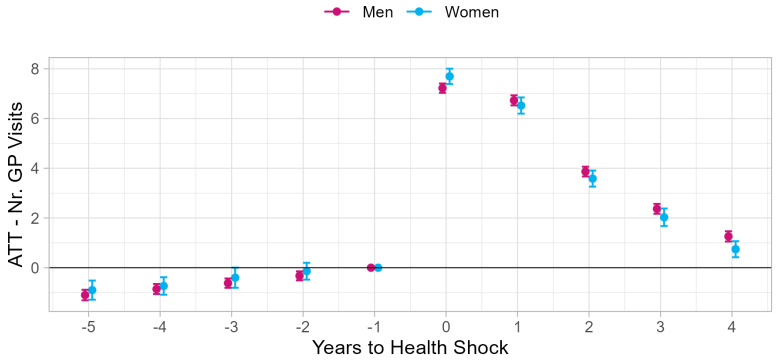
(c) Dynamic ATT for days hospitalized in matched sample.

Figure 23: Matched sample results for Callaway and Sant'Anna (2021) estimator. Dynamic ATT plots for various healthcare measures. Each panel shows the ATT for a specific measure, illustrating the impact of interventions over time. Simultaneous 95% confidence bands obtained through bootstrapping are depicted.

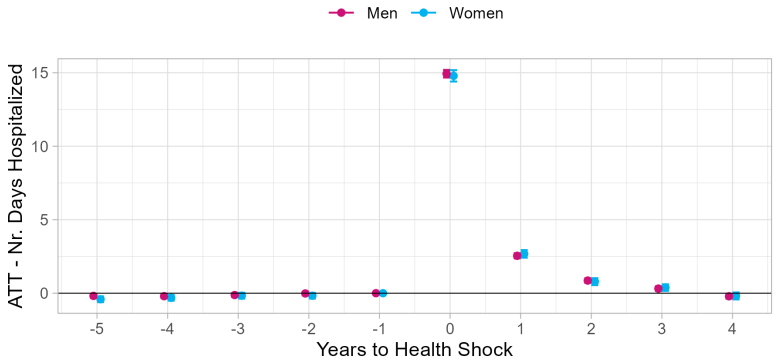
C.3 Dynamic TWFE



(a) Dynamic ATT for statin DDD.



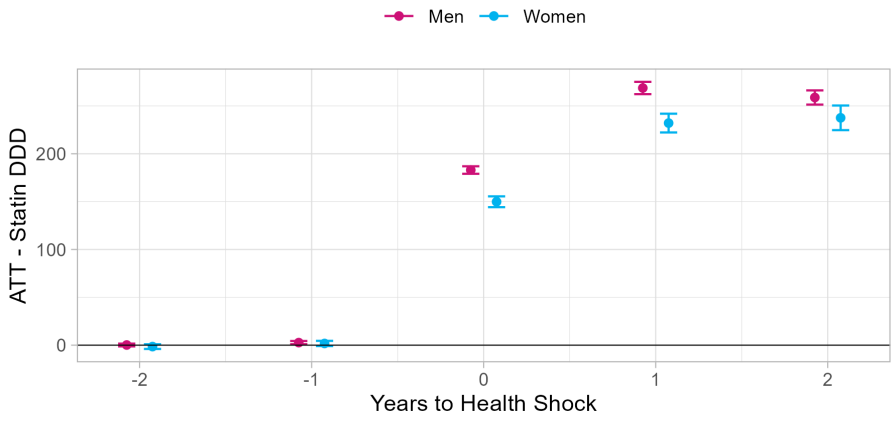
(b) Dynamic ATT for number of GP visits.



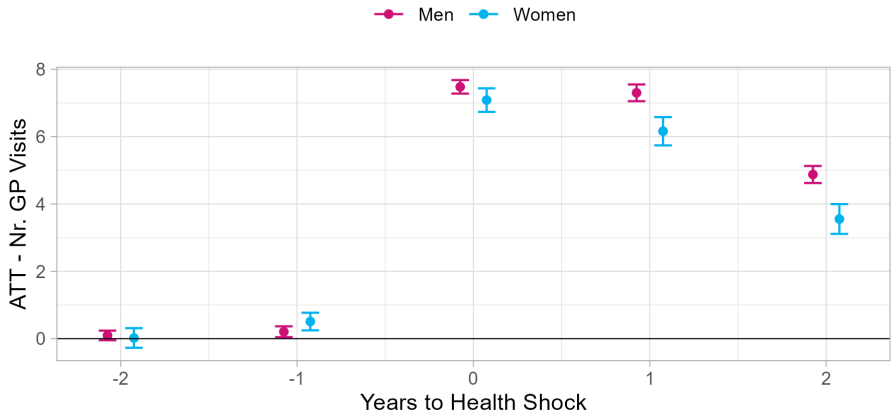
(c) Dynamic ATT for days hospitalized.

Figure 24: Dynamic TWFE Estimates as in Fadlon and Nielsen (2019).

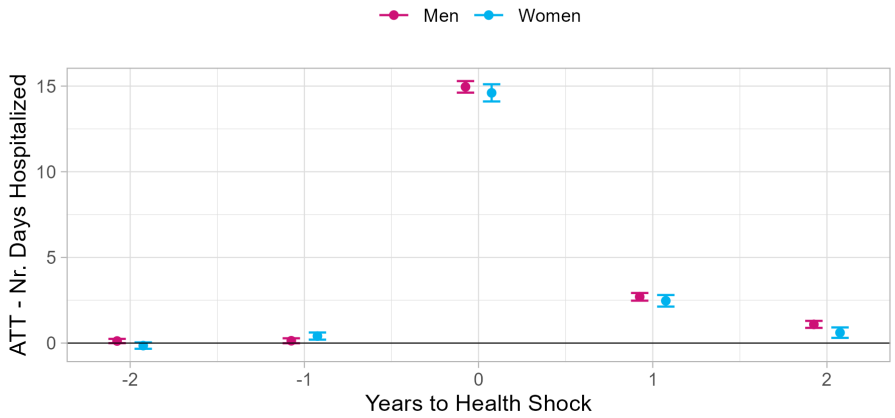
C.4 Robustness check: 3-year window



(a) Dynamic ATT for statin DDD.



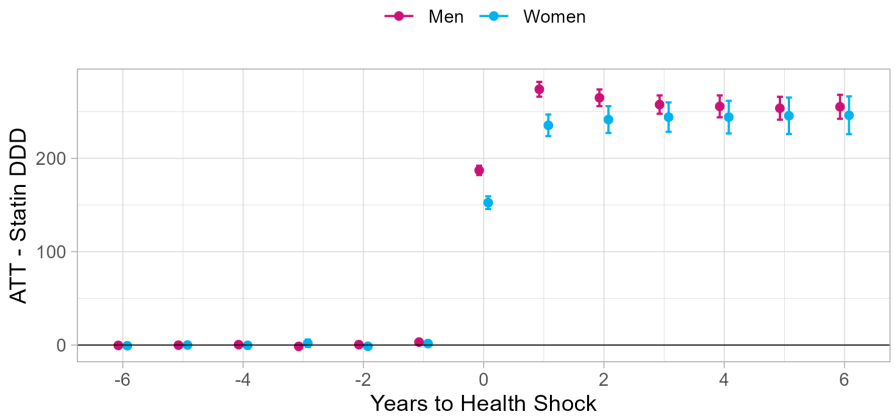
(b) Dynamic ATT for number of GP visits.



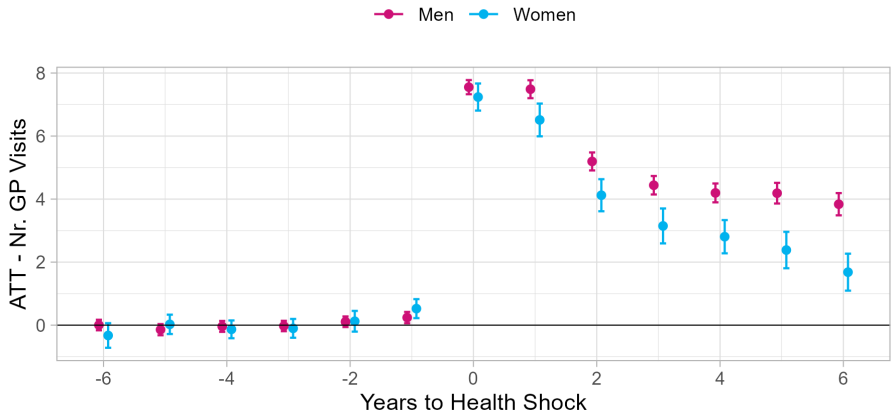
(c) Dynamic ATT for days hospitalized.

Figure 25: Dynamic ATT estimates using a 3-year window.

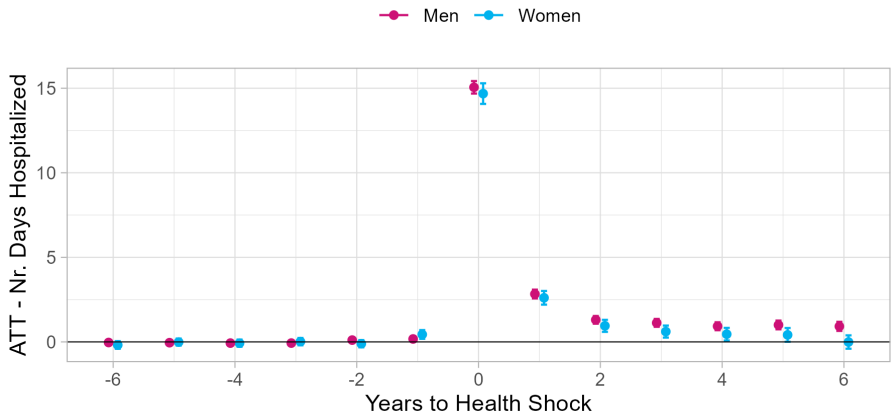
C.5 Robustness check: 7-year window



(a) Dynamic ATT for statin DDD.



(b) Dynamic ATT for number of GP visits.



(c) Dynamic ATT for days hospitalized.

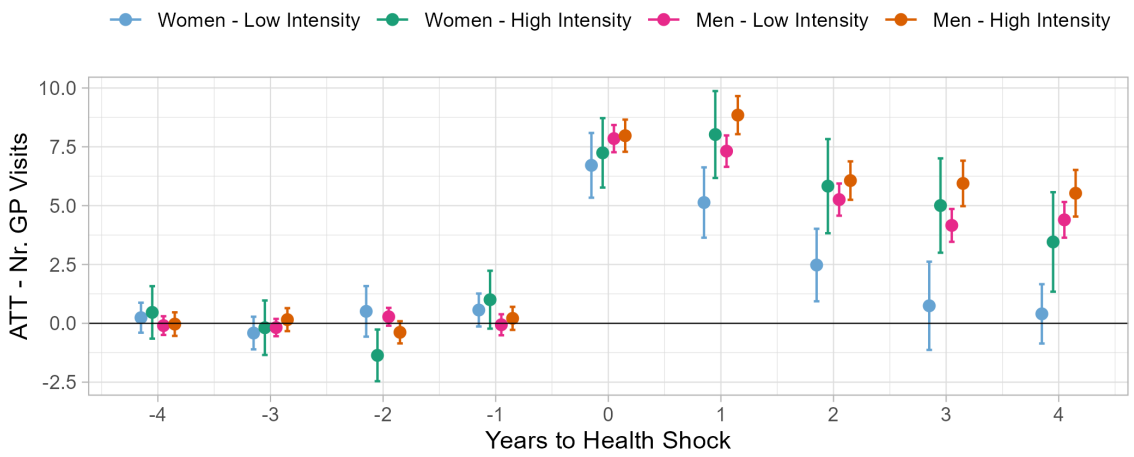
Figure 26: Dynamic ATT estimates using a 7-year window.

D Tables and Figures for Mechanisms

D.1 Intensity of Health Shock



(a) Dynamic ATT for statin DDD.



(b) Dynamic ATT for number of GP visits.



(c) Dynamic ATT for days hospitalized.

Figure 27: Dynamic ATT estimates by intervention intensity group. Each panel corresponds to a different outcome and includes simultaneous 95% confidence bands.

D.2 Balanced Panel

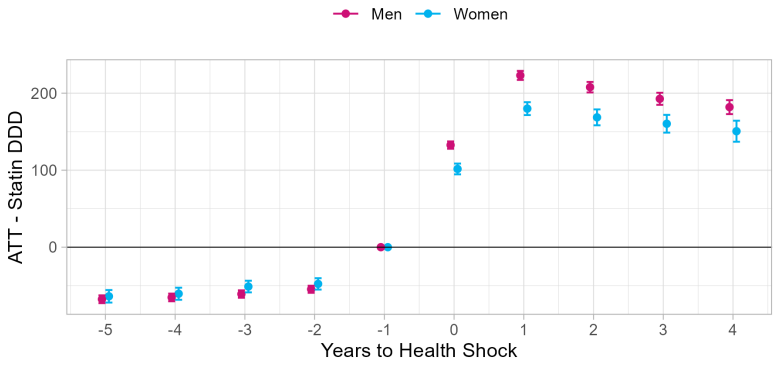


Figure 28: Dynamic ATT for statin DDD.

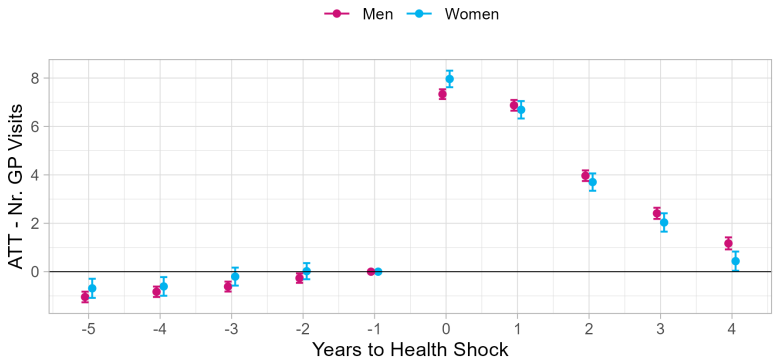


Figure 29: Dynamic ATT for GP Visits.

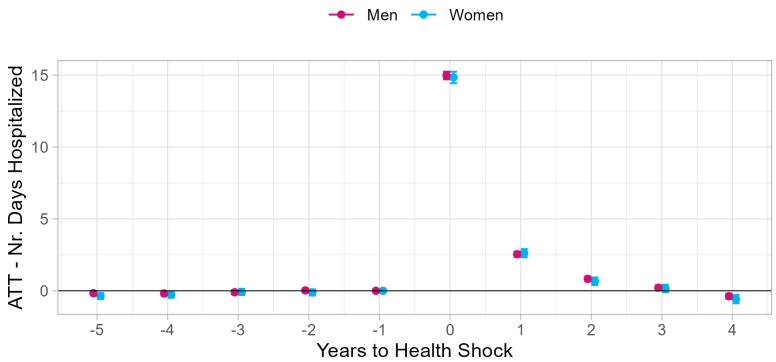
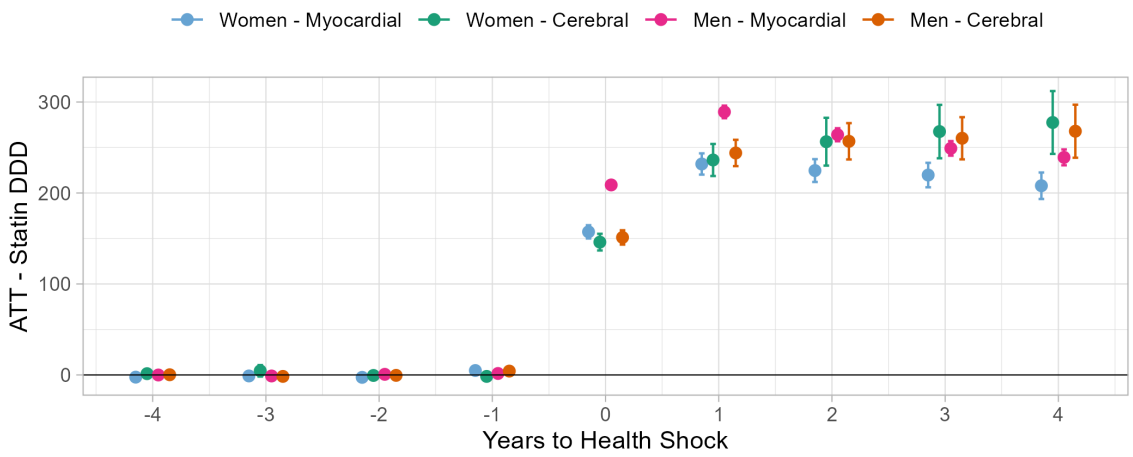


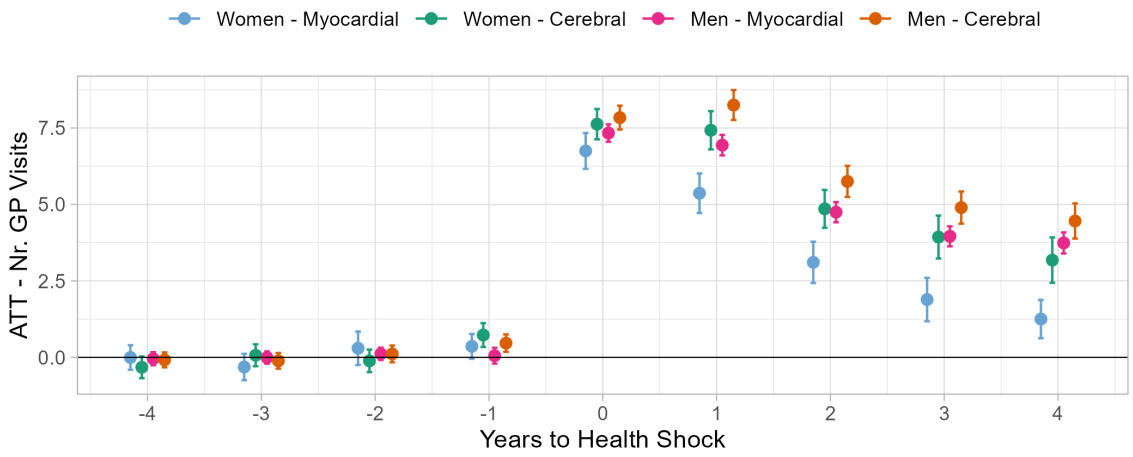
Figure 30: Dynamic ATT for days hospitalized.

Figure 31: Dynamic TWFE Estimates as Fadlon and Nielsen (2019): Balanced Panel. Note that balanced refers to the time window: individuals who are observed over all 5 pre-treatment and all 5 post-treatment periods are included.

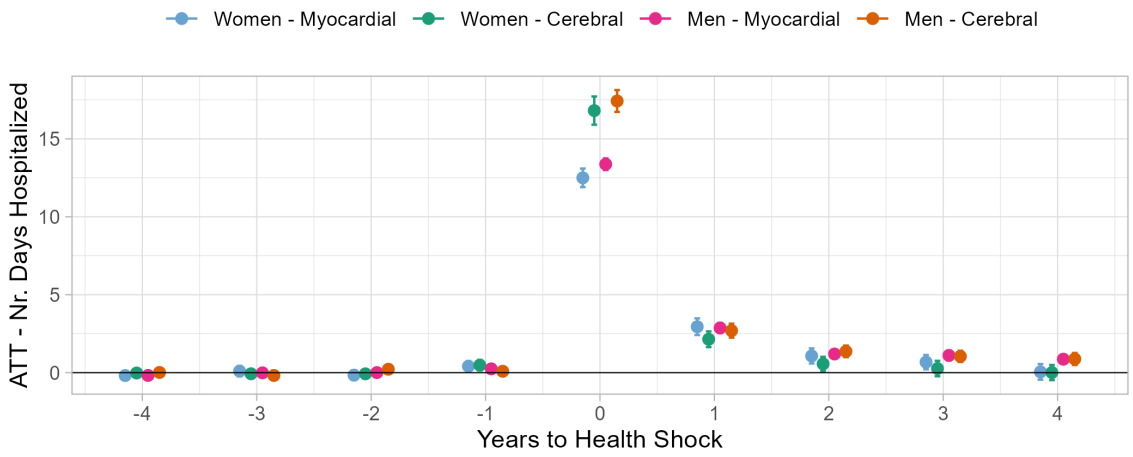
D.3 Results for cerebral versus myocardial infarction



(a) Dynamic ATT for statin DDD.



(b) Dynamic ATT for number of GP visits.



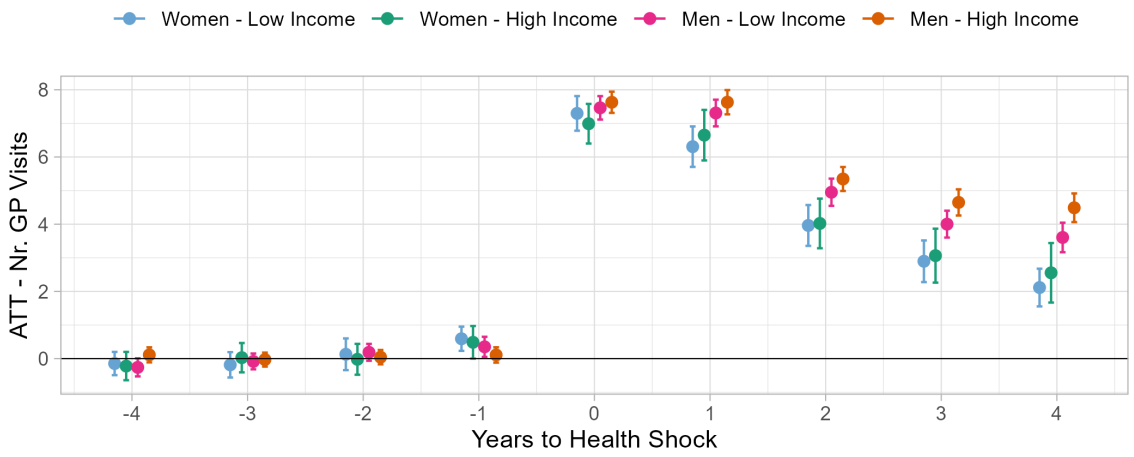
(c) Dynamic ATT for days hospitalized.

Figure 32: Results with Callaway and Sant'Anna (2021) estimator for myocardial and cerebral infarction (the two treatments). Each panel presents dynamic ATT estimates for a different healthcare outcome.

D.4 Results for low- versus high-income individuals



(a) Dynamic ATT for statin DDD.



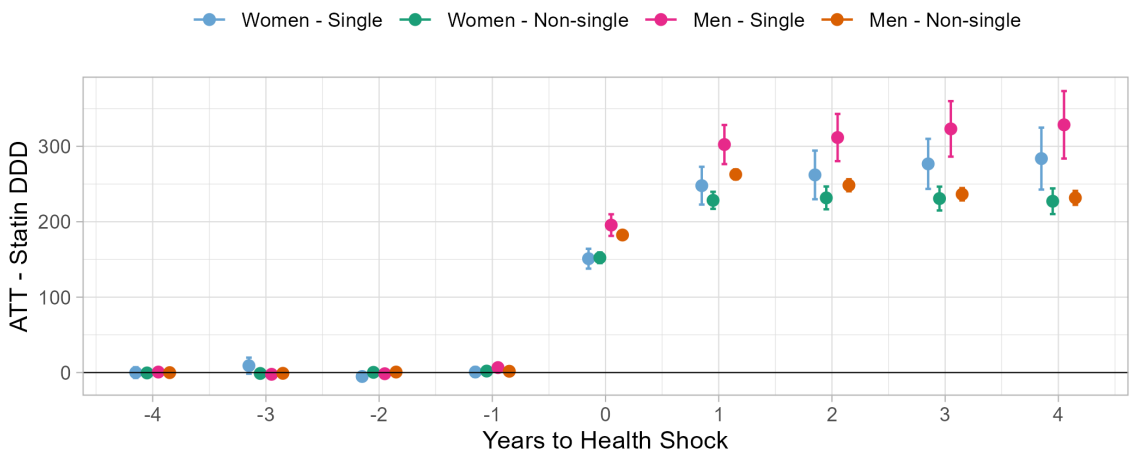
(b) Dynamic ATT for number of GP visits.



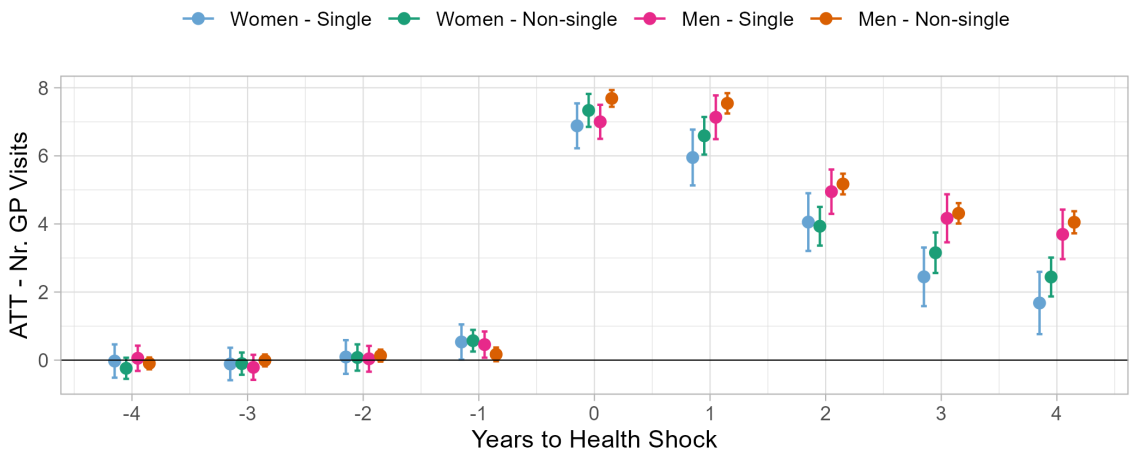
(c) Dynamic ATT for days hospitalized.

Figure 33: Results with Callaway and Sant'Anna (2021) estimator for low-income and high-income individuals (based on median income in full sample).

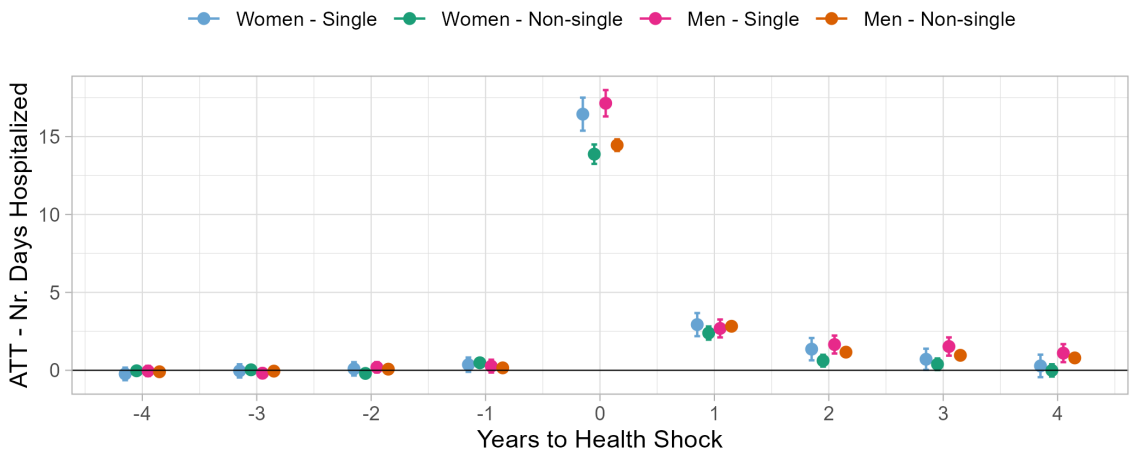
D.5 Results for single versus non-single individuals



(a) Dynamic ATT for Statin DDD.



(b) Dynamic ATT for number of GP visits.



(c) Dynamic ATT for days hospitalized.

Figure 34: Results with Callaway and Sant'Anna (2021) estimator for single and non-single individuals. Each panel displays dynamic ATT estimates with simultaneous 95% confidence bands.

D.6 Statin Pickup

Table 8: Statin Pickup Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	-0.00	-0.02	0.01	24934	0.43
-3	-0.01	-0.02	0.01	26468	0.16
-2	0.00	-0.01	0.02	28110	0.84
-1	0.02	0.00	0.03	28307	0.00
0	1.37	1.34	1.40	30103	0.00
1	2.10	2.05	2.15	28547	0.00
2	2.06	2.01	2.12	25945	0.00
3	2.04	1.98	2.10	23790	0.00
4	2.01	1.95	2.08	21944	0.00
Panel B: Men					
-4	-0.01	-0.03	0.01	12927	0.29
-3	-0.00	-0.02	0.02	13616	0.82
-2	-0.01	-0.04	0.01	14313	0.08
-1	0.04	0.02	0.06	14342	0.00
0	1.26	1.21	1.30	15140	0.00
1	1.99	1.91	2.06	14363	0.00
2	2.03	1.94	2.12	13046	0.00
3	2.06	1.96	2.16	11934	0.00
4	2.02	1.92	2.13	10918	0.00

D.7 Healthcare expenditure

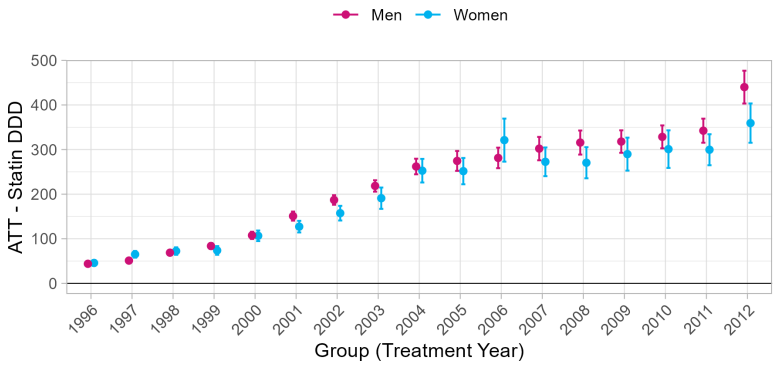
Table 9: Healthcare Expenditure on Statins Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	0.17	-3.72	4.06	12927	0.91
-3	0.78	-3.83	5.38	13616	0.63
-2	-3.54	-7.84	0.77	14313	0.02
-1	2.68	-1.82	7.17	14342	0.10
0	147.61	138.05	157.18	15140	0.00
1	218.18	205.16	231.20	14363	0.00
2	208.16	194.65	221.66	13046	0.00
3	193.05	179.33	206.77	11934	0.00
4	178.60	164.04	193.16	10918	0.00
Panel B: Men					
-4	0.92	-2.54	4.38	24934	0.47
-3	-1.61	-4.99	1.77	26468	0.19
-2	-1.52	-4.89	1.85	28110	0.21
-1	2.79	-0.43	6.01	28307	0.02
0	184.01	176.28	191.74	30103	0.00
1	267.82	257.63	278.02	28547	0.00
2	250.34	239.82	260.87	25945	0.00
3	227.56	216.81	238.31	23790	0.00
4	210.80	199.29	222.31	21944	0.00

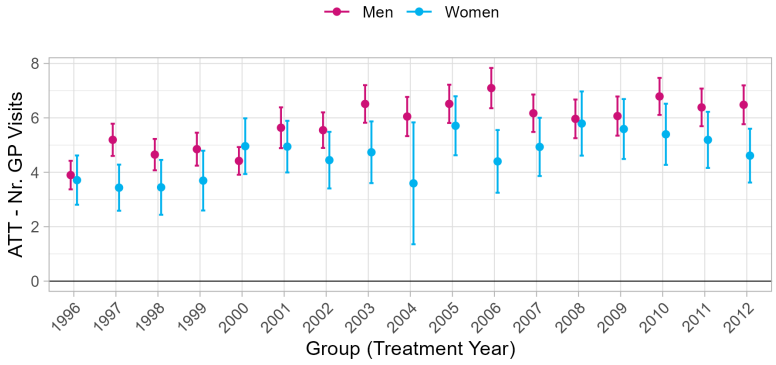
Table 10: Healthcare Expenditure on GP Visits Dynamic ATT

Event Time	Avg ATT	CI-Lower	CI-Upper	No uniq treated	p-value
Panel A: Women					
-4	143.50	-1285.31	1572.32	12927	0.78
-3	94.24	-1119.01	1307.48	13616	0.84
-2	-448.47	-1473.15	576.21	14313	0.23
-1	608.61	-399.79	1617.01	14342	0.10
0	3997.68	2892.60	5102.75	15140	0.00
1	3790.05	2291.64	5288.47	14363	0.00
2	1386.85	223.43	2550.28	13046	0.00
3	1021.10	-207.49	2249.68	11934	0.02
4	683.02	-879.70	2245.74	10918	0.21
Panel B: Men					
-4	-294.08	-994.09	405.94	24934	0.25
-3	-340.93	-947.54	265.68	26468	0.12
-2	342.32	-268.35	953.00	28110	0.13
-1	13.76	-605.79	633.32	28307	0.96
0	4353.89	3751.14	4956.64	30103	0.00
1	4286.78	3597.15	4976.42	28547	0.00
2	3845.77	2977.85	4713.69	25945	0.00
3	2858.35	2082.19	3634.51	23790	0.00
4	2579.34	1424.38	3734.30	21944	0.00

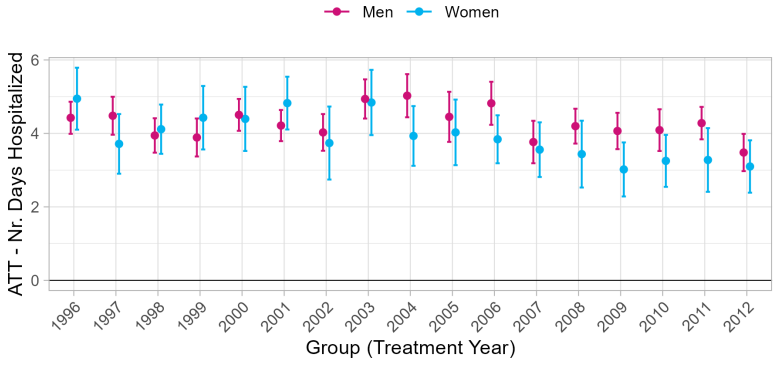
D.8 Cohort-specific (group-specific) ATTs



(a) Group ATT for statin DDD.



(b) Group ATT for GP visits.



(c) Group ATT for days hospitalized.

Figure 35: Group-specific treatment effects for different outcomes. Each panel displays group average treatment effects with associated uncertainty.