

Do “Consumer-Directed” Health Plans Bend the Cost Curve Over Time?

Amelia M. Haviland^a

Matthew D. Eisenberg^b

Ateev Mehrotra^c

Peter J. Huckfeldt^d

Neeraj Sood^e

^a *Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213 and RAND 4570 Fifth Avenue Suite 600 Pittsburgh, PA 15213, United States amelia@andrew.cmu.edu*

^b *Carnegie Mellon University 5000 Forbes Avenue Pittsburgh, PA 15213, United States matthewdeisenberg@cmu.edu*

^c *Harvard Medical School 25 Shattuck Street Boston, MA 02115 and RAND 20 Park Plaza 9th Floor Suite 920 Boston, MA 02116, United States mehrotra@hcp.med.harvard.edu*

^d *University of Minnesota 420 Delaware Street Minneapolis MN 55455, United States huckfeld@umn.edu*

^e *University of Southern California 635 Downey Way Los Angeles, CA 90089, United States nsood@healthpolicy.usc.edu*

Abstract

“Consumer-Directed” Health Plans (CDHPs) combine high deductibles with personal medical accounts and are intended to reduce health care spending through greater patient cost sharing. Prior research shows that CDHPs reduce spending in the first year. However, there is little research on the impact of CDHPs over the longer term. We add to this literature by using data from 13 million individuals in 54 large US firms to estimate the effects of a firm offering CDHPs on health care spending up to three years post offer. We use a difference-in-differences analysis and to further strengthen identification, we balance observables within firm, over time by developing weights through a machine learning algorithm. We find that spending is reduced for those in firms offering CDHPs in all three years post. The reductions are driven by spending decreases in outpatient care and pharmaceuticals, with no evidence of increases in emergency department or inpatient care.

Funding

This study was funded in part by a grant from the National Institute of Aging (NIA) and the NIH Common Fund for Health Economics (Grant R01-AG043850) and a grant from the California Health Care Foundation.

1. Introduction

Consumer-directed health plans (CDHPs), health plans in which high deductibles are paired with tax-advantaged personal medical accounts, are becoming increasingly common. Enrollment in CDHPs in the employer-sponsored market has grown from 4% to 20% over the last five years (KFF and HRET 2014). CDHPs are the dominant type of plan on the individual health market, accounting for nearly four fifths of the plans offered on the federal health insurance exchange. CDHPs are promoted as a strategy for reducing health care spending and some have suggested that the increasing prevalence of CDHPs may have contributed to the recent slowdown in US health care spending (Cutler and Sahni 2013; Haviland et al. 2012).

While the theory and empirical evidence regarding the impact of CDHPs on spending in the short term is clear, the longer-term impacts are less certain. Conceptually, higher deductibles and the associated increase in out-of-pocket costs suggest that patients will reduce their health care spending in the short term. Not surprisingly, there is robust empirical evidence that CDHPs effectively incentivize patients to change their health care use and reduce costs in the first year after introduction (e.g. Buntin et al. 2011; Lo Sasso et al. 2010). However, the effects of CDHPs on spending in the longer term are a priori ambiguous and the empirical evidence on the long term effect of CDHPs on spending is limited (Fronstin et al., 2013; Borah et al. 2011; Wharam et al. 2011). In this paper, we address this gap in the literature by describing a theoretical framework for long term impacts and empirically evaluating the long term effects of CDHPs on health care spending. We do the latter by analyzing up to 3 years of data post-CDHP offer from 54 large employers, half of which offered a diverse set of CDHPs to their employees. By investigating a diverse set of firms, we are able to analyze heterogeneities in long term CDHP impacts by CDHP plan structure.

Research on CDHPs must address the observational nature of the data; neither this paper, nor any other in the last two to three decades, has run an experiment and randomly assigned health insurance plans with variable cost sharing. The options for identifying the effect of CDHPs as they are currently offered are limited. One option some have pursued is to compare those enrolling in CDHPs to those working for the same firms but enrolling in other plans with lower cost sharing. This avoids firm level selection but is subject to individual level selection bias. An alternative used by a few study teams is to focus only on full-replacement firms, those that choose to offer only CDHP(s) to all their employees. This approach strategically avoids individual level selection bias but limits generalizability and introduces firm level selection bias. In contrast, we take a previously unexplored middle road and compare the cost trends for all individuals at firms offering CDHPs to cost trends for individuals at similar firms not offering CDHPs. This is an intent-to-treat design as the focus is on individuals offered CDHPs and not on individuals enrolled in CDHPs. Since the decision to offer CDHPs is made by the employer, this intent-to-treat design eliminates any individual level selection bias. The design also likely reduces any firm level selection bias as offering a CDHP alongside other plans is a moderate step relative to full replacement with a CDHP.

In addition, we implement a difference in difference identification strategy allowing us to control for time invariant firm level differences. Given this identification strategy, the key assumption required to identify the impact of CDHP offer is that firms choosing to offer a CDHP (treatment firms) in the study period would, on average, have had the same time trend in costs as firms choosing not to offer, had the treatment firms not offered CDHPs. Although we cannot directly test this assumption, we mitigate any violations by controlling for time-varying changes in employee composition through a novel application of a machine learning technique in order to

develop weights that balance firms' health plan enrollee composition over time. The estimates from this intent to treat analysis are straightforward to interpret as the effect of offering CDHPs on health care spending.

At the firm level, we find that CDHP offer is associated with an approximately 5 percent reduction in total health care spending in each of the three years after CDHPs were introduced relative to cost growth observed for non-offering employers. The long term decreases in spending are focused in outpatient care and drugs and there is little impact on inpatient or emergency department spending. If these effects are due only to changes in health care spending among those enrolled in CDHPs, they imply local average treatment effects for those enrolled in CDHPs of an approximately 15 percent reduction in total spending in each the first three years. Differences in impacts by CDHP plan structure are not statistically significant. However, consistent with our hypotheses, the pattern of the point estimates suggests that the impact of CDHPs is greater when paired with HSAs (versus HRAs) and when employers make smaller account contributions.

The remainder of the paper begins with an outline of the conceptual framework surrounding the long term effects of CDHPs and a brief review of related literature, followed by descriptions of the data and variables, presentation of analytic methods, results, and discussion.

2. Conceptual Framework

2.1 Optimality of Care Choices Resulting in Short Term Cost Reductions

Whether CDHPs continue to keep health care costs down in the long term depends on two unknown factors. First is the nature of the health care choices patients make in the short term to reduce costs. At one extreme all patients make optimal value-based health care decisions leading

to sustainable cost reductions in the long term. At the other extreme patients indiscriminately reduce care, and in particular they forgo needed care, leading to deteriorating health and higher health care spending in the long term. The second unknown factor is whether CDHP financial incentives continue to be effective over time, as balances in personal medical accounts may grow.

What we know about these two factors is limited. CDHP enrollees may cut back on necessary or high-value care because of imperfect information on insurance coverage, health care price, and health care quality or myopia that leads consumers to heavily discount the future consequences of current action. Empirical evidence on the nature of cost-saving changes in health care use is mixed but suggests there is some cause for concern. Twenty-five years ago, the RAND Health Insurance Experiment (HIE) (Newhouse 1993) found that increased cost sharing caused patients to reduce both necessary and unnecessary care. More recent observational studies have similarly provided evidence that CDHP enrollees reduce some necessary preventive and chronic care (Haviland et al. 2011; Charlton et al. 2011; Chen et al. 2010; Huckfeldt et al. 2015). And two recent studies asking CDHP enrollees about the key features of their coverage found that the vast majority of enrollees did not know the deductible did not apply to certain high value care (Reed et al. 2012; Reed et al. 2009).

The nature of health care choices made to reduce spending in the short term, and their impacts on long term spending are likely to differ by type of care. The RAND HIE found different price elasticities for different types of services (e.g. -0.14 for hospital care and -0.31 for outpatient care) (Manning et al. 1987). These differences can be driven by relatively better or worse price and quality information availability and perceptions of the current and future value of different types of care. Specific patterns of differential impacts by type of care can thus inform

whether CDHPs enrollees are forgoing necessary care. For instance, if a sharp decline in outpatient spending in the short term is followed by increased ED use and inpatient spending in the long term, this would be suggestive of people cutting back on necessary care.

Regarding the continued effectiveness of CDHP financial incentives two recent reports provide some information regarding account balances over time. One summarizes balances of 1.5 million HSAs held at five large banks (AHIP 2014). This summary shows balances increasing on average from about \$1,200 when opened to \$2,700 by the third year. However, in 2012, they show more than 60% of HSAs end the year with balances of less than \$1,000. The second study uses survey data from members of an internet panel and finds that 47% of HRA and HSA account holders combined had balances less than \$1,000 in 2008 (Fronstin, 2014). It is unknown how representative these findings are.

2.2 Effect of Learning

While substantial barriers exist for patients in making optimal value-based health care decisions, these may be mitigated by CDHP-offering employers making health care decision tools available, particularly if these improve over time. Greater familiarity with these tools has been demonstrated to allow CDHP enrollees to more effectively shop for lower cost providers (Whaley et al. 2014). In addition, over time, CDHP enrollees may gain other skills and knowledge relevant to making high value health care choices. This could include becoming more familiar with the plan design and coming to understand that most preventive care is excluded from the deductible (Buntin et al. 2011; Reed et al. 2012). With this knowledge, CDHP enrollees may stop forgoing necessary care, making spending reductions more plausibly sustainable.

2.3 Plan Structure

Enrollees' responses to CDHPs as time goes on are likely to depend on plan features that determine the extent of enrollee cost exposure. Differences in such features may underlie the mixed findings in the limited literature on long term CDHP impacts (see next section for a summary of this literature). These features include type of personal medical account, Health Reimbursement Account (HRA) versus a Health Savings Account (HSA), the level of the deductible, and the level of employer contribution to the account. HRAs and HSAs are similar in that contributions are pre-tax, both employer and employee may make contributions, and funds roll over from year-to-year. There are, however, important differences. First, employers must contribute funds to HRAs, retain ownership of the account, and typically keep any balance when an employee leaves the firm. HSAs, in contrast, are owned by the employee, funds may be used tax-free for health care costs at any time including retirement, and funds may be withdrawn for other purposes but with tax consequences. Thus HSAs provide a stronger incentive to save. The variability in employer account contribution and the size of the deductible will also impact the enrollee's exposure to the increased cost sharing. Our data allow us to investigate the potentially different impacts of these common CDHP benefit designs over time.

3. Related Literature

There have been a number of studies examining the short term impacts of offering a CDHP in the employer sponsored health insurance setting. An excellent review of this work (Bundorf, 2012) concludes there is robust evidence that CDHPs usually lead to reduced spending growth in the first year they are offered by employers. The effect sizes have consistently been found to be between 5 (Losasso et al. 2010) and 24 percent (Haviland et al. 2011). These first year effects were mostly driven by changes in pharmaceutical (Parente et al. 2008) and outpatient spending (Buntin et al. 2011). Below, we focus on the small literature investigating longer term impacts.

There are four studies assessing impacts of CDHPs on spending over more than a one year time frame. One case study compared an employer which fully replaced traditional plans with two CDHP options to a second employer that continued to offer traditional plans over a 4 year time horizon (Fronstin et al., 2013). The CDHPs were HSA eligible and the employer deposited \$700/1300 in employee's accounts for individuals/families. In the first year of offer, the authors found reductions in all components of health care spending except inpatient stays (but including emergency department visits), relative to the comparison firm. However, at four years after offer, employees in the intervention firm had lower spending in only pharmaceuticals and laboratory tests, but not any other aspect of inpatient or outpatient care.

The other three studies assess impacts of CDHPs on two years of spending. Borah and colleagues (2011) also studied a single intervention firm and, contrary to Fronstin et al. (2014), found no impact on spending in either the first or second year after offer relative to a comparison firm. The remaining two studies differ in two ways. First, they both make use of multiple CDHP-offering firms. Second, they use those enrolling in non-CDHPs at the CDHP-offering firms as their comparison group. LoSasso et al. (2010) study data from one insurer and hundreds of firms where the CDHPs offered were paired with HSAs and had an average deductible of \$1,500. The employer account contribution and plan-specific deductible was not noted. CDHP enrollment was associated with a 5-7 percent decrease in spending relative to individuals at the same firms who did not enroll in a CDHP (they used a subset of 1-to-1 propensity score matched individuals in sensitivity analyses). They report that more of the reduction occurs in the first rather than the second year but do not provide a statistical significance test of the difference. Charlton et al. (2011), also compared CDHP enrollees to non-CDHP enrollees within the same firms and find relative decreases in spending in both years.

Finally there are four studies that do not measure the impact of CDHPs on spending but focus on utilization. Wharam et al. (2011) and Wharam et al. (2013), compare hospital utilization for many small firms, each of which fully replaced their prior HMO plan with a CDHP, to a set of small firms that continued to offer only the HMO two years after offer of CDHP. The CDHPs offered have a range of deductibles (some as low as \$500) and neither the distribution of these deductibles nor employer account contribution levels are reported. They find that for CDHP enrollees emergency department use is lower in both years but the reduction in inpatient stays is only lower in the first year. Because these are full replacement firms these studies are not subject to bias due to individual level CDHP enrollment selection. We would likely expect their utilization results to generalize to other small firms offering a single health insurance plan to their employees, particularly in the same region and industries.

The other two utilization studies, Kozhimannil et al. (2013) and Reddy et al. (2014), make use of data from one firm, Harvard, to compare utilization of those choosing to enroll in an HMO to that of CDHP enrollees over two years. In both years CDHP enrollees had fewer office visits (for both higher and lower priority conditions), laboratory tests, and emergency department use than HMO enrollees. They also find reductions in inpatient stays in the first year followed by increases in the second year.

Thus, findings on the long term impact of CDHPs on spending are inconclusive with four studies to date with different findings: similar reductions in the first two years, no reductions in some or all types of spending by the second/fourth year, or no impact in any year. However, due to the different settings, CDHP designs, comparison groups, and identification strategies, it is not necessarily surprising that results have varied. The utilization studies find reductions in ED use

and a variety of outpatient care in both CDHP years and reductions in inpatient visits in the first year only, with one study showing a significant increase in the second year.

The prior work has a number of important limitations. The case studies have just one intervention firm which limits generalizability, particularly as the findings of these studies do differ. The only study to assess impacts beyond two years is one of these case studies. The studies that compare CDHP enrollees to non-CDHP enrollees at CDHP-offering employers are subject to potential individual level selection bias. These studies use regression adjustment on observed covariates and (in one case) propensity-score matching to address potential bias, but whether such adjustment is sufficient is unclear as there are likely unobserved factors affecting both plan selection and future health care use.¹ Studies using employers fully replacing their non-CDHP offerings with a CDHP may be more likely to suffer from firm level selection bias.² Each case study shows the impacts of a CDHP with a single set of design features and the multi-firm studies did not have access to plan level information regarding the design features of the CDHPs. Hence, it is unknown whether the differing results in the literature may be due to the shifting CDHP design features across, and within, studies.³ Finally, none of the studies make statistical comparisons of first year vs. longer term impacts.

¹ Prior literature finds that when CDHPs are offered as an option alongside traditional plans, those who enroll in the CDHP tend to be considerably healthier and a bit younger and wealthier than those who enroll in other plans (McDevitt et al. 2014).

² Particularly for small employers, this could arise if large increases in health care costs lead to the switch to CDHPs but also tend to be followed by lower costs merely by regression to the mean. For large employers it has been fairly uncommon to offer only CDHPs and this decision may be coupled with other unique firm features affecting health care costs.

³ The treatment effect being estimated, the impact of CDHPs on spending, may not be the same across papers. This could be due to either the CDHPs in question *or the plans the comparison groups are enrolled in* having substantially different designs. CDHPs are preferred-provider organization (PPO) plans and in all four spending studies the comparison groups have also been enrollees of PPOs, just those with low deductibles. Some studies instead use health maintenance organization plans (HMOs) as the comparator. Prior literature has suggested that cost growth is slower in HMOs than PPOs (for example Health Affairs, 2000, Baker, Cantor, Long, Marquis) and thus

We add to this literature and improve over past limitations in three ways. First, we study the impacts of CDHPs on health care spending for three years using data from more than 20 large CDHP-offering employers and almost five million employees and dependents located throughout the U.S. None of the four multi-firm studies have gone beyond two years of impacts and two did not have access to data on spending. Second, by using an intent-to-treat design and benchmarking the lower deductible PPO-offering comparison firms to national trends we avoid bias due to individual level selection into CDHPs and ensure that any impacts are attributable to differences in cost sharing rather than differences in plan type. Third, we are able to investigate potential heterogeneous treatment effects due to extent of CDHP cost sharing. In comparison, prior studies have focused on either a single CDHP design or multiple, unspecified CDHP designs.

4. Data

The data was constructed by our research group specifically to study CDHPs across a wide range of employers, plan types, and geographic areas. The data include enrollment and claims data from 2003 through 2007 from 54 large U.S. employers. These data are supplemented by detailed insurance benefit design, geocoded socio-demographic information, and firm-level survey data. In total the data includes 13 million person years of data (~5 million from CDHP-offering firms and ~8 million from non-CDHP offering firms) in hundreds of health plans offered by all major and many of the smaller health insurance carriers in the U.S.

The employers were chosen in two ways. The first group includes 26 employers that were identified and recruited to join the study because they reported offering a CDHP at some point

this comparison may yield different results due to the difference in comparison group. Fewer than 14% of those with employer sponsored insurance are enrolled in HMOs whereas 57% are in PPOs (KFF 2014).

between 2004 and 2007. These firms were clients of Towers Watson and were selected to include a range of CDHP benefit designs, proportions of employees enrolling in those CDHPs, geographic regions, and industries. Two of the recruited firms had very low rates of CDHP adoption in only 2007 (less than 3.0% take-up). These two firms were switched to the comparison group. The remaining 28 firms were selected from firms contributing data to the MarketScan data product produced by Thomson Reuters. These 28 firms were selected to roughly match the geographic region, size, and industry distribution of the recruited set of employers. See Appendix table A1-1 for full information on all sample restrictions.

The key independent variable of interest is whether a firm offered CDHPs. For this analysis, we define a firm as offering a CDHP in a given year if they offered a CDHP *and* greater than 5% of their employees were in a CDHP. We use this information to define three treatment indicator variables: first year of CDHP offer, second year of CDHP offer, and third year of CDHP offer. We also create parallel indicator variables designating the first and second years prior to CDHP offer.

We characterized the relative strength of the cost sharing incentives of the CDHPs with three variables. One is an indicator for type of personal medical account the CDHP offered – HSA versus HRA. Second is based on the employer contribution to the personal medical accounts where low and high are defined by a median split across offering employers (\$500). The third variable describes the effective deductible, defined as the deductible minus employer account contribution. Again this was defined as low or high by the median split (\$1,250). These indicators are interacted with the main CDHP offer indicators in some models. Deductible levels were calculated from the claims data for all plans. For CDHPs, deductible values and other benefit design features (HSA vs. HRA, employer contribution), were confirmed via surveys and

interviews with human resources executives at each CDHP offering firm. For the local average treatment effect (LATE) analyses described below we also create an indicator variable designating whether or not each individual in CDHP offering firms in years of offer is enrolled in a CDHP.

The outcome variables are constructed from the claims data. The main outcome was total annual health costs per person which was the total of amounts paid, by all payers including enrollee, for health care services in each plan year. In addition to total health care costs, costs are also divided into four mutually exclusive categories: outpatient, pharmaceutical, inpatient, and emergency department.

Individual level covariates are obtained from the enrollment data accompanying the claims: Census region interacted with living in a metropolitan statistical area (MSA), age, age squared, family type, and sex. In addition the following zip-code level Census data were linked based on geocoded employee address: median household income, % unemployed, % college degree, % high school degree, % Black, % non-Hispanic White. The Census region by MSA indicator variables are included as geographic fixed effects. Firm and calendar-year fixed effect indicators are created from the claims and enrollment data.

5. Methods

5.1. Econometric Model

We seek to identify the average effect of employer CDHP *offer* in the first, second and third years after CDHPs are introduced by an employer. We use an employer level difference-in-differences empirical approach, estimating the differences in cost growth over time between treatment (CDHP offering firms) and control firms.

To model total spending, we use a generalized linear model (GLM) with gamma-type errors and a log link to account for the skew in healthcare costs and the truncation at zero. We chose the gamma log specification through a model selection process where, using only firm years in which CDHPs were not offered, we estimated GLMs with either an identity link or log link and with either constant variance, proportional to the mean variance, or proportional to the mean squared variance and assessed the mean square error (MSE), mean absolute prediction error (MAPE), Akaike information criterion (AIC), and Bayesian information criterion (BIC)⁴.

Our baseline specification to identify the intent-to-treat (ITT) effect is:

(1)

$$y_{ift} = f(\alpha_0 + \beta^{-1}Pre2_t + \beta^1Post1_t + \beta^2Post2_t + \beta^3Post3_t + X_{it}'\delta + \gamma_tYear_t + \theta_fFirm_f + \varepsilon_{it})$$

where, y is a measure of costs for individual i in firm f in year t , the *Pre*- and *Post*- variables indicate year relative to first CDHP offer (with *Pre1*, the baseline year before CDHP offer, as the omitted category), X is a vector of individual demographics, zip code characteristics, and geographic fixed effects, $Year$ is a vector of calendar year fixed effects, and $Firm$ is a vector of firm fixed effects. The unit of analysis is the person-year. We include individuals in the models when they are enrolled in a health plan for one or more full plan years.⁵ We cluster standard errors at the firm level.

The coefficients of interest are the difference-in-differences (DiD) estimators, β^1 , β^2 , and β^3 . They indicate how much lower or higher average total health care cost growth is in each of the three years after CDHP offer relative to what it would have been without CDHP offer, after controlling for time invariant firm characteristics and secular trends. The coefficient on the *Pre2*

⁴Fit statistics from model selection procedure can be found in table A1-2.

⁵ Table A1-1 shows impacts of these sample restrictions on our analytic sample.

indicator is included to test for differences in cost growth between treatment and control firms prior to CDHP offer in the treatment firms. The key identifying assumption is that treatment firms would have had the same time trends as control firms in the absence of CDHP offer by treatment firms. In the next section, we describe a novel machine learning based approach that mitigates the possibility that this identifying assumption is violated.

We estimate (1) for total annual health care costs and separately on health care costs for each of the four types of care (outpatient, inpatient, emergency department, or prescription drugs). To test for differential effects by strength of CDHP incentive, we estimate (1) with *Post1*, *Post2*, and *Post3* interacted with a) personal medical account type (HSA or HRA), b) employer contribution to personal medical accounts (above or below median contribution), and c) effective deductible (above or below median). In these additional models, the coefficients of interest are the treatment interaction terms and the total effects by type of treatment.

We also estimate the impact of CDHP enrollment, rather than offer, under the additional assumption of no spillovers of the CDHPs onto employees at CDHP offering firms who are not CDHP enrollees. To estimate this local average treatment effect (LATE), researchers often turn to an instrumental variables approach where treatment offer is an instrument for enrollment (Finkelstein et al. 2012). This two stage least squares approach is not feasible in our setting as we are estimating a non-linear model with three binary endogenous regressors and three binary instruments. Instead, we estimate the LATE by dividing the DiD estimates (Table 3) by the average take-up rates in the relevant firm year.

5.2. A Machine Learning Approach to Achieve Consistent Within Firm Composition Over Time

The firms in our study are large, most with multiple locations, and many with a significant proportion of lower income employees. Due to employee turnover and firm acquisitions, the employees and dependents of each firm change over time. As a result, our estimates may capture both the effects of CDHPs and changes in composition if the firm fixed effects and a simple specification of the time varying person-year level covariates are insufficient to address year-to-year fluctuations in the composition of employees. For example, it is possible that younger employees leave control firms at a higher rate than treatment firms. CDHPs' lower premiums may be particularly attractive to younger employees who on average have lower incomes and relatively little need for health care, which may lead to higher retention of these employees. If this is the case then we would not be able to distinguish the effects of differential retention of young employees from the true CDHP effect.

One approach for dealing with attrition and other changes in composition would be to investigate a continuously enrolled cohort in CDHP and control firms over the entire sample period. However, a continuously enrolled population may differ in observed and unobservable ways from those with incomplete enrollment over the sample period; this is particularly true given the long time span we examine. Instead we use a novel approach that reduces potential bias in the difference-in-differences estimates by keeping a close approximation to the (weighted) joint distribution of covariates of individuals receiving coverage through each firm constant over time. Because the procedure balances the full joint distribution, it removes reliance on the particular specification of covariates in the outcome model. The balancing weights stand in for the set of potentially complex non-linear interactions required to obtain balance in the joint distribution of covariates over time for every firm. In this sense, we ensure that the firm fixed effects 'work' in that they are representing sets of individuals whose joint distribution of

observed covariates is time-invariant. To our knowledge, this is the first paper to identify and control for this source of potential bias in a more systematic way than firm fixed effects and/or simply specified regression control for individual level observed covariates.

Our approach proceeds as follows. For each treatment firm we assign weights to those present in each non-baseline year such that, after weighting, the joint distribution of covariates for each non-baseline year matches that observed in the firm's baseline year (year prior to CDHP offer). Similarly, we assign weights to the individuals in each year in each control firm to match the joint distribution of covariates for individuals in that firm observed in 2005 (the middle of the study period as there is no baseline year defined by initiation of treatment). The models used to create the weights were estimated separately for each employer and each pair of years (the baseline and a single non-baseline year). As a result, the cost growth we observe after weighting removes the effects of changes in observed socio-demographics, zip code characteristics, or geographic location of employees and their dependents from year to year. This makes the CDHP effect 'doubly-robust' as if either the models predicting year for each pair of years for each firm or the cost model specification of the individual level covariates is correct, and the observed covariates are sufficient, then the treatment effect estimate will be unbiased (Kang and Schafer 2007).

To construct the weights, we use a statistical machine learning methodology, generalized boosted regression (implemented in a streamlined version of the R package TWANG⁶). Using generalized boosted regression in this context is preferable to the more commonly estimated Logit model for two reasons. First, generalized boosted regression fits highly flexible models

⁶ <http://cran.r-project.org/web/packages/twang/vignettes/twang.pdf>.

incorporating complex interactions of the covariates, leading to weights that produce better balance on the full joint distribution rather than just the marginal for each variable individually (McCaffrey et al., 2004). Second, this method produces a distribution of weights that are less extreme and hence less able to cause variance inflation, a well-known problem with the Logit (Lee et al. 2011). Generalized boosted regression is an iterative procedure where the stopping rule is based on the degree of success in creating covariate balance. At each iteration, the procedure searches over simple regression tree models, each fit to the residuals from the model in the prior iteration, and selects the one that best improves the fit of the model to the data. Regression trees describe the relationship between a dependent variable and covariates by partitioning observations into regions defined by values of the covariates. That new regression tree adjustment is added to the model at the prior iteration to create the new model. The final model is the additive sum of the collection of simple regression tree models each one providing the best improvement at its iteration (McCaffrey et al. 2004). There are different metrics used to establish balance for the stopping rule and while the standardized difference in means is common, we chose to use the KS statistic, a measure of the extent of differences in the cumulative distribution functions of two variables or sets of variables, as the criteria for stopping the process in order to balance the full distributions of covariates, not just their means. The methodology also allowed us to identify situations in which changes over time within firm were too fundamental to be balanced and we dropped such firm-years from the analysis based on there being any covariate, which had a KS statistic greater than 0.05.⁷

⁷ See table A1 for the effect of this and all other data restrictions on the resulting analytic sample

We achieved excellent balance on covariates⁸ within firm over time⁹. Appendix Figures A1-3 through A1-6 plot the distribution of the balance statistics for control and treatment firms that were retained. As the figures demonstrate, balance was good to start with in many instances and the methodology greatly reduced the differences in standardized effect sizes and KS statistics when they were present. We apply these weights in all descriptive statistics and models.

5.3. Sensitivity of results to balanced panel of firms

As described above, the firms in our data offer CDHPs in different calendar years while the data window is fixed in calendar time. In the models, we include the largest number of pre- and post-CDHP years available in the data. This means we estimate the post-offer effects using an unbalanced panel (i.e., different firms identify each year's effect). To be more specific, the estimated first-year effect is identified by 24 firms, the second year effect by 20 of those firms, and the third year effect by 13 of those firms. Thus in our main specification, any differences in the second or third year effect relative to the first year effect could be the result of changes in CDHP effects over time and/or changes in firm composition. In order to distinguish between these sources of change, we re-estimate the main models twice, first restricting to only those firms with at least two years of data following CDHP offer and second restricting to only those firms with at least 3 years of data following CDHP offer.

6. Description of the Sample and the Treatment

⁸ We balanced on gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), % White (zip), 9 census region dummies and a dummy for being in an MSA.

⁹ The generalized boosted regression procedure has tuning parameters: we set the maximum regression tree level at 2, the shrinkage at 0.01, and the bagging fraction at 1.

Table 1 presents descriptive statistics for the weighted sample, stratified by treatment status. The roughly five million treatment individuals look fairly similar on observables to the eight million control individuals. The enrollees at CDHP offering firms are slightly older, more likely to be without children, and reside in zip codes with a lower median household income.

Figure 1 demonstrates that the weighted cost growth in control firms is very similar to that seen in the private insurance market nationally over this time period (CMS, 2013), suggesting strong external validity. Figure 2 shows that the cost growth for control firms is also similar to that for the subset of treatment firms for which we have two years of data prior to the CDHP offer. This supports the key assumption of the difference-in-difference identification strategy. While not an assumption required for our identification strategy, the average cost *levels* in the baseline year are also similar: \$3,051 (SE = \$472) for control firms and \$3,284 (SE = \$518) for treatment firms.¹⁰

Figure 3 and Table 2 summarize the nature of the CDHP offer. Averaged across all employees and dependents of CDHP offering firms, offering a CDHP is associated with a 2-3 fold increase in average deductibles (Figure 3). In contrast, average deductibles increase smoothly and quite gradually in the control firms. Average deductibles increase from the 1st to 3rd year post offer among the treatment firms because of increases in the proportion of employees enrolling in CDHPs (Table 2).

The remaining columns of Table 2 summarize the key features of the CDHPs being offered in each treatment year. In the first two years of offer, the mix of HSAs and HRAs is approximately 50/50, average employer account contributions are approximately \$150 and

¹⁰ Standard errors are calculated as the standard deviation of the firm means divided by the square root of the number of firms.

average effective deductibles are close to \$1,000. We have data for 13 of the 24 treatment firms for a third year post offer. These firms have a higher proportion of HRAs than HSAs and more generous employer account contributions. For additional information on changes in firm composition over time see Appendix Table A3-1.

7. Results

7.1. Intent-to-treat effect

Relative to control firms and the pre-offer years, firms offering a CDHP had an estimated 6.6, 4.3, 3.4 percent lower annual spending in the first three years respectively ($p < 0.05$ for each difference) (Table 3, column 1). In addition, the CDHP second pre-year trend coefficient is insignificant, supporting the assumption of a parallel trend in the absence of CDHP offer. In section 6.4 below, we discuss whether the post CDHP offer impact is changing over time controlling for differential composition of treatment firms over time.

The CDHP effect varies considerably across spending category. Relative to non-offering firms, annualized spending growth on pharmaceuticals is 5 to 9.5 percentage points lower in the three years after firms offer CDHPs ($p < 0.01$) and spending growth on outpatient care is 3.0 to 6.8 percentage points lower in the first three years though the estimate loses statistical significance in the third year ($p < 0.05$ in first two years). In contrast, for inpatient cost growth, we have only marginally statistically significant evidence of lower spending relative to non-offering firms in the first two years of CDHP offer ($p < 0.10$) while the third year estimate is non-significant and very close to zero. Finally, we do not detect any differences in cost growth for emergency department (ED) care in any of the first three years of CDHP offer although, due to high variance in ED spending, estimates are imprecise.

7.2. Local average treatment effect

The results in Table 4 indicate how large the impact on cost growth would be for those *enrolled* in CDHPs if they were the only ones contributing to the reduced cost growth, i.e. if there were no spillovers to employees offered but not enrolled in CDHPs. These estimates range from 7 to 22 percentage point reductions in health care cost growth in the first three years of enrollment in a CDHP ($p < .005$, standard errors range from 3 to 9 percentage points).

In columns three and four (Table 4), we estimate an additional intent to treat model including only the five firms who have a high CDHP take-up.¹¹ For these firms any contributions of spillovers to CDHP offer effects would be limited. As expected, the intent-to-treat estimates are larger, ranging from 9 to 13 percentage point reductions in cost growth ($p < 0.01$). However, the local average treatment effect estimates are similar to those for the full sample (none of the effects are significantly different from the full sample estimates). This finding supports the hypothesis that the reductions in cost growth may be attributable primarily to CDHP enrollees rather than spillovers to those enrolled in other plans.

7.3. Differential Impact of CDHP by Degree of Cost Sharing Incentives

In Table 5, we consider whether CDHP design features lead to differential impact on spending. Each row displays the estimated first, second, and third year CDHP offer effects stratified on whether an HRA or HSA was offered, employer account contribution, and effective deductible. These are calculated using post-estimation from models with CDHP offer interaction terms (see Appendix Tables A2-4, A2-5, and A2-6 for those results). There are no statistically

¹¹ Five firms met the high take-up criteria of being above the 75thtile of take-up in all available post years. The mean take-up rates in these firms were 90%, 92%, and 95% in the first, second, and third post years, respectively.

significant differences in CDHP effects across any of these three specifications. However, the pattern of results suggests greater cost reductions that last longer for firms offering CDHPs with HSAs (rather than HRAs), with lower employer contributions to personal medical accounts, and with higher effective deductibles.

7.4. Testing sensitivity to balanced firm panel

The models in which we restrict the treatment group to only include firms with at least 2 years of data post CDHP offer or with at least 3 years of data post CDHP offer are shown in Appendix Tables A3-2 and A3-3. When we restrict to firms with at least 2 years of data post CDHP offer, the year 1 post-offer effects on total costs and costs of each component of health care are very similar to the full sample results (Table A3-2). This is not surprising as only four firms are lost. The 20 firms with at least 2 years of post CDHP offer data are similar to the full sample of 24 firms in CDHP take-up, HRA/HSA mix, and employer contribution to personal medical accounts (Table A3-1). In contrast, the impact in the first and second years of CDHP offer for the 13 firms with three years of data post offer are smaller (closer to zero) and less often statistically significant compared with the full sample of 24 firms. Compared to the full sample, the 13 firms with 3 years of post-CDHP offer data were more likely to pair the plans with HRAs versus HSAs and to give greater employer account contributions.

While the point estimates for Years 1, 2, and 3 post-offer are monotonically decreasing with time since offer, the change in firm composition just discussed complicates any interpretation of this pattern. With the full sample (Table 3), we cannot reject that the CDHP offer effects are the same in each of the three years post offer ($0.09 < p < 0.34$). This is because the third year estimates are available for only 13 out of 24 CDHP firms. These 13 firms offered CDHPs with

systematically weaker incentives than those first offered later in the study period. Thus the change in point estimates overtime could reflect this change in firm composition.

However, when we restrict the sample to just a balanced panel of the 13 firms, we again find that the CDHP effect point estimates monotonically decrease overtime. Again the differences between the years are not statistically significant. Overall based on this analysis we conclude that there is weak evidence that the cost saving effects of CDHPs diminish overtime and moderate evidence that they do not increase over time.

8. Summary and Discussion

With health care costs continuing to grow faster than GDP (CMS 2014; BEA 2014) it is critical to understand the effectiveness of cost-reduction strategies. Prior literature has established that CDHPs reduce spending in the short term. However, the longer term impacts are less clear. There has been concern that CDHP enrollees will decrease their use of necessary care and this will result in increased spending in the long term due to greater complications.

This study substantially adds to our knowledge on the long term cost impacts of CDHPs. We estimated spending trends for three years across over 13 million people across the country in an analysis estimating CDHP impacts without the threat of individual level selection bias. We find that health care cost growth among firms *offering* a CDHP is significantly lower in each of the first three years after offer. This result suggests that, at least at large employers, the impact of CDHPs persists and is not just a one-time reduction in spending. However, an important caveat is that the decrease in spending may be smaller in year 3 compared to year 1 post-offer. Recognizing that the differences are not statistically significant, these results are suggestive and consistent with a decreasing impact of CDHPs over time.

The decreases in total spending growth observed are primarily due to reductions in spending on outpatient care and pharmaceuticals. In contrast, by the third year there are no differences in either emergency department or inpatient spending. There are several potential explanations for this differential impact depending on whether reductions in costs are achieved through price shopping, switching to higher-value treatment options, or blanket reductions in care.

Pharmaceutical spending is ideally suited for learning over time as chronic medications are purchased regularly and price information is fairly accessible (Huckfeldt et al. 2015). Also, generic drugs, where available, provide a clearer signal regarding value than most treatment options. Some patients may also believe that taking their medications less regularly has little health consequence, although research has shown that cost-sharing induced reductions in pharmaceutical use can lead to increased hospitalizations (Chandra et al. 2010). In contrast, emergency department care and inpatient care may be less amenable to any of the three mechanisms for reducing costs. It is difficult to obtain price information and in many instances the care is emergent making it impossible to shop for care. In addition, the incentives to reduce spending might be limited as the cost of one inpatient episode will typically be greater than the deductible. Outpatient care is intermediate between these two extremes. Outpatient physician visits tend to be repeated more than inpatient care but less than pharmaceutical purchases, perceptions of the harms of reducing care are likely to be similarly intermediate, and price and quality information is difficult to obtain.

Our results on the differential impact of CDHP design features on spending are not statistically significant. However, our estimates are consistent with what theory would suggest: CDHPs with larger financial incentives are associated with greater and more long lasting reductions in spending than CDHPs with smaller financial incentives. As families increasingly

enroll in CDHPs with a variety of design features, further research is needed to determine which structures are most beneficial.

The magnitude of the first year effects we observe are consistent with most prior research investigating the short-term effect of CDHPs. Similar to LoSasso et al. (2010), Charlton et al. (2011), and Fronstin et al. (2013) and contrary to Borah et al. (2011), we find reductions in health care cost growth beyond the first year of CDHP offer. Contrary to Wharam et al. (2011, 2013), Kozhimannil et al. (2013), and Reddy et al. (2014), we do not detect any reductions in emergency department spending over three years¹². Our findings on inpatient spending are consistent with Wharam et al. (2011, 2013). We find reduced spending in the first two of three years post offer and they detect reductions in inpatient utilization for the first of two years. In contrast, in their Harvard case study, Kozhimannil et al. (2013) and Reddy et al. (2014), find increases in inpatient use in the second year. The differences between our study and these particular prior studies may be due to their use of HMO rather than low deductible PPO enrollees as the comparison group, individual level selection in the Harvard studies or the difference in firm sizes for the Wharam studies.

Our final contribution is methodological, the novel application of a machine learning technique to mitigate potential violations of the identification assumption (that treatment units would have had the same trend in outcomes over time as controls if they had not offered the treatment). This procedure reduces sensitivity of results to the regression specification for covariates in outcome models, strengthens the effectiveness of fixed effects (such as firm fixed

¹² Although several of these studies are on ED use, not costs.

effects) in reducing potential bias in difference in difference designs, and is relatively accessible for economists without a background in machine learning.

The results presented here are limited to large employers and therefore may not extend to Medicaid beneficiaries, the individual or small group market, or to the health insurance exchanges where, on average, deductibles and out of pocket maximums are higher and/or enrollees have fewer financial resources. While the firms in this study were specifically selected to have lower income employees, all families had at least one adult working full time with benefits so they are typically better off than families not offered employer sponsored insurance.

In summary, in the first large multi-employer study to investigate long term CDHP spending impacts we find reductions in health care cost growth in all three years post CDHP offer and do not detect increases in any component of health care spending. These findings do not support either the concern that decreases in spending will be a one-time occurrence or that short-term decreases in spending with a CDHP will result in increases in spending in the long term due to complications of forgone care. We cannot rule out either of these concerns developing over an even longer time frame.

References

- Americas Health Insurance Plans. 2014. "An Analysis of Health Savings Account Balances, Contributions, and Withdrawals in 2012." Center for Policy and Research.
- Borah, B., ME Burns, and N. D. Shah. 2011. "Assessing the Impact of High Deductible Health Plans on Health-Care Utilization and Costs: A Changes-in-Changes Approach." *Health Economics* 20(9): 1025-42
- Bundorf, K., 2012. "Consumer Directed Health Plans: Do They Deliver." The Robert Wood Johnson Foundation Research Synthesis Report No. 24
- Buntin, M., A. Haviland, R. McDevitt, and N. Sood. 2011. "Health Care Spending and Preventive Care in High-Deductible and Consumer-Directed Health Plans." *American Journal of Managed Care* 17(3): 222-30
- Chandra, A., J. Gruber, and R. McKnight. 2010. "Patient Cost-Sharing and Hospitalization Offsets in the Elderly." *American Economic Review* 100(1): 193-213.
- Charlton, M.E., T. Levy, R.R. High, J.E. Schneider, and J.M. Brooks 2011. "Effects of Health Savings Account Eligible Plans on Utilization and Expenditures." *American Journal of Managed Care* 17(1): 79-86.
- Chen, S., RA. Levin, and JA. Gartner. 2010. "Medication Adherence and Enrollment in a Consumer-Driven Health Plan." *American Journal of Managed Care* 16(2) 43-50.
- CMS. 2013. "Table 1 National Health Expenditures; Aggregate and Per Capita Amounts, Annual Change and Percent Distribution: Selected Calendar Years 1960-2012." Centers for Medicare and Medicaid Services. <http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/tables.pdf>
- Cutler, D. and N.R. Sahni 2013. "If Slow Rate of Health Care Spending Growth Persists." *Health Affairs* 32(5): 841-50.
- Finkelstein, A., S. Taubman, B. Wright, M Bernstein, J. Gruber, J. Newhouse, H. Allen, K. Baicker, and The Oregon Health Study Group. 2012. "The Oregon Health Insurance Experiment: Evidence from the First Year." *The Quarterly Journal of Economics* 127(3): 1057-1106.
- Fronstin, P. (2014). "Health Savings Accounts and Health Reimbursement Arrangements: Assets, Account Balances, and Rollovers, 2006-2013." *EBRI Issue Brief* no. 395.
- Fronstin, P., M.-J. Sepulveda, and M.C. Robebuck. 2013. "Medical Utilization and Adherence in a Health Savings Account-Eligible Plan." *American Journal of Managed Care* 19(2): 400-07.

Haviland, A., M.S. Marquis, R.D. McDevitt, and N. Sood. 2012. "Growth of Consumer Directed Health Plans to One-Half of All Employer-Sponsored Insurance Could Save \$57 billion Annually." *Health Affairs* 31(5): 1009-15.

Haviland, A., N. Sood, R. McDevitt, and S.M. Marquis 2011. "How Do Consumer Directed Health Plans Affect Vulnerable Populations." *Forum for Health Economics & Policy*. 14(2) (Online): 1558-9544.

Huckfeldt, P., A. Haviland, A. Mehrotra, Z. Wagner, and N. Sood, 2015. "Patient Responses to Incentives in Consumer-Directed Health Plans: Evidence from Pharmaceuticals. Working Paper.

Kang, J., Schafer, J. 2007. "Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean From Incomplete Data." *Statistical Science*. 22(4): 523-39.

KFF, HRET. 2014. "Employer Health Benefits: 2013 Summary of Findings." The Kaiser Family Foundation and Health Research and Education Trust

Kozhimannil, K. B., M. R. Law, C. Baluer-Patterson, F. Zhang, J.F. Wharam. 2013. "The Impact of High Deductible Health Plans on Men and Women: An Analysis of Emergency Department Care." *Medical Care* 51(8): 639-45.

Lee, B.K., J. Lessler, and E.A. Stuart. 2011. "Weight Trimming and Propensity Score Weighting." *PLoS ONE* 6(3): e18174 doi: [10.1371/journal.pone.0018174](https://doi.org/10.1371/journal.pone.0018174).

Lo Sasso, A., L.A. Helmchen, and R.Kaestner. 2010. "The Effects of Consumer-Directed Health Plans on Health Care Spending." *The Journal of Risk and Insurance* 77(1): 85-103.

Manning, W. G., J. P. Newhouse., N. Duan, E. Keeler, B. Benjamin, A. Leibowitz, M.S. Marquis, J. Zwanziger. 1987. "Health Insurance and the Demand for Medical Care." *The American Economic Review* 77(3): 251-77.

McCaffrey, D.F., G. Ridgeway, and A.R. Morral. 2004. "Propensity Score Estimation With Boosted Regression for Evaluating Causal Effects in Observational Studies." *Psychological Methods* 9(4) 403-25.

McDevitt, R., A. Haviland, L. Laudenberger, R. Lore, M. Eisenberg, N. Sood. 2014. "Risk Selection into Consumer-Directed Health Plans: An Analysis of Family Choices within Large Employers." *Health Services Research*. 49(2): 609-27.

Newhouse, JP. *Free for All? Lessons from the RAND Health Insurance Experiment*. Cambridge, MA: Harvard University Press, 1993.

Parente, S., Feldman R. 2008. "Effects of a Consumer-Driven Health Plan on Pharmaceutical Spending and Utilization." *Health Services Research*. 43(5).

Reddy, S., D. Ross-Degnan, A. Zalaysky, S.B. Soumerai, and J.F. Wharam. 2014. "Impact of a High-deductible Health Plan on Outpatient Visits and Associated Diagnostic Tests." *Medical Care* 52(1): 86-92.

Reed, M., V. Fung, M. Price, R. Brand, N. Benedetti, S.F. Derose, J.P. Newhouse, and J. Hsu. 2009. "High-Deductible Health Insurance Plans: Effort to Sharpen a Blunt Instrument." *Health Affairs* 28(4): 1145-1154. doi: 10.1377/hlthaff.28.4.1145

Reed, M.E., I. Graetz, V. Fung, J.P. Newhouse, and J. Hsu. 2012. "In Consumer-Directed Health Plans, A Majority of Patients were Unaware of Free or Low-Cost Preventive Care." *Health Affairs* 31(12): 2641-2648. doi: 10.1377/hlthaff.2012.0059

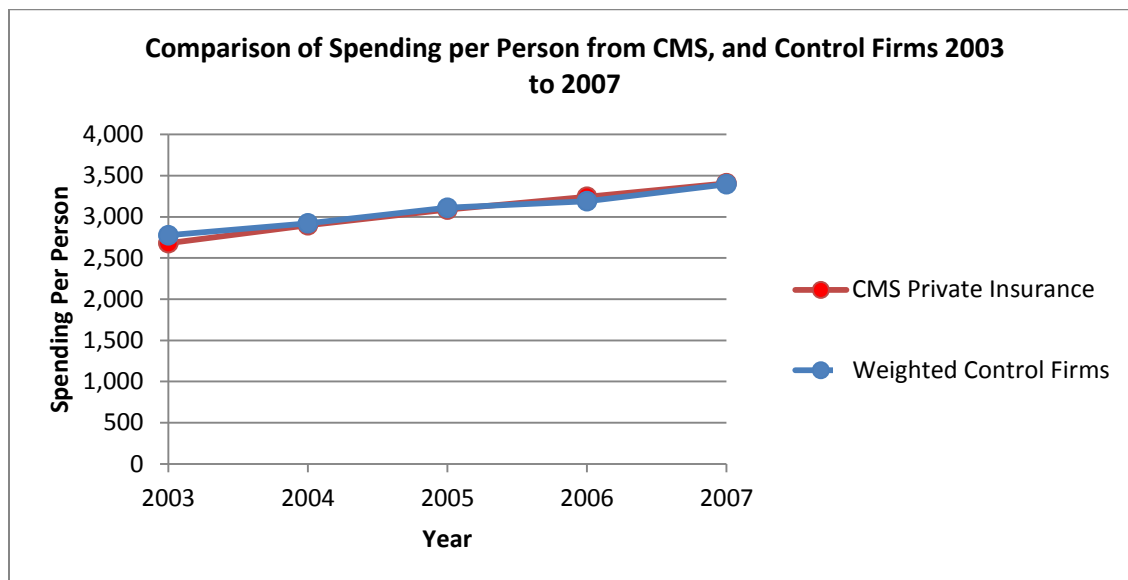
Whaley, C., J. S. Chafen, S. Pinkard, G. Kellerman, D. Bravata, R. Kocher, N. Sood. 2014. "Association Between Availability of Health Service Prices and Payments for These Services." *Journal of the American Medical Association* 312(16) 1670-76.

Wharam, J.F., B.E. Landon, F. Zhang, S.B. Soumerai, D. Ross-Degnan. 2011. "High-Deductible Insurance: Two-Year Emergency Department and Hospital Use."

Wharam, J. F., F. Zhang, B E. Landon, S. B. Soumerai, and D. Ross-Degnan. 2013. "Low-Socioeconomic-Status Enrollees In High-Deductible Plans Reduced High-Severity Emergency Care." *Health Affairs* 32(8): 1398-1406.

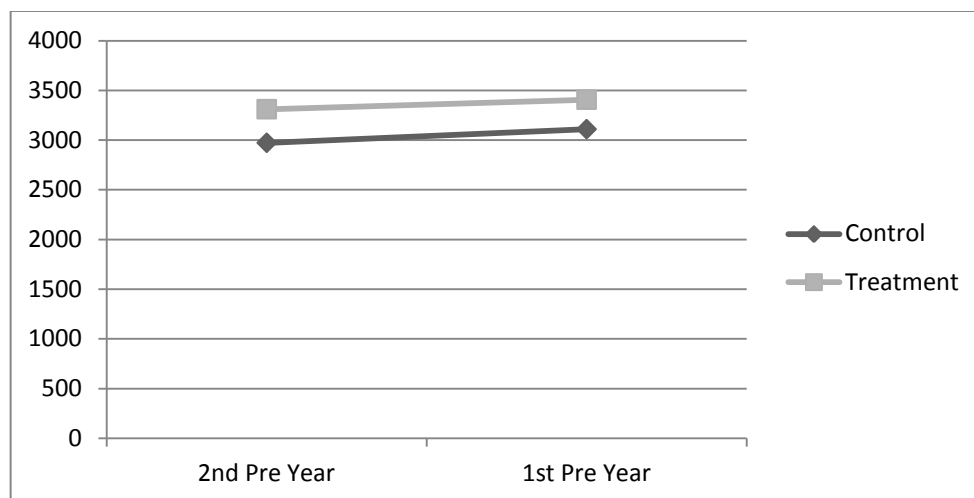
Figures

Figure 1: Control Firm Cost Trends Compared to National Data for Private Insurance



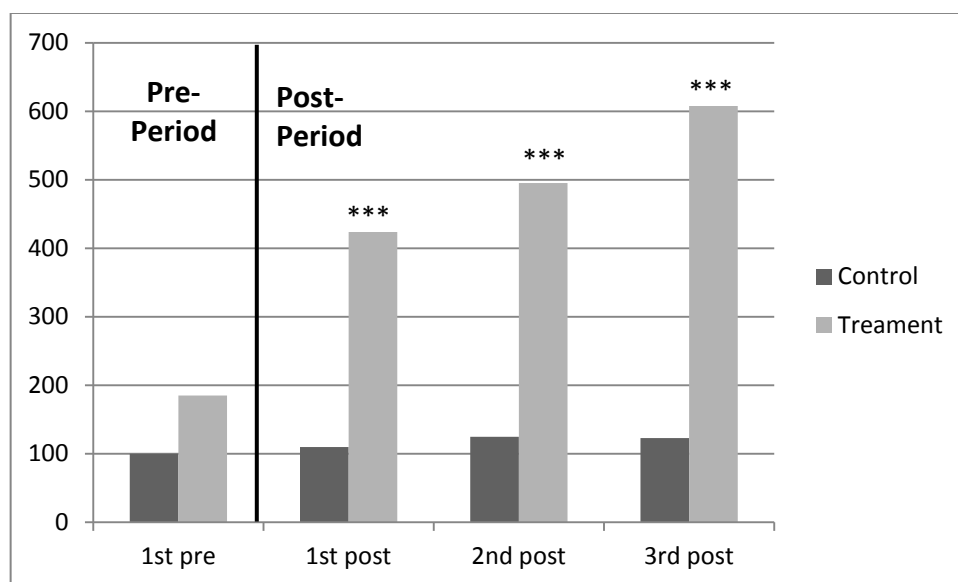
Notes: Source for CMS data is available at <http://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/tables.pdf>. This is the NHE Tables PDF. The specific table is Table 21 Medicare and Private Health Insurance; Per Enrollee Expenditures and Annual Percent Change, Calendar Years 1969-2011.

Figure 2: Comparison of Cost Trends Prior to CDHP Offer



Notes for 2a: The treatment line is mean total costs across the subset of treatment firms with 2 pre CDHP offer years available. The control line is the weighted average of average annual control firm costs, weighted by the distribution of calendar years within a CDHP offer year for the corresponding treatment firms. For example, the 1st Pre Year for the treatment firms is a combination of calendar years: 8% 2003, 68% 2004, and 24% 2005. To obtain the control mean for this '1st Pre Year', we take a weighted average of the control firm costs in 2003, 2004, and 2005 with weights of 0.08, 0.68, and 0.24, respectively.

Figure 3: Average Individual Deductible by Treatment Year



Notes: The treatment bars are the mean individual deductible for all employees at treatment firms. For those enrolled in CDHPs at treatment firms, the mean deductibles are \$1426, \$1492, and \$1396 for the first, second, and third post year, respectively. Take up rates were 19.4%, 21.4%, and 34.6% in the first, second, and third post years. The control bars are the weighted average of average annual control firm individual deductibles, weighted by the distribution of calendar years within a CDHP offer year for the corresponding treatment firms (for example of weighted average calculation, see notes to figure 2). In a model including calendar year fixed effects and with standard errors clustered by firm, differences in deductibles between the treatment and control firms were statistically insignificant in the first pre-offer year ($p = 0.20$) and statistically significant in all three years post CHDP offer ($p < 0.001$).

Tables

Table 1: Description of the sample at the person-year level

| VARIABLES | Controls (N=10,100,357) | | Treatment (N=4,908,955) | |
|------------------------|----------------------------|--------|----------------------------|--------|
| | Mean | SD | Mean | SD |
| Demographics | | | | |
| Age | 30.70 | 17.50 | 32.69 | 17.98 |
| Female | 0.504 | 0.500 | 0.501 | 0.500 |
| Single Adult | 0.133 | 0.340 | 0.135 | 0.341 |
| Married, No Kids | 0.137 | 0.344 | 0.146 | 0.353 |
| Adult Male with Kids | 0.0466 | 0.211 | 0.0498 | 0.217 |
| Adult Female with Kids | 0.0642 | 0.245 | 0.0516 | 0.221 |
| Married, with Kids | 0.617 | 0.486 | 0.6023 | 0.490 |
| Median HH Income | 52,120 | 18,915 | 51,373 | 17,249 |
| % Unemployed | 0.0470 | 0.0278 | 0.0471 | 0.0286 |
| % College | 0.282 | 0.162 | 0.255 | 0.148 |
| % High School | 0.565 | 0.112 | 0.588 | 0.101 |
| % Hispanic | 0.0793 | 0.132 | 0.0670 | 0.119 |
| % Black | 0.0950 | 0.156 | 0.102 | 0.182 |
| % White | 0.776 | 0.216 | 0.787 | 0.227 |
| New England | 0.0544 | 0.227 | 0.0228 | 0.149 |
| Middle Atlantic | 0.109 | 0.311 | 0.0515 | 0.221 |
| East North Central | 0.285 | 0.451 | 0.428 | 0.495 |
| West North Central | 0.0535 | 0.225 | 0.0890 | 0.285 |
| South Atlantic | 0.187 | 0.390 | 0.106 | 0.307 |
| East South Central | 0.0651 | 0.247 | 0.0526 | 0.223 |
| West South Central | 0.130 | 0.336 | 0.127 | 0.333 |
| Mountain | 0.0525 | 0.223 | 0.0223 | 0.148 |
| Pacific | 0.0620 | 0.241 | 0.0416 | 0.200 |
| Unknown Location | 0.00285 | 0.0533 | 0.0602 | 0.238 |
| d(MSA) | 0.875 | 0.331 | 0.819 | 0.385 |

Notes: Weighted means and standard deviations for the entire sample, all years (2003-2007) pooled. Statistics are calculated based on the weights described in section 4.2.

Table 2: Summary of Take-Up and Benefit Design Features of CDHPs Offered by Treatment Firms

| | All Treatment Firm Individuals | | Treatment Firm Individuals Enrolled in CDHPs | | | |
|---------------|-----------------------------------|---------|---|-------|-------------------------|-------------------------|
| | N | Take-up | HRA | HSA | Account Contribution | Effective Deductible |
| 1st Post Year | 1,269,190 | 0.194 | 0.592 | 0.408 | 157.2 (267.7) | 991.8 (359.9) |
| 2nd Post Year | 1,002,859 | 0.214 | 0.426 | 0.574 | 154.2 (290.2) | 988.6 (414.2) |
| 3rd Post Year | 480,535 | 0.346 | 0.724 | 0.276 | 262.2 (318.5) | 821.7 (306.9) |

Notes: Summary for all available treatment firm years. Effective deductible is actual individual deductible minus the account contribution for an individual.

Table 3: GLM Models of CDHP Offer on Costs

| VARIABLES | (1) Total Costs | (2) Outpatient Costs | (3) Drug Costs | (4) Inpatient Costs | (5) ED Costs |
|-----------------------------------|------------------------|-------------------------|------------------------|------------------------|---------------------|
| Treat*(2 nd Pre Year) | -0.000640 (0.0160) | -0.00726 (0.0144) | 0.0128 (0.0192) | 0.00945 (0.0236) | -0.0369 (0.0421) |
| Treat*(1 st Post Year) | -0.0655** (0.0256) | -0.0675** (0.0264) | -0.0612*** (0.0173) | -0.0638* (0.0331) | -0.0337 (0.0833) |
| Treat*(2 nd Post Year) | -0.0434*** (0.0145) | -0.0405** (0.0204) | -0.0536*** (0.0151) | -0.0536* (0.0322) | 0.0620 (0.115) |
| Treat*(3 rd Post Year) | -0.0339** (0.0142) | -0.0303 (0.0194) | -0.0949*** (0.0236) | 0.00197 (0.0322) | 0.0412 (0.0673) |
| Proportion of Total Costs | 1.00 | 0.53 | 0.23 | 0.20 | 0.04 |
| Observations | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 |
| Number of Firms | 54 | 54 | 54 | 54 | 54 |

Notes: Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. We include a full set of firm fixed effects and year fixed effects. Controls include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and an indicator for being in an MSA.

Table 4: Intent-to-treat and Treatment on the Treated Results (Assuming No Spillovers) for Total Costs

| | (1) | (2) | (3) | (4) |
|-----------------------------------|------------------------|------------------------|------------------------|------------------------|
| | All Firms | | High Take-up Firms | |
| | ITT | LATE | ITT | LATE |
| Treat*(1 st Post Year) | -0.0655** (0.0256) | -0.2226** (0.087) | -0.1300*** (0.0236) | -0.1452*** (0.0264) |
| Treat*(2 nd Post Year) | -0.0434*** (0.0145) | -0.1353*** (0.0452) | -0.119*** (0.0200) | -0.1300*** (0.0219) |
| Treat*(3 rd Post Year) | -0.0339** (0.0142) | -0.0691** (0.0289) | -0.0852*** (0.0208) | -0.0899*** (0.0220) |

Notes: Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Local average treatment estimates (LATE) are calculated by taking the coefficient given in table 2a and dividing by average CDHP take up in that treatment year. Standard errors were derived using the delta method and the Stata command *nlcom*. Five firms met the high takeup criteria of being above the 75thtile of take-up in all available post years. The mean take-up rates in these firms were 90%, 92%, and 95% in the first, second, and third post years, respectively.

Table 5: Treatment Effects, Stratified by CDHP Account Features

| | <u>Account Type</u> | |
|----------------------|-----------------------------|--------------------------|
| | HRA | HSA |
| 1 st Post | -0.0517*** (0.0177) | -0.0705** (0.0337) |
| 2 nd Post | -0.0287* (0.0169) | -0.0504** (0.0197) |
| 3 rd Post | -0.0217 (0.0152) | -0.0508*** (0.0183) |
| | <u>Account Contribution</u> | |
| | At <Median Account | At > Median Account |
| 1 st Post | -0.0958* (0.0507) | -0.0343** (0.0157) |
| 2 nd Post | -0.0544** (0.0268) | -0.0353 (0.0251) |
| 3 rd Post | -0.0411 (0.0347) | -0.0239 (0.0149) |
| | <u>Effective Deductible</u> | |
| | At <Median Effective Ded | At > Median Efection Ded |
| 1 st Post | -0.0317* (0.0169) | -0.0949** (0.0477) |
| 2 nd Post | -0.0364 (0.0255) | -0.0522** (0.0233) |
| 3 rd Post | -0.0230 (0.0146) | -0.0463 (0.0301) |

Notes: Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Median account contribution is \$500. Median effective deductible is \$1,250. Full model estimates available in appendix table A2-4 and A2-5. Estimated using the STATA command *margins, eydx(*)*. While different, none of the CDHP effects are significantly different from each other across account type.

Appendices

Appendix A1: Documentation of Preliminary Analysis Steps

Table A1-1: Sample Restrictions

| Restriction Type | Restriction | % of the Sample Dropped |
|-------------------|--|-------------------------|
| Analysis Criteria | Person-years with less than one plan year of continuous enrollment | 34% |
| Analysis Criteria | Person-years in a control firm and enrolled in a CDHP; in a treatment year, enrolled in a CDHP before firm offered; in a control firm with a high deductible; in a treatment firm with a high deductible before firm offered | 0.4% |
| Data Anomalies | Person-years with missing plan information, irregular family structures | 7.8% |
| Data Anomalies | Person-years with very high costs | .00117% |
| Data Anomalies | Firms for which we were unable to achieve adequate within firm balance on covariates over time | 4.4% |

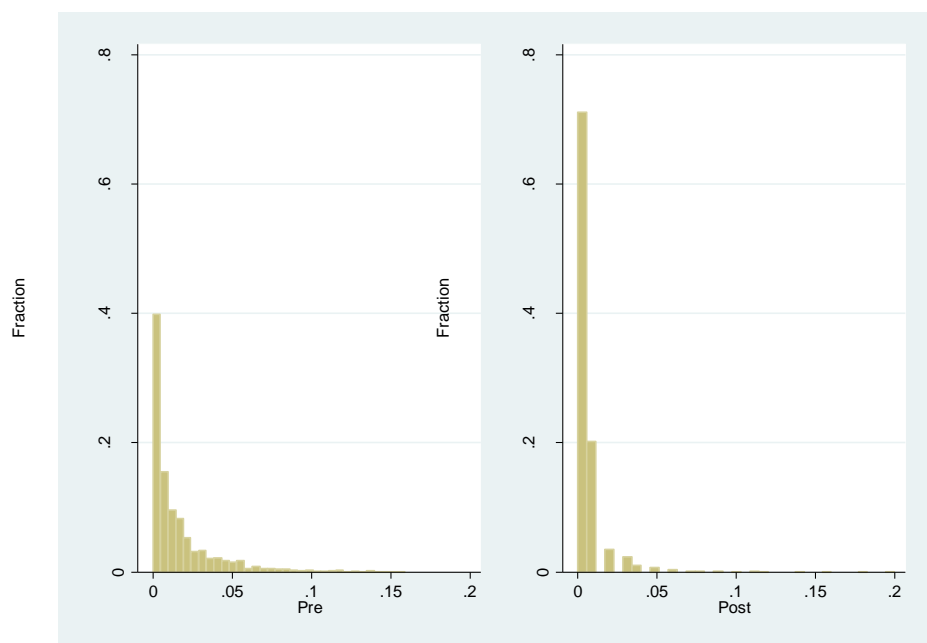
Table A1-2: Fit Statistics From Candidate GLM Models

| | (1) | (2) | (3) |
|----------------|------------|-------------|------------|
| | Gamma-Log | Poisson-Log | Normal-Log |
| Pre + Control† | | | |
| <i>AIC</i> | 17.99 | 7191 | 21.52 |
| <i>BIC</i> | -1.49e+08 | 7.46e+10 | 1.07e+15 |
| Full Sample | | | |
| <i>AIC</i> | 17.40 | 7096 | 20.80 |
| <i>BIC</i> | -1.980e+08 | 9.590e+10 | 1.380e+15 |

† Model selection was carried out using model runs including data for all years for control firms and data only for pre-CDHP offer years for treatment firms. We provide the full sample results for completeness.

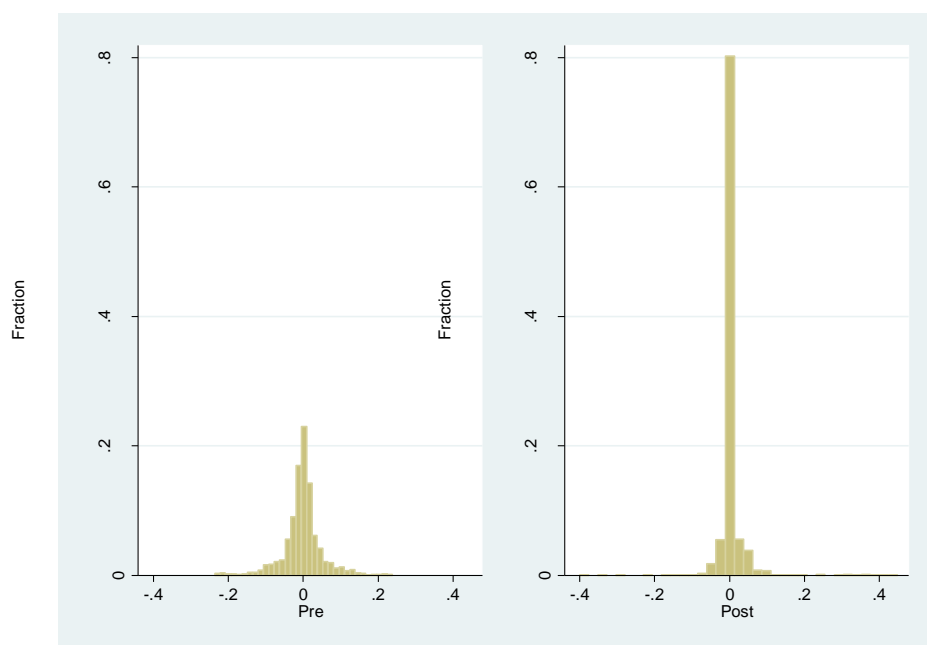
Notes: These fit statistics come from the full total cost model from table 3 estimated with different GLM assumptions.

Figure A1-3: Histogram of KS Statistics, Control Firms, Before and After Weighting



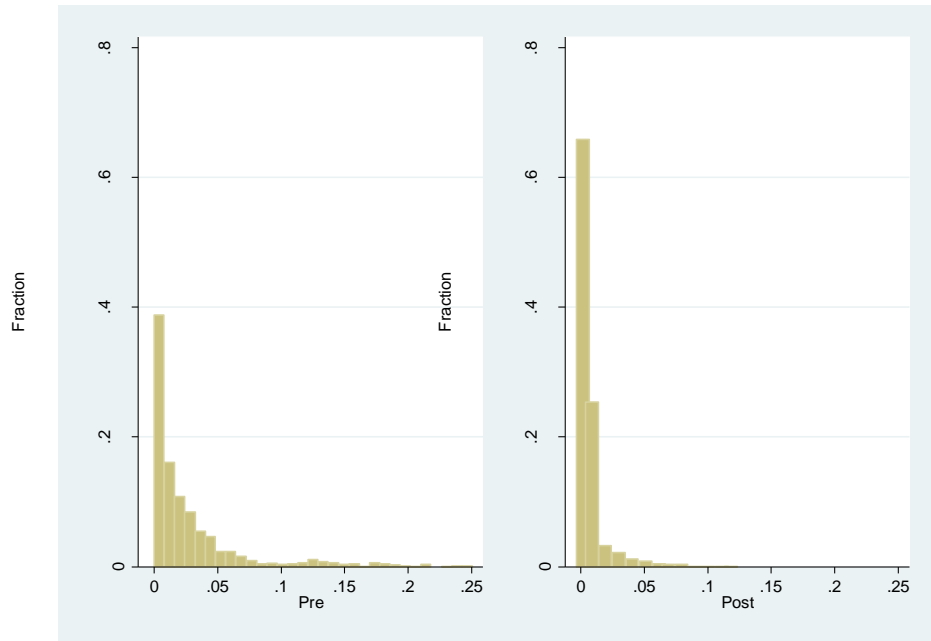
Notes: This figure presents the distribution of KS statistics across all control firms, all years, and all covariates. Pre weighting, the 99%tile was 0.16 and post weighting it was 0.1.

Figure A1-4: Histogram of Standardized Effect Size Differences, Control Firms, Before and After Weighting



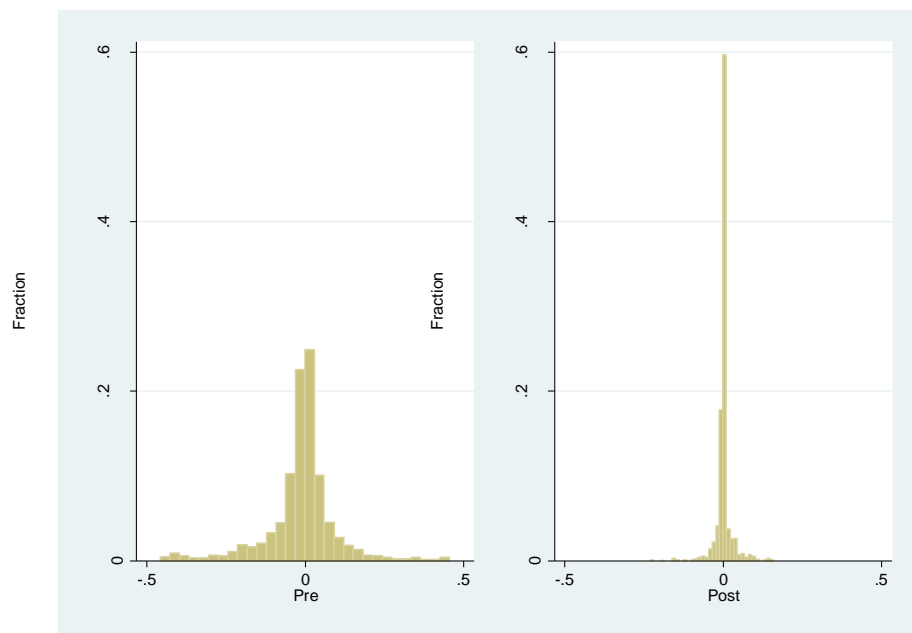
Notes: This figure presents the distribution of KS statistics across all control firms, all years, and all covariates. Pre weighting, the 99%tile was 0.24 and post weighting it was 0.14.

Figure A1-5: Histogram of KS Statistics, Treatment Firms, Before and After Weighting



Notes: This figure presents the distribution of KS statistics across all treatment firms, all years, and all covariates. Pre weighting, the 99%tile was 0.26 and post weighting it was 0.06.

Figure A1-6: Histogram of Standardized Effect Size Differences, Treatment Firms, Before and After Weighting



Notes: This figure presents the distribution of KS statistics across all treatment firms, all years, and all covariates. Pre weighting, the 99%tile was 0.46 and post weighting it was 0.10.

Appendix A2: Additional Results Related to Tables 3 and 4

Table A2-1: Full Model Results from Table 3

| VARIABLES | (1) Total Costs |
|-----------------------------------|---------------------------|
| Treat*(2 nd Pre Year) | -0.000640 (0.0160) |
| Treat*(1 st Post Year) | -0.0655** (0.0256) |
| Treat*(2 nd Post Year) | -0.0434*** (0.0145) |
| Treat*(3 rd Post Year) | -0.0339** (0.0142) |
| Female | 0.281*** (0.0126) |
| Age | 0.00880*** (0.00162) |
| Age Squared | 0.000316*** (2.47e-05) |
| Married, No Kids | 0.120*** (0.0184) |
| Adult Male with Kids | -0.0866*** (0.0231) |
| Adult Female with Kids | 0.0427** (0.0167) |
| Married with Kids | 0.155*** (0.0372) |
| Median HH. Income (Zip) | 4.22e-07 (3.08e-07) |
| % Unemployed (Zip) | 0.165 (0.135) |
| % College (Zip) | -0.183*** (0.0507) |
| % High School (Zip) | -0.226*** (0.0733) |
| % Hispanic (Zip) | -0.0638 (0.0748) |
| % Black (Zip) | -0.0135 (0.0681) |
| % White (Zip) | 0.159** (0.0677) |
| Middle Atlantic | -0.124*** (0.0414) |
| East North Central | -0.0457 (0.0355) |

| | |
|---------------------------|-----------------------|
| West North Central | -0.0485 (0.0396) |
| South Atlantic | -0.00888 (0.0380) |
| East South Central | 0.00494 (0.0427) |
| West South Central | 0.0118 (0.0448) |
| Mountain | -0.00627 (0.0505) |
| Pacific | -0.139** (0.0602) |
| Unknown Location | -0.144** (0.0729) |
| d(MSA) | -0.127*** (0.0321) |
| Middle Atlantic*d(MSA) | 0.171*** (0.0444) |
| East North Central*d(MSA) | 0.138*** (0.0331) |
| West North Central*d(MSA) | 0.130*** (0.0395) |
| South Atlantic*d(MSA) | 0.0936** (0.0381) |
| East South Central*d(MSA) | 0.101** (0.0414) |
| West South Central*d(MSA) | 0.179*** (0.0423) |
| Mountain*d(MSA) | 0.0455 (0.0570) |
| Pacific*d(MSA) | 0.233*** (0.0466) |
| 2004 | 0.0688*** (0.0135) |
| 2005 | 0.144*** (0.0118) |
| 2006 | 0.182*** (0.0140) |
| 2007 | 0.256*** (0.0129) |
| Observations | 13,553,830 |

Table A2-2: Marginal Effects (in \$ Terms)

| VARIABLES | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|-----------------------------------|----------------------|-------------------------|------------------------|-------------------|----------------------|
| Treat*(2 nd Pre Year) | -2.034 (50.80) | -12.27 (24.25) | 6.129 (15.30) | -4.521 (5.164) | 9.309 (13.93) |
| Treat*(1 st Post Year) | -208.2** (81.55) | -114.0** (44.61) | -41.35* (21.48) | -4.129 (10.21) | -44.50*** (12.59) |
| Treat*(2 nd Post Year) | -138.1*** (46.19) | -68.47** (34.52) | -34.73* (20.90) | 7.590 (14.06) | -38.93*** (10.97) |
| Treat*(3 rd Post Year) | -107.7** (45.31) | -51.21 (32.76) | 1.279 (20.86) | 5.051 (8.252) | -68.96*** (17.23) |
| Observations | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Calculated from the estimates in Table 3. Marginal effects estimated using the Stata command *margins, dydx(*)* where these are dollar differences between the reported row and the omitted baseline (1st pre year).

A2-3: Fraction Enrolled in CDHP by Account Feature Groups each Treatment Year

| | 1 st post Year | 2 nd post Year | 3 rd post Year |
|------------------------------------|---------------------------|---------------------------|---------------------------|
| <i>Account Type</i> | | | |
| HRA | 0.238 | 0.230 | 0.307 |
| HSA | 0.176 | 0.202 | 0.562 |
| <i>Account Contribution</i> | | | |
| At <Median Account | 0.188 | 0.247 | 0.759 |
| At >Median Account | 0.202 | 0.181 | 0.239 |
| <i>Effective Deductible</i> | | | |
| At <Median Effective Ded | 0.248 | 0.223 | 0.357 |
| At > Median Effective Ded | 0.143 | 0.203 | 0.279 |

Table A2-4: Cost Models with Account Type Interactions

| VARIABLES | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|-----------------------------------|------------------------|-------------------------|------------------------|---------------------|------------------------|
| Treat*(2 nd Pre Year) | -0.00318 (0.0164) | -0.0102 (0.0152) | 0.00433 (0.0245) | -0.0321 (0.0396) | 0.0150 (0.0195) |
| Treat*(1 st Post Year) | -0.0517*** (0.0177) | -0.0473** (0.0209) | -0.0811** (0.0406) | 0.00628 (0.0535) | -0.0501* (0.0262) |
| Treat*(2 nd Post Year) | -0.0287* (0.0169) | -0.0229 (0.0202) | -0.00948 (0.0385) | -0.0187 (0.0827) | -0.0743*** (0.0260) |
| Treat*(3 rd Post Year) | -0.0217 (0.0152) | -0.0169 (0.0208) | 0.0214 (0.0258) | 0.0543 (0.0788) | -0.104*** (0.0261) |
| Treat*HSA*(1 st Post) | -0.0188 (0.0264) | -0.0272 (0.0415) | 0.0201 (0.0574) | -0.0481 (0.1203) | -0.0135 (0.0321) |
| Treat*HSA*(2 nd Post) | -0.0218 (0.0264) | -0.0258 (0.0366) | -0.0677 (0.0518) | 0.124 (0.180) | 0.0319 (0.0333) |
| Treat*HSA*(3 rd Post) | -0.0291 (0.0223) | -0.0252 (0.0375) | -0.0609 (0.0468) | -0.123 (0.107) | 0.0305 (0.0391) |
| Demographics | Yes | Yes | Yes | Yes | Yes |
| Year Dummies | Yes | Yes | Yes | Yes | Yes |
| Location | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 |
| Number of Firms | 54 | 54 | 54 | 54 | 54 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. Demographics include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and a dummy for being in an MSA.

Table A2-5: Cost Models with Account Contribution Interactions

| | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|---|-----------------------|-------------------------|------------------------|---------------------------------|-----------------------------------|
| Treat*(2 nd Pre Year) | -0.00271 (0.0180) | -0.00809 (0.0164) | 0.00314 (0.0277) | -0.0353 (0.0445) | 0.0118 (0.0200) |
| Treat*(1 st Post Year) | -0.0958* (0.0507) | -0.0844 (0.0523) | -0.103 (0.0635) | -0.167** (0.0764) | -0.0878*** (0.0297) |
| Treat*(2 nd Post Year) | -0.0544** (0.0268) | -0.0420 (0.0385) | -0.112* (0.0610) | 0.137 (0.213) | -0.0514*** (0.0141) |
| Treat*(3 rd Post Year) | -0.0411 (0.0347) | -0.0437 (0.0375) | 0.0169 (0.0795) | 0.0752 (0.123) | -0.114*** (0.0397) |
| Treat*1 st Post Year*Above Median Account Contribution | 0.0615 (0.0530) | 0.0350 (0.0531) | 0.0728 (0.0725) | 0.289* (0.156) | 0.0556* (0.0327) |
| Treat*2 nd Post Year*Above Median Account Contribution | 0.0191 (0.0390) | 0.00121 (0.0449) | 0.110 (0.0745) | -0.145 (0.226) | -0.00659 (0.0298) |
| Treat*3 rd Post Year*Above Median Account Contribution | 0.0173 (0.0389) | 0.0201 (0.0414) | -0.000876 (0.0864) | -0.0143 (0.148) | 0.0283 (0.0469) |
| Demographics | Yes | Yes | Yes | Yes | Yes |
| Year Dummies | Yes | Yes | Yes | Yes | Yes |
| Location | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Proportion of Total Costs | 1.00 | 0.53 | 0.20 | 0.04 | 0.23 |
| Observations | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 |
| Number of Firms | 54 | 54 | 54 | 54 | 54 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. Demographics include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and a dummy for being in an MSA.

Table A2-6: Cost Models with Effective Deductible Contribution Interactions

| | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|---|----------------------|-------------------------|------------------------|-----------------------------------|-------------------------------------|
| Treat*(2 nd Pre Year) | -0.00198 (0.0171) | -0.00806 (0.0157) | 0.00651 (0.0255) | -0.0371 (0.0433) | 0.0120 (0.0198) |
| Treat*(1 st Post Year) | -0.0317* (0.0169) | -0.0503*** (0.0141) | -0.0215 (0.0401) | 0.143 (0.128) | -0.0274** (0.0138) |
| Treat*(2 nd Post Year) | -0.0364 (0.0255) | -0.0353 (0.0250) | -0.0296 (0.0370) | 0.0208 (0.0723) | -0.0551** (0.0247) |
| Treat*(3 rd Post Year) | -0.0230 (0.0146) | -0.0250 (0.0202) | 0.0327 (0.0265) | 0.0688 (0.0820) | -0.0955*** (0.0227) |
| Treat*1 st Post Year*Above Median Effective Deductible | -0.0632 (0.0505) | -0.0323 (0.0506) | -0.0807 (0.0699) | -0.322** (0.160) | -0.0625** (0.0317) |
| Treat*2 nd Post Year*Above Median Effective Deductible | -0.0158 (0.0359) | -0.0114 (0.0447) | -0.0507 (0.0621) | 0.0797 (0.211) | 0.00142 (0.0303) |
| Treat*3 rd Post Year*Above Median Effective Deductible | -0.0232 (0.0312) | -0.00848 (0.0343) | -0.115 (0.0747) | -0.0505 (0.104) | 0.0332 (0.0596) |
| Demographics | Yes | Yes | Yes | Yes | Yes |
| Year Dummies | Yes | Yes | Yes | Yes | Yes |
| Location | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Proportion of Total Costs | 1.00 | 0.53 | 0.20 | 0.04 | 0.23 |
| Observations | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 | 13,553,830 |
| Number of Firms | 54 | 54 | 54 | 54 | 54 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. Demographics include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and a dummy for being in an MSA.

Appendix A3: Models to Check Firm Composition Over Time

Table A3-1: Plan Descriptives by Treatment Cohort

| All Firms | N | Takeup | HRA | HSA | Account Contribution |
|----------------------------------|-----------|--------|-------|--------|----------------------|
| 1st post Year | 1,269,190 | 0.194 | 0.141 | 0.0792 | 157.2 (267.7) |
| 2nd Post Year | 1,002,859 | 0.214 | 0.091 | 0.122 | 154.2 (290.2) |
| 3rd Post Year | 480,535 | 0.346 | 0.251 | 0.0954 | 262.2 (318.5) |
| Firms with at least 2 Post Years | | | | | |
| 1st post Year | 980,554 | 0.185 | 0.105 | 0.0795 | 126.3 (264.9) |
| 2nd Post Year | 1,002,859 | 0.214 | 0.091 | 0.122 | 154.2 (290.2) |
| 3rd Post Year | 480,535 | 0.346 | 0.251 | 0.0954 | 262.2 (318.5) |
| Firms with 3 Post Years | | | | | |
| 1st post Year | 412,142 | 0.253 | 0.24 | 0.0127 | 290.2 (337.8) |
| 2nd Post Year | 453,573 | 0.270 | 0.191 | 0.0786 | 290.7 (350.6) |
| 3rd Post Year | 480,535 | 0.346 | 0.251 | 0.0954 | 262.2 (318.5) |

Table A3-2: Cost Models Restricted to Treatment Firms That Offer at Least Two Years (20 Treatment Firms)

| VARIABLES | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|-----------------------------------|------------------------|-------------------------|------------------------|---------------------|------------------------|
| Treat*(2 nd Pre Year) | 0.0162 (0.0144) | 0.00412 (0.0155) | 0.0242 (0.0222) | 0.0268 (0.0302) | 0.0333* (0.0197) |
| Treat*(1 st Post Year) | -0.0770** (0.0324) | -0.0759** (0.0326) | -0.0933** (0.0407) | -0.0872 (0.0655) | -0.0593*** (0.0222) |
| Treat*(2 nd Post Year) | -0.0428*** (0.0162) | -0.0399* (0.0221) | -0.0587* (0.0352) | 0.0597 (0.109) | -0.0486*** (0.0154) |
| Treat*(3 rd Post Year) | -0.0359** (0.0175) | -0.0317 (0.0216) | -0.0085 (0.0357) | 0.0262 (0.0612) | -0.0908*** (0.0243) |
| Demographics | Yes | Yes | Yes | Yes | Yes |
| Year Dummies | Yes | Yes | Yes | Yes | Yes |
| Location | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 12,808,248 | 12,808,248 | 12,808,248 | 12,808,248 | 12,808,248 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. Demographics include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and a dummy for being in an MSA.

Table A3-3: Cost Models Restricted to Treatment Firms That Offer at Least Three Years (13 Treatment Firms)

| VARIABLES | (1) Total Costs | (2) Outpatient Costs | (3) Inpatient Costs | (4) ED Costs | (5) Drug Costs |
|-----------------------------------|-----------------------|-------------------------|------------------------|---------------------|------------------------|
| Treat*(1 st Post Year) | -0.0369** (0.0176) | -0.0357** (0.0169) | -0.0520 (0.0452) | 0.0201 (0.0548) | -0.0387 (0.0242) |
| Treat*(2 nd Post Year) | -0.0175 (0.0147) | -0.0156 (0.0149) | 0.0207 (0.0334) | -0.0299 (0.0799) | -0.0566** (0.0253) |
| Treat*(3 rd Post Year) | -0.0218* (0.0126) | -0.0189 (0.0162) | 0.0308 (0.0285) | 0.0208 (0.0746) | -0.0926*** (0.0237) |
| Demographics | Yes | Yes | Yes | Yes | Yes |
| Year Dummies | Yes | Yes | Yes | Yes | Yes |
| Location | Yes | Yes | Yes | Yes | Yes |
| Firm Fixed Effects | Yes | Yes | Yes | Yes | Yes |
| Observations | 10,851,860 | 10,851,860 | 10,851,860 | 10,851,860 | 10,851,860 |

Clustered (at firm level) standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. GLM Models estimated with a Gamma distributional assumption and log-link. Demographics include gender, age, age squared, six level family type, median household income (zip), % unemployed (zip), % college degree (zip), % high school degree (zip), % Hispanic (zip) % Black (zip), and % White (zip). Location controls include full interactions between 9 census region dummies and a dummy for being in an MSA.