# All in a Deductible's Work:

# Cost-Sharing, Chronic Conditions, and Health

Saumya Chatrath*

November 30, 2021

**Abstract**

This paper measures the health effects of a large, exogenous increase in the plan deductible on patients with diabetes and high cholesterol. I find that an on average if-random increase in the individual deductible of $1,000 does not affect the health of the average patient with these conditions. However, when I divide patients on the basis of their underlying disease severity, I find that the deductible increase worsens the health of patients that were initially in control of their condition. Among diabetics with initial glycemic control, the deductible change causes a 2.3 percentage point, or a 19% increase relative to the mean in the probability of having a high HbA1c value ($> 7.5\%$). Among patients with stable baseline cholesterol, the deductible change causes a 0.5 percentage point, or 12.5%, increase relative to the mean in the probability of having a high cholesterol value ($> 160$ mg/dl). I investigate the mechanisms that drive these results and find that it is driven by the differential impact of the deductible increase on primary care use, and in short-term reductions of high value prescription drugs. These results have implications for optimal plan design and programs that target chronic disease management.

*I am grateful for the advice and support of my dissertation committee David Cutler, Michael Chernew, and Benjamin Sommers. For their helpful comments, I would also like to thank participants at the Harvard Health Care Policy Department's Health Economics Seminar and participants at the 2021 ASHEcon Conference. I would also like to acknowledge the kind support and grants from the Health Policy and Insurance Group in the Department of Population Medicine at Harvard University

# 1 Introduction

Estimating the health effects of increases in cost-sharing has long been a challenging empirical issue and remains far from being a settled topic. The implications of this open question are enormous. Insurers and employers have increased deductibles, coinsurance rates, and copays in insurance plan design to control health care spending growth on the premise that population health will be unaffected. Can cost-sharing increases that are designed to restrain over-use of medical services lead to worse health outcomes amongst patients? To answer this question, I use quasi-experimental methods to estimate the causal impact of exogenous increases in medical deductibles on clinical health outcomes. So far, the majority of research on this topic, described in detail in Section 2, has been unable to estimate proximate clinical effects of plan design changes. These null findings have provided justification for insurers and employers to further shift the cost-burden of insurance plans onto patients. Consistent with this, I also obtain null estimates when I focus my analysis on the general population. A novel finding of this paper is that there is heterogeneity by baseline health in the negative impact of cost-sharing increases. Contrary to what we would expect, patients with better underlying health are the ones that bear the brunt of negative consequences.

My study period, which extends from 2004 to 2017, provides a rapidly evolving backdrop for these types of questions. I describe developments in insurance plan design, and in the treatment of chronic conditions in Section 3. In my analysis, I leverage a large data set of medical claims and laboratory values to quantify the effect of a large change in a plan's medical deductible on the health of patients with diabetes and hyperlipidemia (i.e. high cholesterol). Section 5 describes the data and analytic sample that I use. Like the plurality of individuals in U.S., the patients in my sample are commercially insured via their employers. I focus on patients with these pre-specified chronic conditions for several reasons. First, health care spending in the United States is highly skewed. Five percent of the population accounts for half of its health expenditure (Bernard et al., 2012). Patients suffering from these conditions are some of the most expensive to treat. In fact, one study estimated

that care for people with diagnosed diabetes accounts for one in four health care dollars in the U.S. (Peterson, 2018). Second, I can use unbiased markers of health among this population to estimate the impact of plan design changes. These conditions require ongoing maintenance and monitoring by both patients and physicians. I leverage biomarker values from laboratory test results to measure whether the severity of these conditions is affected by large cost-sharing changes. Finally, I can examine sub-groups of patients that might be particularly vulnerable to deductible increases including those that live in low income areas and those that are sicker on average.

My main empirical analysis, explained in Section 4, examines non-elderly beneficiaries with employer sponsored insurance that were continuously enrolled in their plans for at least two or more plan years, and for whom I have biomarker values in the pre- and post deductible change periods. These beneficiaries all work at employers that only offer one Optum plan per plan year. This provides a setting where the deductible change is independent of enrollees' preferences, and so selection is less of a concern. I use a stacked differences-in-differences framework to compare the evolution of biomarker values in the "treatment" group - patients working at employers that increased their individual & family deductibles by $\geq$ \$500, and $\geq \$1,000$ respectively - to the "control" group of patients working at employers that did not change their deductibles. As I alluded to earlier, a novel contribution of my paper is that I can test the impact of plan design changes *conditioned on baseline disease severity*. To do so, I divide patients into two groups - those with stable control of their conditions prior to the deductible increase, and those without. I define diabetic patients with stable baseline glycemic control as having an A1c level less than 7.5% and hyperlidemia patients with stable cholesterol as having LDL levels less than 160 mg/dl. For both conditions patients with initial values above these levels are classified as having an uncontrolled baseline case of the disease. I then test the impact of the deductible change within each sub-group separately. In practice, this means that I separately estimate the impact of the price shock on treated patients vs. control patients that are observably similar with respect to baseline health. I

repeat this analysis for patients living in low income and in high income zip-codes.

Section 6.2 presents my primary set of results. I find that the deductible increase has no impact on the management of diabetes and hyperlipidemia among the general population of patients with these conditions. When I separate out the impact by baseline disease severity, I find that there are negative health consequences for patients that had stable control over their condition in the pre-period. When compared to patients in the control group with equivalent initial lab results[1], patients in the treatment group experience a statistically significant deterioration in their health as signalled by an increase in their biomarker values. Diabetes and high cholesterol patients in the treatment group have a 19% and 13% increase relative to the mean in the probability of having a high A1c and high LDL level respectively. An equivalent effect is not seen for patients that were observably sicker in the pre-period. Surprisingly, I find that there is no differential impact of the intervention on the sub-group of patients that live in low-income and high-income zip codes; both groups are mostly unaffected by the deductible increase. To understand these results better, I dig into potential mechanisms in Section 6.3. To do so, I compare the evolution of trends in medical utilization to see whether there is a differential impact of the deductible increase on categories of medical use among sub-groups. I tested three major hypotheses. First, whether the results we observe are driven by a selection in testing i.e. if only the sickest patients in the treatment group choose to get tested. Second, whether trends in the use of high value prescription drugs explain the above. Third, whether we can attribute it to patterns of primary care use. I do not find evidence for the first theory as the deductible increase does not affect the number of tests taken. Instead I find that among patients with initial glycemic or cholesterol control, patients in the treatment group reduce their visits to the PCP by 2-4%. In the short-term i.e. the quarters immediately following the deductible increase there is also a dip in prescriptions for oral anti-diabetic drugs (for diabetes patients), and statins (for high cholesterol patients). Changes in the use of both these measures corresponds to the timing of

[1]Those with A1c levels $< 7.5\%$, or LDL levels $< 160$ mg/dl

the spike in high biomarker values. This analysis has important ramifications for employers and insurers that prioritize chronic disease management programs among their enrollees.

# 2 Background & Literature

My research is situated at a turning point in the demand-side literature in health care. For many years health economists contributed to a vast literature that empirically quantified moral hazard in health care; this is the notion that health insurance lowers the marginal cost-of-care for an individual which causes health care use to rise (Pauly 1968, Cutler and Zeckhauser 2000). If patients manage to accurately assess the benefits of care weighed against the costs of care, this phenomenon can induce more efficient care delivery patterns. In other words, if we assume that patients optimize, then health benefits equal cost-sharing at the margin. For many years this was the assumption that underpinned welfare calculations, and policy recommendations (Feldstein, 1973). Unfortunately, researchers are increasingly finding that patients misvalue care. This phenomenon known as "behavioral hazard", finds that patients frequently make mistakes in their cost-benefit calculations (Baicker et al., 2015). These mistakes come in the form of being present-biased, underweighting non-salient symptoms, and overweighting immediate & hassle costs of care among others (Laibson 1997; Osterberg and Blaschke 2005; Newhouse 2006).

The field of health economics has grown increasingly interested in understanding the health impacts of cost-sharing changes particularly as they pertain to vulnerable populations i.e. patients that are lower income and/or sicker than the general population. Thus far, while researchers have exhaustively documented the causal impact of cost-sharing on medical utilization, attempts to tie these changes in utilization to health outcomes have been more elusive. Several studies have characterized the type of care that patients cut back on, and found that both high and low-value services are equally affected (Newhouse and Group 1993, Baicker et al. 2015, Brot-Goldberg et al. 2017). Other studies have used utilization outcomes

that serve as proxies for health deterioration. For example, Chandra et al. (2010) found that savings from reduced outpatient and prescription drug use due to copay increases are offset by increases in hospitalizations (which serves as a proxy for adverse health effects). Studies that focused on diabetes patients found that a switch to a high deductible insurance plan caused them to delay seeking care (Wharam et al., 2018), and to increase the frequency of their visits to the emergency room (Wharam et al., 2017).

A critical reason for the shift that I mentioned is because a majority of the studies that have the strongest experimental claim to estimating the causal impact of cost-sharing on health have mostly found null effects on the general population. The RAND Health Insurance Experiment, a seminal study that randomized families to plans with different tiers of cost-sharing, did not find the health of the average patient enrolled in their study was adversely affected. They did however find evidence of health improvements among lower income patients that were enrolled in the free care plan (Manning et al., 1987). Similarly, the Oregon Health Insurance Experiment (OHIE) which evaluated the effect of a lottery for Medicaid enrollment did not, "detect significant changes in measures of physical health including blood pressure (systolic or diastolic), cholesterol (HDL or total), glycated hemoglobin, or a measure of 10-year cardiovascular risk" (Finkelstein et al., 2012). It is worth noting that both these experiments have sample size constraints in the size of the populations of lower income and chronically ill patients respectively. This might impede the ability of researchers to detect health deteriorations. In the RAND HIE the sub-group of lower income patients assigned to each plan was between 580-780 patients. Meanwhile, in the OHIE the sample of patients with a pre-lottery diagnosis of diabetes was 872 patients. In a follow up analysis for the OHIE the authors acknowledge that, "the relatively small sample sizes, particularly for those with pre-existing diabetes diagnoses, mean that we cannot rule out clinically meaningful effects because of relatively large standard errors" (Allen and Baicker, 2021). Experiments that focused exclusively on vulnerable populations have had more success in estimating the causal impact of cost-sharing changes. Choudhry et al. (2011) examined the

impact of randomly assigning patients discharged after a heart attack to zero copays for prescription drugs. They found improvements in health with their secondary outcomes as the rates of revascularization were significantly reduced in the full coverage group.

A recent study by Chandra et al. (2021) notes that RAND and the OHIE do not have large enough sample sizes to measure rare adverse health outcomes such as mortality. In their study they use a quasi-experimental design combined with machine learning methods to find that Medicare recipients that face a large increase in their prescription drug coinsurance rates reduce their use of drugs which causes an increase in their mortality rate. Theirs is one of the few studies to directly tie increases in cost-sharing to negative health outcomes. My analysis aims to add to our understanding of the direct harms of cost-sharing.

# 3 Context

My study period extends from 2004-2017. During this time three concurrent trends are relevant to my analysis. Firstly, the burden of health care costs increasingly shifted to patients during this period via increases in plan cost-sharing features including deductibles, coinsurance rates, and copayments. Secondly, there is a marked rise in the population of patients that suffer from the chronic conditions of interest. Finally, patents for several highly valuable cholesterol-lowering medications including Zocor (in 2006), and Lipitor (in 2011) expired during these years. The patent for metformin, the recommended first-line therapy for Type 2 diabetes, expired in 2000 prior to the start of my study period. Taken together, a fascinating picture emerges in the backdrop of my study period; the population of patients with chronic diseases is growing amid an effort to shift health care costs to consumers. Meanwhile, patent expirations for highly valuable medications expand access to cheaper generics which might enable patients to control their conditions better. I explore these trends in further detail, and the implications of their confluence in this section.

Rapidly rising health care costs in the United States have caused employers to shift the

cost burden of care onto their employees. The Kaiser Family Foundation (KFF), which publishes annual reports documenting the generosity of employer-sponsored coverage, finds that the average annual deductible for single coverage tripled from $414 in 2004, to $1,505 in 2017 far outpacing the rate of inflation (KFF, 2019). Additionally, the percentage of workers enrolled in a plan with an annual deductible for single coverage increased from 51% in 2004, to 81% in 2017. Another striking development is the rise of the high deductible health plan (HDHP), defined as plans with at least $1,000 in single coverage, and $2,000 in family coverage. While these account for very few plans in 2004, they account for about 28% of all employer-sponsored plans in 2017. In summary, there is a decline in the generosity of employer-sponsored insurance during my study period. I leverage these changes to investigate their effect on vulnerable patients i.e. individuals that are higher risk, and/or live in lower-income areas.

On the patient side, these changes to commercial insurance are occurring during a time when the chronically ill population was increasing as a percentage of the workforce. The Center for Disease Control and Prevention (CDC) estimates that the prevalence of diabetes increased by three percentage points during my study period (from 5.3% in 2004 to 8.2% in 2017). The picture is a bit more complicated with hyperlipidemia. While overall cholesterol levels have declined in the population (Carroll et al., 2013) this is due to the increased use of statins. Among Americans 40 years and older, statin use increased from 17.9% of the population in 2003 to 27.8% of the population in 2013. However, increases in the obesity rate and a change in the guidelines for identifying and treating high cholesterol in 2013 (Ziaeian et al. (2016)) means that there was a growth in the underlying population of patients identified as having "high cholesterol". I plot average monthly A1c and LDL levels in my sample in Figure 1; during the sample period average A1c levels remained relatively flat around 7.4%, while average LDL levels dropped by about 9% (from 115 mg/dl to 105 mg/dl).

Finally, patent expirations of highly valuable medications, and a growing awareness of the benefits of value-based insurance design (Chernew et al., 2007) means more patients,

particularly those with employer sponsored insurance, have access to effective therapies that allow them to manage their conditions. I plotted the average out-of-pocket payment for metformin & insulin (for diabetic patients), and statins (for high cholesterol patients) in Figure 2. I find that the average monthly OOP spending for a 30-day supply of these drugs remains relatively flat for metformin (around $11-13 per month), increases for insulin (from $20 to $35 per month), and falls precipitously for statins (from $28 to <$12 per month).

My study examines whether large increases in cost-sharing, specifically deductible increases in health plans, affect patient management of their conditions. I measure disease management using laboratory values for A1c levels for diabetes, and LDL cholesterol levels for hyperlipidemia.

# 4 Empirical Strategy

In this analysis, I measure the impact of large deductible increases on the management of diabetes and hyperlipidemia among commercially insured patients. I use a differences-in-differences specification to model the impact on a range of laboratory, and medical outcomes. Across all my specifications, the "treatment" group includes individuals that work at employers that increased their plan deductible, while the "control" group comprises individuals working at employers that did not change their deductibles. These models control for underlying differences across individuals and for variation over time in the outcomes across all individuals by the use of fixed effects. I use within individual, over-time variation in the deductible to identify a "treatment effect". Researchers commonly use this approach in order to reduce the bias caused by time-invariant unobserved confounders i.e. factors that simultaneously correlate with deductible increases, and the outcome measures (Smith, 2021). A visual test of the parallel trends assumption can be seen in the graphs for the dynamic differences-in-differences model that uses leads and lags.

## 4.1 Research Design Assumption

My identifying assumption is that absent the deductible change, medical utilization and laboratory outcomes among treatment and control groups would change similarly. In order for this parallel trends assumption to hold, the timing of the deductible change must be exogenous. To ensure this I limit the types of employers that I include in my sample. Typically, employees choose an insurance plan from the range of options offered by their employers. In such a situation, we expect that patients with predictably higher health care expenses, for example, patients with known chronic conditions, would choose more generous coverage that might be less likely to change the deductible. Also if employees are offered a range of plans that differ, then they are much more likely to switch between plans in response to a discrete jump in cost-sharing. To limit this type of selection bias in my sample, I focus my analysis on employers that only offer one Optum plan per plan year. This strategy allows me to exploit the impact of changes in plan cost-sharing that are uncorrelated with the health of individuals.

## 4.2 Two-Way Fixed Effects Model

Since employers change their deductibles at different times, treatment adoption is staggered across individuals. In such settings, the two-way fixed effects estimator is commonly used. I start by running the following model with individual fixed effects $i$ to account for fixed differences across individuals, and calendar-year quarter $q$ fixed effects to account for overall trends over time.[2] The coefficient $\alpha_1$ in Equation 1 estimates the treatment effect of the deductible change on range of outcomes. Standard errors are clustered at the individual level:

$$y_{iq} = \alpha_0 + \alpha_1 Post * Treat + \gamma_i + \pi_q + \varepsilon_{iq} \tag{1}$$

[2]In this context, each calendar-year quarter is given a unique value ($q = 1, 2, ..., 72$) starting with the first quarter of 2000 where $q = 1$, up to the fourth quarter of 2017 where $q = 72$.

In addition to the above, I also estimate event-study models that allow the treatment effect to vary over time instead of averaging the treatment effect over all the post-treatment periods. Such models allow me to analyze pre-treatment trends, and to provide a visual companion to the static model. By flexibly incorporating time in the model using leads and lags we can see the change in the outcome in the treatment vs. control group over time relative to the omitted quarter (I use two quarters before the deductible change as the omitted quarter for reasons that are better explained below):

$$y_{iq} = \alpha_0 + \sum_{t=-8, t\neq -2}^{7} \alpha_t \mathbf{1}\{q - E_i = t\} + \gamma_i + \pi_q + \varepsilon_{iq} \tag{2}$$

In the model above $y_{iq}$ is the outcome for individual $i$ in calendar-year quarter $q$. The key regressors are a series of dummy variables $\mathbf{1}\{q - E_i = t\}$ that take the value of one for each event-time quarter, where event-time is defined for each individual relative to the quarter in which their cost-sharing changed ($E_i$). As with Equation 1 above, I include individual, and calendar-year quarter fixed effects, with standard errors clustered at the individual level. A new and emerging literature has pointed out several issues with the two-way fixed effects estimator that I detail below. In light of these concerns I primarily present results from a cleaner model that is not subject to these biases.

## 4.3 Issues with TWFE

Recent developments in quasi-experimental methodologies have shed more light on issues with differences-in-differences specifications. This new literature is primarily concerned with rollout designs i.e. when different groups get treated at different times which is the setting for my analysis. Goodman-Bacon (2021) found that the treatment effect derived from the standard two-way fixed effects differences-in-differences model could give us biased estimates when there is staggered adoption of treatment. His central finding is that the DD estimator obtained by exploiting variation across units that receive treatment at different times is a

variance weighted average of all possible two-group/ two-period DD estimators in the data. Callaway and Sant'Anna (2021), Borusyak et al. (2021), and Sun and Abraham (2021) are just a few examples of papers that validate this finding. These papers all provide alternative estimators that can be used to estimate unbiased treatment effects.

Baker et al. (2021) simulated different conditions under which staggered differences-in-differences regressions are run and provide an overview of the situations where they do or do not hold. They find that if groups treated at different times have (1) treatment effects that vary in size (eg. a different slope change for earlier vs. later treated units) or (2) dynamic treatment effects i.e. treatment effects that change over time then the standard TWFE estimator gives us incorrect results. Both of these scenarios are plausible in my data, and so the recommendations given by the authors of these studies are very relevant to my analysis.

While all the papers referenced above differ in their approaches to constructing alternative estimators, one common feature they have is is that they all calculate separate treatment effects for each *cohort* that is treated. The papers then each provide suggestions on how to aggregate the treatment effects for each cohort. In this analysis, a cohort is comprised of patients in the treatment group that all experienced an increase in their cost-sharing in the same calendar-year quarter (eg. $q = 30$), while an *event* refers to the deductible increase that occurs in $q = 30$. I will use this terminology when discussing the stacked regression model in the next section. To test whether my main TWFE specification results hold, I provide results from an additional model that circumvents these issues. While my initial results will compare the estimates obtained by the two models, the majority of my results are from the stacked regression model.

## 4.4 Stacked Regression Model

The stacked regression model is based on work done by Cengiz et al. (2019), Baker et al. (2021). They key to this method is to create a separate dataset for each "event" i.e. quarter that there is a deductible change. The dataset for each "event" consists of the "cohort" of

patients in the treatment group (that experienced a qualifying deductible increase during that event), and contemporaneous control patients (that did not experience a deductible increase). So to continue the example above, for the event in $q = 30$, individuals that experienced a deductible increase that quarter will be in the treatment group, and individuals that do not experience a deductible increase are in the control group. Furthermore, I limit the time window to be two years before and two years after treatment. It is worth noting that my data is not a balanced panel. As patients join firms, or quit the jobs in the middle of the year they will enter and exit my sample. I impose the condition that I must observe every patient for two or more plan years however (even if they are not enrolled for the duration of the plan year).

Once I have constructed each separate dataset, and stacked them together, I can run the standard differences-in-differences specification with a slight adjustment. The modification here is that we have to saturate the unit and time fixed effects with indicators for each stacked dataset. The modified model I run is:

$$y_{iqe} = \beta_0 + \beta_1 Post * Treat + \mu_{ie} + \lambda_{qe} + \varepsilon_{iqe} \tag{3}$$

where $i$ denotes individual, $q$ denotes calendar-year quarter, $e$ denotes a specific event, and standard errors are clustered at the individual level. Analogously, to provide a visual companion to the stacked regression model, I run a model with a full set of lags and leads:

$$y_{iqe} = \beta_0 + \sum_{t=-8, t\neq -2}^{7} \beta_t \mathbf{1}\left[q - E_i = t\right] + \mu_{ie} + \lambda_{qe} + \varepsilon_{iqe} \tag{4}$$

I omit the quarter that is two quarters before an individual experiences a change in cost-sharing i.e. $t = -2$. This is done to capture the baseline difference between individuals for whom the event does, and does not occur *before* they were aware of any future cost-sharing changes. If I use the quarter before the change in cost sharing ($t = -1$) as my omitted category it could bias my estimates as patients might change their behavior in anticipation

of a change in their future cost-sharing. In the above specification, $\beta_1$ is the effect one quarter after the change in cost-sharing *relative* to two quarters before the change. I do not define event time for control individuals as they are enrolled in plans that do not experience a cost-sharing change.

Since I align all events in relative event-time, and impose a stricter criteria for control groups, this method helps to reduce the bias described in the previous section. More specifically, this method only uses "never-treated" patients as controls, and as such does not bias my estimates from using "already-treated" patients as controls. I compare estimates obtained from both models on laboratory outcomes for the full sample. However, as I mentioned earlier, my main results will primarily be from the stacked model.

# 5 Data

I use administrative and claims data from Optum's de-identified Clinformatics Data Mart Database, which includes over 46 million unique members who were enrolled in employer sponsored commercial plans offered by a large national insurer. These data include enrollment information, plan design features[3], basic demographics (age, gender), and all medical and pharmacy claims. I link patients to zip-code level socio-demographic characteristics using data from the American Community Survey (ACS). A unique feature of this data is that I have information on laboratory test results for a portion of my sample[4]. These lab values are key to my analysis as I use them as outcomes for disease management, and I use them to construct various subgroups of interest.

[3]This includes a series of variables on the level of various plan characteristics including individual and family deductible levels, coinsurance rates, and copays for

[4]The data vendor has contracted with a national laboratory testing chain to link patients to clinical test results. I only have lab values for patients that get tested at this chain; about 40% of my sample has at least one lab test result.

## 5.1 Sample Selection

I employ several cleaning steps described in Table 1 to obtain my larger sample. Among other criteria, patients in my data all work at employers that only offer one plan per plan year, and are enrolled for at least two plan years. A common concern with measuring the effect of cost-sharing on utilization is adverse selection as we expect that predictably sick individuals will choose plans that have lower levels of cost-sharing. To mitigate this concern, my final sample includes employers that offer only one insurance plan per plan year. As I previously mentioned, this allows me to isolate the variation in plan cost-sharing that is uncorrelated with the health of the individuals. My cleaning steps leave me with a general population sample of over 2.8 million patients. I then use the medical and prescription drug claims of these patients to identify individuals with Type II diabetes, and hyperlipidemia. I follow the methodology employed by CMS' Chronic Condition Warehouse to classify patients into the two conditions. The CCW uses site of care information, along with diagnosis codes, and prescription drug fills to link patients to specific chronic conditions. Using their methods I identify 170,005 patients with diabetes, and 362,570 patients with hyperlipidemia.

Since a lot of my analysis relies on biomarkers, and I do not have access to this data for the full population, I must restrict my sample further. In Table 2 I describe the restrictions that I use to assemble the analytic samples for my primary and secondary outcomes. My primary outcomes examine the effect of the deductible increase on lab values which is why I require at least one lab test result in the pre- and in the post-periods. The secondary outcomes focus on measures of medical utilization, and I require that patients have at least one lab test result in the pre-period. This allows me to maximize the sample size that I leverage to estimate treatment effects. The implicit assumption here is that the medical results using the larger sample can be extrapolated to the smaller sample. I directly test this assumption and find that it holds. Table 3 provides counts for the final analytic samples used; about 15% of the total sample is used for the primary outcomes, and 25-30% is used for the secondary outcomes.

Once my analytic samples are assembled, I examine whether there are observable differences between the treatment and control groups. In Figure 3 I plot the distribution of initial biomarker values in the two groups, and use the Kolmogorov-Smirnov test for equality of distribution functions. For diabetics, the distribution of initial A1c values is the same. However for high cholesterol patients, initial LDL values are slightly higher in the treatment group. This is confirmed by balance table t-tests in Tables 5 - 6. I also find that the treatment group is more likely to be female, and to live in zip-codes with lower median household income. Despite this the median households incomes are well above the national figures. Finally, there are zero or very minimal differences in age between the two groups.

## 5.2 Subgroup Analysis

I first test the effect of deductible increases on the general population of patients with chronic conditions. There are two additional groups of higher risk patients that are of interest to insurers, and policymakers. These include (1) higher risk patients i.e. individuals with uncontrolled diabetes or hyperlipidemia (2) lower income patients. I will now describe how I construct each of these sub-groups of interest.

### By Initial Disease Severity

As I have access to lab values in my analytic samples, I can test whether a deductible increase worsens the management of chronic conditions by using an unbiased measure of initial disease severity - baseline biomarkers. For each patient I create a binary flag that indicates whether they have a high baseline lab value using pre-defined thresholds. For diabetes patients this is an A1c level $\geq 7.5\%$, and for hyperlipidemia patients an LDL level $\geq 160$ mg/dl. To estimate whether the treatment exacerbates patients with uncontrolled conditions, I constrain the sample (so both treatment and control groups have high baseline values) and re-run my main specifications. I run an equivalent analysis for patients with low baseline values.

**By Income**

I also conduct a sub-group anaysis by household income. I use median household income available at the 5-digit zipcode level from the American Community Survey (ACS) to classify patients as living in low income ($\leq \$40,000$), or high income ($\geq \$100,000$) areas. Unfortunately I only have this data at the 5-digit zipcode level, and cannot obtain more granular estimates than this.

## 5.3 Intervention

My sample consists of commercially insured patients with diabetes, and hyperlipidemia. The intervention of interest is an individual deductible change of $\geq \$500$ and/or a family deductible change of $\geq \$1,000$. I impose an additional restriction on the treatment group; the pre-treatment individual deductible must be $\leq \$1,000$, and the family deductible must be $\leq \$2,000$. This is because deductibles larger than these amounts are commonly used to denote high deductible health plans (KFF, 2019).

## 5.4 Outcomes

I study a set of laboratory and medical utilization outcomes listed in Table 4. I also provide a detailed description of how I constructed each outcome in the Data Appendix. All outcomes are aggregated for each individual to the quarter level. Since I have the full universe of claims for each individual I infer that quarters with missing medical or prescription claims data are quarters where patients did not incur any utilization. All of the results that I report include coefficient estimates from Equations 1 - 4, and their corresponding t-statistics.

## 5.5 Robustness Checks

I conduct some robustness checks to test the validity and scope of my results. First, I vary the size of the deductible increase to see whether my primary outcomes (biomarker values)

exhibit similar results with different sizes of the intervention. As with my main results, I continue to impose the restriction that the pre-treatment individual deductible must be $\leq \$1,000$, and the family deductible must be $\leq \$2,000$. I then test the impact of a "small" change[5] and a "large" change.[6] Second, I vary the threshold that divides patients into "low" and "high" initial value groups. For patients with diabetes the new threshold that I use is 7% A1c, and for patients with high cholesterol it is 130 mg/dl. A full set of robustness results are included in Section 12 in the Appendix.

# 6 Results

## 6.1 First-Stage

The average size of the deductible change is an increase of about $1,050 for the individual deductible and $2,300 for the family deductible as can be seen in Tables 7 - 8. As a result, out-of-pocket spending rises by approximately $60–75 per patient; this represents a 22 - 32% increase in spending relative to the mean. There is a seasonal trend in the out-of-pocket spending measure that reflects the non-linear nature of health insurance plans. In the quarter of the deductible increase, there is a noticeable jump in out-of-pocket spending incurred by patients that drops in subsequent quarters. This happens because as patients use medical services over the year, they spend towards their deductible and eventually hit the limit. After reaching their deductibles, patients are typically responsible for a smaller share of additional medical costs.

Interestingly, when I examine the treatment effect on out-of-pocket spending broken up by initial disease severity, I find slightly different trends in the two conditions. These results are in the Appendix in Figure A1. Among diabetes patients, those with low baseline A1c values have a smaller increase in their out-of-pocket spending amounts relative to the mean when I

[5] An individual deductible change of $\geq \$250$ and/or a family deductible change of $\geq \$500$

[6] An individual deductible change of $\geq \$1,000$ and/or a family deductible change of $\geq \$2,000$

compare them to those with high baseline values (a relative increase of 15% vs. 35%). There is an equivalent split seen in the hyperlipidemia sample as out-of-pocket spending increased by about 30% for the low baseline group vs. 40% for the high baseline group.

## 6.2 Primary Outcomes

My primary outcomes include A1c (%) and LDL (mg/dl) biomarker values, and binary flags to indicate uncontrolled levels of these outcomes for patients with diabetes (A1c$>$ 7.5%) and hyperlipidemia (LDL$>$ 160 mg/dl) respectively. I use a slightly more restrictive sample for these results as I do for the rest of the medical results (see Table 3). I find that a large exogenous increase in the deductible does not affect the average A1c values for diabetes patients (Figure 6), or LDL levels for high cholesterol patients (Figure 9) among the general population of patients with these diseases. Among hyperlipidemia patients, I do find that there is a 0.7 percentage point increase in the probability of having a high LDL flag which represents an 9% increase in the likelihood of this flag relative to the mean. However, for the most part I obtain null results. We can see this in both the difference-in-difference model estimates and in the event-study graphs.

I then divide the sample with respect to a patient's baseline biomarker value, and I find that patients that have stable control of their conditions before the deductible change are the ones that are adversely affected while those that have an uncontrolled baseline case of the condition are not affected. So in other words, the deductible increase worsens the health of patients that were previously in better health. When I examine the sub-sample of diabetic patients with stable baseline A1c values, I find that the treatment group experiences a 0.08 percentage point increase in their A1c levels relative to the control group. This translates to a 1.2% increase relative to the mean A1c value. HbA1c tests are blood tests that are used to measure a patient's blood sugar level over the past three months (CDC). As we can see in Figure 7 there is a spike in A1c levels two quarters after the deductible increase. Since this test is a moving average of each patient's glucose level over the past few months it is

consistent to see a lag in the outcome. To give a sense of the magnitude of this impact, one retrospective cohort study found that, "each percentage point reduction in hemoglobin A1c was associated with a 37% decrease in risk for microvascular complications, and a 21% decrease in the risk of death related to diabetes" (Stratton et al., 2000). I also find that these patients experience a 2.3 percentage point increase in the probability of a high A1c value, i.e. a 19% increase relative to the mean. Meanwhile, patients with high baseline A1c values do not experience a statistically significant change in their A1c levels, or in the probability of having a high A1c value relative to the control group. The coefficient for this group is actually negative, which could raise concerns about regression to the mean. Since I am comparing observably similar patients at baseline (high initial value "treatment" patients vs. high initial value "control" patients), there would have to be more regression to the mean in the former for this to a concern. Since there is no medical reason to expect this - it is not a concern. What we are likely observing is a ceiling effect on the "high initial value" sub-group. Since they already have poor control over their conditions, there is less scope for additional health consequences that are negative.

The results by initial disease severity are fairly consistent for patients with hyperlipidemia (Figure 10). Within the sub-group of patients with stable baseline cholesterol levels, I do not find a statistically significant impact of the treatment on LDL values. I do, however, find that these patients experience a 0.5 percentage point increase in the probability of having a high LDL value, i.e. a 12.5% increase relative to the mean. This is a similar magnitude to the result I found for the stable diabetes sub-sample. Once again, patients with unstable baseline control do not experience a statistically significant change in their LDL levels, or in the probability of having a high LDL flag relative to the control group. To verify that my results aren't driven by outliers I plot the unadjusted trends in A1c and LDL values in Appendix Figure A2. From a visual inspection it appears that the parallel trends assumption in the pre-period holds. So we can be reasonably confident that the treatment effect is due to the deductible increase.

I then test whether there is a differential effect on sub-groups split up on the basis of zip-code level income. Since I do not have data on individual level household income, I am using zip-code level median household income as a proxy. I find that the deductible increase has no effect on biomarker values of treatment vs control patients living in low-income and in high income zip codes. The differences-in-differences coefficient is small, and not significant. In the event study graphs for diabetes (Figure 8) there is evidence of a spike in the high A1c flag 5 quarters after the intervention among patients living in low income zip-codes. This spike however quickly dissipates. No similar effect is seen among patients living in high income zip-codes. Since I obtain mostly null results for the income sub-groups I do not further investigate differential patterns in their utilization. My outcomes provide suggestive evidence that patients that would have been considered low-risk at the outset are the ones that bear the health effects of large deductible changes. In the next section, I examine the mechanisms that might lead to these changes in biomarker values.

## 6.3 Mechanisms

In this section, I dig deeper into the mechanisms to explain the discrepancy in the impact of the deductible increase on disease management between the subgroup of patients with stable and unstable initial conditions. To this end, I investigate whether there are differential trends across the subgroups in their post-intervention medical utilization that might explain the primary outcomes. I continue to use differences-in-differences & event study methods to test various theories about what drives the divergence in the biomarkers across my subgroups.

### Theory #1: Selection into A1c or LDL testing

For all these analyses, I am comparing the effect of the exogenous deductible change within the subgroup of patients that have either low or high initial biomarker values. It is possible that the adverse outcomes that I observe in the previous section are an artefact of selection into testing. In other words, it is possible that among patients with low initial A1c or

LDL values, the patients that choose to get tested post-change are those with a noticeable worsening of their condition. I can directly test whether this hypothesis is true by looking at the impact of the change on the number of A1c tests per quarter conducted. If there is a marked reduction in the number of these tests conducted post-intervention among the treated group relative to the control group, this is a sign that this theory could be true. As I previously mentioned, I only have biomarker values for a portion of patients, but I have access to the full universe of medical claims incurred; this enables me to measure how often each patient received a test. I use Current Procedural Terminology (CPT) codes to identify when a patient received a HbA1c test (CPT 83036, 83037), or an LDL test (80061, 80053, 83704). The latter were mostly conducted via lipid panels that measured several values, such as total, and HDL cholesterol as well.

My analyses reveal that the treatment does not affect the number of A1c or LDL tests conducted by either sub-sample (Tables 15 - 16). This is a somewhat surprising result, as the vast cost-sharing literature has provided ample evidence that patients respond to the point-of-service costs of health care. In the corresponding event study graphs, I find a small decline in testing the quarter after the deductible change in the "low initial value" group. There is however a quick recovery back to baseline. Among patients with "high initial values" testing remains relatively flat.

There are several reasons that the total number of A1c or LDL tests might not be impacted by deductible increases. First, my sample population appears to be higher income than the general U.S. population (based on median household income in their zip-codes). So perhaps they are not as liquidity constrained, and value these tests more than the cost of an incremental test. Second, these patients are commercially insured via their employers. Many large, commercial insurers employ aggressive chronic condition management programs that carve out certain high-value tests and prescriptions from the cost-sharing schedule. In fact, I directly test this by examining the impact of the intervention on the out-of-pocket spending associated with these tests. To do this, I add up all out-of-pocket costs incurred

by patients on the date that they receive a test. I do so because such a measure gives us a clearer picture of the "full cost" that a patient incurs when obtaining the test. I find that on average, patients spend between $4-6 more post-change to obtain an A1c test. With such a small effect size it is not surprising that test volume is unaffected.

**Theory #2: High Value Prescription Drug Use**

The second theory I test is whether there is a differential treatment effect on prescription drug use by sub-group. For each condition I study the impact of the treatment on specific therapies that have been deemed high value by the medical community.

In order to identify high value prescription drugs, I use keyword searches in generic ingredients, and branded names as well as prescription drug classifications. For diabetic patients, I look at prescriptions of metformin[7], oral anti-diabetic agents (OADs), insulin, and diabetes supplies[8]. For high cholesterol patients I look at prescriptions for common statins including atorvavastin (Lipitor), fluvavastin (Lescol), lovavastin (Altoprev), pitavastin (Livalo), pravavastin (Pravachol), rosuvavastin (Crestor), and simavastin (Zocor).

Metformin has long been recommended as the first-line therapy for the treatment of Type 2 diabetes mellitus (Marín-Peñalver et al., 2016). In fact, it is the most commonly prescribed antihyperglycemic prescription drug in the United States. A meta-analysis of over 2,000 clinical trial results for for metformin indicate that it reduces HbA1c levels by 0.6 to 0.95 percentage points (Hirst et al., 2012). The larger umbrella of oral antidiabetic agents (OADs) includes six classes of drugs (including metformin) that are used to control blood glucose levels in Type 2 diabetics. A systematic review on the effect of OADs on A1c levels found that the benefits of initiating an OAD agent are most apparent in the first 4-6 months of therapy. Since I identify diabetic patients in my sample using a combination of medical and prescription claims, it is safe to assume that the majority of these patients are already being treated for their conditions. So we should not expect to see effect sizes as large

[7]Brand names include Glucophage, Fortamet, Glumetza, and Riomet

[8]These include testing strips, meters, lancets, and insulin syringes/needles.

as the ones in clinical trials on A1c levels. The theory that I am testing here is that if one of the sub-groups differentially cuts back on metformin or OAD prescriptions it would cause average glucose levels to rise, and we would see a corresponding rise in A1c values.

The differences-in-differences results indicate that within the sub-group of patients with baseline control over their diabetes, the deductible increase does not affect prescriptions for high value drugs. The event study results in Figure 13 show a different story. In the plots we can see that among beneficiaries in the "low initial value group" - in the two quarters immediately following treatment there is a statistically significant decline in metformin consumption by 2 days supplied or a 6 percent relative to the mean. Equivalently, there is a significant decline in OAD consumption by 5 days supplied or a 9% relative to the mean. Both types of prescriptions are soon restored to pre-treatment levels. The timing of the dip in OAD utilization matches up with the rise in A1c levels. Compared to this, among beneficiaries in the "high initial value group" - we see that metformin prescriptions are permanently affected post-treatment. Among this sample, the treatment results in 1.6 fewer days supplied of metformin prescriptions, which is a 4.6 percent decline relative to the mean. This effect persists up to 8 quarters after treatment. There is a null effect on OAD days supplied in this group. Taken together, this is suggestive evidence that patients with high initial values reduce their use of the first-line therapy in favor of other OADs. This potentially accounts for the null health effects among this group as well. I also plot results for insulin, and diabetes supplies prescriptions in Figure 14. There is evidence of anticipation of the intervention (the spike in consumption the quarter before the change i.e. t=-1), but no discernible pattern in the post-period for these outcomes.

For high cholesterol patients, statins are high-value drugs that control their conditions. A meta-analysis of clinical trials of statins found that these drugs reduced mean LDL cholesterol by 19.4 mg/dl (or 1.08 mmol/L) (Collaborators, 2012). When I test the impact of the intervention on statin days supplied, I do not find evidence that the two sub-groups I study are differentially affected by a large deductible change. The difference-in-difference model

coefficients are negative, but the magnitude of these coefficients is small and not statistically significant. In the event study plots in Figure 15 we can see evidence of treatment anticipation. Patients with controlled cholesterol in the baseline period anticipate the treatment by increasing their prescription fills in the quarters prior to the deductible increase. The magnitude of this jump is about 1.5 days supplied per patient. The spike in statin fills is followed by a reduction in the two following quarters. Consumption eventually smooths out to pre-intervention levels by four quarters or one year after the change. The event study plots of the two sub-groups largely mirror each other; we do not see evidence that there is a differential impact of the deductible increase on either sub-group.

The magnitude of the prescription drug results from my analyses do not match the existing literature (Goldman et al., 2004). The reason for this could be because (a) these plans carve prescription drugs out of the deductible and so they are not subject to price increases (b) insurers and employers deliberately keep cost sharing for high value drugs low to encourage use. In either case, if the cost to employees is not appreciably increasing, it is consistent with economic theory that use of these drugs does not meaningfully change. In Appendix Figures A4 - A6 I test the impact of the treatment on the OOP spending on these drugs and find effect sizes ranging from \$0-4. These are relatively minor increases and could explain the results that we see. In summary, I find evidence that the deductible increase causes a temporary drop in prescriptions for OADs and statins. There is a differential pattern by sub-group for the former, but not for the latter.

**Theory #3: Primary Care Office Visits**

The third, and final theory pertains to disruptions in primary care office visits. Both conditions studied require continuous monitoring, and ongoing management. Lifestyle modifications are also recommended as a first-line intervention in the management of these diseases including weight loss, dietary modification, and aerobic exercise (Schofield et al. (2016)). Primary care physicians (PCPs) typically counsel patients on chronic condition management;

when conditions worsen, they intervene by modifying the drugs prescribed, changing dosing regimens, or even referring patients to seek specialized help from specialists.

For the diabetes population, patients with low initial A1c values in the treatment group reduce office visits to their primary care physicians by 0.036; this translates to a 4 percent reduction relative to the mean. From the event study plots in Figure 16 we see that there is once again evidence of treatment anticipation within this sub-group. Prior to the deductible change, there is an increase in visits to the PCP by these patients, presumably to receive treatment and prescription fills, before their deductible increases. This is an 11% spike relative to the sample mean. Subsequently, there is a decrease in visits to PCPs for the next three quarters before recovering to pre-treatment levels. The pattern of utilization is completely different for patients with high initial A1c values. Compared to the former sub-group, these patients do not experience a sustained decline in visits in subsequent quarters. Their differences-in-differences point estimate is also negative, but it is not statistically significant. From the event study plots we can see that trends in visits to PCPs remained similar between the treatment and control group in the right panel of Figure 16.

A similar story emerges amongst high cholesterol patients. The sub-group with low initial values anticipates treatment, and increases visits to their PCPs by 0.05 or 6% relative to the mean prior to the deductible change. Visits to their primary care providers fall for the next three quarters before recovering in the fourth quarter. The treatment group reduces overall visits to PCPs by 0.014 which translates to a 1.6% reduction in the number of visits made. Within the sub-group of patients with high initial values, the treatment effect of the deductible on PCP visits is negative but not significant.

In short, for both diabetes and hyperlipidemia I find a differential trend between the sub-sample of patients with stable baseline biomarkers, and those without. The former group reduces their primary care office visits while the latter group does not. For diabetics, the timing of the decline in PCP visits matches up exactly with the decline in prescriptions for oral anti-diabetic agents. Doctors are in charge of renewing prescriptions, so it makes

intuitive sense that when patients cut back on visits they also cut back on drugs. This is suggestive evidence that PCPs play a vital role in managing chronic conditions, and that by foregoing this care patients with stable initial conditions worsen their health.

# 7 Robustness

I conduct several robustness checks in Section 12 in the Appendix to test the validity of my primary results. For these checks I present the results that are broken out by initial disease severity. First, I vary the size of the deductible increase that qualifies as the "intervention" to see whether there is a gradient to the magnitude of the effects that I observe. Second, I vary the size of the threshold that I use to divide patients into "high" and "low" initial value groups.

I test the effect of a few different sized deductible increases. The "small" change represents an increase in the individual and family deductibles by $\geq \$250$ and $\geq \$500$ respectively; the "large" change represents an increase in the individual and family deductibles by $\geq \$1,000$ and $\geq \$2,000$ respectively. As with my main specification, I continue to impose the condition that the pre-treatment individual deductible must be $\leq \$1,000$, and the family deductible must be $\leq \$2,000$. In Tables A12 - A15 I find that my primary outcomes broken out by initial disease severity remain quite similar in magnitude and significance. Among the sub-sample of diabetic patients with stable baseline A1c values, the treatment group experiences a 0.07 p.p. or 1% increase in their A1c levels relative to the mean. They also experience a 2.1-2.2 p.p. or 18% increase in the probability of a high A1c level relative to the mean. For the equivalent sub-sample of high cholesterol patients, I find that the deductible changes cause a 0.3 - 0.5 p.p. or 8-13% increase in probability of having a high LDL value relative to the main specification. The results for the former group remain statistically significant while the latter are not. Also, consistent with my main results there isn't a statistically significant effect of the deductible increase on the sub-sample of patients that did not initially have control over

their conditions. It is surprising that the magnitude of the effects that I detect remain stable in the face of such different size deductible changes. This is suggestive evidence that the size of the change does not matter as much as the fact that the deductible was changed in the first place.

Next, I vary the threshold that that I use to divide patients into "high" and "low" initial value groups. I change the threshold from 7.5% to 7% for diabetic patients, and from 160 mg/dl to 130 mg/dl for high cholesterol patients. In Figure A8 I find that diabetes patients with low initial values experience a 0.06 p.p. or 0.9% increase in their A1c values relative to the mean. More notably, these patients also experience a 3 p.p. or 33% increase in the probability of a high A1c level relative to the mean. This is a much larger effect size than I what I observe in my main specification. Equivalently, in Figure A9 I find that hyperlipidemia patients with stable initial conditions experience a 0.7 point or 0.75% increase in their cholesterol levels relative to the mean.

I ran the above robustness checks to test the stability of my primary outcomes. I unfortunately do not have enough data on biomarker values to conduct a placebo test where I simulate a deductible change in the pre-period. I find that my primary outcomes remark remarkably stable across these tests. While the magnitude of the effects that I observe vary slightly, they are all mostly in the same direction and significance as before.

# 8 Conclusion

In this analysis I estimated the impact of large deductible changes in employer-sponsored insurance on the management of diabetes and hyperlipidemia. A novel finding of my paper is that changes in cost-sharing adversely impact patients that were lower risk prior to the treatment. To understand why, I analyze the mechanisms through which this result is borne out. I find suggestive evidence that this result is driven by a reduction in primary care office visits, and among diabetics a contemporaneous dip in high value prescriptions. This

analysis adds to the evidence that primary care centered treatment is critical in managing the stability of chronic conditions.

There are several limitations to my analysis. The patients in my setting are higher income, and have access to relatively generous insurance. This is evident in the results that I describe where I do not find large increases in OOP spending for certain valuable services and prescription drugs despite the intervention being a large deductible increase. Also, since these patients are generally higher income they might not be as liquidity constrained as the general populations. My sample is also limited to patients that work at employers that only offer one Optum plan per plan year. These tend to be smaller employers as large employers are more likely to offer plan choice (KFF, 2019). The generalizability of these results then depends on whether employees that work at smaller employers are similar to those that work at larger employers. Patients with employer sponsored insurance are also healthier than those on Medicare, and Medicaid (KFF, 2013). So these results are likely an underestimate of the treatment effect on populations without commercial insurance.

There are also several important implications to my analysis. Insurers and self-insured employers should be proactive about providing support to their chronically ill members in the event of cost-sharing changes. Adjustments in insurance plan design should be concurrently paired with extra support for disease management. My results also surprisingly indicate that there isn't a gradient of health effects that correlate with the size of the cost-sharing change. So even if employers are making small increases to deductibles there will likely be health effects on chronically ill patients unless they are thoughtful about the implementation of these shifts. As researchers continue to investigate the health consequences of cost-sharing changes, my analysis suggests that they should do so on the basis of pre-treatment or baseline health. Finally, my results underscore the importance of consistent primary care treatment in chronic disease management.

# References


Allen, Heidi and Katherine Baicker (2021) "The Effect of Medicaid on Care and Outcomes for Chronic Conditions: Evidence from the Oregon Health Insurance Experiment," Working Paper 29373, National Bureau of Economic Research, 10.3386/w29373.

Baicker, Katherine, Sendhil Mullainathan, and Joshua Schwartzstein (2015) "Behavioral Hazard in Health Insurance *," *The Quarterly Journal of Economics*, 130 (4), 1623–1667, 10.1093/qje/qjv029.

Baker, Andrew, David F. Larcker, and Charles C. Y. Wang (2021) "How Much Should We Trust Staggered Difference-In-Differences Estimates?" SSRN Scholarly Paper ID 3794018, Social Science Research Network, Rochester, NY, 10.2139/ssrn.3794018.

Bernard, Didem, Cathy Cowan, Thomas Selden, Liming Cai, Aaron Catlin, and Stephen Heffler (2012) "Reconciling Medical Expenditure Estimates from the MEPS and NHEA, 2007," *Medicare & Medicaid Research Review*, 2 (4), mmrr.002.04.a09, 10.5600/mmrr.002.04.a09.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2021) "Revisiting Event Study Designs: Robust and Efficient Estimation," *arXiv:2108.12419 [econ]*, - (-), 1–54.

Brot-Goldberg, Zarek C., Amitabh Chandra, Benjamin R. Handel, and Jonathan T. Kolstad (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, 10.1093/qje/qjx013.

Callaway, Brantly and Pedro H. C. Sant'Anna (2021) "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225 (2), 200–230, 10.1016/j.jeconom.2020.12.001.

Carroll, Margaret D., Brian K. Kit, David A. Lacher, and Sung Sug Yoon (2013) "Total and High-Density Lipoprotein Cholesterol in Adults: National Health and Nutrition Examination Survey, 2011-2012," *NCHS data brief*, 132, 1–8.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer (2019) "The Effect of Minimum Wages on Low-Wage Jobs*," *The Quarterly Journal of Economics*, 134 (3), 1405–1454, 10.1093/qje/qjz014.

Chandra, Amitabh, Evan Flack, and Ziad Obermeyer (2021) "The Health Costs of Cost-Sharing,"Technical Report w28439, National Bureau of Economic Research, Cambridge, MA, 10.3386/w28439.

Chandra, Amitabh, Jonathan Gruber, and Robin McKnight (2010) "Patient Cost-Sharing and Hospitalization Offsets in the Elderly," *The American Economic Review*, 100 (1), 193–213, 10.1257/aer.100.1.193.

Chernew, Michael E., Allison B. Rosen, and A. Mark Fendrick (2007) "Value-Based Insurance Design," *Health Affairs*, 26 (Supplement 2), w195–w203, 10.1377/hlthaff.26.2.w195.

Choudhry, Niteesh K. et al. (2011) "Full Coverage for Preventive Medications after Myocardial Infarction," *New England Journal of Medicine*, 365 (22), 2088–2097, 10.1056/NEJMsa1107913.

Collaborators, Cholesterol Treatment Trialists' (CTT) (2012) "The Effects of Lowering LDL Cholesterol with Statin Therapy in People at Low Risk of Vascular Disease: Meta-Analysis of Individual Data from 27 Randomised Trials," *Lancet*, 380 (9841), 581–590, 10.1016/S0140-6736(12)60367-5.

Cutler, David M. and Richard J. Zeckhauser (2000) "Chapter 11 - The Anatomy of Health Insurance," in Culyer, Anthony J. and Joseph P. Newhouse eds. *Handbook of Health Economics*, 1 of Handbook of Health Economics, 563–643: Elsevier, 10.1016/S1574-0064(00)80170-5.

Feldstein, Martin S. (1973) "The Welfare Loss of Excess Health Insurance," *Journal of Political Economy*, 81 (2), 251–280.

Finkelstein, Amy, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group (2012) "The Oregon Health Insurance Experiment: Evidence from the First Year*," *The Quarterly Journal of Economics*, 127 (3), 1057–1106, 10.1093/qje/qjs020.

Goldman, Dana P., Geoffrey F. Joyce, Jose J. Escarce, Jennifer E. Pace, Matthew D. Solomon, Marianne Laouri, Pamela B. Landsman, and Steven M. Teutsch (2004) "Pharmacy Benefits and the Use of Drugs by the Chronically Ill," *JAMA*, 291 (19), 2344–2350, 10.1001/jama.291.19.2344.

Goodman-Bacon, Andrew (2021) "Difference-in-Differences with Variation in Treatment Timing," *Journal of Econometrics*, 225 (2), 254–277, 10.1016/j.jeconom.2021.03.014.

Hirst, J. A., A. J. Farmer, R. Ali, N. W. Roberts, and R. J. Stevens (2012) "Quantifying the Effect of Metformin Treatment and Dose on Glycemic Control," *Diabetes Care*, 35 (2), 446–454, 10.2337/dc11-1465.

KFF (2013) "Medicaid Enrollees Are Sicker and More Disabled than the Privately-Insured," March.

——— (2019) "Employer Health Benefits Annual Survey Archives."

Laibson, David (1997) "Golden Eggs and Hyperbolic Discounting," *The Quarterly Journal of Economics*, 112 (2), 443–477.

Manning, Willard G., Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz (1987) "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review*, 77 (3), 251–277.

Marín-Peñalver, Juan José, Iciar Martín-Timón, Cristina Sevillano-Collantes, and Francisco Javier del Cañizo-Gómez (2016) "Update on the Treatment of Type 2 Diabetes Mellitus," *World Journal of Diabetes*, 7 (17), 354–395, 10.4239/wjd.v7.i17.354.

Newhouse, Joseph P. (2006) "Reconsidering the Moral Hazard-Risk Avoidance Tradeoff," *Journal of Health Economics*, 25 (5), 1005–1014, 10.1016/j.jhealeco.2006.06.003.

Newhouse, Joseph P. and Insurance Experiment Group (1993) "Free for All?: Lessons from the RAND Health Insurance Experiment,"Technical report, Harvard University Press.

Osterberg, Lars and Terrence Blaschke (2005) "Adherence to Medication," *New England Journal of Medicine*, 353 (5), 487–497, 10.1056/NEJMra050100.

Pauly, Mark V. (1968) "The Economics of Moral Hazard: Comment," *The American Economic Review*, 58 (3), 531–537.

Peterson, Matthew (2018) "Economic Costs of Diabetes in the U.S. in 2017," *Diabetes Care*, 41 (5), 917–928, 10.2337/dci18-0007.

Schofield, Jonathan D., Yifen Liu, Prasanna Rao-Balakrishna, Rayaz A. Malik, and Handrean Soran (2016) "Diabetes Dyslipidemia," *Diabetes Therapy*, 7 (2), 203–219, 10.1007/s13300-016-0167-x.

Smith, Laura Barrie (2021) "The Effect of Nurse Practitioner Scope of Practice Laws on Primary Care Delivery," *Health Economics*, 1 (n/a), 1–21, 10.1002/hec.4438.

Stratton, I. M., A. I. Adler, H. A. Neil, D. R. Matthews, S. E. Manley, C. A. Cull, D. Hadden, R. C. Turner, and R. R. Holman (2000) "Association of Glycaemia with Macrovascular and Microvascular Complications of Type 2 Diabetes (UKPDS 35): Prospective Observational Study," *BMJ (Clinical research ed.)*, 321 (7258), 405–412, 10.1136/bmj.321.7258.405.

Sun, Liyang and Sarah Abraham (2021) "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects," *Journal of Econometrics*, 225 (2), 175–199, 10.1016/j.jeconom.2020.09.006.

Wharam, J. Frank, Christine Y. Lu, Fang Zhang, Matthew Callahan, Xin Xu, Jamie Wallace, Stephen Soumerai, Dennis Ross-Degnan, and Joseph P. Newhouse (2018) "High-Deductible Insurance and Delay in Care for the Macrovascular Complications of Diabetes," *Annals of Internal Medicine*, 169 (12), 845, 10.7326/M17-3365.

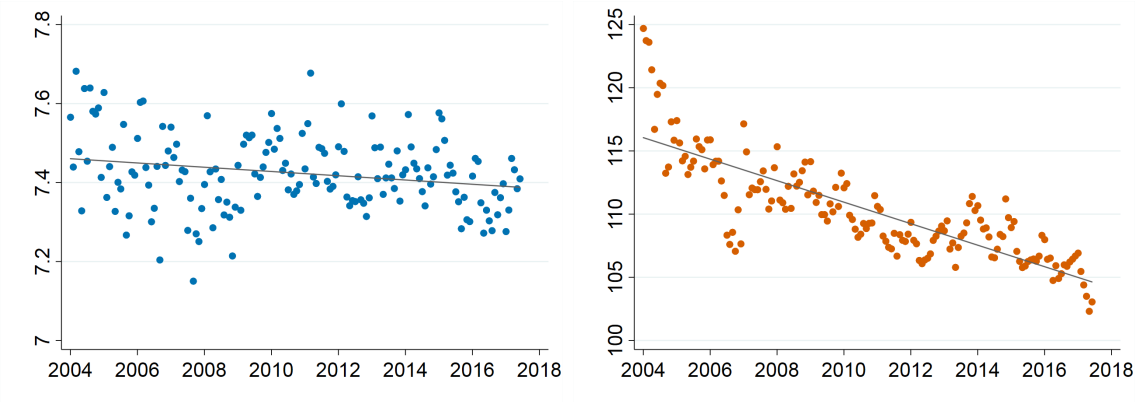
Wharam, J. Frank, Fang Zhang, Emma M. Eggleston, Christine Y. Lu, Stephen Soumerai, and Dennis Ross-Degnan (2017) "Diabetes Outpatient Care and Acute Complications Before and After High-Deductible Insurance Enrollment: A Natural Experiment for Translation in Diabetes (NEXT-D) Study," *JAMA Internal Medicine*, 177 (3), 358, 10.1001/jamainternmed.2016.8411.

Ziaeian, Boback, John Dinkler, Yuanlin Guo, and Karol Watson (2016) "The 2013 ACC/AHA Cholesterol Treatment Guidelines: Applicability to Patients with Diabetes," *Current diabetes reports*, 16 (2), 13, 10.1007/s11892-015-0703-5.

# 9 Tables and Figures

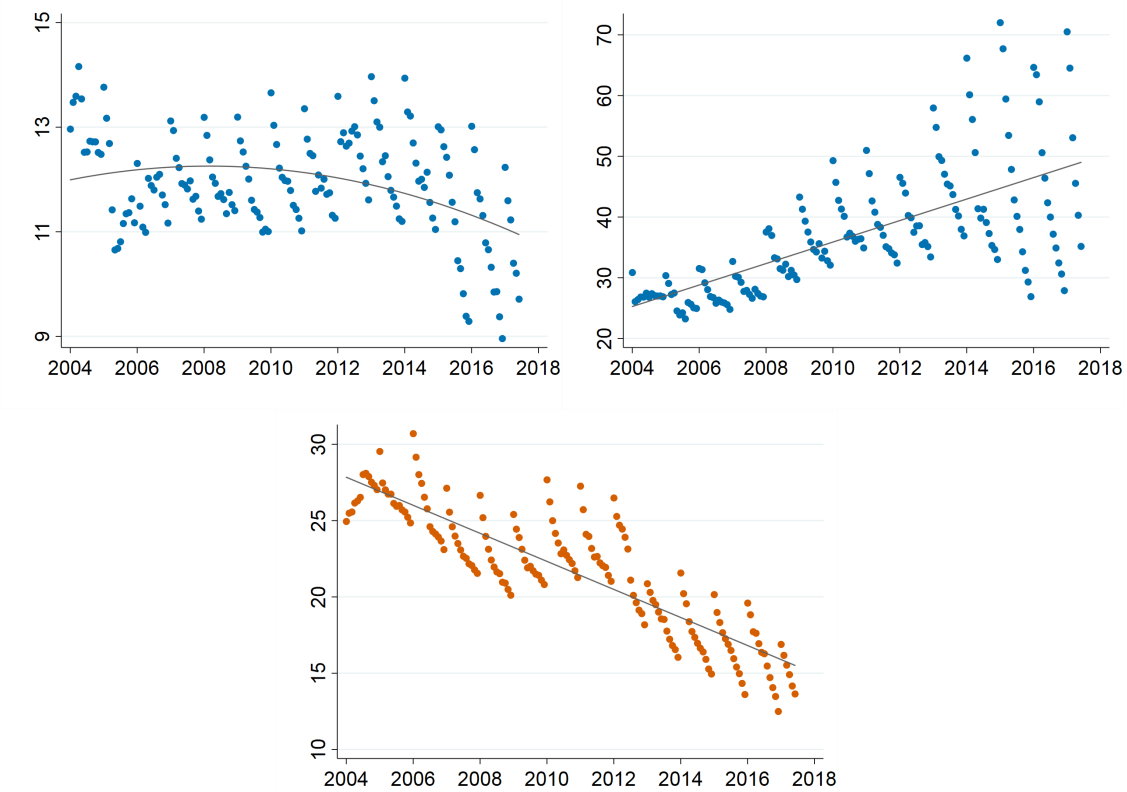
## 9.1 Summary Stats

Figure 1: Average Monthly A1c and LDL values




Note: These figures are plots of average monthly biomarker values of A1c (left), and LDL (right) for patients in my sample with diabetes and high cholesterol respectively.

Figure 2: Average Monthly OOP Spending ($)




Note: These figures are plots of average monthly out-of-pocket spending for metformin (top left), insulin (top right), and statins (bottom) among patients in my sample.

Table 1: Sample Selection

| | Enrollees | Employers |
|---|---|---|
| Initial Sample | 46,705,097 | 192,895 |
| Plan design information available in data | 27,558,405 | 190,964 |
| Employers offer one plan per plan year | 9,802,938 | 161,497 |
| Minimum of ten enrollees per employer-plan year | 8,853,628 | 77,336 |
| Enrollees are continuously enrolled for two or more plan years | 2,896,935 | 40,243 |
| Sample of Patients with Diabetes | 170,005 | 21,242 |
| Sample of Patients with Hyperlipidemia | 362,570 | 29,192 |

Note: These are the cleaning steps that I employ to obtain my general population sample. The sample of patients with diabetes and hyperlipidemia is further restricted pending availability of lab test results.

Table 2: Sample Criteria

| | | Criteria |
|---|---|---|
| Primary Outcomes | Lab Result | $\geq 1$ in pre & post periods |
| | Time | -2 years to +2 years |
| Secondary Outcomes | Lab Result | $\geq 1$ in pre-period |
| | Time | -2 years to +2 years |

Note: I describe the criteria that I use to construct the analytic samples for my primary outcomes (biomarkers), and secondary outcomes (plan design and medical utilization measures).The criteria for the former is more stringent because in order to measure the impact of the deductible change on biomarkers, it is necessary to have at least one value in the pre-period, and at least one value in the post period.

Table 3: Sample Counts

| | Condition | Treatment | Control | Total |
|---|---|---|---|---|
| Primary Outcomes | Diabetes | 1,881 | 23,981 | 25,862 |
| | Hyperlipidemia | 5,064 | 59,609 | 64,673 |
| Secondary Outcomes | Diabetes | 2,750 | 41,518 | 44,268 |
| | Hyperlipidemia | 7,495 | 102,374 | 109,869 |

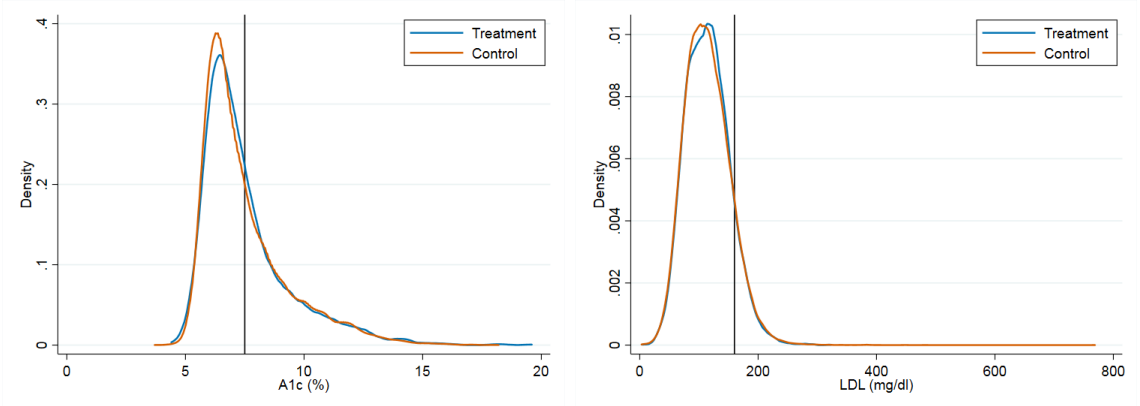
Note: After deploying the sample criteria described in Table 2, I am left with the above group sizes. Primary outcomes include laboratory test results, and secondary outcomes include measures of medical utilization such as counts of lab tests, physician office visits, and prescription drugs days supplied.

Table 4: Outcomes

| | Category | Condition | Variable |
|---|---|---|---|
| Treatment | First-Stage | Both | Deductible (indv & family), OOP spending ($) |
| Primary Outcomes | Lab Results | Diabetes | A1c (%), High A1c ($\geq 7.5$) |
| | Lab Results | Hyperlipidemia | LDL (mg/dl), High LDL ($\geq 160$) |
| Secondary Outcomes (claims-based mechanisms) | Office Visits | Both | PCP, Specialist |
| | Office Visits | Diabetes | Diabetes Specialist |
| | Office Visits | Hyperlipidemia | Cardiologist |
| | Prescription Drugs | Both | Branded, Generic |
| | Prescription Drugs | Diabetes | Metformin |
| | Prescription Drugs | Hyperlipidemia | Statins |
| | Lab Tests | Diabetes | Number of A1c tests |
| | Lab Tests | Hyperlipidemia | Number of LDL tests |
| | Hospital-Based | Both | ED visits, Inpatient stays |

Note: This table is a comprehensive list of outcomes that I study. I use Current Procedural Terminology or CPT codes to identify A1c tests (CPT: 83036, 83037), and LDL tests (CPT: 80061, 80053, 83704). "Both" indicates that the outcome applies to both conditions that I study. "Diabetes Specialist" includes endocrinologists, cardiologists, and nephrologists. I have chosen to only include a subset of these outcomes in the results of this paper; a full set of outcomes is available in the Supplementary Results Appendix.

Figure 3: Distribution of Biomarkers




Note: I plot the distributions of biomarker values for patients with diabetes (left), and hyperlipidemia (right) in the pre-period. I also include reference lines for the divide between "high" and "low" initial values. This is at 7.5% A1c for diabetics, and 160 mg/dl for high cholesterol patients.

Table 5: Diabetes Balance Table

| Variable | Control group | Treatment group | Difference |
|---|---|---|---|
| Age | 51.138 | 51.137 | -0.001 |
| | (10.686) | (10.138) | (0.255) |
| Female (%) | 45.248 | 48.379 | 3.130*** |
| | (49.775) | (49.987) | (1.192) |
| ACS Income ($) | 64,134.23 | 60,151.37 | -3,982.87*** |
| | (30,304.75) | (28,487.52) | (729.29) |
| First A1c Value | 7.489 | 7.533 | 0.044 |
| | (1.816) | (1.890) | (0.044) |
| Observations | 23,981 | 1,881 | 25,862 |

Note: This table tests whether diabetes patients in the treatment and control groups are observably similar at baseline. The "Difference" column reports t-tests of the difference in means between "treatment" and "control" groups. Significance stars are at the 1, 5, and 10% levels.

Table 6: Hyperlipidemia Balance Table

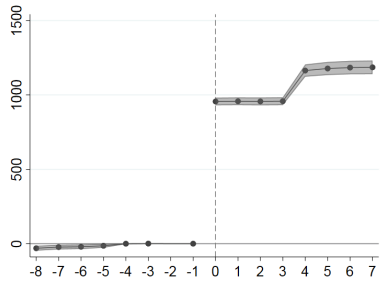
| Variable | Control group | Treatment group | Difference |
|---|---|---|---|
| Age | 51.651 | 51.244 | -0.406*** |
| | (9.791) | (9.420) | (0.143) |
| Female (%) | 43.332 | 46.169 | 2.837*** |
| | (49.554) | (49.858) | (0.726) |
| ACS Income ($) | 68,065.59 | 63,845.08 | -4,220.51*** |
| | (32,471.25) | (29,603.85) | (476.62) |
| First LDL Value | 114.430 | 115.875 | 1.445*** |
| | (38.212) | (37.932) | (0.559) |
| Observations | 59,609 | 5,064 | 64,673 |

Note: This table tests whether hyperlipidemiapatients in the treatment and control groups are observably similar at baseline. The "Difference" column reports t-tests of the difference in means between "treatment" and "control" groups. Significance stars are at the 1, 5, and 10% levels.

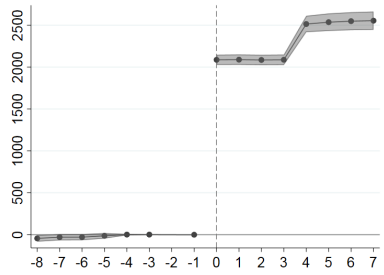
## 9.2 First-Stage

Table 7: Diabetes Sample First-Stage

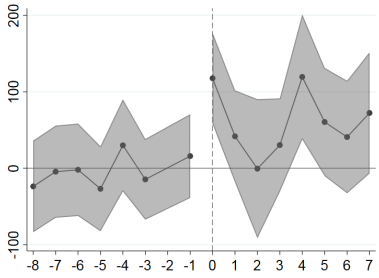
| | Individual Deductible | Family Deductible | Out-of-Pocket |
|---|---|---|---|
| Post*Treat | 1040.554*** | 2252.284*** | 58.644*** |
| | (70.720) | (64.316) | (3.871) |
| Mean | 939.86 | 2124.21 | 256.56 |
| Obs | 4,079,239 | 4,079,239 | 4,079,537 |




(a) Diabetes - Indv Ded



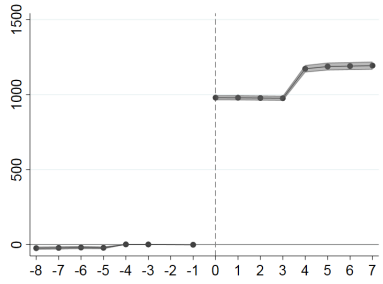

(b) Diabetes - Fam Ded



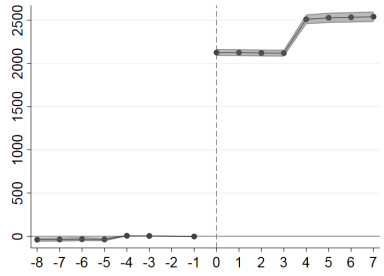

(c) Diabetes - OOP

Table 8: Hyperlipidemia Sample First-Stage

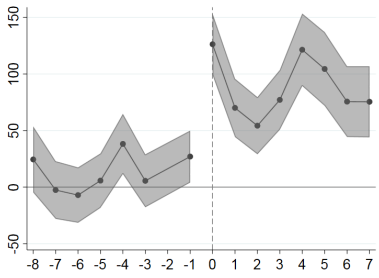
| | Individual Deductible | Family Deductible | Out-of-Pocket |
|---|---|---|---|
| Post*Treat | 1060.908*** | 2283.278*** | 74.488*** |
| | (116.018) | (108.129) | (14.833) |
| Mean | 907.06 | 2058.38 | 228.26 |
| Obs | 10,986,248 | 10,986,248 | 10,986,824 |




(a) Hyperlipidemia - Indv Ded




(b) Hyperlipidemia - Fam Ded



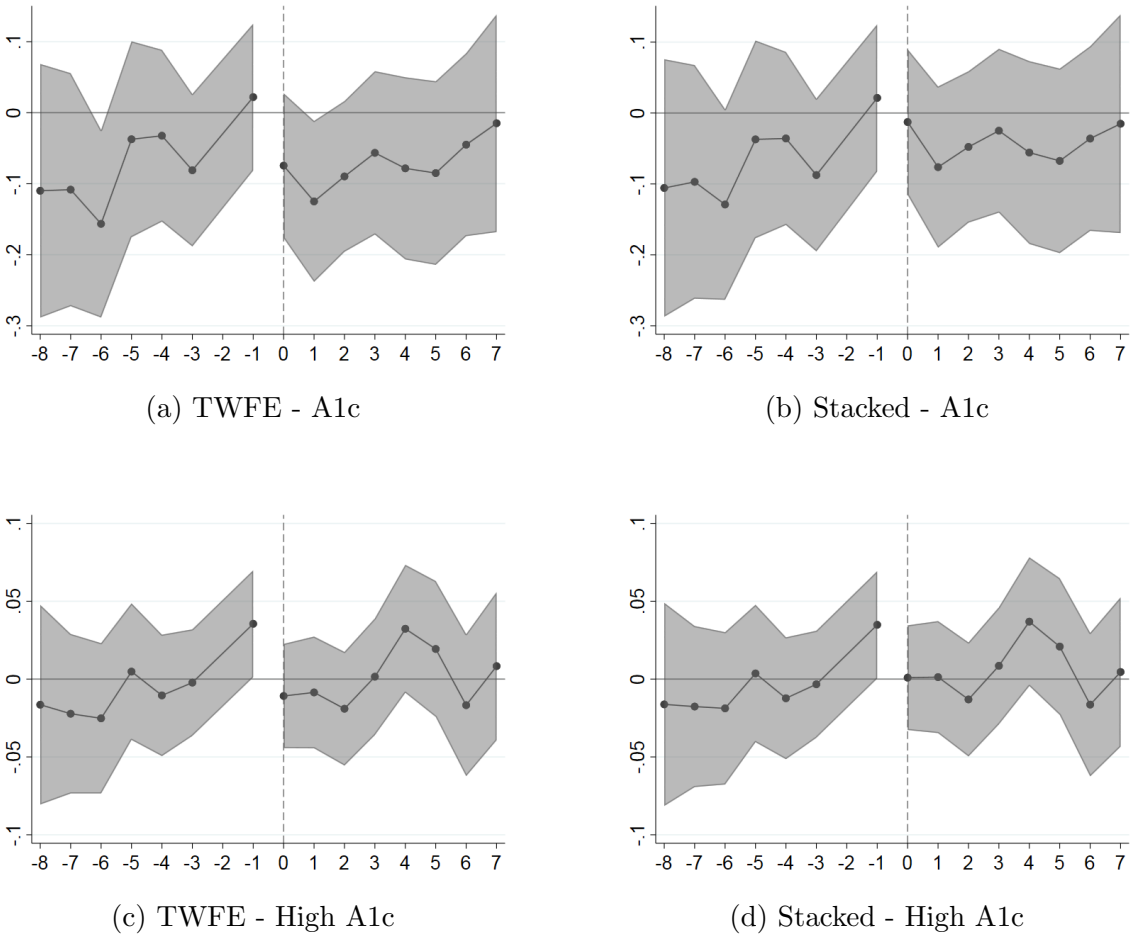

(c) Hyperlipidemia - OOP

## 9.3 Primary Outcomes - Diabetes

Table 9: A1c Results for Full Sample

| | A1c level | | High A1c> 7.5% | |
|---|---|---|---|---|
| | TWFE | Stacked | TWFE | Stacked |
| Post*Treat | -0.041 | -0.004 | -0.006 | 0.001 |
| | (-1.341) | (-0.125) | (-0.604) | (0.135) |
| Mean | 7.38 | 7.35 | 0.34 | 0.33 |
| Obs | 124,992 | 700,139 | 124,992 | 700,139 |

Figure 6: Event Study Graphs




(a) TWFE - A1c

(b) Stacked - A1c

(c) TWFE - High A1c

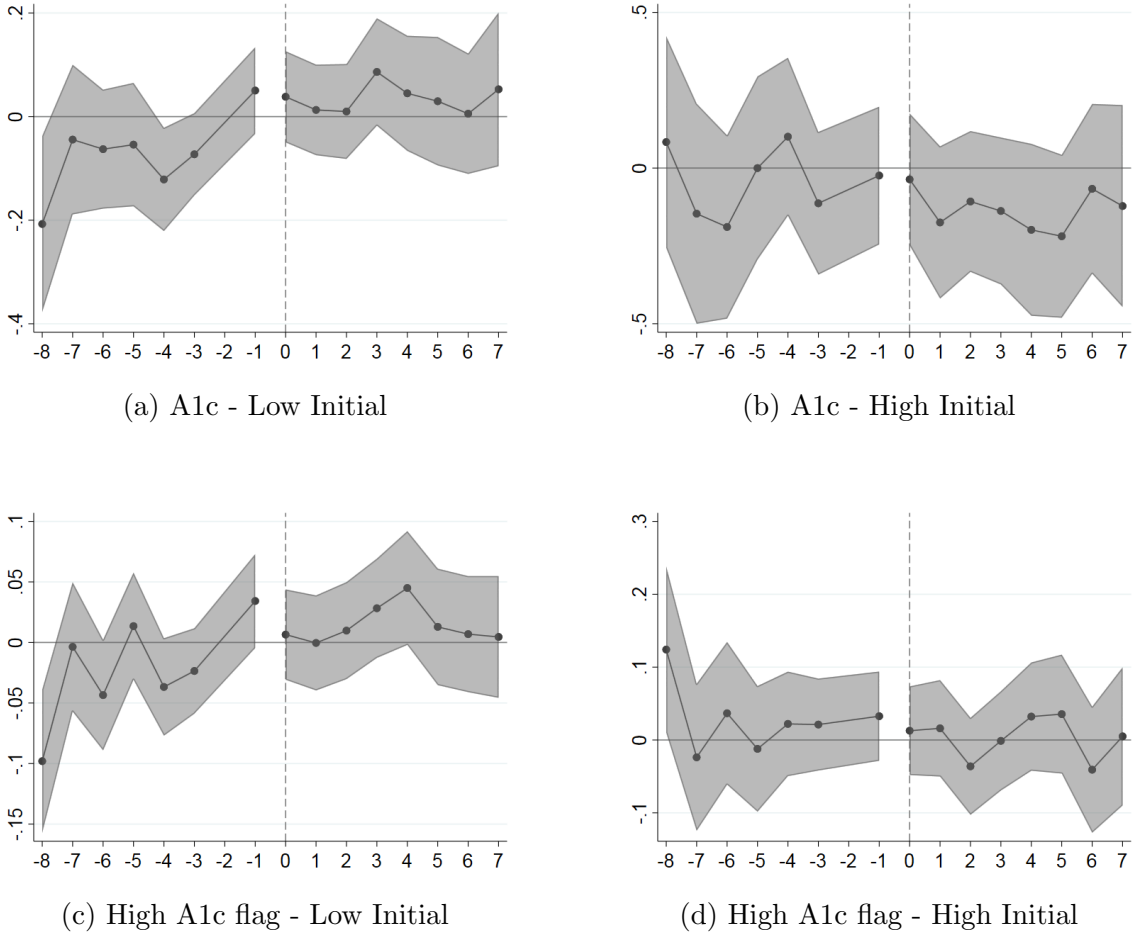
(d) Stacked - High A1c

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\alpha_t$ from Equation 2 (left) & $\beta_t$ from Equation 4 (right) for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

Table 10: A1c Results by Initial Disease Severity

| | A1c Value | | High A1c> 7.5% | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.077*** | -0.098 | 0.023** | -0.018 |
| | (2.613) | (-1.624) | (2.287) | (-1.063) |
| Mean | 6.68 | 8.52 | 0.12 | 0.69 |
| Obs | 445,221 | 254,918 | 445,221 | 254,918 |

Figure 7: Event Study Graphs




(a) A1c - Low Initial

(b) A1c - High Initial

(c) High A1c flag - Low Initial

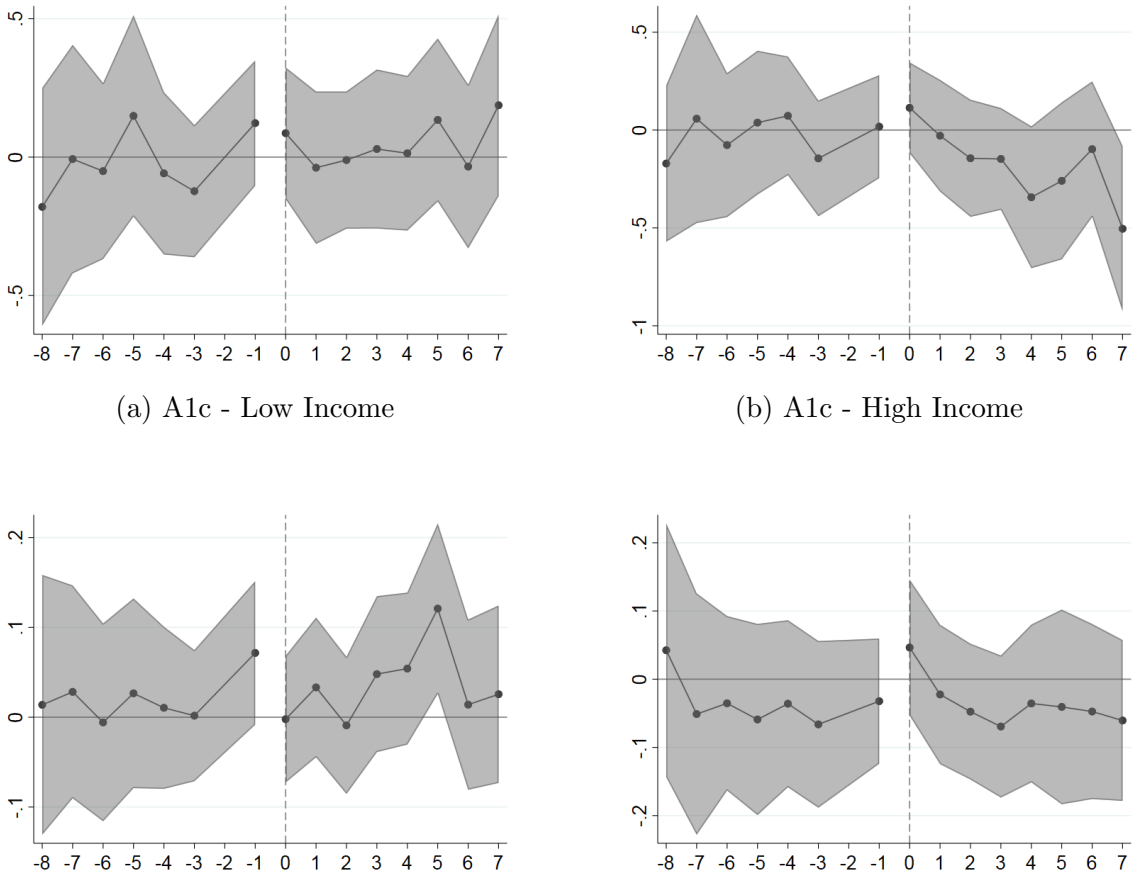
(d) High A1c flag - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on patients enrolled in plans that change ("treatment") relative to patients enrolled in plans that do not ("control") with the quarter 2 quarters before the change as the base period.

Table 11: A1c Results by Income

| | A1c Value | | High A1c> 7.5% | |
|---|---|---|---|---|
| | Low Income | High Income | Low Income | High Income |
| Post*Treat | 0.043 | -0.101 | 0.008 | 0.004 |
| | (0.580) | (-1.382) | (0.420) | (0.149) |
| Mean | 7.35 | 7.35 | 0.33 | 0.33 |
| Obs | 693,717 | 692,770 | 693,717 | 692,770 |

Figure 8: Event Study Graphs




(a) A1c - Low Income

(b) A1c - High Income

(c) High A1c flag - Low Income

(d) High A1c flag - High Income

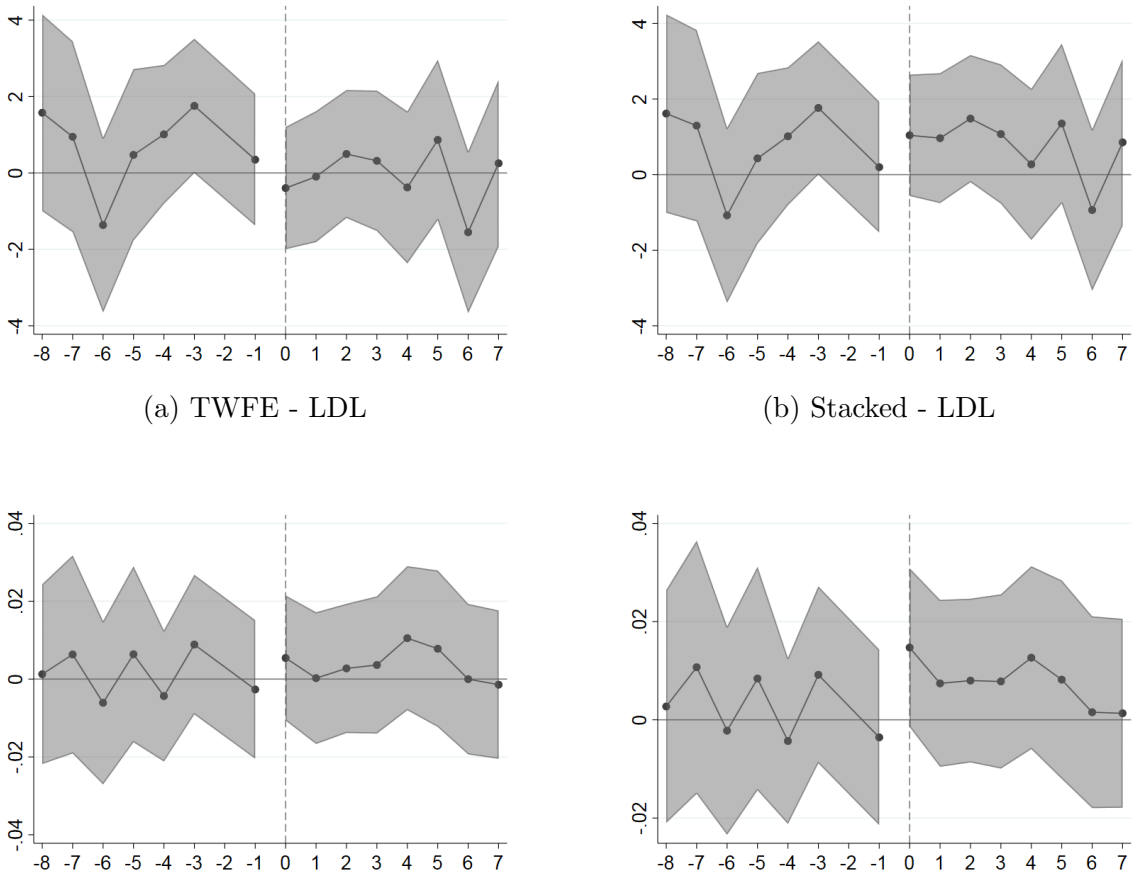
Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on patients enrolled in plans that change ("treatment") relative to patients enrolled in plans that do not ("control") with the quarter 2 quarters before the change as the base period.

## 9.4 Primary Outcomes - Hyperlipidemia

Table 12: LDL Results for Full Sample

| | LDL level | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | TWFE | Stacked | TWFE | Stacked |
| Post*Treat | -0.670* | 0.226 | 0.003 | 0.007* |
| | (-1.649) | (0.549) | (0.723) | (1.834) |
| Mean | 107.79 | 106.85 | 0.08 | 0.08 |
| Obs | 269,224 | 1,567,005 | 269,224 | 1,567,005 |

Figure 9: Event Study Graphs




(a) TWFE - LDL

(b) Stacked - LDL

(c) TWFE - High LDL

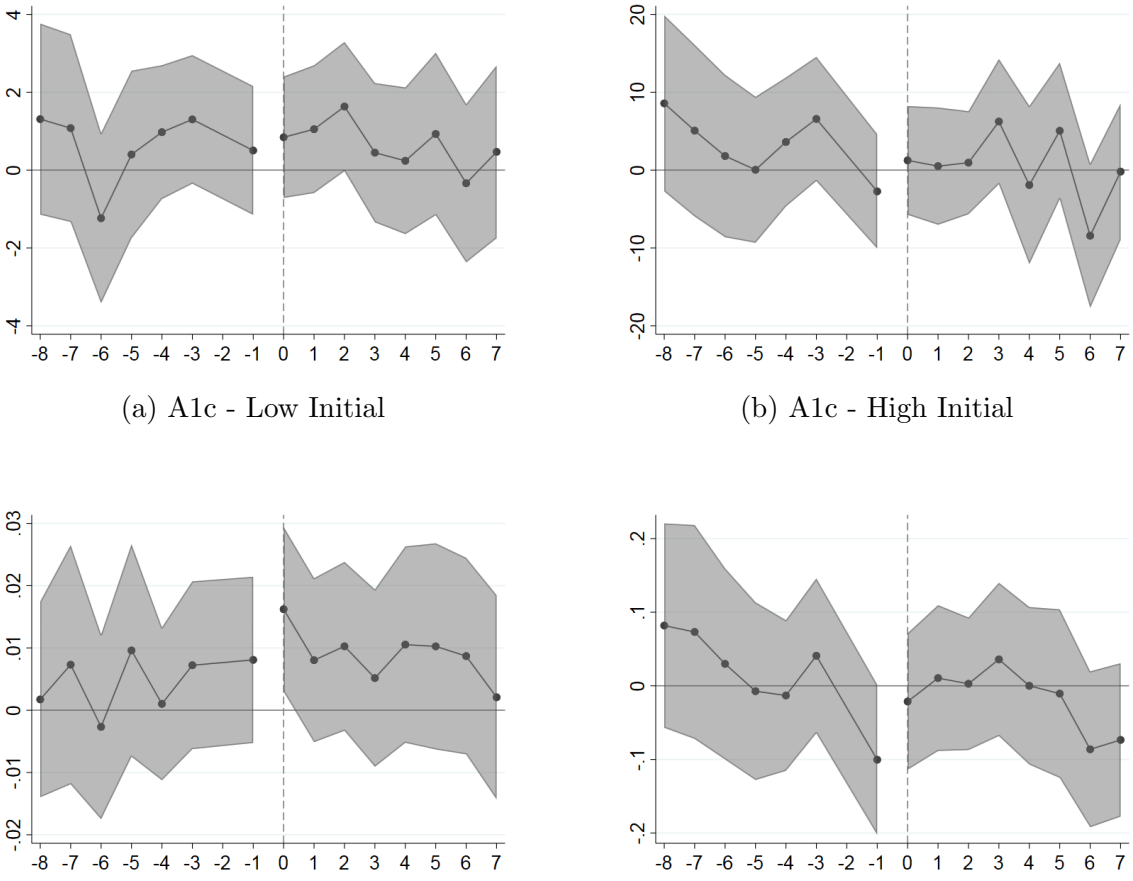
(d) Stacked - High LDL

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\alpha_t$ from Equation 2 (left) & $\beta_t$ from Equation 4 (right) for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

Table 13: LDL Results by Initial Disease Severity

| | LDL Value | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.188 | -0.929 | 0.005* | -0.002 |
| | (0.476) | (-0.531) | (1.815) | (-0.075) |
| Mean | 102.12 | 145.17 | 0.04 | 0.42 |
| Obs | 1,395,193 | 171,812 | 1,395,193 | 171,812 |

Figure 10: Event Study Graphs




(a) A1c - Low Initial

(b) A1c - High Initial

(c) High A1c flag - Low Initial

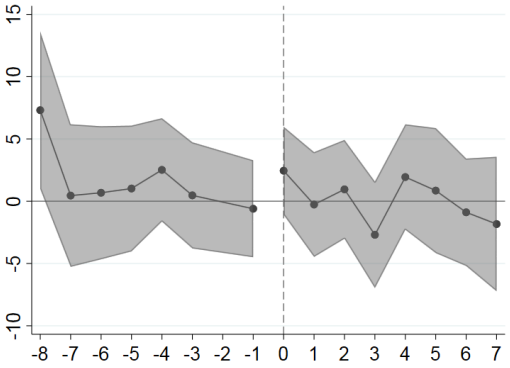
(d) High A1c flag - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on patients enrolled in plans that change ("treatment") relative to patients enrolled in plans that do not ("control") with the quarter 2 quarters before the change as the base period.

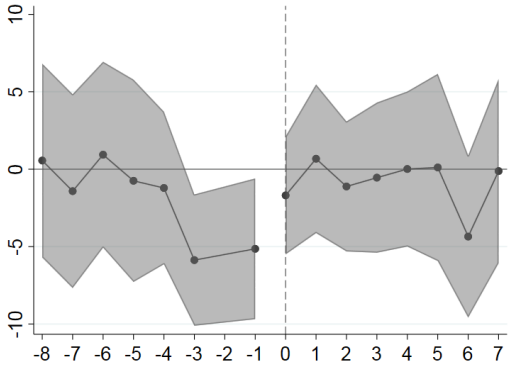
Table 14: LDL Results by Income

| | LDL Value | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | Low Income | High Income | Low Income | High Income |
| Post*Treat | -0.674 | 1.609 | 0.000 | 0.016 |
| | (-0.725) | (1.574) | (0.035) | (1.367) |
| Mean | 106.84 | 106.84 | 0.08 | 0.08 |
| Obs | 1,551,547 | 1,550,411 | 1,551,547 | 1,550,411 |

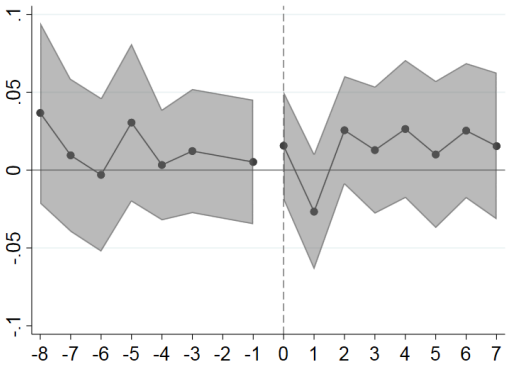
Figure 11: Event Study Graphs



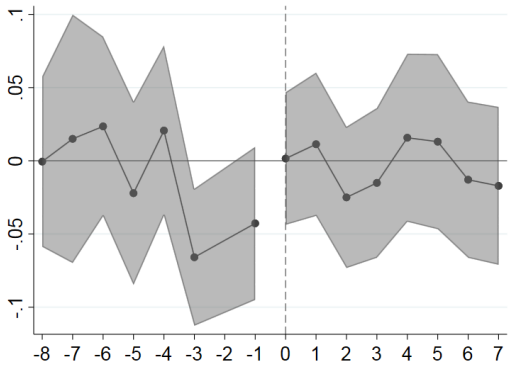

(a) A1c - Low Income




(b) A1c - High Income




(c) High A1c flag - Low Income




(d) High A1c flag - High Income

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on patients enrolled in plans that change ("treatment") relative to patients enrolled in plans that do not ("control") with the quarter 2 quarters before the change as the base period.

# 10 Mechanisms

## 10.1 Number of Tests

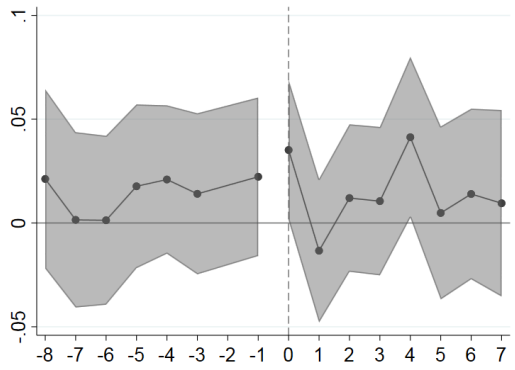
Table 15: Diabetes

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 0.001 | -0.011 |
| | (0.126) | (-1.039) |
| Mean | 0.43 | 0.47 |
| Obs | 2,590,202 | 1,489,335 |

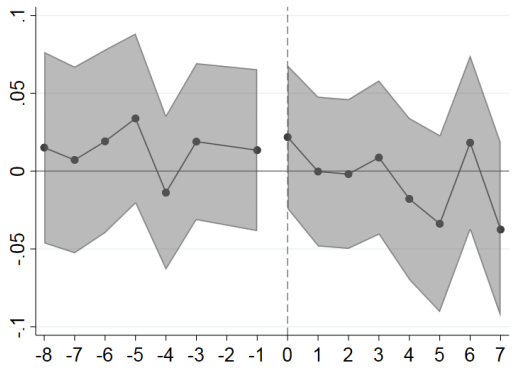
Table 16: Hyperlipidemia

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.008 | -0.015 |
| | (-1.240) | (-0.855) |
| Mean | 0.61 | 0.54 |
| Obs | 9,603,569 | 1,383,255 |

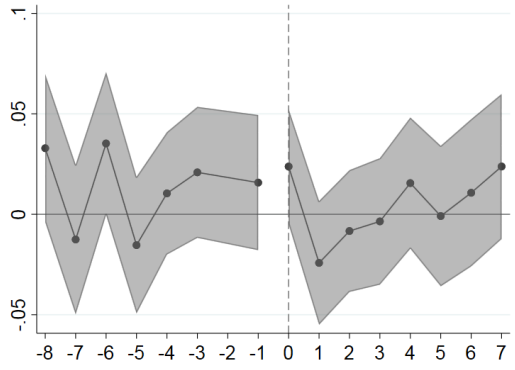
Figure 12: Event Study Graphs



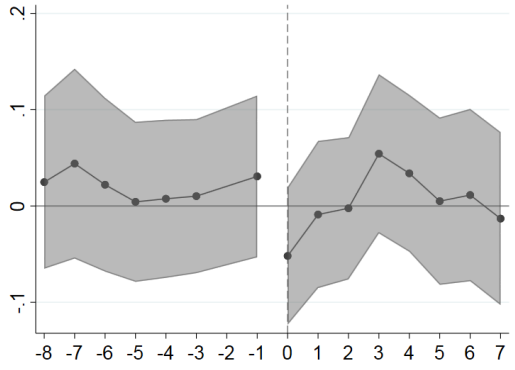

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial




(d) Hyperlipidemia - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

## 10.2 High Value Prescriptions

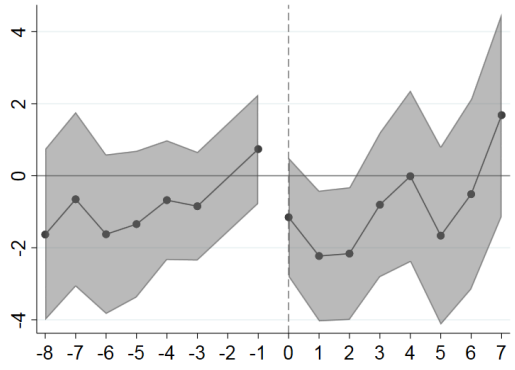
Table 17: Metformin

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.500 | -1.643* |
| | (-0.703) | (-1.862) |
| Mean | 32.85 | 35.10 |
| Obs | 2,590,202 | 1,489,335 |

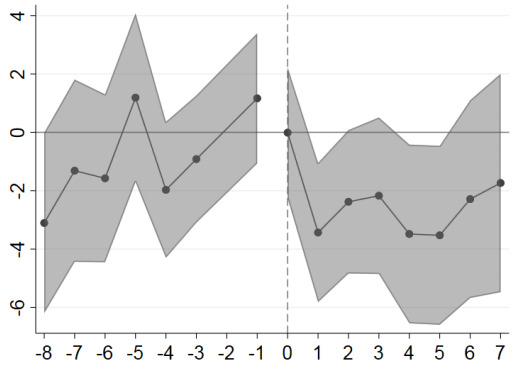
Table 18: OAD

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.124 | 0.663 |
| | (-0.124) | (0.404) |
| Mean | 56.17 | 73.19 |
| Obs | 2,590,202 | 1,489,335 |

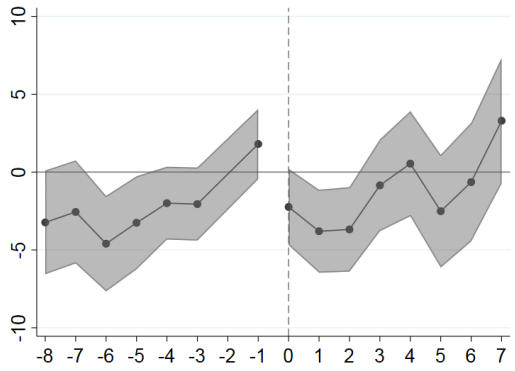
Figure 13: Event Study Graphs



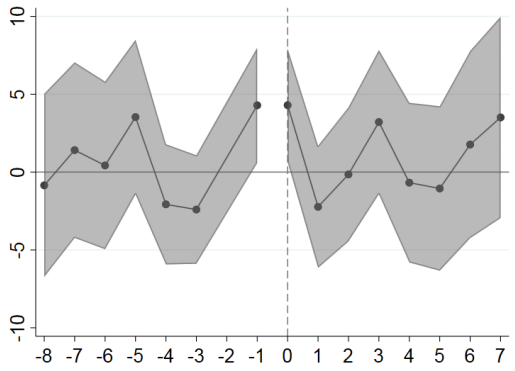

(a) Metformin - Low Initial




(b) Metformin- High Initial




(c) OAD - Low Initial




(d) OAD - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

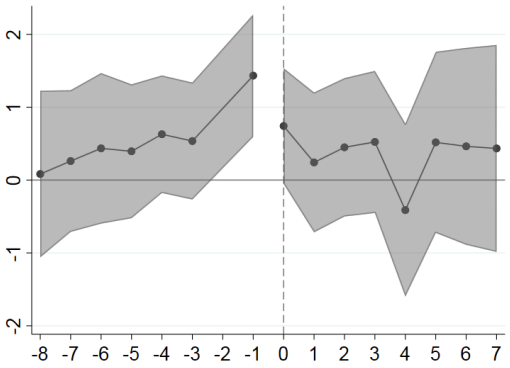
Table 19: Insulin

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.137 | 0.180 |
| | (-0.425) | (0.219) |
| Mean | 7.26 | 25.68 |
| Obs | 2,590,202 | 1,489,335 |

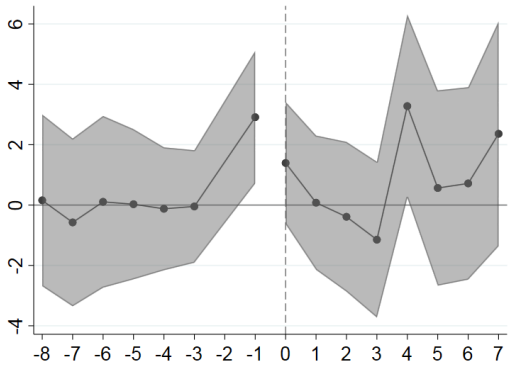
Table 20: Diabetes Supplies

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.016 | 0.337 |
| | (-0.031) | (0.454) |
| Mean | 14.02 | 20.58 |
| Obs | 2,590,202 | 1,489,335 |

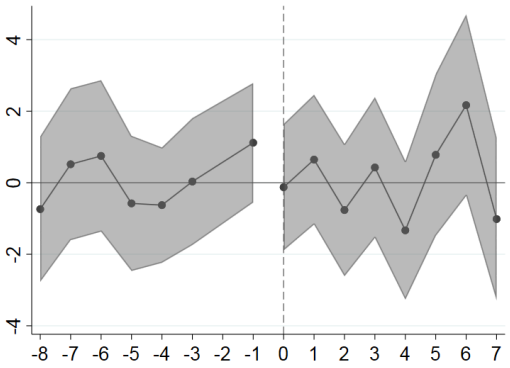
Figure 14: Event Study Graphs



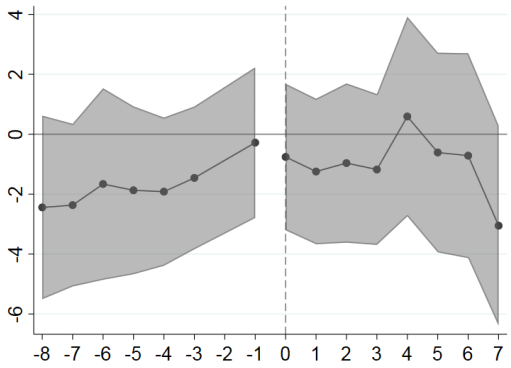

(a) Insulin - Low Initial




(b) Insulin - High Initial




(c) Diabetes Supplies - Low Initial



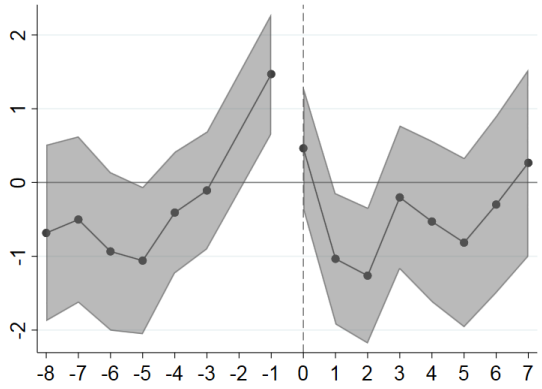

(d) Diabetes Supplies - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

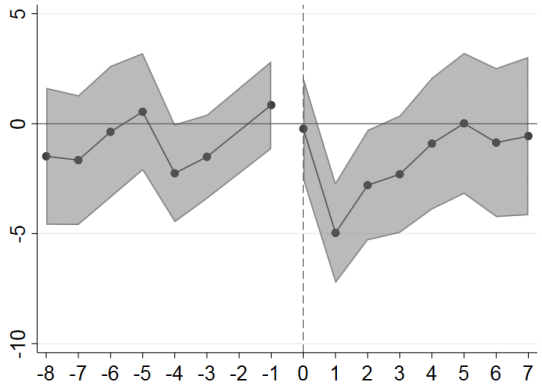
Table 21: Statins Days Supplied

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.324 | -1.136 |
| | (-1.040) | (-1.302) |
| Mean | 34.19 | 24.55 |
| Obs | 9,603,569 | 1,383,255 |

Figure 15: Event Study Graphs




(a) Low Initial




(b) High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

## 10.3 Primary Care Office Visits

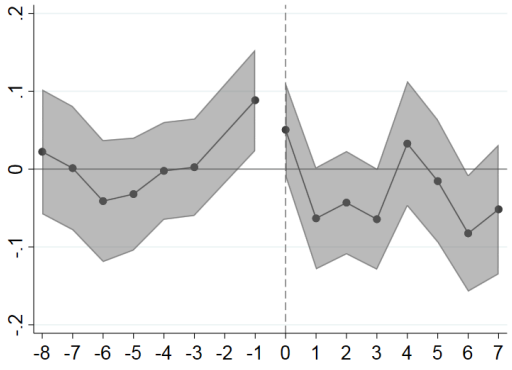
Table 22: Diabetes

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.037** | -0.025 |
| | (-2.145) | (-1.271) |
| Mean | 0.89 | 0.87 |
| Obs | 2,590,202 | 1,489,335 |

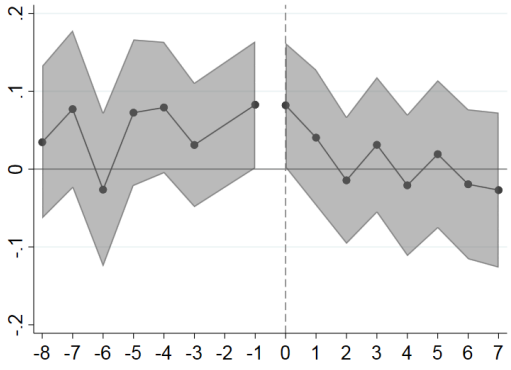
Table 23: Hyperlipidemia

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.014* | -0.026 |
| | (-1.869) | (-1.320) |
| Mean | 0.83 | 0.79 |
| Obs | 9,603,569 | 1,383,255 |

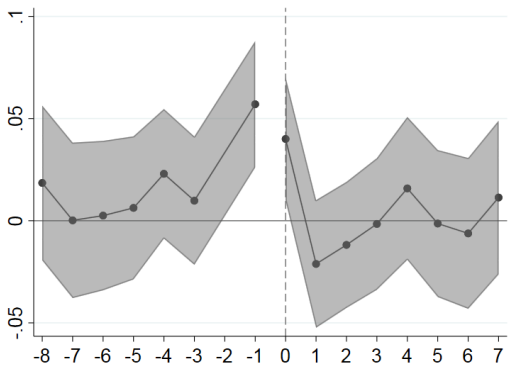
Figure 16: Event Study Graphs



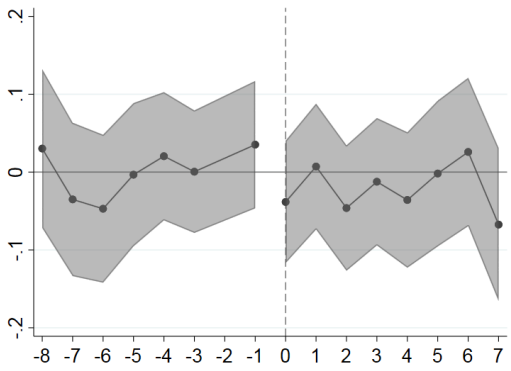

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial




(d) Hyperlipidemia - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

# 11 Appendix

## 11.1 Out-of-Pocket Spending

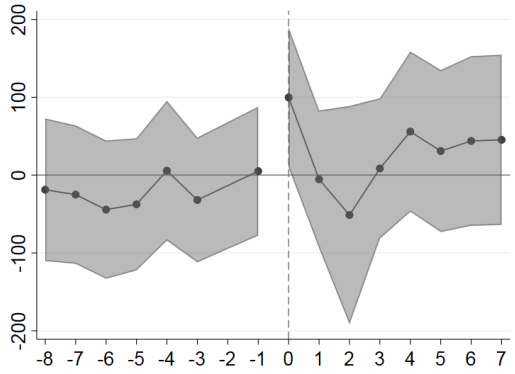
Table A1: Diabetes

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 40.113* | 89.222*** |
| | (1.792) | (5.783) |
| Mean | 262.24 | 246.69 |
| Obs | 2,590,202 | 1,489,335 |

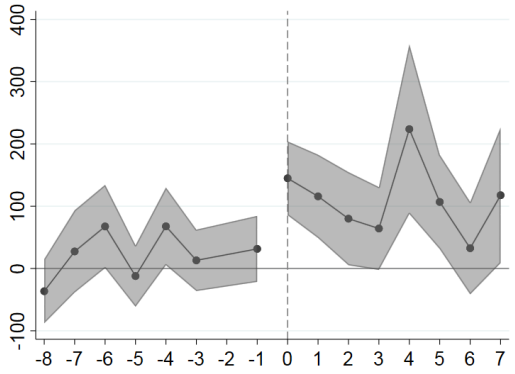
Table A2: Hyperlipidemia

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 72.988*** | 84.913*** |
| | (13.316) | (7.178) |
| Mean | 230.63 | 211.84 |
| Obs | 9,603,569 | 1,383,255 |

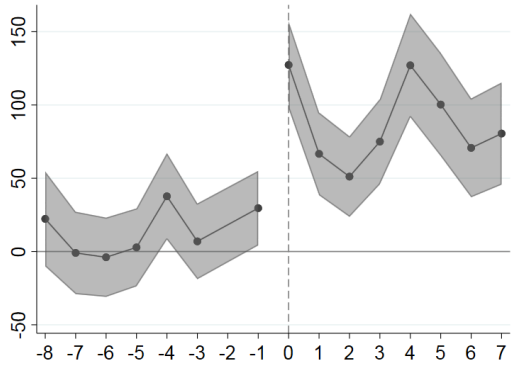
Figure A1: Event Study Graphs



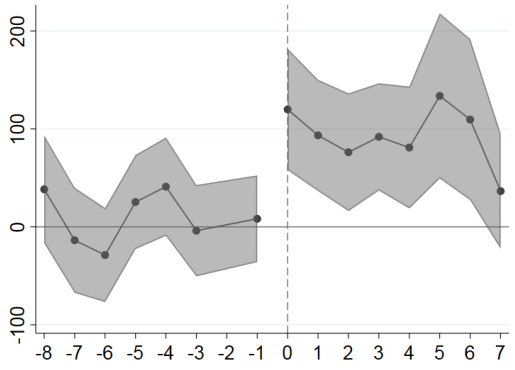

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial

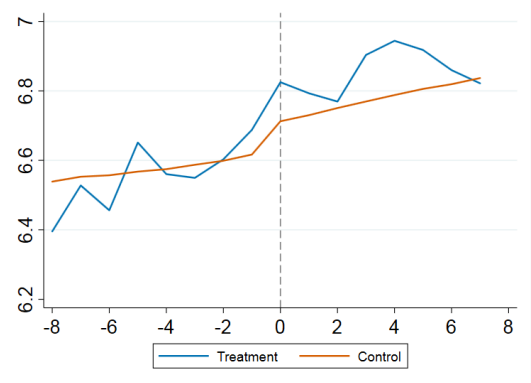



(d) Hyperlipidemia - High Initial

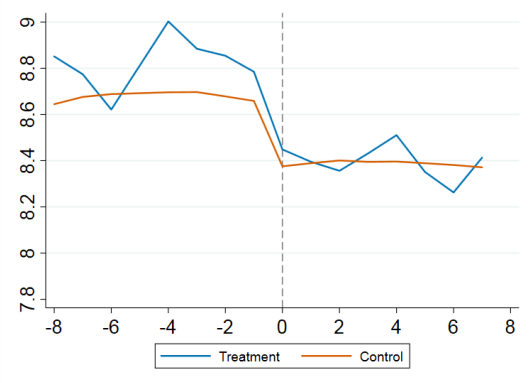
Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

## 11.2 Unadjusted Biomarker Trends

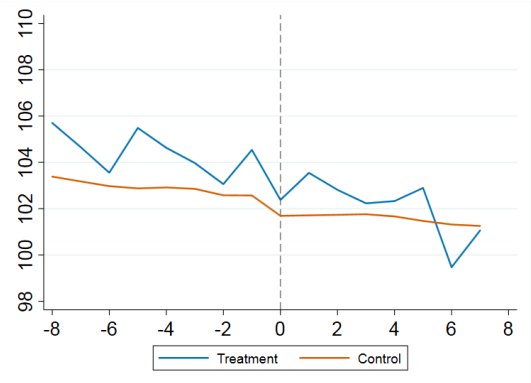
Figure A2: Raw Plots by Initial Disease Severity



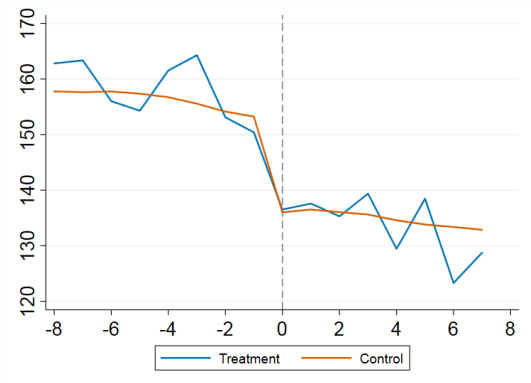

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial




(d) Hyperlipidemia - High Initial

Note: Figure shows raw unadjusted plots of mean A1c values for diabetes patients (top panel), and mean LDL values for hyperlipidemia patients (bottom panel) separated out by "treatment" and "control" patients. This allows us to visually inspect whether the parallel trends assumption in the pre-period holds.

## 11.3 OOP Results for A1c or LDL Tests

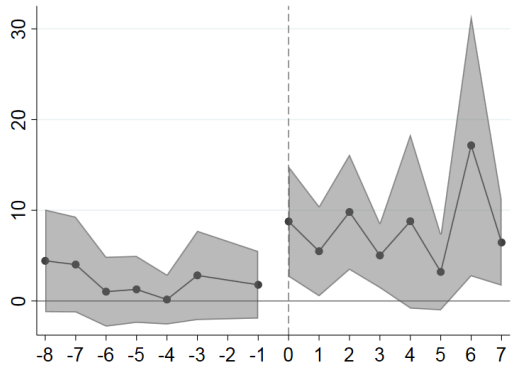
Table A3: Diabetes

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 6.175*** | 4.973*** |
| | (3.972) | (2.715) |
| Mean | 14.75 | 15.55 |
| Obs | 2,590,202 | 1,489,335 |

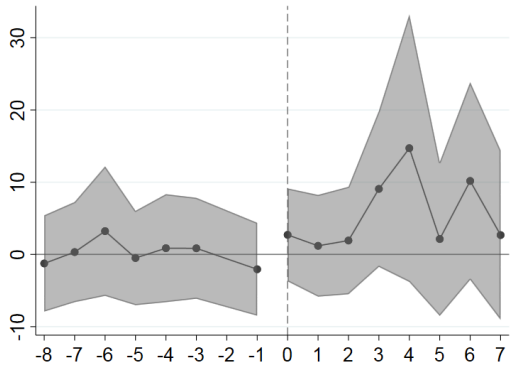
Table A4: Hyperlipidemia

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 4.142*** | 5.061* |
| | (4.656) | (1.884) |
| Mean | 16.25 | 14.39 |
| Obs | 9,603,569 | 1,383,255 |

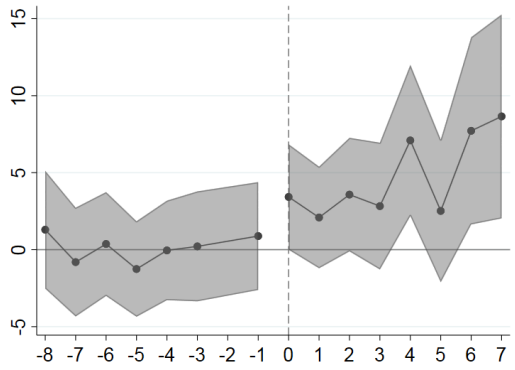
Figure A3: Event Study Graphs



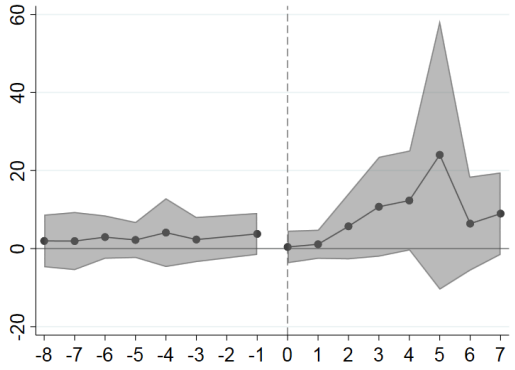

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial




(d) Hyperlipidemia - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

## 11.4 OOP Results for High Value Prescription Drugs

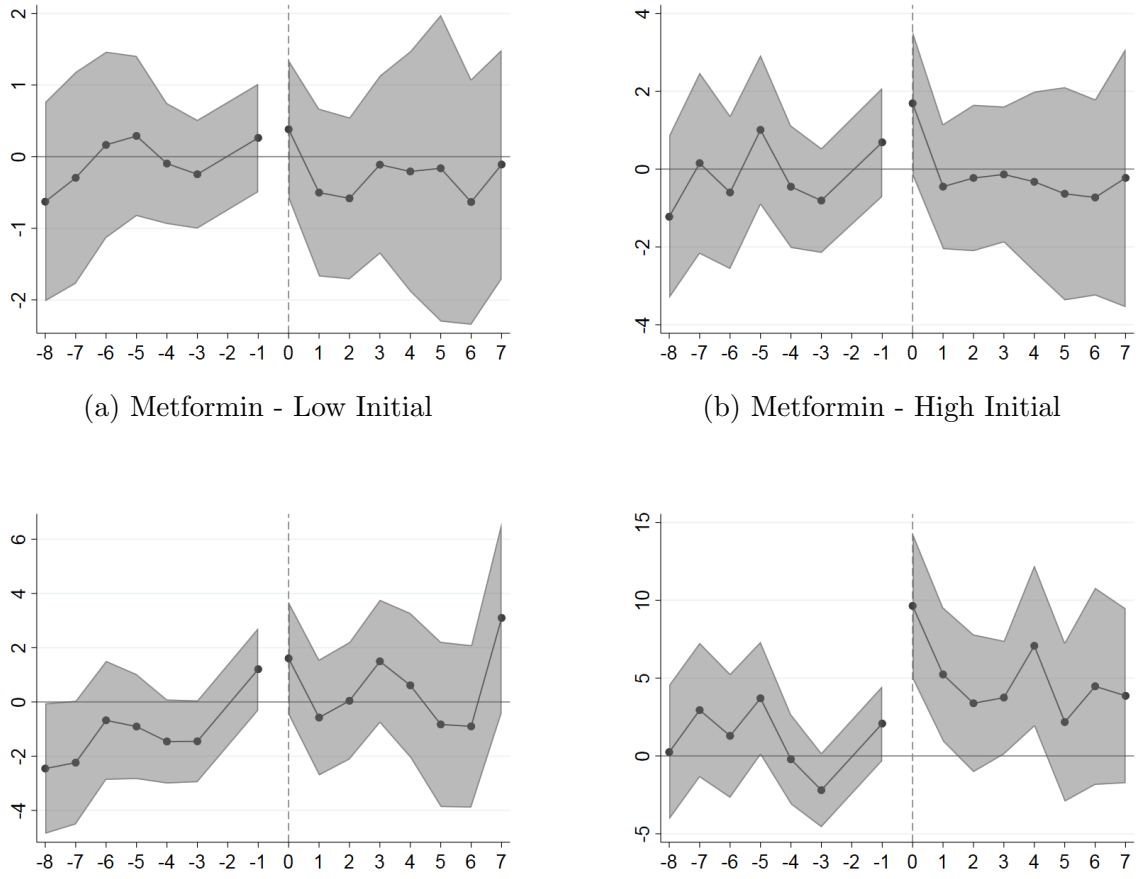
Table A5: Metformin

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | -0.180 | 0.132 |
| | (-0.377) | (0.191) |
| Mean | 12.41 | 14.56 |
| Obs | 2,590,202 | 1,489,335 |

Table A6: Oral Anti-Diabetic Drugs OOP

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 1.370* | 4.619*** |
| | (1.665) | (2.954) |
| Mean | 29.56 | 41.55 |
| Obs | 2,590,202 | 1,489,335 |

Figure A4: Event Study Graphs



(a) Metformin - Low Initial

(b) Metformin - High Initial

(c) OAD - Low Initial

(d) OAD - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

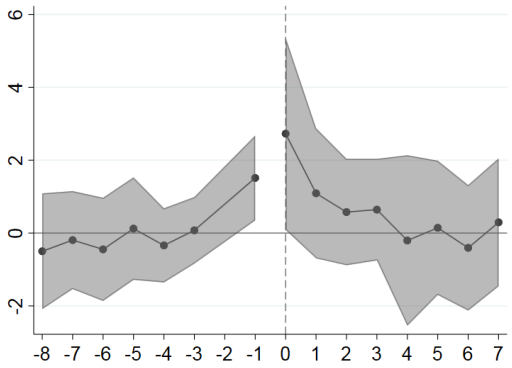
Table A7: Insulin OOP

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 0.690 | 3.911*** |
| | (1.158) | (2.885) |
| Mean | 8.17 | 30.65 |
| Obs | 2,590,202 | 1,489,335 |

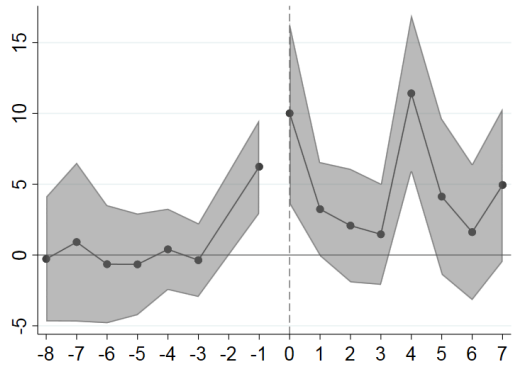
Table A8: Diabetes Supplies OOP

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 0.729* | 0.680 |
| | (1.849) | (1.090) |
| Mean | 6.99 | 10.70 |
| Obs | 2,590,202 | 1,489,335 |

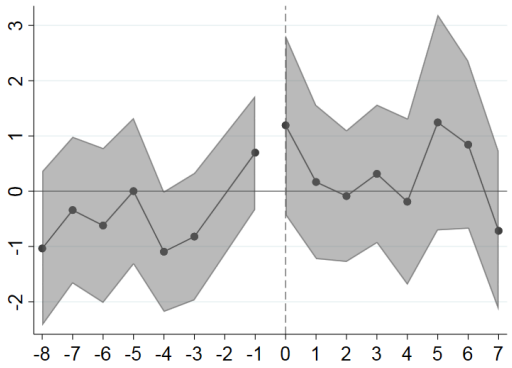
Figure A5: Event Study Graphs



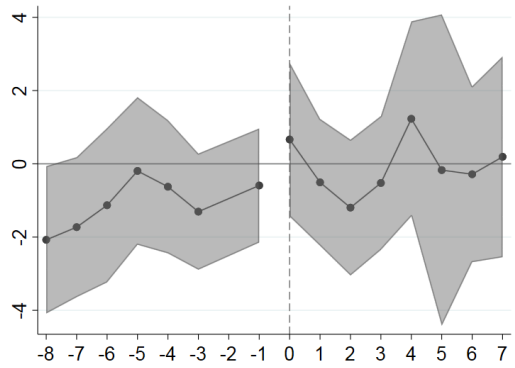

(a) Insulin - Low Initial




(b) Insulin - High Initial




(c) Diabetes Supplies - Low Initial



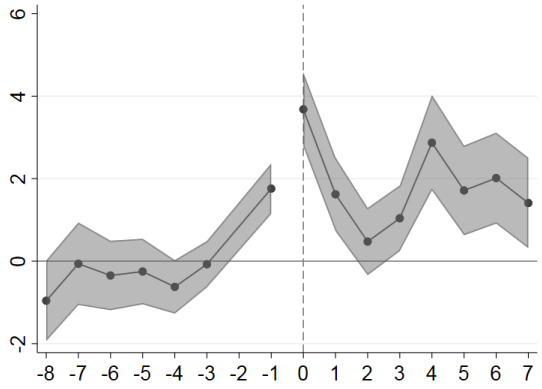

(d) Diabetes Supplies - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

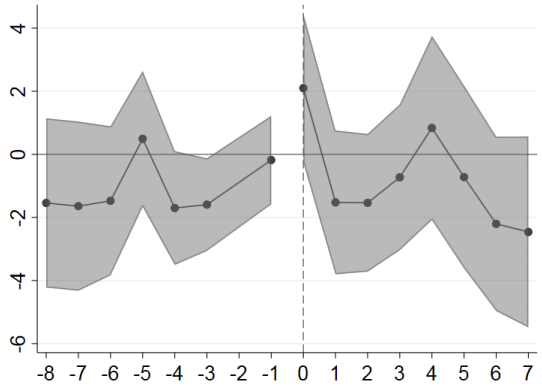
Table A9: Hyperlipidemia - Statins OOP

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 1.820*** | 0.292 |
| | (5.748) | (0.371) |
| Mean | 20.65 | 15.25 |
| Obs | 9,603,569 | 1,383,255 |

Figure A6: Event Study Graphs




(a) Low Initial




(b) High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

## 11.5 OOP Results for PCP Visits

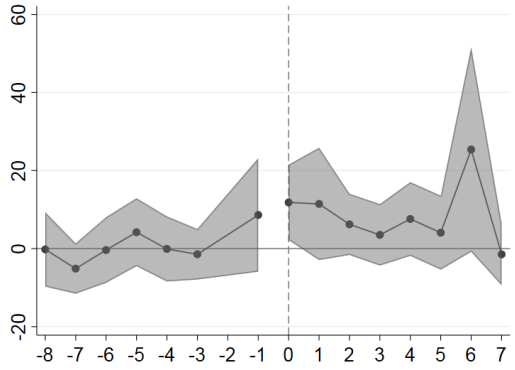
Table A10: Diabetes

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 7.464*** | 8.874*** |
| | (2.631) | (2.688) |
| Mean | 37.20 | 36.86 |
| Obs | 2,590,202 | 1,489,335 |

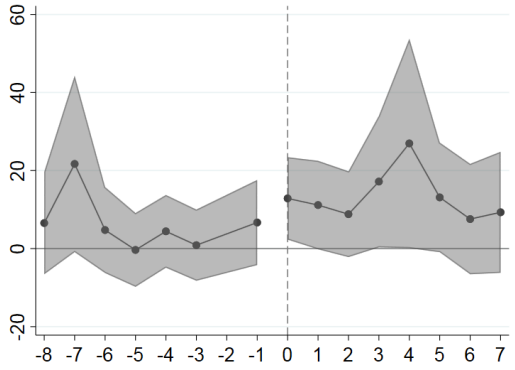
Table A11: Hyperlipidemia

| | Low Initial | High Initial |
|---|---|---|
| Post*Treat | 10.982*** | 7.722*** |
| | (9.282) | (3.174) |
| Mean | 34.05 | 31.78 |
| Obs | 9,603,569 | 1,383,255 |

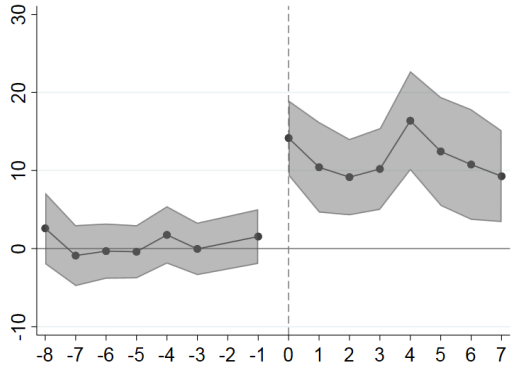
Figure A7: Event Study Graphs



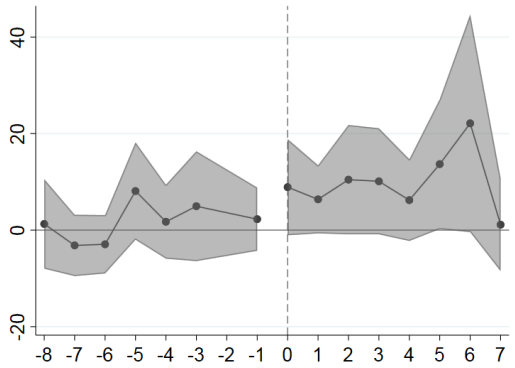

(a) Diabetes - Low Initial




(b) Diabetes - High Initial




(c) Hyperlipidemia - Low Initial




(d) Hyperlipidemia - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

# 12 Robustness

## 12.1 Vary Treatment Size

**Diabetes**

Table A12: "Small" Deductible Change

| | A1c Value | | High A1c> 7.5% | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.067*** | -0.068 | 0.021** | -0.017 |
| | (2.595) | (-1.303) | (2.338) | (-1.140) |
| Mean | 6.68 | 8.52 | 0.12 | 0.69 |
| Obs | 446,689 | 255,872 | 446,689 | 255,872 |

Note: Table shows differences-in-differences estimates around the time a "low" deductible change goes into effect. This is a change in the individual deductible of $\geq$ \$250, and/or a change in the family deductible of $\geq$ \$500. The results in the table are the coefficient $\beta_1$ from Equation 3. It shows the effect of the deductible increase on biomarker values for "treatment" patients relative to "control" patients.

Table A13: "Large" Deductible Change

| | A1c Value | | High A1c> 7.5% | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.067 | -0.121 | 0.022* | -0.013 |
| | (1.588) | (-1.563) | (1.664) | (-0.603) |
| Mean | 6.68 | 8.52 | 0.12 | 0.69 |
| Obs | 442,906 | 253,629 | 442,906 | 253,629 |

Note: Table shows differences-in-differences estimates around the time a "high" deductible change goes into effect. This is a change in the individual deductible of $\geq \$1,000$, and/or a change in the family deductible of $\geq \$2,000$. The results in the table are the coefficient $\beta_1$ from Equation 3. It shows the effect of the deductible increase on biomarker values for "treatment" patients relative to "control" patients.

## Hyperlipidemia

Table A14: "Small" Deductible Change

| | LDL Value | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.022 | -1.229 | 0.003 | -0.014 |
| | (0.061) | (-0.780) | (1.310) | (-0.740) |
| Mean | 102.13 | 145.19 | 0.04 | 0.42 |
| Obs | 1,399,410 | 172,337 | 1,399,410 | 172,337 |

Note: Table shows differences-in-differences estimates around the time a "low" deductible change goes into effect. This is a change in the individual deductible of $\geq$ \$250, and/or a change in the family deductible of $\geq$ \$500. The results in the table are the coefficient $\beta_1$ from Equation 3. It shows the effect of the deductible increase on biomarker values for "treatment" patients relative to "control" patients.

Table A15: "Large" Deductible Change

| | LDL Value | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.019 | 0.516 | 0.005 | -0.012 |
| | (0.038) | (0.220) | (1.305) | (-0.424) |
| Mean | 102.12 | 145.17 | 0.04 | 0.42 |
| Obs | 1,387,757 | 170,950 | 1,387,757 | 170,950 |

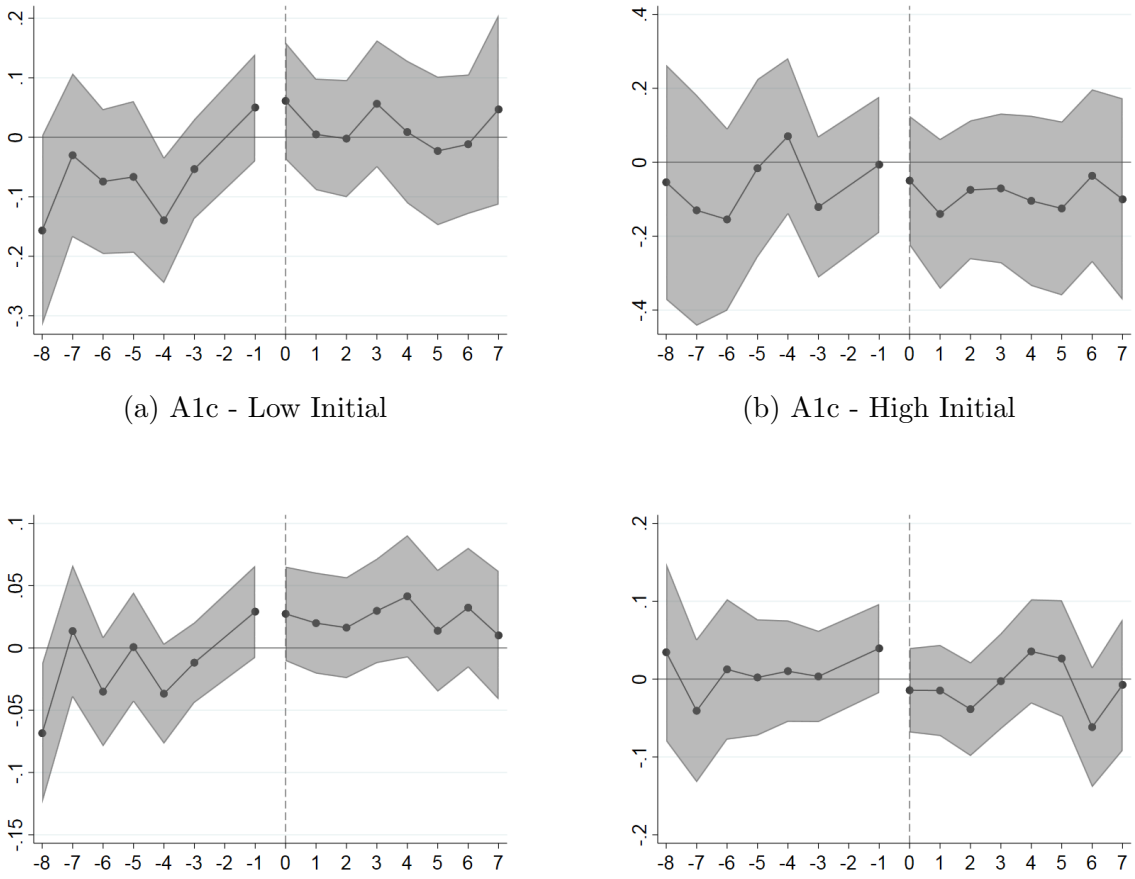
Note: Table shows differences-in-differences estimates around the time a "high" deductible change goes into effect. This is a change in the individual deductible of $\geq \$1,000$, and/or a change in the family deductible of $\geq \$2,000$. The results in the table are the coefficient $\beta_1$ from Equation 3. It shows the effect of the deductible increase on biomarker values for "treatment" patients relative to "control" patients.

## 12.2 Vary Threshold Size

Table A16: Diabetes

| | A1c Value | | High A1c> 7.5% | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.060** | -0.053 | 0.030*** | -0.024 |
| | (2.173) | (-0.978) | (3.022) | (-1.511) |
| Mean | 6.55 | 8.28 | 0.09 | 0.60 |
| Obs | 374,602 | 325,537 | 374,602 | 325,537 |

Figure A8: Event Study Graphs



(a) A1c - Low Initial

(b) A1c - High Initial

(c) High A1c - Low Initial

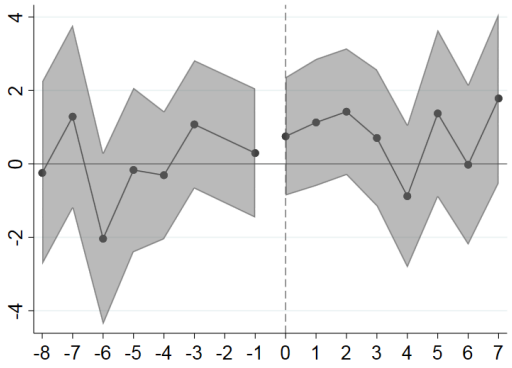
(d) High A1c - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.

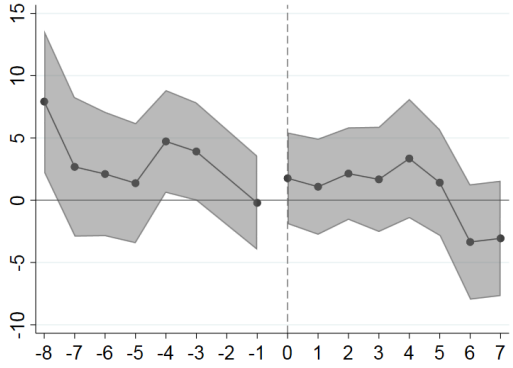
Table A17: Hyperlipidemia

| | LDL Value | | High LDL> 160 mg/dl | |
|---|---|---|---|---|
| | Low Initial | High Initial | Low Initial | High Initial |
| Post*Treat | 0.712* | -1.151 | 0.003 | 0.015 |
| | (1.698) | (-1.363) | (1.004) | (1.378) |
| Mean | 95.30 | 133.74 | 0.02 | 0.21 |
| Obs | 1,096,385 | 470,620 | 1,096,385 | 470,620 |

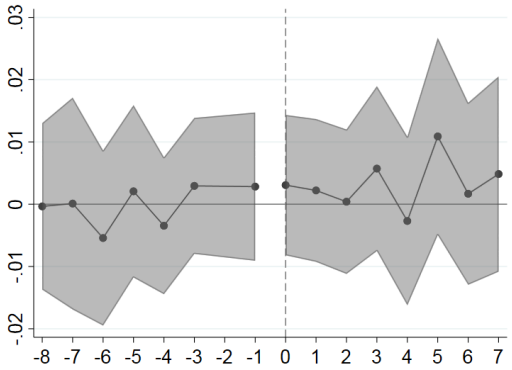
Figure A9: Event Study Graphs



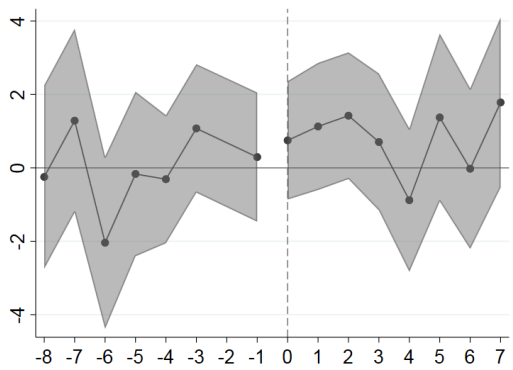

(a) LDL - Low Initial




(b) LDL - High Initial




(c) High LDL - Low Initial




(d) High LDL - High Initial

Note: Figure shows event studies around the time a deductible change goes into effect (marked by a dashed grey line). The data points plotted are coefficients $\beta_t$ from Equation 4 for each quarter around the change. They show the effect on "treatment" vs. "control" patients relative to 2 quarters before the change.