What Does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics
Appendix for Online Publication

Zarek C. Brot-Goldberg, Amitabh Chandra, Benjamin R. Handel, Jonathan T. Kolstad

December 14, 2016

Abstract: This appendix supplements the main text with additional analyses and robustness checks. It is organized following the order of the main article to provide for easy navigation. The main article discusses much of the material in this appendix in the context of the primary results. Please contact the authors of this study with any further questions about these additional analyses.

A1 Robustness Checks of Primary Sample Construction

The main sample we use throughout the paper is constructed so as to ensure we can analyze long-term trends in spending. We constructed a similar sample using weaker restrictions to show that our sample restrictions are innocuous in terms of their effects on the final result. Our primary sample is restricted to only include employees who were enrolled in a health insurance plan at the firm for all years between t_{-4} and t_1, the entire span of our data. Our alternate sample is only restricted to employees who were enrolled between t_{-2} and t_0, which includes employees who may have left the firm in t_1, or joined it in t_{-4} or t_{-3}. Summary statistics for our main sample and this alternative are given in the first two columns of Table A1. This new sample includes approximately 8,000 additional employees and 10,000 additional dependents. These excluded employees are relatively younger, and have smaller families (mostly those employees who joined the first during t_{-4} or t_{-3}), but the overall mix of ages among them and their dependents changes only slightly. Most importantly, the distribution of health spending is nearly identical.

Another concern with our approach is that, since employees were aware of the policy change well in advance, they might make the decision to leave the firm in advance of being required to switch into a health insurance plan with cost-sharing. To test this, we plot the hazard rate of employees and their dependents exiting the firm's insurance coverage for each month in our data. We see no meaningful change in employee exit either around the announcement date for the plan switch (October of year t_{-3}) or the implementation date (January of year t_0). There is some incremental dependent attrition at the implementation date, but not enough to meaningfully impact our main results.

For those few who do exit in response to the change, one might expect them to be relatively sicker, which might induce a selection bias into our results. To examine this, we look at employees who exited the firm in t_{-1}, the year before the change. Summary statistics for this group of 1,153 employees are given in the third column of Table A1. This group of employees and their dependents does differ somewhat on demographic variables. Moreover, on average, this group spends approximately \$700 more in t_{-1} than individuals in our main sample. However, this difference seems to be driven by the upper tail of a small number of individuals, as the medians of the two spending distributions are nearly identical, and the 75th percentiles are different by a minor amount.

Given these similarities, we feel comfortable using our main sample restrictions throughout the paper.

A2 Intertemporal Substitution Analysis

In our analysis, we measure the extent to which employees increase spending in t_{-1} above expectations by substituting care that would otherwise have been obtained in the future if not for the policy change. To measure this 'excess mass', we first try to predict from prior years what spending would have been during t_{-1}, then measure the disparity. We run a regression as described in the main text in Section 3, for which the results are given in Table A2. We then calculate the excess mass as the difference between the true mean monthly individual spending amount and the predicted level. This measurement of excess mass is given in Table A3.

We note that, starting in December, excess mass is positive and high for December, November, and October (the three months with the largest excess mass among months in t_{-1}), before it drops down to nearly zero in September. There are some other outlier months across t_{-1} (March and August both have unusually high spending levels), however, as shown in Figure A2, the number of claims in those months is fairly reasonable relative to the trend. Careful investigation of those months (which cannot be shown due to individual privacy issues) uncovers that spikes in mean spending in those two months are primarily driven by a very small handful of unusually high-cost consumers. We take these combined trends as evidence that the majority of intertemporal substitution behavior is coming from care substituted into the last three months of t_{-1}.

One issue is that deviations from trend can occur both because of intertemporal substitution, as well as because of some nonzero draw of the unobservable idiosyncratic error term, $\bar{\epsilon}_t$. To account for our uncertainty over this term, we construct a confidence interval around our excess mass computation. We note that the mean squared error (MSE) of a regression is a consistent estimator of the variance of $\bar{\epsilon}$ in our model. Assuming that errors are not serially correlated, the standard deviation of the sum of the error terms for October, November, and December is $\sqrt{3 \cdot MSE}$, which in our case is approximately equal to 26.16. We multiply this term by 1.96 to get the 95% confidence interval for excess mass used in Table 4.

A3 Treatment Effect Standard Errors

We compute the standard errors for all estimates presented in Section 3 with a block bootstrap method. We take a sample (with replacement) of N individuals from our primary sample, where N is the number of individuals in the sample, including their spending levels for both t_{-1} and t_0. We then compute the percent change in spending between t_{-1} and t_0 for this new sample. Importantly, to compute standard errors for our anticipatory-spending-adjusted estimates, we generate an 'excess mass' estimate for this new sample (in the method described in Appendix A2), and then use it to adjust our estimates accordingly. We repeat this procedure 1000 times, retaining 1000 sampled estimates of the treatment effect. We take the square root of the variance of these 1000 estimates, and use that as our estimate of the standard error of our treatment effect estimates.

A4 Elasticity Estimates

A typical metric used to compare price sensitivity estimates in medical spending is the arc elasticity of total medical spending with respect to the price consumers face. As discussed in Aron-Dine et al. (2013), describing a non-linear insurance contract by one price is an oversimplification, since consumers face many potential true marginal prices throughout the contract and also face different marginal prices based on their respective health risks. The notion that it is difficult for one price to represent an insurance contract for a population is supported in our Section 5 analysis, which shows that consumers face very different prices throughout the year and that they respond to spot prices instead of true expected marginal prices.

Nevertheless, for comparison purposes, in Tables III and IV we present the semi-arc elasticity of total medical spending with respect to price:

$$\frac{(q_{t_0} - q_{t_{-1}})/(q_{t_0} + q_{t_{-1}})}{(p_{t_0} - p_{t_{-1}})/2}$$

Here, q_t is mean individual total medical spending in year t, and p_t is the single 'price' of insurance coverage for the population in year t. We follow the literature here, and take the single price of the HDHP in t_0 to be the proportion of medical spending that consumers in the overall population would have paid for if t_{-1} medical spending occurred under the HDHP plan design. This is .219 in the primary sample in our setting. The price of the PPO in t_{-1} is 0 since consumers do not pay anything for health care on the margin in the PPO. We note that while most of the literature uses arc elasticity rather than semi-arc elasticity, when the price change in question starts from zero price, arc elasticity just represents the % quantity change so is not a satisfactory descriptive statistic.[1] The semi-arc elasticity represents the change in quantity, normalized by the baseline quantity, divided by the change in price.[2]

[1] The arc elasticity in our context would be $\frac{(q_2-q_1)/(q_2+q_1)}{(p_2-p_1)/(p_2+p_1)}$. If p_1 is 0, then the denominator of this fraction always equals 1, and so the arc elasticity really only gives the arc change in quantity, regardless of the magnitude of the price change.

[2] In general, as with the arc-elasticity measure, one might want to normalize the price change as well to reflect differences in scale (e.g. comparing changes of \$5 to \$10 versus \$5000 to \$10000). In our setting, this is not an issue because we define price as the share of firm-wide costs that fall on the employee, following past work on moral hazard (see e.g. Manning et al. (1987)). Since this percentage is a relative measure already, this scaling issue does not arise when using the semi-arc elasticity measure.

As Tables III and IV reveal, the semi-arc elasticity for our primary causal treatment effect estimate lies in the range [-0.59, -0.69], averaging over both post-period years, while those from the other approaches in these tables lie between -0.57 and -1.32. We compare these estimates to two of the main estimates cited in the RAND Health Insurance Experiment, which compare two pairs of consumer groups: (i) those with 100% or 84% actuarial value plans or (ii) those with 84% or 69% actuarial value plans.[3] We use statistics from Keeler and Rolph (1988) to compute semi-arc elasticities of -2.11 and -2.26 respectively for these two estimates. Our semi-arc elasticity estimates range between one-quarter and one-half of those for RAND. Though, by this metric, consumers are less price sensitive in our setting, we note that the economic magnitudes of our treatment effect estimates are still substantial (regardless of the elasticity measures / comparison) and that there are many potentially important differences between our setting and the RAND setting.

A5 Early Switcher Difference-In-Differences

Our primary sample includes individuals who were in the PPO prior to the required switch, and thus those that were actively required to join the HDHP in t_0. As discussed in Section 2, approximately 85% of consumers at the firm fall into this category and were required to switch into the HDHP. In this section, we use consumers who voluntarily switched to the HDHP earlier, in either t_{-2} or t_{-1}, as a control group for the treatment effect analysis just described. By incorporating an additional control group, we estimate a differences-in-differences specification where we compare the change in spending from t_{-1} to t_0 in our primary sample, where consumers were required to switch plans, to the control group where consumers were enrolled in the HDHP in both years. We focus on the t_{-1} - t_0 two-year period for this analysis to remove confounds that could manifest over longer time horizons: as shown in the earlier analysis, t_{-2} statistics are similar to t_{-1}, and t_0 similar to t_1.

Figure A3 plots the mean individual monthly spending from t_{-4} - t_1 for (i) our primary sample (ii) individuals who switched to the HDHP at the beginning of t_{-2} (6,255 individuals) and (iii) individuals who switched to the HDHP at the beginning of t_{-1} (5,528 individuals). We note that the early switcher samples are balanced, in the sense that employees are present from t_{-4} - t_1, and that prior to joining the HDHP these employees and their dependents were enrolled in the PPO.

The figure clearly illustrates that early switchers are, on average, healthier than those in our primary sample who are required to switch for t_0. In addition, the figure shows a relative drop for mean spending for t_{-2} switchers in t_{-2}, for t_{-1} switchers in t_{-1}, and for t_0 required switchers in t_0. Figure A4 plots median spending over time for these different cohorts, and shows a similar pattern with much less noise since the median is a more robust statistic.

The fact that early switchers are healthier suggests that, in order to use them as a meaningful comparison group for the primary sample, we need to form a modified primary sample that matches the population of early switchers based on health status. For this analysis, we pool the two groups of early switchers (t_{-2} and t_{-1}) since we will be analyzing the spending change from t_{-1}-t_0. To measure health status in a predictive sense, we leverage the Johns Hopkins ACG software, which assigns each individual a predictive score, based on their past year of detailed claims data, for the upcoming health year. This score reflects the type of diagnoses that an individual had in the past year, along with their age and gender, rather than relying on past expenditures alone.[4]

[3] The 84% actuarial value contract has a 25% coinsurance rate up to an out-of-pocket maximum of $1000 while the 69% actuarial value plan has a 95% coinsurance up to a $1000 out-of-pocket maximum.

[4] See e.g. Handel (2013), Handel and Kolstad (2015) or Carlin and Town (2009) for a more in depth explanation of predictive ACG measures and their use in economics research. See http://acg.jhsph.org/index.php/the-acg-system-advantage/predictive-models for further technical details on these predictive algorithms.

We quantify the health status of early switchers with the observed distribution of individual-level ACG health status predictions for the year t_{-1}. We characterize this distribution with ventiles (20 equal sized buckets) of this predictive score, and weight the primary sample observations to match this distribution. Each ventile has, by definition, 5% of the early switcher sample. Thus, if 8% of the primary sample is contained in one of the early switcher ventiles, those individuals are weighted by $\frac{.05}{.08} = \frac{5}{8}$ in the weighted primary sample. We construct weights in this manner across the health status distribution to match the primary sample to the early switcher sample based on health status.

Figure A5 plots mean monthly individual-level spending for the pooled sample of early switchers and for our health-status weighted primary sample through t_0. The figure clearly illustrates that, prior to the switch in t_{-1}, when the two samples are in different plans, the HDHP consumers spend approximately 25% less than PPO consumers. In t_0, when both groups are in the HDHP, they spend almost identically (which also indicates successful matching on health status). Column 4 in Table III presents the quantitative difference-in-differences t_{-1}-t_0 spending reduction due to the HDHP switch implied by this figure:

$$[\bar{y}^{WPS}_{AS,t_0} - \bar{y}^{WPS}_{AS,t_{-1}}] - [\bar{y}^{ES}_{CPI,t_0} - \bar{y}^{ES}_{CPI,t_{-1}}]$$

Here, $\bar{y}^{S}_{M,T}$ refers to mean individual spending in year T under model M for sample S. Model AS refers to the model with both anticipatory spending and age/CPI adjustments. Model CPI refers to the model adjusting for age/CPI adjustments.[5] Sample WPS refers to the weighted primary sample, while sample ES refers to the early switcher sample.

A6 Truven MarketScan Difference-in-Differences

Truven Data. In Section 3, we use data from Truven Analytic's MarketScan commercial claims database both as a control group, and to construct weights for an externally valid estimate. In this appendix section, we describe the data in more detail, and display an alternative version of the above exercises where we use Truven data with linked income as part of our matching procedure.

The Truven MarketScan database is a nationally representative individual-level commercial database that collects health insurance claims from a number of large insurers across the U.S., and includes data for both the insurance policyholder and their dependents. Much like our own firm's data, it includes claim-line-level data on the universe of medical visits and prescription drug usage for those individuals whose data it collects. It includes identifiers for the insurance carrier, the employer (where applicable), and the specific plan the individual is enrolled in, although no details are given about the cost-sharing characteristics of that plan. It also includes basic demographics, including as age and gender.

We restrict the Truven sample we use to individuals receiving private health insurance (i.e., not Medicaid or Medicare) during the years t_{-1} and t_0, as well as to only individuals who live in the state where the majority of our firm's workers reside. This leaves us with roughly 600,000 individuals in this sample in each year. For each individual, we compute their total incurred spending for each month by adding up the allowed expenditures claimed by them for both medical visits and prescription drugs. As in our primary sample, we deflate spending in the Truven sample according to the medical CPI given by the BLS to account for medical price inflation. This data is used in the analysis of Section 3.

We account for the vastly different demographics between the Truven data and the employees and their dependents in our dataset. Normally, one would account for this in a regression framework by using demographics as control variables. However, such a regression would produce the average treatment effect, averaged across both our primary sample and the Truven sample. Relative to this framework, our baseline treatment effects are measuring the average treatment effect on the treated (ATOT), since they average only

[5] We adjust for anticipatory spending in the weighted primary sample, which switches for t_0, and not for the early switcher sample, which remains in the HDHP over these two years. Even if there is some anticipatory spending for some HDHP consumers in December in a given year, it should be the same cross-sectionally (detrended) in t_{-1} and t_0.

for our firm's employees and dependents, so the estimates would be incomparable. Therefore, we instead follow, e.g., Bitler et al. (2006), and, rather than use control variables in a regression, we reweight the Truven control group with propensity scores so that it, demographically, resembles our main firm's data. We take a nonparametric approach to constructing propensity scores, by dividing our main sample into cells based on age (in 5-year bins, with a cell for those aged 65 and older) and gender. We compute the proportion of individuals in our main sample in each cell, and then weight each observation in the Truven data by its equivalent cell weight in the main sample data. We then compute the treatment effect as:

$$ATOT = [\bar{y}_{t_0}^{PS} - \bar{y}_{t_{-1}}^{PS}] - [\bar{y}_{t_0}^{WTS} - \bar{y}_{t_{-1}}^{WTS}]$$

where

$$\bar{y}_t^{PS} = \frac{\sum_{i \in PS} y_{i,t}}{\sum_{i \in PS} 1}, \qquad \bar{y}_t^{WTS} = \frac{\sum_{i \in TS} w_{i,t} y_{i,t}}{\sum_{i \in TS} w_{i,t}}$$

and $w_{i,t}$ is the number of individuals in the primary sample in time t with the same age and gender cell as individual i. This procedure produces the estimate given in Column 6 of Table IV. Note that, for the primary sample, spending is adjusted for anticipatory spending (which is irrelevant for the Truven sample individuals), using the procedure given in Appendix A2.

We also seek to provide an externally-valid estimate, that acknowledges that our sample has a demographic composition quite unlike that of the rest of the United States. To do so, we reverse the procedure above, instead reweighting our primary sample according to the demographics of the Truven sample. We use the same age and gender cells, and so our estimate represents, in this context, the average treatment effect on the untreated, as such:

$$ATOU = [\bar{y}_{t_0}^{WPS} - \bar{y}_{t_{-1}}^{WPS}] - [\bar{y}_{t_0}^{TS} - \bar{y}_{t_{-1}}^{TS}]$$

where

$$\bar{y}_t^{WPS} = \frac{\sum_{i \in PS} \tilde{w}_{i,t} y_{i,t}}{\sum_{i \in PS} \tilde{w}_{i,t}}, \qquad \bar{y}_t^{TS} = \frac{\sum_{i \in TS} y_{i,t}}{\sum_{i \in TS} 1}$$

and $\tilde{w}_{i,t}$ is the number of individuals in the Truven sample in time t with the same age and gender cell as individual i.

Because Truven samples in a way that is nationally representative, we consider this treatment effect estimate to roughly approximate that of the U.S. under-65 privately insured population. This result is given in Column 5 of Table IV.

Income-Linked Analysis. As noted, the firm we analyze employs workers at wages that well exceed national averages. Therefore, finding a matching control group based on income is quite challenging. For about half of the policyholders in the Truven data, we are able to link their claims to a secondary dataset from Experian that includes the policyholder's annual income. For this analysis (seen below), we further restrict the sample to only those making at least $75,000 per year. This is because the lowest income category in the dataset for the firm we study is "less than $75,000," and so we are unable to accurately match the incomes of anyone making below that amount in our firm.[6] These restrictions leave us with around 30,000 individuals in each year for our secondary Truven sample.

Figure A6 replicates the graph from the right side of Figure II with this secondary sample. Much like in that graph, the income-linked Truven sample has spending levels below that of our sample, although

[6]Moreover, only 6% of employees at our firm are in this bracket, while the vast majority of employees within the Truven data are, so a very small portion of our data would have an excessively high weight.

the gap is far smaller. This is likely due to the fact that we restrict this secondary sample to only high-income individuals, who have higher purchasing power and are likely also employed by firms with similarly-generous health benefits. However, the PPO plan offered by the firm we study was extraordinarily generous even among high-income employers, explaining why the gap still remains. To quantify the the relative spending reduction, we follow the same procedure given above to compute an estimate of the ATOT for our primary sample firm. However, in this analysis, we further subdivide the demographic cells by income, using eight brackets of $25,000, beginning with the bracket $75,000-$100,000 and ending with a bracket including all of those who make over $250,000. Column A2 in Table A4 gives our ATOT estimate, which we bound between -18% and -23.7%, which is slightly higher than our primary estimates but lower than the estimates from our other Truven-derived control group.

We also replicate the procedure generating an externally valid estimate on this new control group, given in Column A1 of Table A4. We get an estimate bounded between -2.1% and -6.7%, unusually lower than our other primary estimates. This likely comes from the fact that the most common income bracket in the Truven sample is the lowest one, whereas the same bracket comprises of only a miniscule amount of our primary sample, so a small number of individuals are given high weights. In our primary sample, as given in Table A5, the lowest income bracket is also the least responsive, thereby depressing our estimate.

A7 Additional Analysis of Treatment Effect Heterogeneity

In this section, we present a number of figures and graphs that provide more detail on heterogeneity in spending trends across a variety of categories. First, we expand on Figure I in the text, by presenting Figures A7 and A8. In these figures, we break down the highest quartile of ACG score into four subgroups, and show that we can observe spending responses to the policy change broadly even across the top end of the sickness distribution. Figure A8 in particular shows that even in the 99th percentile of expected health risk, the median individual-level spending is reduced in the years following the change, despite the fact that individuals in this risk bracket should have no incentive to do so. Figure A9 first defines medical claims into categories based on the service location where medical care was received, and then plots spending in each of these categories over the entire timespan of our data. We see sharp reductions in office and emergency room visits, outpatient hospital care, and preventive care, with no real change in mental health spending or inpatient hospital care. Figure A10 breaks down spending cutbacks for prescription drugs, showing that cuts come from both branded and generic drugs.

The treatment effects estimated for these spending breakdowns, as well as others, is given in Table A5. We use the methods developed in Section 3 to estimate The table presents estimates comparing t_{-1} spending to t_0 spending for parsimony: t_{-1} to t_1 comparisons are similar and included in Table A8. We present three sets of treatment estimates: raw single-difference estimates (in Column 1), the same estimates adjusting for aging and health care CPI growth (Column 2), and our preferred estimates, which adjust for those factors as well as anticipatory spending (Column 3). The sickest quartile of individuals, who spend on average $12,335 in t_{-1}, reduce spending by between 18-22% under our preferred treatment effect measures. These treatment effects are slightly larger for the ex ante health status quartiles 1 (healthiest), 2, and 3 respectively, though off much lower pre-treatment spending bases.[7] [8] The table also presents these results for consumers categorized

[7]The health status quartile treatment effect analysis fixes the quartiles based on predictive indices for t_{-1}, but allows consumers to switch between those quartiles from one year to the next. This means that the cross-sectional health status quartile populations change over time, but the definition of a quartile in terms of health status remains the same. This is why the % of consumers in each quartile is slightly different than 25%.

[8]We note that the average of these health status quartile treatment effects, weighted by total spending, is slightly larger than the treatment effect presented for the entire population in Table III. In the raw spending and age/CPI-adjusted only treatment effects, this difference is because the quartiles have slightly different mixtures of health status *within the health status range for the quartile* over the years. For the anticipatory spending adjusted estimates, this difference could also come from the fact that anticipatory spending regressions /adjustments are done separately for each quartile. In Table A9, also in this section, we present some additional versions of this analysis, intended for robustness, where health status quartiles are defined as true quartiles on a

by number of documented chronic conditions entering a given calendar year, revealing surprisingly limited heterogeneity on this dimension.

Table A5 also documents heterogeneous treatment effects by (i) consumer demographics and (ii) type of location the medical service was performed at. One notable result is that spending reductions for dependents are limited (12%) and there are no anticipatory spending shifts for this group, suggesting that parents may be less willing to economize on care or shift care for their children. Spending reductions do not seem to vary much by age and, most surprisingly, income. In particular, the lowest spending bracket has the lowest respond, despite potentially having the large incentives to cut back conditional on health risk.

From our service location analysis, one notable result is that spending is reduced across all eight of these broad spending categories, and that the effects have a fairly narrow range of a 6% CPI adjusted reduction (mental health) to a 25% reduction (ER spending). This is somewhat surprising, since some categories seem more elective (e.g. physician office visits, 18% reduction) and others seem less elective (e.g. inpatient, 13% reduction). Notably, consumers reduce spending for both branded drugs (20%) and generic drugs (19%). In addition, spending on services that are classified as preventive is reduced by 10%. This is especially striking since (i) these services are all free to consumers under the HDHP (as mandated under the ACA) and (ii) these are services that may prevent higher spending and poor health in the future.

Table A6 displays our 'excess mass' calculations used to compute Column 3, constructed as described in Appendix A2. The first column shows the final excess mass calculation used in Table A5, while the second column gives the standard error for that calculation. The last three columns break down the excess mass for each month used in the data. We can see that most of the excess mass is driven by above-trend spending in December t_{-1}, as nearly every category of spending results in a positive excess mass calculation for that month.

Table A7 presents standard errors for the treatment effect estimates in Table A5. We construct standard error estimates by block bootstrap the same way we do for our primary estimates, as described in Section 3. In our block bootstrap procedure for demographic subsamples (for example, the effect for spouses), we first take the relevant demographic subsample, and *then* perform our bootstrap procedure, so that the number of individuals in that subsample is held fixed.

Finally, Table A9 presents an alternate version of our ACG quartile analysis from Table A5. In the initial analysis, we allow ACG scores for a given individual to vary over time in order to measure the treatment effect. In this table, we instead fix an individual's ACG score at one point (using their score constructed using either t_{-2} or t_{-1} claims data), and calculate their treatment effect over time. This method can suffer from mean reversion, where consumers with high scores previously due to chance may look as though they decrease spending later, which is why we do not use it for our main analysis. Presented here, we can see some evidence of this mean reversion, although it is not very strong relative to our treatment effects.

A8 Additional Analysis of Price Shopping

We do a number of robustness checks on our analysis of consumer price shopping. The first is that we verify that the rankings of prices across providers within a class of procedures is constant over time. To do so, for each procedure-year pair, we assign each provider in our restricted provider-procedure-year set a ranking according to their price for that procedure-year. We then calculate Spearman's rank correlation coefficient for each consecutive pair of years. The result from this exercise is given in Table A10. For nearly all pairs, the coefficient is very strong, over 0.92 for all year pairs. We view this as evidence supporting our modeling assumption that the rankings are approximately constant.

We additionally perform a version of our price shopping analysis on new employees. The key reason for doing so is because a lack of price shopping in the short run that we observe in our data may be driven by pre-existing relationships between consumers and providers. These relationships may make it difficult to

year to year basis, though the ACG index boundaries of each quartile may change .

switch to a new provider, even if the previous provider is more expensive. We do this by taking the claims of new employees in t_{-1} and t_0. We use claims from these employees only for the year in which they were a new employee, and we compare these two cross-sections in the same way we compared pairs of years in our main analysis. The results are given in Table A11. Again, we see no evidence for price shopping, instead finding slight increases in prices achieved. The primary driver of differences in spending for new employees, as in our main sample, is quantity reductions.

In Table A12, we show which of our decomposition pieces have positive value for the 30 top procedures by dollars spent. In Table A13, we disaggregate this, displaying the values for each of the 30 procedures. Due to space concerns, we present the decomposition only between t_{-1} and t_0. It is clear to see that very few procedures seem to exhibit meaningful consumer price shopping.

In Tables VII and VIII, we presented decompositions for high and low value care, for the year pairs t_{-1}-t_0 and t_{-3}-t_{-2}. Tables A15, A16 and A17 present these same decompositions for all year pairs. Table A19 replicates the analysis performed in Table A5 for these precise definitions of high and low value care.

A9 Additional Analysis of Reduction in Preventive Care

In Section 4 we investigated the nature of spending reductions for several different kinds of health care services. Preventive care services are of specific interest, because they are generally considered to be low cost, high value services that policymakers would like to encourage consumers to use: there is a range of past research, summarized nicely in Baicker et al. (2015) that shows ample evidence for underconsumption of preventive care by consumers. Under the Affordable Care Act, a baseline set of preventive services are required to be offered free of charge by insurers, with the intention of encouraging consumers to take up such services (see, e.g., Kaiser Family Foundation (2011)). Moreover, in this vein, there are many current supply-side policies (such as pay-for-performance bonus programs) that are implemented to incentivize medical providers to more effectively deliver preventive services.

Our analysis find that consumers reduce the quantity of general preventive services consumed by 7.5% in year t_0, the first year post-switch, and by a further 5.2% in the second year post-switch. We find that in the first year post-switch, the quantity of services that are considered to be preventive with a prior diagnosis is reduced by 12.2%. Given that these are considered to be cost-effective services, and that *these services are free* both before and after the switch to high-deductible health care, it is interesting to better understand exactly how consumers are reducing consumption of these services.

There are several potential reasons for why consumers reduce preventive services in our study:

1. Consumers (and potentially their providers) may not understand that preventive services are free, and instead think that their cost has gone up along with all other services. In this case, consumers would reduce preventive service consumption in response to a perceived increase in price.

2. Consumers may first choose whether to visit their provider to consume a bundle of services, some of which are free and some not, and then decide whether to consume preventive care. In this case, the extra cost of the bundled services may discourage consumers from going to their providers, and lead to lower consumption of preventative care. Relatedly, some consumers may not think about preventive care, and only consume such services during office visits when encouraged to by providers. In that scenario, a reduction in office visits would also lead to a reduction in preventive care consumption.

We investigate these potential explanations by studying whether reductions in preventive care occur on the extensive margin (fewer visits to primary care providers) or the intensive margin (fewer preventive services consumed conditional on a provider visit). If consumers consume the same amount of preventive care conditional on making an office visit, this suggests that they are not reacting heavily to a perceived price increase in preventive care, and instead going to their providers less because of the costs of other bundled services. If consumers reduce preventive care on the intensive margin, conditional on visiting their provider, this suggests that they are responding to a perceived price increase. We feel that this decomposition provides useful evidence for distinguishing between these hypotheses, even though there are some subtleties in mapping the hypotheses above to this extensive-intensive margin decomposition,.

To conduct this analysis, we used four different methods for classifying a primary care provider office visit. In this appendix, we focus on what we view as the best two methods, and briefly describe the other two at the end of this section. Crucially, all methods yield similar results.

The first method looks in each month and determines whether an individual had a CPT code or ICD-9 code that specifically signified a primary care office visit.[9] For this method, we split the year up into months and for each individual develop an indicator of whether they had a primary care visit in a given month.

The second method defines primary care providers by specialty listed in the data (this field is populated for all providers). We consider providers listed under either (i) family medicine (ii) preventive medicine, general or (iii) internal medicine to be primary care providers. Then, any time a patient sees such a provider in a month, we classify the patient as having made a visit to a primary care provider in that month.

For each of these two methods, we study intensive margin preventive care use in two distinct ways. Our primary methods looks at dollars spent on preventive care per office visit. Our second method looks at a binary indicator variable of whether any preventive services were consumed during the office visit.

For each approach, we investigate both for our primary sample and for the price shopping decomposition sample (restricted to the main company location). Since the results are similar, here we only present the results for the price shopping sample, to be consistent with the tables in Section 4. Additionally, we run the analysis separately for services that are generally considered preventive, and also for services that are only considered preventive with a prior diagnosis.

Table A20 presents the results for our approach that measures the intensive margin of preventive care based on $ per provider visit spent on that care. When provider visits are measured according to ICD-9 and CPT codes, the number of provider visits (extensive margin) decreases by 12.1% from t_{-1} to t_0 (143,887 to 126,406). For general preventive care, spending per visit increases from \$62.57 to \$64.79 (3.5%) per visit. For preventive care with a prior diagnosis, spending per visit decreases from \$117.18 to \$114.59 (2.1%). Thus, for this approach, the reduction in preventive spending can be attributed almost entirely to the extensive margin, i.e. fewer office visits, rather than doing fewer preventive services per office visits.

When office visits are measured according to provider specialty, the overall results are quite similar. The number of office visits declines from 63,121 to 54,218 (14.1%) from t_{-1} to t_0. On the intensive margin, spending per visit for general preventive care increases by 3.3% while for preventive care with a prior diagnosis it decreases by 1.8%. A similar emphasis on the extensive margin holds for different classification methods we use for provider visits, suggesting that this result is robust to different definitions of the extensive margin.

Table A21 presents this decomposition when we measure preventive care use with a binary indicator of whether any preventive care was done in a given month (rather than $ per visit). The results are similar: almost all reductions come at the extensive margin. The extensive margin statistics are the same for this case as with the prior case, the only difference is with the measurement of the intensive margin. The change in the intensive margin spending for the ICD-9 / CPT office visit classification is 3.0% for general preventive care and -0.7% for preventive care with a prior diagnosis. Similar results hold for the second classification (provider specialty) showing that the extensive margin impact is similar regardless of the method chosen for measuring preventive care use. Figure A11 shows the entire density of intensive margin spending for general preventive care for the two definitions of an office visit. The density of $ per visit spent is very close together in each case, suggesting that almost all of the action is on the extensive margin in terms of reduced preventive quantities and spending.

[9] There are many such codes, and the list we used is available upon request.

Table A21 also investigates this decomposition for specific preventive services, including mammographies, colonoscopies, and urinalysis. These cases highlight the difference in extensive vs. intensive margin effects across services. For the physician specialty classification of office visits, mammographies per visit declines by 6.7%, urinalysis per visit declines by 0.4%, and colonoscopies per visit declines by 35.0%. Thus, the urinalysis effect is almost entirely related to the extensive margin, there is some intensive margin mammography effect, though not a large one, and the intensive margin effect for colonoscopies is substantial, indicating that that is one margin consumes / providers are clearly responding to in the treatment year. We caution that results for specific services like colonoscopies should be viewed in light of any changing guidelines that occur over time for how physicians should prescribe them, though there is no national downward trend in those services.

Taken together, these results suggest that consumers are reducing preventive care consumed primarily at the extensive margin: when they actually visit a provider they're doing almost the exact same level and type of preventive care as previously, but, they are doing less preventive care in general because they are reducing their doctor visits in general. This suggests that consumers are reducing preventive services consumed in large part because those services are bundled with other, costly, services for provider visits, and that patients are reducing the number of times they consume such bundles. The results work against the hypothesis that these spending reductions are coming primarily because consumers perceive higher prices for preventive services as well, and are purposefully reducing the consumption of such services when they go to the provider. Note that our results do not perfectly test these hypotheses, since consumers could be reducing visits because they perceive the prices of preventive care to be higher. However, the use of preventive care conditional on an office visit suggests that there is no substitution at all once a visit is scheduled, making it unlikely that consumers are reducing visits primarily because of perceived preventive care prices.

In analysis not reported here we perform a range of robustness checks. First, we investigate two additional methods for classifying office visits (i) medical events where service location is denoted Office and (ii) identifying providers as primary care providers based on claims that patients who visit them have (then applying this designation of primary care provider to all claims affiliated with that provider. We also test for preventive spending that doesn't fall under our different office visit definitions, and note that most spending classified as preventive falls under our office visit definitions. We also repeat the analysis for the primary sample in the main text (not restricted to the main company region). All of these contingencies point in the same direction: consumers are reducing office visits overall, but not reducing the amount of preventive care done per office visit.

A10 Additional Analysis of Responses to Non-Linear Contract

We present versions of our descriptive analysis of employee responses to the non-linear structure of the HDHP, where we instead use single employees, or employees with only a single dependent, in Figures A13 and A14. These figures replicate the analysis shown in Figures III, IV, and A12 in the text for those populations. Incremental spending for the next month and for the rest of the following year is given for employee-month combinations in a given tier of the HDHP in t_0. These figures provide results that are qualitatively similar in nature to those for employees with two or more dependents.

A11 LASSO Results

To demonstrate further that variation in end of year price does not explain spending differences, we turn to a method originally employed by Backus et al. (2015). We restructure our prior regression model (with all three prices) as a penalized linear model, specifically a LASSO model,[10] and estimate the model for different values for the coefficient constraint. As the LASSO coefficient size constraint binds more tightly, the solution algorithm will be forced to set some coefficients to zero. We use a stepwise regression model to focus on the set of constraint values that make the algorithm remove a variable from the model. It will begin with those variables that least explain variation in health spending. We think of this as a data-driven way to characterize the 'importance' of each of the price variables in explaining health spending choices. Furthermore, by estimating a penalized regression we can flexibly capture correlations between dependent variables, an advantage in our setting as different price measures are all based on a mapping from measures of health and spending over time.

Figure A18 presents the results of this exercise for the key price coefficient of interest: spot price, expected, end-of-the-year marginal price and last years end-of-the-year marginal price. These results are based on t_0 and t_1 respectively. The coefficients at the far right represent the unconstrained OLS regression; the far left represents the completely constrained LASSO model (where all coefficients are set to zero), with points in between representing constraint levels between these two extremes.

As the constraint binds (moving from right to left), the coefficients on the expected end-of-year marginal price variables are the first set to zero, implying that they are relatively unimportant for explaining the variation. In t_0 and t_1 we see the most important factor, both in terms of effect size and the fact that it remains different from zero as the penalty function gets vary large (steps go to 0), is spot price of 1. In t_0 we see some impact of the 4th quartile of the E[EOY Marginal Price] though the magnitude is far smaller. A similar result occurs for last years marginal price of .1 in the t_0 plot. For t_1 the results are quite similar for spot price of 1: it is the most significant in terms of longevity as well as in magnitude. Together these results lend further evidence, using an alternate empirical approach that flexibly allows the price response to fit the data, that primary driver of the behavioral response is for those under the deductible.

[10]LASSO is equivalent to OLS (a linear model minimizing squared residuals) with an additional constraint on the sum of the absolute values of the coefficients.

A12 Appendix Tables and Figures

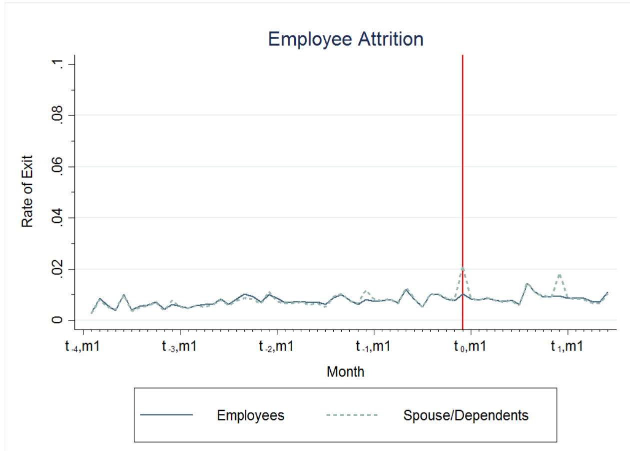

Figure A1: This figure plots the monthly hazard rate of exit from the firm's insurance coverage over time for employees and their dependents.

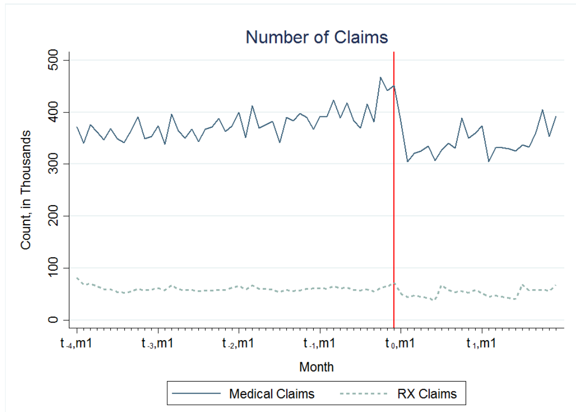

Figure A2: This figure plots total number of monthly claims, both for medical claims and prescription drug claims over time, for our primary sample.

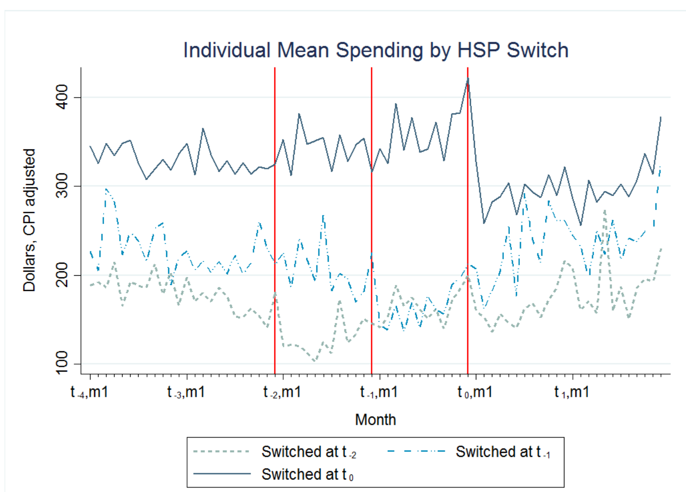

Figure A3: This figure plots mean monthly spending over time for consumers who (i) are in our primary sample (and thus were required to switch to the high-deductible plan in t_0) (ii) those who elected to switch early to the HDHP in t_{-1} and (iii) those who elected to switch early to the HDHP in t_{-2} (and stayed in that plan over time).

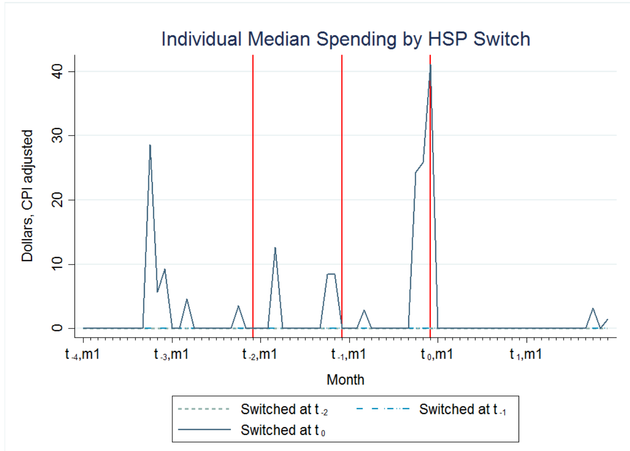

Figure A4: This figure plots median monthly spending over time for consumers who (i) are in our primary sample (and thus were required to switch to the high-deductible plan in t_0) (ii) those who elected to switch early to the HDHP in t_{-1} and (iii) those who elected to switch early to the HDHP in t_{-2} (and stayed in that plan over time).

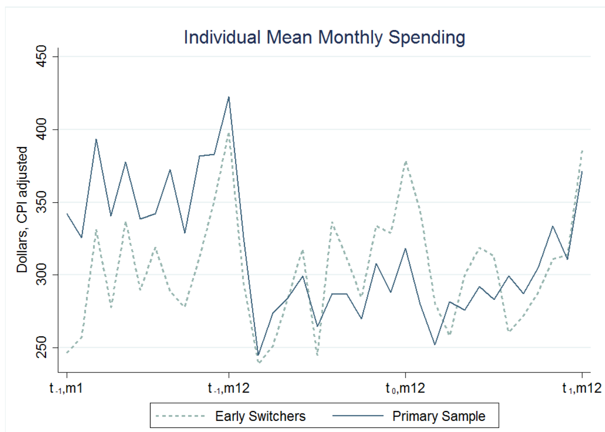

Figure A5: This figure plots mean monthly individual spending over time for consumers who (i) are in our pooled sample of early switchers and (ii) are in our weighted primarily sample through t_0, matched to the early switcher sample based on the health status distribution.

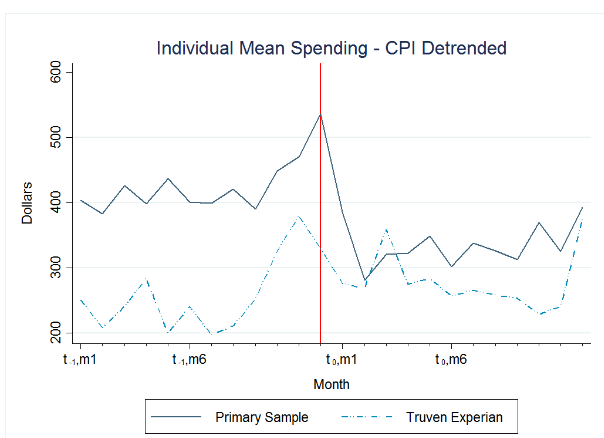

Figure A6: This graph presents spending for our primary sample alongside spending for the weighted control group formed from Truven MarketScan data with linked income fields.

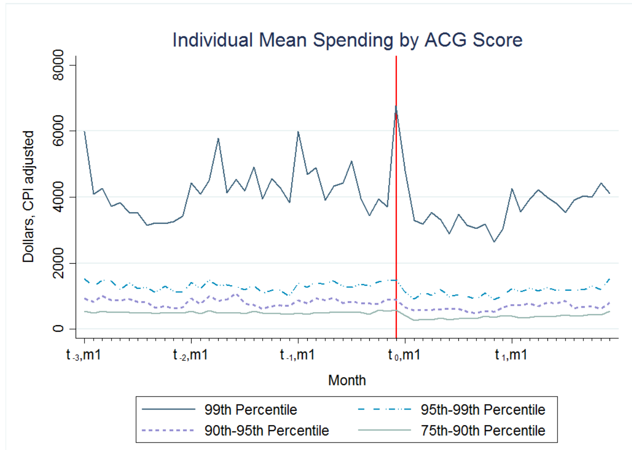

Figure A7: This figure plots adjusted mean spending for individuals in a given month, by ACG predictive health index bin (the index is calculated at the beginning of each calendar year). This graph divides individuals in the top quartile of the ACG index into smaller subgroups.

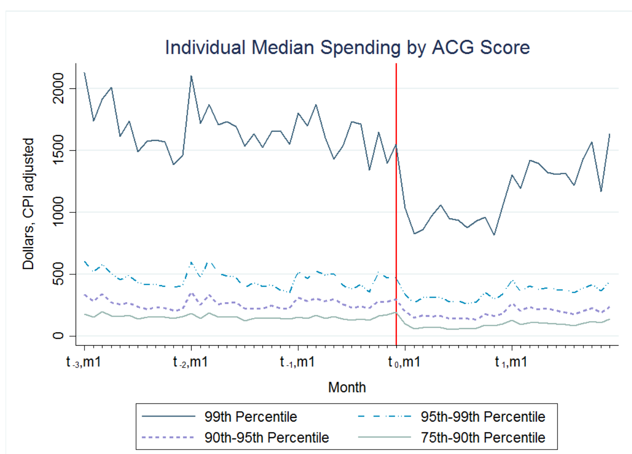

Figure A8: This figure plots adjusted median spending for individuals in a given month, by ACG predictive health index bin (the index is calculated at the beginning of each calendar year). This graph divides individuals in the top quartile of the ACG index into smaller subgroups.

Figure A9: This figure plots mean medical spending for individuals in a given month, by the type of care, both adjusted and unadjusted for age and price trends. These categories are mutually exclusive, except for Preventive.

Figure A10: This figure plots mean prescription drug spending for individuals in a given month, for brand and generic drugs, both adjusted and unadjusted for age and price trends.

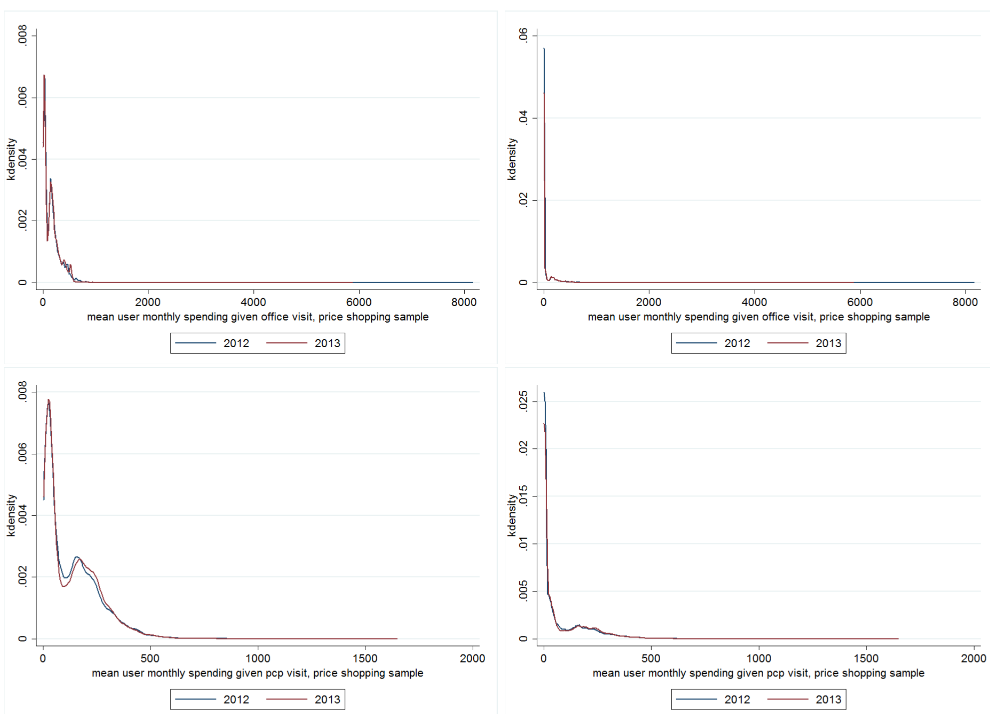

Figure A11: This figure presents per visit spending densities for general preventive care services, before and after the required switch to high deductible care. The top section of the figure presents these densities for our first definition of an office visit (defined by CPT codes) while the bottom presents these densities for our second definition, based on provider specialty. The left half of the figure presents densities conditional on spending great than 0 on preventive in a visit, the right half presents densities including zero spending on preventive care.

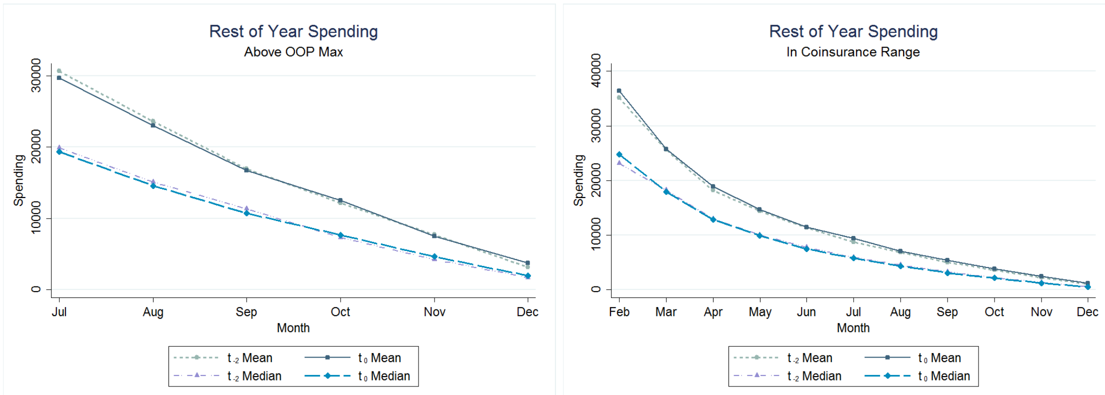

Figure A12: The left panel in this figure shows incremental spending for the rest of the year, for families who have passed the out-of-pocket maximum by the start of a given month in t_0, compared to t_{-2} incremental spending for equivalent quantiles of pre-period consumers. The right panel presents the analogous figure for families who start a given calendar year month in the coinsurance arm of the HDHP (and matched t_{-2} consumers).

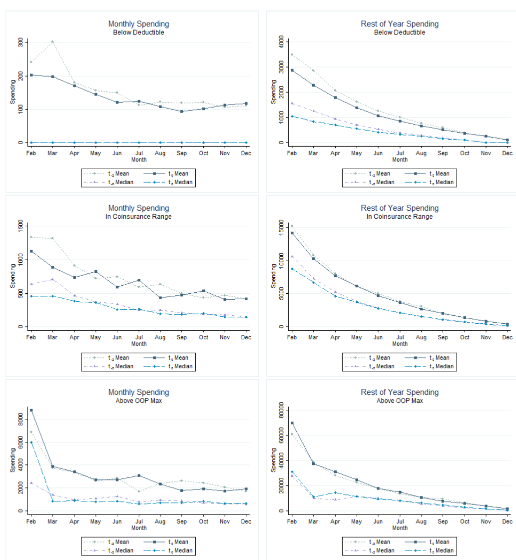

Figure A13: This figure shows incremental spending for employees who have passed the out-of-pocket maximum by the start of a given month in t_0, for single employees. The left side of the figure studies incremental spending for the next month, while the right side studies incremental spending for the rest of the year. This t_0 incremental spending is compared to t_{-2} incremental spending for the equivalent quantiles of consumers based on total yearly spending up to month m, M_m.

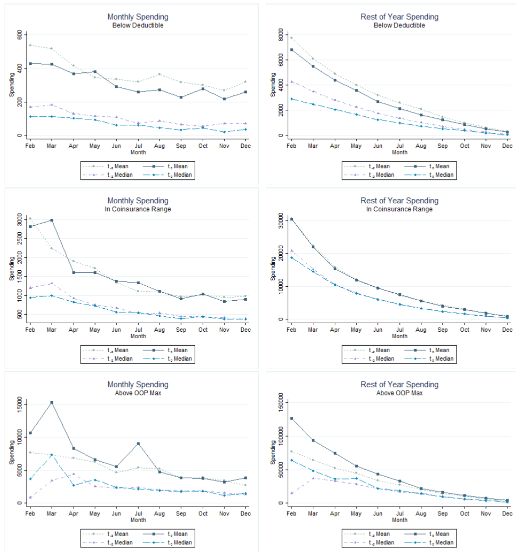

Figure A14: This figure shows incremental spending for employees who have passed the out-of-pocket maximum by the start of a given month in t_0, for employees with one dependent.

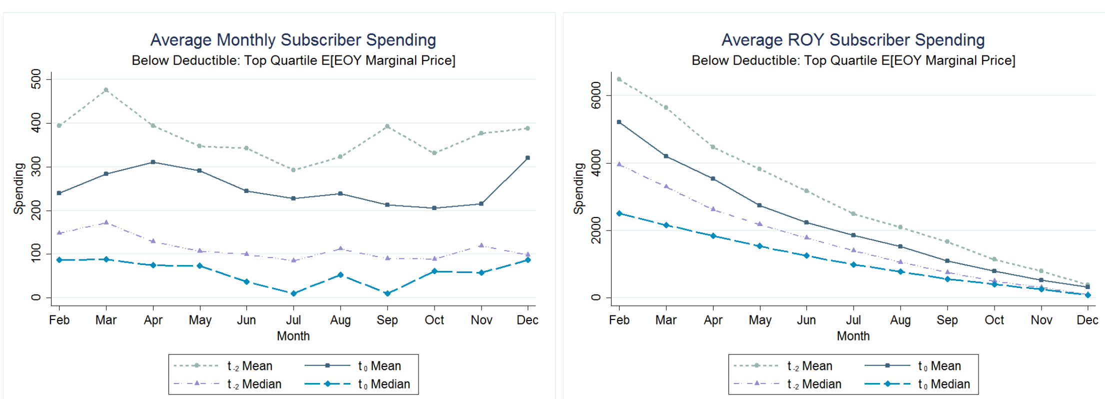

Figure A15: This figure shows incremental spending for employees who have passed the out-of-pocket maximum by the start of a given month in t_0, for families with the highest quartile of shadow price.

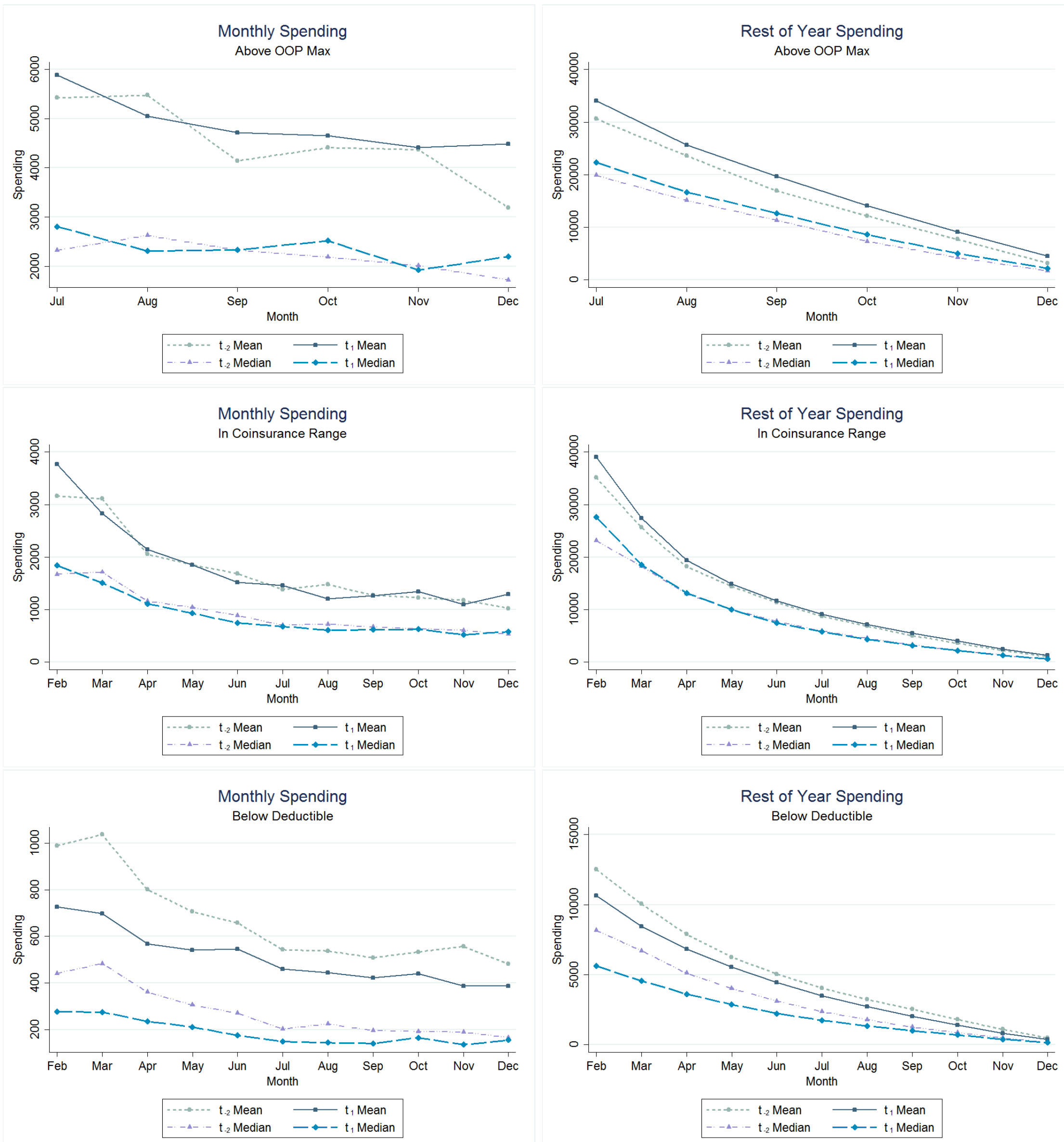

Figure A16: This figure presents descriptive results for t_1, comparing incremental spending in that year by plan arm to spending by equivalent quantiles of consumers in t_{-2}. These figures are directly analogous to those presented earlier in this section, describing how incremental spending in t_0 compares to that in t_{-2}. The left panels present incremental spending for the next month conditional on start of month plan arm, while the right panels present incremental spending for the rest of the year.

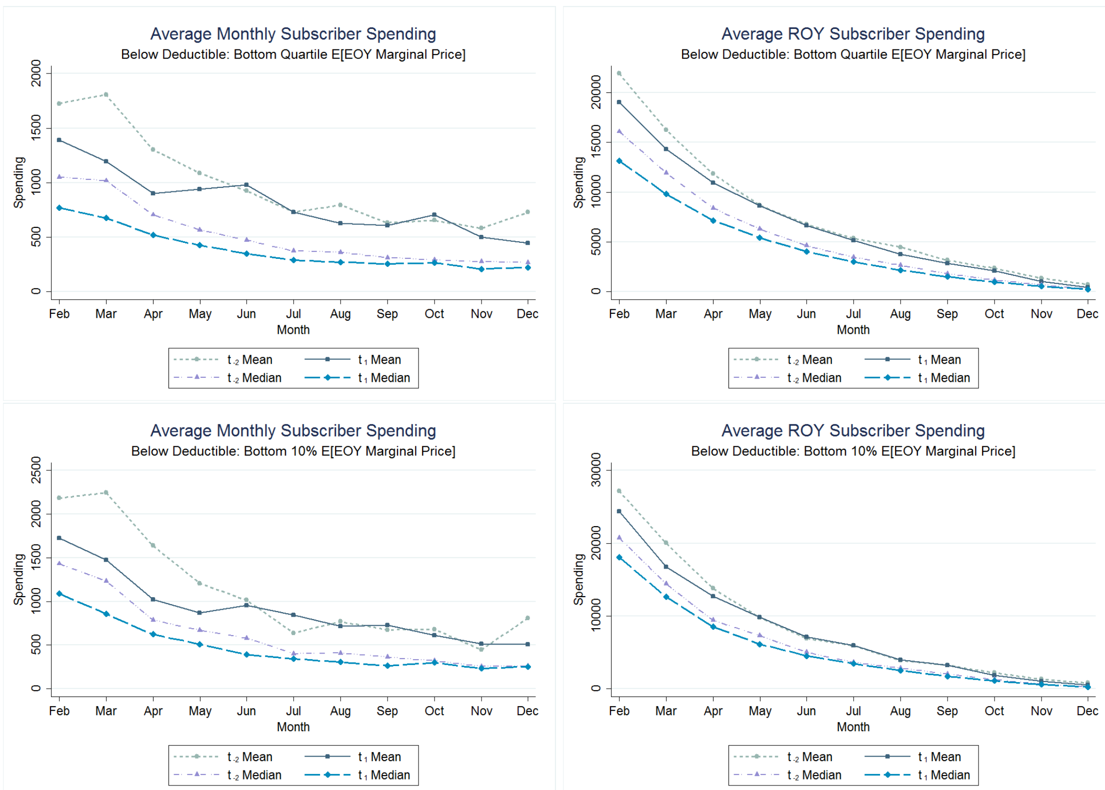

Figure A17: This figure presents descriptive results for t_1, and examines how predictably sick consumers under the deductible at the beginning of a month reduce incremental spending. These figures are directly analogous to those presented earlier in this section, describing how incremental spending in t_0 compares to that in t_{-2}. The left panels present incremental spending for the next month conditional on start of month plan arm, while the right panels present incremental spending for the rest of the year.

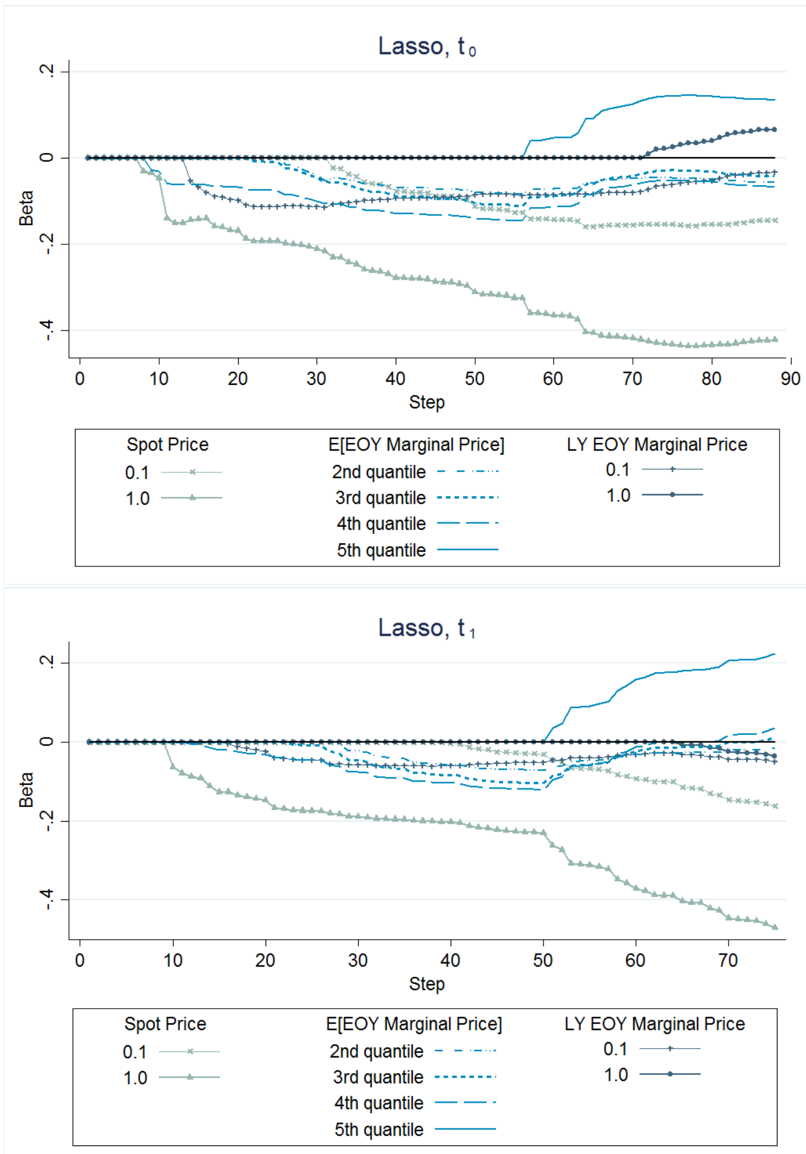

Figure A18: This figure presents our results from the LASSO procedure described in the text. Each step denotes the point where (moving from right to left) a variable is removed from the regression (i.e., its coefficient is set to zero).

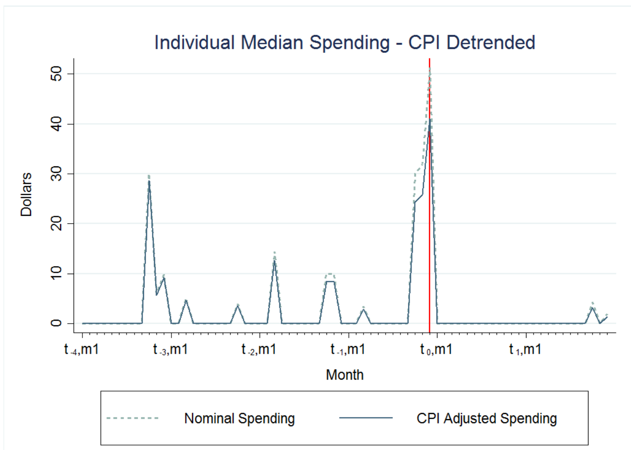

Figure A19: This figure plots median monthly spending for individuals in our primary sample from t_{-4}-t_1, both adjusted and unadjusted for age and price trends.

Sample Demographics	Primary Sample	Alternate Sample	Employees Exiting in t_{-1}
N - Employees	22,719	31,042	1,153
N - Emp. & Dep.	76,759	95,224	3,180
Enrollment in PPO in t_{-1}	100%	100%	100%
Gender - Emp. & Dep. **% Male**	51.4%	48.8%	41.4%
Age, t_{-1} - Employees			
18-29	4.3%	7.0%	5.9%
30-54	91.4%	88.2%	77.0%
$\geq$ 55	4.3%	4.8%	6.4%
Age, t_{-1} - Emp. & Dep.			
$<$ 18	36.1%	33.2%	24.8%
18-29	8.8%	9.6%	10.9%
30-54	52.0%	48.9%	42.0%
$\geq$ 55	2.8%	2.9%	3.9%
Income, t_{-1}			
Tier 1 ($<$ $100K)	7.3%	7.6.8%	9.7%
Tier 2 ($100K-$150K)	64.7%	65.0%	59.0%
Tier 3 ($150K-$200K)	22.6%	20.1%	15.9%
Tier 4 ($>$ $200K)	4.7%	4.2%	2.6%
Family Size, t_{-1}			
1	16.1%	18.4%	15.2%
2	17.9%	18.7%	32.4%
3+	65.9%	62.9%	52.4%
Individual Spending, t_{-1}			
Mean	$5,223	$5,375	$5,921
25th Percentile	$631	$645	$533
Median	$1,795	$1,817	$1,796
75th Percentile	$4,827	$4,890	$5,151
95th Percentile	$18,810	$19,141	$21,986
99th Percentile	$52,360	$53,239	$59,481

Table A1: This table presents summary demographic statistics for (i) our primary sample, which is restricted to employees present over the time horizon t_{-4}-t_1, and their dependents; and (ii) an alternate sample, which is only restricted to employees present over the time horizon t_{-2}-t_0. When relevant, statistics for the primary sample are presented for the year t_{-1}.

Regression Results	
Variable	Coefficient
Months Since Jan. in Year t_{-4}	0.442
February	-32.37
March	15.28
April	-11.07
May	-11.90
June	-5.87
July	-32.34
August	-20.96
September	-31.93
October	-19.79
November	-22.54
December	-27.71

Table A2: This table presents coefficients from the regression model used to measure excess mass.

Excess Mass	
Month	Excess Mass
December	85.83
November	41.57
October	37.83
September	-2.15
August	20.91
July	12.21
January to June (average)	0.34

Table A3: This table presents the computed excess mass for each month in the second half of t_{-1}.

HDHP Switch Differences-in-Differences Analysis	Model	
	(A1) Ext. Validity Truven-Income Weighted DID	(A2) Truven-Income-Control DID
% Decrease, t_{-1}-t_0	[-2.1%, -6.7%]	[-18.0%, -23.7%]
Semi-Arc Elasticity	[-0.10,-0.34]	[-0.90,-1.18]

Table A4: This table details the treatment effect of the required HDHP switch under two specifications described in the accompanying text: (i) external validity difference-in-differences using weights derived from income-linked Truven MarketScan data and (ii) Income-linked Truven control group difference-in-differences.

Heterogeneous HDHP Spending Impact				Treatment Effect		
				(1)	(2)	(3)
	Group %	Spending %	t_{-1} Mean Spending	Nominal Spending	CPI	Anticipatory Spending
Age 0-17	36.26	24.29	3465.65	-0.07	-0.11	-0.11*
Age 18-29	8.81	7.59	4442.77	-0.15	-0.19	-0.19*
Age 30-54	51.99	62.08	6164.59	-0.19	-0.23	[-0.13,-0.18]
Age 55+	2.92	5.95	11051.14	-0.11	-0.15	[-0.05,-0.11]
Income \$0-100K	6.30	6.91	5701.99	-0.03	-0.07	[-0.00, -0.04]
Income \$100-150K	63.04	62.98	5209.86	-0.13	-0.17	[-0.08, -0.13]
Income \$150-200K	24.93	24.20	5026.86	-0.15	-0.18	[-0.15, -0.17]
Income \$200K+	5.73	5.91	5340.94	-0.12	-0.15	[-0.09,-0.12]
Employee	33.47	35.77	5532.76	-0.20	-0.23	[-0.12,-0.18]
Spouse	23.92	35.12	7495.02	-0.16	-0.20	[-0.10,-0.16]
Dependent	42.61	29.11	3570.33	-0.08	-0.12	-0.12*
ACG Quartile 1**	28.51	9.74	1643.56	-0.25	-0.28	-0.28*
ACG Quartile 2**	23.83	12.15	2824.78	-0.39	-0.41	[-0.39,-0.40]
ACG Quartile 3**	23.53	21.45	4564.50	-0.36	-0.38	[-0.33,-0.36]
ACG Quartile 4**	24.13	56.66	12335.85	-0.21	-0.25	[-0.18,-0.22]
ACG Top 1%**	0.79	9.33	66606.47	-0.25	-0.28	-0.28*
0 Chronic Cond.	62.78	38.34	3202.64	-0.15	-0.19	[-0.16,-0.18]
1-2 Chronic Cond.	33.13	47.38	7240.37	-0.18	-0.22	[-0.18, -0.20]
3+ Chronic Cond.	4.19	14.18	19093.34	-0.13	-0.17	[-0.05,-0.12]
Inpatient Hosp.		16.53	863.48	-0.09	-0.13	[-0.07,-0.11]
Outpatient Hosp.		18.07	944.15	-0.13	-0.17	[-0.06,-0.12]
ER		3.11	162.40	-0.21	-0.25	-0.25*
Office Visit		7.61	397.86	-0.15	-0.18	[-0.13,-0.16]
RX		16.91	883.62	-0.16	-0.19	[-0.15,-0.17]
RX - Brand		12.23	638.82	-0.16	-0.20	[-0.16,-0.18]
RX - Generic		4.05	211.62	-0.15	-0.19	[-0.19,-0.19]
Mental Health		9.45	493.86	-0.02	-0.06	-0.06*
Preventive		9.50	496.28	-0.06	-0.10	[-0.05,-0.08]
Other		22.94	1198.07	-0.26	-0.29	[-0.17,-0.24]

*Anticipatory spending estimate itself is negative or not significant from 0

**Quartile definition constant, population shifts across quartiles each year. Mixture of health status within quartile bounds differs from year to year.

Table A5: This table summarizes our descriptive evidence for the heterogeneous treatment effects of the required HDHP switch. For parsimony, the tables presents the estimates from t_{-1}-t_0: see the Appendix for the estimates comparing t_{-1} to t_1. The table presents the results for different (i) demographics (ii) health status measures and (iii) types of health services. The first column reports the % of people within a given demographic group or health status group for categories (i) and (ii), and the % of total spending a given service spending is for category (iii). The second column reports average mean individual yearly spending for categories (i) and (ii), and average mean individual spending for each type of service for category (iii). The second through fourth columns present, for each respective framework, the % change in spending (for each demographic group, or type of service) as a result of the required HDHP switch from t_{-1} to t_0.

Excess Mass Calculation					
			Individual Month Calculations		
	Total Excess Mass	Standard Error	October	November	December
Age 0-17	-85.51	12.09	-26.65	-43.50	-15.37
Age 18-29	-33.24	38.13	-20.89	-2.70	-9.65
Age 30-54	253.49	8.65	42.24	61.23	150.01
Age 55+	525.20	78.48	110.05	68.57	346.58
Income 0-100K	201.84	29.77	99.47	28.29	74.08
Income 100-150K	190.07	15.36	43.67	52.99	93.41
Income 150-200K	71.60	21.73	0.20	19.47	51.93
Income 200K+	126.37	23.98	51.14	28.09	47.14
Employee	243.51	9.75	46.09	46.36	151.06
Spouse	308.67	19.70	53.90	89.33	165.44
Dependent	-91.79	13.15	-32.01	-41.88	-17.90
ACG Quartile 1	0.12	7.72	-3.15	2.18	1.09
ACG Quartile 2	42.49	11.94	-9.33	18.68	33.14
ACG Quartile 3	101.35	11.69	29.46	-13.83	85.72
ACG Quartile 4	446.90	26.67	77.45	107.11	262.34
ACG Top 1%	139.48	664.99	-945.06	-1068.03	2152.57
0 Chronic Conditions	56.33	9.10	9.13	14.57	32.63
1-2 Chronic Conditions	118.64	16.04	10.94	5.75	101.94
3+ Chronic Conditions	985.15	65.44	102.65	165.03	717.47
Inpatient Hosp.	25.89	8.79	9.80	1.81	14.27
Outpatient Hosp.	48.37	3.70	8.05	15.95	24.38
ER	-1.40	0.69	-1.64	-1.20	1.44
Office Visit	12.48	1.02	2.56	4.04	5.88
RX	18.87	1.47	0.94	5.54	12.39
RX - Brand	11.93	1.05	-0.39	3.50	8.83
RX - Generic	1.82	0.58	0.06	0.35	1.42
Mental Health	-5.58	1.96	2.30	-4.63	-3.25
Preventive	11.52	1.15	1.96	3.58	5.99
Other	61.34	2.44	14.58	18.56	28.20

Table A6: This table gives the excess mass calculations (with their associated standard error) for each category of individual spending, calculated as detailed in Appendix A2. These excess mass calculations are used in the construction of the final column of Table A5.

	(1) Nominal Spending	(2) CPI	(3) Anticipatory Spending
Age 0-17	4.0	3.8	[4.3, 0.8]
Age 18-29	3.3	3.1	[3.6, 1]
Age 30-54	1.2	1.2	[1.3, 0.3]
Age 55+	9.4	9.1	[10, 1.6]
Income 0-100K	17.4	16.8	[17.1, 1.3]
Income 100-150K	2.6	2.5	[2.9, 0.4]
Income 150-200K	1.9	1.9	[2.1, 0.4]
Income 200K+	4.0	3.8	[4.3, 0.8]
Employee	2.3	2.2	[2.4, 0.4]
Spouse	1.8	1.7	[2, 0.5]
Dependent	3.3	3.2	[3.6, 0.6]
ACG Quartile 1	3.8	3.7	[4.2, 0.5]
ACG Quartile 2	2.7	2.6	[2.8, 0.4]
ACG Quartile 3	2.8	2.7	[3, 0.3]
ACG Quartile 4	1.6	1.6	[1.7, 0.3]
ACG Top 1%	5.8	5.6	[6.5, 1.1]
0 Chronic Conditions	2.1	2.0	[2.3, 0.4]
1-2 Chronic Conditions	2.1	2.0	[2.3, 0.3]
3+ Chronic Conditions	3.9	3.7	[4, 0.7]
Inpatient	7.6	7.3	[7.8, 1.4]
Outpatient Hosp.	2.2	2.1	[2.5, 0.5]
ER	1.3	1.3	[1.4, 0.4]
Office Visit	0.4	0.4	[0.4, 0.1]
RX	1.1	1.1	[1.2, 0.2]
RX - Brand	1.5	1.5	[1.6, 0.3]
RX - Generic	0.7	0.6	[0.7, 0.2]
Mental Health	2.4	2.3	[2.5, 0.5]
Preventive	0.9	0.8	[0.9, 0.2]
Other	1.1	1.1	[1.2, 0.3]

Table A7: This table presents the standard errors for the treatment effects given in Table A5.

Heterogeneous HDHP Spending Impact						
				Treatment Effect		
				(1)	(2)	(3)
	Group %	Spending %	t_{-1} Mean Spending	Nominal Spending	CPI	Anticipatory Spending
Age 0-17	34.41	22.83	3465.65	-0.03	-0.11	-0.11*
Age 18-29	8.39	7.13	4442.77	-0.07	-0.15	-0.15*
Age 30-54	49.45	58.37	6164.59	-0.12	-0.19	[-0.09,-0.14]
Age 55+	2.65	5.60	11051.14	-0.07	-0.15	[-0.04,-0.09]
Income 0-100K	6.09	6.64	5701.99	-0.02	-0.10	[-0.01,-0.06]
Income 100-150K	61.34	61.19	5209.86	-0.09	-0.17	[-0.08,-0.12]
Income 150-200K	24.50	23.58	5026.86	-0.07	-0.14	[-0.11,-0.13]
Income 200K+	5.31	5.43	5340.94	-0.08	-0.16	[-0.10,-0.13]
Employee	31.66	33.54	5532.77	-0.07	-0.15	[-0.04,-0.09]
Spouse	22.85	32.79	7495.02	-0.12	-0.20	[-0.10,-0.15]
Dependent	40.38	27.61	3570.33	-0.02	-0.11	-0.11*
ACG Quartile 1	27.21	8.56	1643.56	-0.09	-0.17	-0.17*
ACG Quartile 2	22.63	12.24	2824.79	-0.29	-0.35	[-0.31,-0.33]
ACG Quartile 3	22.36	19.54	4564.51	-0.26	-0.32	[-0.27,-0.29]
ACG Quartile 4	22.69	53.59	12335.85	-0.02	-0.10	[-0.01,-0.06]
ACG Top 1%	0.69	8.80	66606.47	-0.05	-0.13	-0.13*
0 Chronic Conditions	59.76	36.65	3202.64	-0.07	-0.14	[-0.10,-0.12]
1-2 Chronic Conditions	31.34	43.46	7240.37	-0.04	-0.13	[-0.09,-0.11]
3+ Chronic Conditions	3.78	13.83	19093.35	0.02	-0.07	[0.06,0]
Inpatient		16.53	863.48	-0.13	-0.20	[-0.13,-0.16]
Outpatient Hosp.		18.08	944.16	-0.08	-0.15	[-0.03,-0.09]
ER		3.11	162.41	0.12	0.03	0.03*
Office Visit		7.62	397.86	-0.10	-0.18	[-0.10,-0.14]
RX		16.92	883.62	-0.01	-0.09	[-0.04,-0.07]
RX - Brand		12.23	638.83	-0.08	-0.16	[-0.11,-0.14]
RX - Generic		4.05	211.62	-0.17	-0.24	[-0.22,-0.23]
Mental Health		9.46	493.87	0.07	-0.02	-0.02*
Preventive		9.50	496.29	0.01	-0.07	[-0.02,-0.05]
Other		22.94	1198.08	-0.21	-0.27	[-0.15,-0.21]

Table A8: This table summarizes our descriptive evidence for the heterogeneous treatment effects of the required HDHP switch, for estimates giving the effect between t_{-1} and t_1 (compared to Table 5's description of . The table presents the results for different (i) demographics (ii) health status measures and (iii) types of health services. The first column reports the % of people within a given demographic group or health status group for categories (i) and (ii), and the % of total spending a given service spending is for category (iii). The second column reports average mean individual yearly spending for categories (i) and (ii), and average mean individual spending for each type of service for category (iii). The second through fourth columns present, for each respective framework, the % change in spending (for each demographic group, or type of service) as a result of the required HDHP switch from t_{-1} to t_0.

Heterogeneous HDHP Spending Impact				Treatment Effect		
				(1)	(2)	(3)
	Group %	Spending %	t_{-1} Mean Spending	Nominal Spending	CPI	Anticipatory Spending
t_{-2} Quartile 1	23.86	7.59	1636.85	-0.26	-0.29	[-0.28,-0.28]
t_{-2} Quartile 2	23.64	11.53	2592.70	-0.33	-0.36	[-0.33,-0.35]
t_{-2} Quartile 3	23.60	20.03	4412.69	-0.37	-0.39	[-0.35,-0.37]
t_{-2} Quartile 4	23.74	54.78	12051.12	-0.22	-0.25	[-0.16,-0.21]
t_{-1} Quartile 1	32.29	10.99	1752.40	-0.24	-0.27	[-0.26,-0.27]
t_{-1} Quartile 2	24.49	14.74	3209.34	-0.38	-0.40	[-0.34,-0.37]
t_{-1} Quartile 3	19.07	19.15	5174.46	-0.36	-0.39	[-0.32,-0.35]
t_{-1} Quartile 4	18.99	49.05	13617.06	-0.20	-0.24	[-0.15,-0.20]

Table A9: This table measures heterogeneous treatment effects by ACG quartile in two alternative ways.

Years	Rank Correlation
t_{-4}-t_{-3}	0.9363
t_{-3}-t_{-2}	0.9370
t_{-2}-t_{-1}	0.9275
t_{-1}-t_0	0.9321
t_0-t_1	0.9371

Table A10: This table gives Spearman's rank correlation coefficient for provider rankings in prices for a given procedure across year pairs in our data.

	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$	$\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$
All Claims	-10.4%	1.3%	1.6%	-16.5%
Preventive w/ Diagnosis	-7.5%	1.8%	0.7%	-10.2%
Preventive Always	3.3%	6.8%	0.6%	-6.5%
Imaging	-22.2%	-0.1%	4.5%	-22.4%

Table A11: This table analyzes price shopping behavior, comparing new employees at the firm in t_{-1} to new employees in t_0.

Total Spending Change Decomposition High Spend Procedures				
	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$	$\Delta PS_{t+1,t}$	$\Delta QE_{t+1,t}$*
% of top 30 w/ Positive Value				
t_{-3}-t_{-2}	80%	63%	43%	73%
t_{-2}-t_{-1}	80%	70%	63%	80%
t_{-1}-t_0	13%	53%	60%	17%
t_0-t_1	77%	37%	57%	80%

*We only present $\Delta QE_{t+1,t}$, given that these results are for one procedure at a time.

Table A12: This table presents the proportion of positive % changes for each part of the spending decomposition, for all 30 of the medical procedures that the firm and its employees spent the most money on, for year pairs from t_{-3} to t_1. The decomposition for the spending change from $t_{-1} - t_0$ is presented for each of these 30 procedures in Table A13 in Appendix A8.

	% Total Spend	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$	$\Delta PS_{t+1,t}$	$\Delta QE_{t+1,t}$
Routine Vaginal Birth (59400)	2.7%	-13.6%	-15.4%	1.4%	0.4%
Infliximab, 10mg (J1745)	2.6%	24.1%	10.2%	-2.6%	16.6%
MRI, Brain (70553)	2.0%	-6.1%	4.7%	-1.8%	-9.0%
Surgical Pathology, Skin (88305)	2.0%	-9.1%	-1.7%	-2.9%	-4.5%
Routine Cesarean Section Birth (59510)	1.9%	-19.1%	-16.8%	-0.1%	-2.2%
CT Scan, Abdomen and Pelvis (74177)	1.9%	-35.1%	-11.2%	-3.5%	-20.5%
Mammography Screening (G0202)	1.5%	-7.6%	0.3%	1.1%	-8.9%
Anesthesia for Vaginal Birth (01967)	1.3%	-15.4%	-1.0%	1.0%	-15.4%
Colonoscopy, with Biopsy (45380)	1.3%	-28.3%	2.6%	0.6%	-31.6%
MRI, Hip/Knee/Ankle (73721)	1.3%	-24.8%	1.2%	2.3%	-28.4%
Upper Gastrointestinal Endoscopy (43239)	1.2%	-24.2%	2.6%	1.1%	-27.9%
Colonoscopy, Diagnostic (45378)	1.1%	-28.5%	0.5%	2.2%	-31.2%
Wart Removal (17110)	1.1%	-24.9%	2.9%	0.7%	-28.4%
Foot, Molded Insert (L3000)	1.1%	-60.3%	2.0%	1.4%	-63.7%
Transvaginal Echography (76830)	1.0%	-21.5%	2.2%	-0.3%	-23.4%
Globulin, 500mg (J1561)	1.0%	49.7%	99.7%	0.0%	-50.0%
Pegfilgrastim, 6mg (J2505)	0.9%	28.0%	-1.2%	7.7%	21.4%
Fetal Non-Stress Test (59025)	0.8%	-11.5%	-4.7%	-8.5%	1.7%
Trastuzumab, 10mg (J9355)	0.8%	16.5%	-19.1%	0.2%	35.4%
Disposable Contact Lens (S0500)	0.7%	-5.9%	3.1%	4.7%	-13.7%
Laparoscopic Cholecystectomy (47563)	0.7%	-27.2%	4.3%	-3.4%	-28.1%
Ultrasound (76817)	0.7%	-17.8%	-5.7%	1.7%	-13.8%
Blood Count Test (85025)	0.7%	-5.0%	-1.7%	5.0%	-8.4%
Ultrasound (76811)	0.7%	-24.4%	-2.2%	1.2%	-23.3%
Echography of Pregnant Uterus (76805)	0.7%	-23.5%	-3.2%	-1.0%	-19.3%
Chest X-Ray (71020)	0.6%	-24.3%	5.7%	0.0%	-30.0%
Ultrasound (76801)	0.6%	-23.1%	0.4%	-0.6%	-22.9%
CT Scan, Abdomen and Pelvis (74176)	0.6%	-34.0%	-26.5%	13.1%	-20.6%
Thyroid Stimulating Hormone (84443)	0.6%	-8.3%	-2.3%	1.5%	-7.5%
MRI, Lumbar (72148)	0.6%	-26.6%	10.6%	-5.4%	-31.8%

Table A13: This table presents the results for our decomposition of the total reduction in medical spending between t_{-1} and t_0, for the top 30 procedures by firm-wide spending.

Prescription Drug Spending Change Decomposition				
	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$ + $\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$	$\Delta QS_{t+1,t}$
t_{-4}-t_{-3}	10.1%	6.4%	3.6%	0.1%
— Brand (38.8%)	10.5%	14.0%	-3.0%	-0.5%
— Generic (61/2%)	16.3%	5.2%	10.5%	0.6%
t_{-3}-t_{-2}	6.6%	5.3%	1.2%	0.1%
— Brand (35.3%)	7.5%	13.1%	-4.9%	0.7%
— Generic (64.7%)	8.3%	1.1%	7.1%	0.1%
t_{-2}-t_{-1}	4.2%	-0.2%	4.5%	-0.1%
— Brand (32.9%)	7.1%	6.7%	0.3%	0.1%
— Generic (67.1%)	-4.1%	-10.4%	6.9%	-0.6%
t_{-1}-t_0	-21.3%	-4.3%	-17.8%	0.8%
— Brand (28.7%)	-20.7%	13.6%	-30.3%	-4.0%
— Generic (71.3%)	-22.4%	-12.0%	-11.8%	1.4%
t_0-t_1	13.9	5.3%	8.1%	0.5%
— Brand (25.1%)	19.1%	17.5%	1.3%	0.3%
— Generic (74.9%)	-2.7%	-10.2%	8.3%	-0.8%

Table A14: This table presents the results for our spending reduction decomposition, applied to prescription drugs. The numbers in parenthesis in the first column indicate the percentage of drugs used that are brand vs. generic.

Medical Care	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$	$\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$	$\Delta QS_{t+1,t}$
Preventive Care, General					
t_{-4}-t_{-3}	4.0%	3.9%	-2.1%	-5.7%	7.9%
t_{-3}-t_{-2}	4.1%	-1.6%	9.2%	-0.4%	-3.1%
t_{-2}-t_{-1}	1.3%	-6.5%	-0.5%	6.3%	2.0%
t_{-1}-t_0	-0.3%	6.4%	2.1%	-7.5%	-1.3%
t_0-t_1	13.0%	12.6%	4.8%	-5.2%	0.8%
Preventive Care w/ Prior Diagnosis					
t_{-4}-t_{-3}	1.5%	3.0%	-0.8%	-0.4%	-0.3%
t_{-3}-t_{-2}	3.0%	2.4%	-0.7%	0.1%	1.2%
t_{-2}-t_{-1}	13.0%	3.6%	0.8%	7.3%	1.3%
t_{-1}-t_0	-10.6%	2.0%	1.0%	-12.2%	-1.4%
t_0-t_1	10.3%	5.8%	-0.2%	3.8%	0.9%
Preventive Care, Diabetics					
t_{-4}-t_{-3}	11.9%	3.3%	-0.9%	9.5%	0.0%
t_{-3}-t_{-2}	15.9%	-1.9%	2.9%	12.5%	2.4%
t_{-2}-t_{-1}	15.5%	1.2%	4.4%	9.7%	0.2%
t_{-1}-t_0	-1.5%	-2.0%	-0.5%	-1.6%	2.7%
t_0-t_1	0.1%	-16.5%	3.6%	16.4%	-3.3%
Mental Health					
t_{-4}-t_{-3}	15.5%	-0.3%	-0.1%	13.9%	2.0%
t_{-3}-t_{-2}	16.2%	-1.3%	0.0%	14.8%	2.8%
t_{-2}-t_{-1}	7.1%	-5.1%	-0.2%	11.5%	1.0%
t_{-1}-t_0	-2.9%	-1.0%	0.2%	-5.4%	3.4%
t_0-t_1	2.9%	-2.7%	1.3%	1.9%	2.4%
Physical Therapy					
t_{-4}-t_{-3}	16.6%	-0.2%	-0.1%	16.1%	0.7%
t_{-3}-t_{-2}	13.5%	0.9%	3.1%	8.5%	0.9%
t_{-2}-t_{-1}	9.2%	-1.1%	-1.7%	10.5%	1.4%
t_{-1}-t_0	-23.9%	0.3%	7.2%	-29.7%	-1.7%
t_0-t_1	1.4%	1.2%	-1.7%	0.5%	1.4%

Table A15: This table presents our spending change decomposition for types of health care that are potentially of high value to consumers, for medical care, for all year pairs in our data. This table extends Table VII.

Drugs	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$ + $\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$	$\Delta QS_{t+1,t}$
Diabetes Drugs				
t_{-4}-t_{-3}	12.0%	12.1%	0.0%	0.0%
t_{-3}-t_{-2}	29.2%	14.8%	12.6%	1.9%
t_{-2}-t_{-1}	20.1%	4.9%	14.5%	0.7%
t_{-1}-t_{0}	-44.5%	6.7%	-48.0%	-3.2%
t_{0}-t_{1}	175.7%	27.9%	115.6%	32.2%
Statins (for cholesterol)				
t_{-4}-t_{-3}	18.2%	15.1%	2.6%	0.4%
t_{-3}-t_{-2}	14.6%	16.8%	-1.9%	-0.3%
t_{-2}-t_{-1}	-31.3%	-8.8%	-24.7%	2.2%
t_{-1}-t_{0}	-47.1%	-34.3%	-19.6%	6.7%
t_{0}-t_{1}	59.6%	23.1%	29.6%	6.8%
Antidepressants				
t_{-4}-t_{-3}	-13.1%	-5.4%	-8.1%	0.4%
t_{-3}-t_{-2}	12.0%	0.4%	11.6%	0.0%
t_{-2}-t_{-1}	-17.7%	-19.7%	2.6%	-0.5%
t_{-1}-t_{0}	-48.7%	-37.4%	-18.0%	6.8%
t_{0}-t_{1}	4.2%	-30.3%	49.5%	-15.0%
Hypertension Drugs				
t_{-4}-t_{-3}	8.6%	1.5%	7.0%	0.1%
t_{-3}-t_{-2}	16.3%	3.2%	12.7%	0.4%
t_{-2}-t_{-1}	-9.0%	-10.9%	2.2%	-0.2%
t_{-1}-t_{0}	-27.9%	-4.9%	-24.2%	1.2%
t_{0}-t_{1}	14.8%	-22.1%	47.3%	-10.4%

Table A16: This table presents our spending change decomposition for types of health care that are potentially of high value to consumers, for drugs, for all year pairs in our data. This table extends Table VII.

Medical Care	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$	$\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$	$\Delta QS_{t+1,t}$
Imaging					
t_{-4}-t_{-3}	7.5%	5.6%	0.1%	3.1%	-1.3%
t_{-3}-t_{-2}	5.5%	2.7%	-1.9%	6.3%	-1.6%
t_{-2}-t_{-1}	12.4%	0.4%	0.2%	13.5%	-1.7%
t_{-1}-t_0	-19.5%	-0.4%	0.6%	-17.7%	-2.0%
t_0-t_1	-2.3%	-2.3%	3.7%	1.1%	-4.8%
CT Scan for Sinuses w/ Acute Sinusitis					
t_{-4}-t_{-3}	-17.3%	4.7%	-0.3%	-22.0%	0.3%
t_{-3}-t_{-2}	11.3%	0.4%	3.9%	5.2%	1.7%
t_{-2}-t_{-1}	30.9%	-8.1%	-3.3%	43.5%	-1.2%
t_{-1}-t_0	-24.8%	0.6%	1.1%	-26.0%	-0.4%
t_0-t_1	-10.7%	-4.6%	1.8%	-8.0%	0.2%
Back Imaging for Non-Specific Low Back Pain					
t_{-4}-t_{-3}	-7.0%	-0.4%	-4.9%	5.5%	-7.2%
t_{-3}-t_{-2}	22.3%	4.3%	-7.6%	14.5%	11.1%
t_{-2}-t_{-1}	9.9%	-1.6%	-5.8%	17.4%	-0.1%
t_{-1}-t_0	-26.1%	6.9%	-6.8%	-21.3%	-4.9%
t_0-t_1	22.1%	-7.4%	5.0%	11.8%	12.7%
Head Imaging for Uncomplicated Headache					
t_{-4}-t_{-3}	-5.0%	3.6%	5.6%	-12.6%	-1.5%
t_{-3}-t_{-2}	18.0%	0.4%	-1.8%	17.9%	1.4%
t_{-2}-t_{-1}	4.2%	1.3%	-12.6%	15.0%	0.5%
t_{-1}-t_0	-23.9%	-1.0%	6.6%	-30.7%	1.2%
t_0-t_1	-10.3%	-11.0%	0.0%	-0.7%	1.4%
Colorectal Cancer Screening for Patients Under 50					
t_{-4}-t_{-3}	-2.0%	5.2%	-1.1%	-6.0%	-0.2%
t_{-3}-t_{-2}	7.6%	1.3%	5.2%	-3.4%	4.5%
t_{-2}-t_{-1}	47.5%	-1.0%	4.4%	25.1%	19.0%
t_{-1}-t_0	-32.3%	0.7%	-0.8%	-26.2%	-6.0%
t_0-t_1	12.2%	0.1%	2.9%	3.6%	5.6%

Drugs	$\Delta TS_{t+1,t}$	$\Delta PPI_{t+1,t}$ + $\Delta PS_{t+1,t}$	$\Delta Q_{t+1,t}$	$\Delta QS_{t+1,t}$
Antibiotics for Acute Respiratory Inspection				
t_{-4}-t_{-3}	13.5%	4.2%	8.9%	0.4%
t_{-3}-t_{-2}	-4.8%	-5.3%	0.5%	0.0%
t_{-2}-t_{-1}	-34.2%	-18.9%	-18.8%	3.6%
t_{-1}-t_0	-47.8%	-6.2%	-44.4%	2.8%
t_0-t_1	4.4%	-29.1%	47.3%	-13.8%

Table A17: This table presents our spending change decomposition for types of health care that are potentially of low value to consumers, for all year pairs in our data. This table extends Table VIII.

Measure	Definition
High Value Care	
Preventive Care, General	A number of definitions generated by the firm, based on both CPT and ICD-9 codes. Some require a specific ICD-9 and CPT combination. All care under these definitions is free for individuals no matter their plan.
Preventive Care, w / Prior Diagnosis	A subset of the above definition, including only those definitions that require a diagnosis code
Preventive Care, Diabetes	ICD-9 codes E08.X-E11.X, E13.X, 249.X, or 250.X, **and** CPT codes relating to standard diabetic tests, e.g. kidney screening
Physical Therapy	Either ICD-9 code V57.1 **or** one of many relevant CPT procedure codes, primarily those starting in '97'
Diabetes Drugs	A number of NDC codes corresponding to many drug types, e.g. insulin
Antidepressants	A number of NDC codes corresponding to many drug types, e.g. SSRIs
Statins	A number of NDC codes corresponding to different brands
Hypertension Drugs	A number of NDC codes corresponding to many drug types, e.g. ACE inhibitors
Low Value Care	
CT Scan of Sinuses for Acute Sinusitis	ICD-9 codes 461.X or 473.X, **and** CPT codes 70486-70488
Head Imaging for Uncomplicated Headache	ICD-9 codes 307.81, 339.X, 346.X, or 784.0 **and** CPT codes 70450, 70460, 70470, or 70551-70553
Back Imaging for Non-Specific Low Back Pain	One of many ICD-9 codes indicating low back pain, e.g. 721.42 **and** one of many CPT codes indicating imaging on the back, e.g. 72010
Colorectal Cancer Screening for Patients Under 50	One of many CPT codes indicating screening, including endoscopy and blood screening, for any individual with age less than 50 in the month of use
Antibiotics for Acute Respiratory Infection	NDC codes corresponding to use of penicillin, amoxicillin, or azithromycin in months where the individual incurred a medical claim with the ICD-9 codes 46X or 478.8

Table A18: This table provides definitions for how we categorize low-value care. For each measure, we take all medical or prescription drug claims that correspond to the definition given. When ICD-9 codes are mentioned above, it refers to diagnosis codes. When CPT codes are mentioned, it refers to procedure codes. For ICD-9 codes, X is a catch-all, e.g., '461.X' refers to all codes beginning with '461.' When "**and**" is mentioned, it means that the definition requires a claim to have both the corresponding ICD-9 code and the corresponding CPT code.

	Spending %	t_{-1} Mean Spending	Treatment Effect (1) Nominal Spending	(2) CPI	(3) Anticipatory Spending
Low Value Care					
CT Scan for Sinuses w/ Acute Sinusitis	0.05	2.50	-28.34%	-31.19%	[-16.31%,-25.13%]
Head Imaging for Uncomplicated Headache	0.14	7.10	-26.22%	-29.42%	[-27.34%,-28.56%]
Back Imaging for Non-Specific Low Back Pain	0.22	11.62	-21.13%	-24.52%	[-12.25%,-19.24%]
Colorectal Cancer Scrng. for Patients Under 50	0.37	19.45	-30.98%	-33.72%	[-9.52%,-24.07%]
Antibiotics for Acute Respiratory Infection	0.11	5.52	-45.93%	-48.01%	[-67.66%,-54.73%]
All Low Value Care	0.98	51.30	-27.72%	-30.71%	[-20.41%,-26.5%]
High Value Care					
Preventive Care, Diabetes	0.02	0.99	-3.69%	-7.73%	[-5.23%,-6.53%]
Physical Therapy	6.45	337.10	-24.99%	-28.08%	[-24.7%,-26.66%]
Diabetes Drugs	0.54	28.43	-46.66%	-48.28%	[-38.83%,-45.06%]
Antidepressants	0.93	48.82	-46.85%	-48.85%	[-52.48%,-50.08%]
Statins	0.38	19.97	-10.82%	-14.69%	-14.69%*
Hypertension Drugs	0.30	15.55	-33.04%	-35.52%	[-37.41%,-36.26%]
All High Value Care	16.96	885.51	-14.59%	-18.13%	[-18.63%,-18.35%]

*Anticipatory spending estimate negative or not significant from 0

Table A19: This table replicates the analysis from Table A5, applied to the categories of care defined as high value and low value in Section 4.

Preventive Care Decomposition **$ per Primary Care Visit**						
	Visits t_{-1}	Visits t_0	\$ per visit t_{-1}	\$ per visit t_0	Ext. $\Delta\%$	Int. $\Delta\%$
General Preventive Services						
ICD-9 / CPT	143,887	126,406	\$62.57	\$64.79	-12.1%	3.5%
Specialty	63,121	54,218	\$65.46	\$67.68	-14.1%	3.3%
Preventive w/ Prior Diagnosis						
ICD-9 / CPT	143,887	126,406	\$117.18	\$114.59	-12.1%	-2.1%
Specialty	63,121	54,218	\$136.46	\$133.98	-14.1%	-1.8%

Table A20: This table analyzes consumer reductions in preventive care shifting from the pre-period t_{-1} to the treatment year t_0. It decomposes the reduction in preventive care into the extensive margin, measured by monthly office visits, and an intensive margin, measured by $ per office visit spent on preventive care. The tables shows how office visits changed, according to two definitions of office visits described in the text, and documents spending on such care per visit.

Preventive Care Decomposition **Binary Care Indicator**						
	Visits t_{-1}	Visits t_0	Prev. Vis. / Vis. t_{-1}	Prev. Vis. / Vis. t_0	Ext. $\Delta\%$	Int. $\Delta\%$
General Preventive Services						
ICD-9 / CPT	143,887	126,406	0.396	0.409	-12.1%	3.0%
Specialty	63,121	54,218	0.496	0.502	-14.1%	1.3%
Preventive w/ Prior Diagnosis						
ICD-9 / CPT	143,887	126,406	0.768	0.763	-12.1%	-0.7%
Specialty	63,121	54,218	0.792	0.793	-14.1%	-2.4%
Mammography						
ICD-9 / CPT	143,887	126,406	0.0086	0.0074	-12.1%	-14.8%
Specialty	63,121	54,218	0.0170	0.0159	-14.1%	-6.7%
Urinalysis						
ICD-9 / CPT	143,887	126,406	0.0716	0.0694	-12.1%	-3.0%
Specialty	63,121	54,218	0.1234	0.1229	-14.1%	-0.4%
Colonoscopy						
ICD-9 / CPT	143,887	126,406	0.0032	0.0024	-12.1%	-23.4%
Specialty	63,121	54,218	0.0047	0.0031	-14.1%	-35.0%

Table A21: This table analyzes consumer reductions in preventive care shifting from the pre-period t_{-1} to the treatment year t_0. It decomposes the reduction in preventive care into the extensive margin, measured by monthly office visits, and an intensive margin, measured by the proportion of office visits where any preventive care services are consumed. The tables shows how office visits changed, according to two definitions of office visits described in the text, and documents concordent changes in care per visit. We also present this decomposition for three specific preventive services.

Policy Change: Price Impact t_{-1} **Total Spending**					
Coverage Tier	Avg. HDHP Price	% Under Deductible	% Over Ded., Under OOP Max.	% Over OOP Max.	Actuarial Value
0 Dependents	0.428	37.92%	49.16%	12.92%	78.31%
1 Dependent	0.293	23.22%	61.08%	15.70%	76.59%
2+ Dependents	0.201	13.30%	68.40%	18.30%	78.24%
All Tiers	0.249	18.42%	64.46%	17.12%	78.05%

Table A22: This table presents statistics for our primary sample describing the average and marginal price changes resulting from the required HDHP switch. We present the average % of total spending paid, as well as the likelihood of reaching each arm of the non-linear HDHP contract.

Ventile Regression Coefficients		
	Coefficient	
Ventile	Treatment	Treatment X t_1
2	-0.0516	0.0428
	(0.0454)	(0.0440)
3	-0.0409	0.00463
	(0.0475)	(0.0466)
4	-0.148***	0.0346
	(0.0486)	(0.0474)
5	-0.140***	0.0399
	(0.0489)	(0.0476)
6	-0.164***	0.0915*
	(0.0495)	(0.0482)
7	-0.121**	0.0429
	(0.0494)	(0.0482)
8	-0.0780	0.0835*
	(0.0494)	(0.0483)
9	-0.150***	0.0913*
	(0.0502)	(0.0492)
10	-0.0376	0.0119
	(0.0529)	(0.0522)
11	-0.0891*	0.114**
	(0.0536)	(0.0527)
12	-0.100*	0.0760
	(0.0542)	(0.0531)
13	-0.145***	0.187***
	(0.0545)	(0.0534)
14	-0.171***	0.135**
	(0.0552)	(0.0537)
15	-0.000201	0.0884
	(0.0555)	(0.0539)
16	-0.0212	0.0719
	(0.0557)	(0.0542)
17	0.0403	0.129**
	(0.0562)	(0.0543)
18	0.113**	0.0911*
	(0.0564)	(0.0547)
19	0.185***	0.0933*
	(0.0565)	(0.0550)
20	0.151***	0.120**
	(0.0568)	(0.0551)

Table A23: This table presents the coefficients on shadow price ventiles for our non-linear contract price regressions.

Mean Individual Spending By Month		
Month	Mean Spending	Mean Spending, Detrended
t_{-4}, March	352.15	347.91
t_{-4}, June	360.89	351.71
t_{-4}, September	333.98	319.80
t_{-4}, December	358.07	337.26
t_{-3}, March	397.97	365.47
t_{-3}, June	362.47	328.91
t_{-3}, September	351.97	313.95
t_{-3}, December	368.23	324.94
t_{-2}, March	436.87	381.86
t_{-2}, June	412.69	355.13
t_{-2}, September	385.52	327.83
t_{-2}, December	376.79	316.01
t_{-1}, March	471.71	393.43
t_{-1}, June	414.34	338.62
t_{-1}, September	404.86	329.01
t_{-1}, December	526.96	422.53
t_0, March	355.94	282.28
t_0, June	338.97	268.07
t_0, September	372.86	287.69
t_0, December	417.47	322.12
t_1, March	405.21	306.96
t_1, June	386.42	290.04
t_1, September	412.19	307.42
t_1, December	512.89	378.54

Table A24: This table gives mean spending by individuals for a set of months in our data.

Family Shares and Total Spend by HDHP Plan Arm						
	February	April	June	August	October	December
Family Shares						
t_0 Deductible Arm	93.1%	77.4%	61.6%	50.2%	40.5%	33.1%
t_0 Coinsurance Arm	6.5%	21.1%	34.9%	44.1%	51.6%	56.0%
t_0 OOP Maximum Arm	0.4%	1.5%	3.4%	5.6%	7.9%	10.9%
Total Spend ($ million)						
t_0 Deductible Arm	10.44	7.93	4.45	3.37	2.54	1.86
t_0 Coinsurance Arm	3.86	6.84	7.59	8.74	9.76	10.24
t_0 OOP Maximum Arm	0.72	2.02	3.13	4.76	5.59	6.25

Table A25: This table shows the share of families who begin a month in t_0 in a given arm of the non-linear HDHP, as well as total spending by month and plan arm across these families for that month.

Shadow Prices by Plan Arm and Health Status					
	Sickest 10%	Quartile 1 (Sickest)	Quartile 2	Quartile 3	Quartile 4
t_0 **Deductible Arm**					
February	0.06	0.08	0.15	0.31	0.58
April	0.09	0.10	0.17	0.40	0.70
June	0.10	0.10	0.22	0.52	0.80
August	0.10	0.11	0.31	0.67	0.88
October	0.10	0.14	0.51	0.83	0.95
December	0.10	0.19	0.75	0.96	0.99
t_0 **Coinsurance Arm**					
February	–	0.01	0.04	0.06	0.10
April	–	0.03	0.06	0.08	0.10
June	–	0.04	0.08	0.09	0.10
August	–	0.05	0.09	0.10	0.10
October	–	0.07	0.09	0.10	0.10
December	–	0.08	0.10	0.10	0.10

Table A26: This table shows mean t_0 family shadow prices, i.e. true expected end-of-year marginal prices, as a function of (i) their spot price at the start of a month and (ii) where they fall in the distribution of family expected-of-year price, conditional on their spot price.

Price Correlations by Month, t_0-t_1			
	Spot-Shadow	Spot-Prior End	Shadow-Prior End
February	0.285	0.131	0.627
April	0.489	0.229	0.564
July	0.668	0.315	0.513
October	0.798	0.363	0.460
December	0.857	0.381	0.437

Table A27: This table shows the correlation in different non-linear contract prices that we consider in our primary regressions, for months pooled over the treatment years t_0-t_1.