Journal of Health Economics 76 (2021) 102424

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/econbase

Effects of employer-offered high-deductible plans on low-value spending in the privately insured population

Brendan Rabideau[a], Matthew D. Eisenberg[a,b], Rachel Reid[c,d,e], Neeraj Sood[f,g,h,*]

[a] *Department of Health Policy and Management, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, United States*
[b] *Hopkins Business of Health Initiative, Johns Hopkins University, Baltimore, MD, United States*
[c] *RAND Corporation, Boston, MA, United States*
[d] *Division of General Internal Medicine and Primary Care, Brigham and Women's Hospital, Boston, MA, United States*
[e] *Harvard Medical School, Boston, MA, United States*
[f] *Price School of Public Policy, University of Southern California, Los Angeles, CA, United States*
[g] *NBER, United States*
[h] *Schaeffer Center for Health Policy & Economics, University of Southern California, Los Angeles, CA, United States*

ARTICLE INFO

Article history:
Received 20 June 2019
Received in revised form 9 December 2020
Accepted 30 December 2020
Available online 9 January 2021

Keywords:
Low-value care
Deductible
HDHP

ABSTRACT

Enrollment in plans with high deductibles has increased more than seven-fold in the last decade. Proponents of these plans argue that high deductibles could reduce wasteful spending by providing patients with incentives to limit use of low-value services that offer little or no clinical benefit. Others are concerned that patients may respond to these incentives by reducing their use of medical services indiscriminately and regardless of clinical benefit, which may negatively impact health outcomes. This study uses individual-level insurance claims data (2008–2013) and plausibly exogenous changes in plan offerings within firms over time to estimate the intent-to-treat and local-average treatment effects of high-deductible plan offerings on spending on 24 low-value services received in the outpatient setting. We find that firm offer of a high-deductible plan leads to a 13.7% ($5.23) reduction in average enrollee spending on low-value outpatient services and a 5.2% ($105.77) reduction in overall outpatient spending. We also find reductions in spending on measures of low-value imaging and laboratory services. We find some evidence that offering high-deductible plans disproportionately reduces low-value spending relative to overall spending, indicating that deductibles may be a way to incentivize value-based decision making.

1. Introduction

Healthcare costs in the United States (US) continue to rise, prompting providers, insurers, and policymakers to search for solutions to reduce spending without sacrificing quality of care (Hartman et al., 2018). Overtreatment has contributed to increasing spending and has been estimated to cost approximately $200 billion annually without contributing meaningfully to clinical well-being (The Healthcare Imperative, 2010) and thereby represents an opportunity to substantially reduce spending without negatively affecting patient health outcomes. Low-value health care use, a subset of overtreatment, is defined as spending on services that provide little to no clinical ben-

* Corresponding author at: Verna and Peter Dauterive Hall #312, University Park Campus, Los Angeles, CA, 90089, United States.
E-mail address: nsood@healthpolicy.usc.edu (N. Sood).

https://doi.org/10.1016/j.jhealeco.2021.102424
0167-6296/

efit and burdens patients with additional expenses and potential risk (Schwartz et al., 2014). Reducing low-value care use is seen as a promising way to bend the health care cost curve (Berwick and Hackbarth, 2012), however, efforts on this front have been hindered by several barriers. First, identifying which services are low-value given the multitude of clinical contexts is a difficult task that seldom results in unanimous agreement (Beaudin-Seiler et al., 2016). Second, even in instances where there is near consensus on services that constitute low-value care, changing provider and consumer behavior is further complicated by imperfect dissemination of information, patient preferences, medicolegal concerns, and other factors (Gawande et al., 2014; Umscheid, 2013).

One mechanism to potentially reduce low-value health care spending is the creation of demand-side incentives for patients to use care more efficiently through plan designs with greater cost-sharing for patients. In particular, there has been a recent trend in plans with higher deductibles. A deductible is a feature of an insurance plan which constitutes an amount of money that a patient must pay out-of-pocket every year before insurance begins covering expenses. In 2017, 28% of employees with employer-sponsored health insurance were enrolled in a high-deductible health plan – up from 20% in 2015 – as well as 90% of individuals covered through plans purchased from the Patient Protection and Affordable Care Act individual marketplace in 2015 (Claxton et al., 2016, 2017; Dolan, 2016). Proponents of high-deductible plans argue that the high-cost sharing associated with them will encourage consumers to shop around for lower prices, and consider the value of the services that they receive since they have skin in the game. On the other hand, critics argue that, as shown as early as the RAND Health Insurance Experiment, high-cost sharing will lead to broad reductions in care regardless of value (Manning et al., 1987). These indiscriminate reductions may result in adverse health outcomes if patients reduce their use of necessary services. Worryingly, this concern could be amplified for those of lower socioeconomic status (Chernew and Newhouse, 2008).

Two previous studies (Reid et al., 2017; Rosenthal et al., 2018) have explicitly examined the effect of high-deductible plans on low-value care use. Rosenthal et al. used cross sectional data with multivariable regression analysis and found no consistent relationships between high-deductible plans and low-value care use. Reid et al. used longitudinal data with a coarsened exact-match strategy between treatment and control enrollees and found that, while switching from a traditional health plan to a high-deductible plan was associated with decreased low-value spending relative to matched-controls, these reductions were proportional to overall reductions, suggesting that patients were not preferentially reducing low-value service use (Reid et al., 2017). However, both studies were unable to account for unobservable differences in enrollees that opted into high-deductible plans relative to those that stayed in traditional plans, resulting in concerns about selection bias. Understanding the causal effect of high-deductible plans on low-value care use is critical to understanding the role that demand side cost sharing plays in altering expenditures and targeting low-value care as a means to bend the health care cost curve.

The goal of this study is to estimate the causal effect of high-deductible plans on measures of low-value spending. The key empirical challenge of this analysis is the endogeneity of high-deductible plan enrollment. Selection into various health plans is endogenous and prior research has shown that high-deductible plan enrollees tend to be younger, healthier, and have higher income than traditional plan enrollees (Greene et al., 2006; McDevitt et al., 2014; Tollen et al., 2004). To address this issue, we follow two analytic strategies common in the related literature. In the first, we identify a set of employers that never offer a plan with a deductible over \$500 (control group) and a second set of employers that began offering a plan with a deductible of at least \$500 during our study period (treatment group). We then estimate the intent-to-treat effect of offering a high-deductible plan on mean annual enrollee spending. By analyzing the intent-to-treat effect of *offering* a high-deductible plan, we can leverage the firm's plausibly exogenous offer to eliminate any bias stemming from individual-level plan selection (Haviland et al., 2016). Second, we leverage the firm's high-deductible plan offer as an instrument for enrollment in one of these plans, which allows us to estimate the local average treatment effect of enrolling in a high-deductible plan (Eisenberg et al., 2017).

Overall, we find that offering a plan with a high deductible results in reductions in enrollee spending on 24 outpatient low-value services. In particular, we find that a high-deductible plan offer causes a statistically significant decrease of 13.7% in spending on 24 low-value outpatient services, a significant decrease of 21.2% in spending low-value imaging service measures, and significant decrease of 22.2% in spending on low-value laboratory service measures. We find that, relative to reductions in overall spending for each of these categories (5.2%, 17.7%, and 13.6%, respectively), the magnitude of the reductions for low-value spending is significantly larger for outpatient and laboratory services. While no significant difference was found between the reductions in low-value and overall spending for imaging services, the point estimates are qualitatively larger for the low-value services. This suggests that high-deductible plans may do more to promote value-based decision making than previously thought.

In Section II we provide a general background on low-value care use and high-deductible plans, and discuss related literature. Section III describes our data, empirical model, identification strategy, and limitations. In Section IV, we discuss our results, and, in Section V, we conclude and discuss implications.

2. Background

2.1. Trends in low-value care

A growing body of literature has focused on determining which services meet the definition of low-value care and to describe their prevalence. One influential approach has been spearheaded by the American Board of Internal Medicine Foundation and their Choosing Wisely campaign, in which more than 75 physicians and professional societies from a variety of specialties have provided recommendations on a subset of services they deemed low-value (Cassel and Guest, 2012). For example, back imaging (inclusive of X-rays, MRIs, and CT scans) within six weeks of the initial onset of back pain for patients without a warranting condition (such as history of cancer, IV drug abuse, neurological impairment, etc.) would be considered low-value because the procedure exposes patients to ionizing radiation, burdens them with unnecessary costs, and has been shown to lead to worse health outcomes than in patients who do not receive it ("Spinal imaging for acute low-back pain," 2019).

Prior research has taken steps to operationalize guidelines from Choosing Wisely and other similar recommendations from the US Preventive Services Task Force, the National Institute for Health and Care Excellence, and the Canadian Agency for Drugs and Technologies in Health into various low-value service use metrics to explore the prevalence and correlates of this care in observational data settings. One early study estimated that 25% of the 65 and older Medicare population used at least one of the specifically enumerated low-value services in 2009 (Schwartz et al., 2014), while another study found that 7.8% of the 18–64, privately insured population had received similarly defined low-value services in 2013 (Reid et al., 2016). Additionally, studies looking at the use of low-value services in the non-insured and Medicaid populations found that there were no significant differences in the receipt of low-value service relative to similar, privately insured patients (Barnett et al., 2017; Charlesworth et al., 2016), suggesting that the problem of overtreatment is pervasive in the US healthcare system. Variation in the number of services ordered has been observed both regionally (Reid et al., 2016) and between different types of provider organizations (Schwartz et al., 2015), and the prevalence of specific services has been subject to change over time (Carter et al., 2017), suggesting that physician ordering behavior is likely malleable with respect to low-value services. Colla and colleagues (2017) provide a systematic review assessing the various strategies used to reduce low-value care.

2.2. The effect of high-deductible health plans

Several studies have documented that enrollment in a high-deductible plan reduces overall healthcare spending, most of which is attributable to a reduction in the quantity of services received as opposed to patients seeking out less costly services or providers (e.g. Zhang et al., 2018; Brot-Goldberg et al., 2017; Haviland et al., 2016; Sinaiko et al., 2016; Sood et al., 2013; Haviland et al., 2012). However, concerns remain that high-deductible plans will lead to a reduction of both high and low-value services, which could result in adverse health effects for the patient. Studies exploring the relative value of the services that high-deductible plan enrollees tend to forgo have provided mixed results. There is some evidence to suggest that there is an indiscriminate reduction in medical service consumption, even in plans with deductible carve-outs for certain types of high-value care such as preventative screenings. Several studies have found reductions in the use preventative cancer screenings in response to enrolling in a high-deductible plan (Beeuwkes Buntin et al., 2011; Brot-Goldberg et al., 2017; Haviland et al., 2011; Wharam et al., 2011). However, a study with a larger, more representative sample and a longer follow-up period found no reduction in cancer screenings in response to either an employer offer of a high-deductible plan, or to actually enrolling in a high-deductible plan (Eisenberg et al., 2017).

Less attention has been given to the effects of high-deductible plans on low-value services, though two studies to our knowledge have specifically examined this relationship. A cross-sectional study on the association between plan type and use of low-value services found that high-deductible plans were not consistently associated with lower use of low-value services, though this study made no attempt to address endogeneity concerns brought about by heterogeneous selection into various plan types (Rosenthal et al., 2018). Another study found that, while switching from a traditional health plan to a high-deductible plan was causally associated with decreased low-value spending relative to matched-controls, these savings were proportional to overall savings, suggesting that the reductions in spending were indiscriminate (Reid et al., 2017). However, this study was also threatened by endogeneity concerns whereby relatively healthier patients who anticipated lower future medical costs may have opted into a high-deductible plan to take advantage of the lower premiums, making it erroneously appear as though the switch into a high-deductible plan caused the lower spending. While the researchers mitigated these threats by applying a coarsened-exact-match between treatment and control enrollees on a rich array of observable characteristics, the study design was unable to account for biases introduced by unobservable factors.

Our analysis adds to the literature by providing robust estimates of the causal effects of firm offer of a high-deductible plan, as well as enrollment in a high-deductible plan, on low-value spending. Our study is one of only a few that addresses the impact of high-deductible plans on low-value spending specifically, and our design reduces the biases from endogenous enrollment in a high-deductible plan relative to the other similar studies. Additionally, our sample is both larger and more representative than the most similar existing study, allowing us to measure the effect of high-deductible plans more precisely and generalize more confidently. The present analysis uses data from 30 large, U.S. firms which represent millions of enrollees in both the treatment and control arms, whereas Reid et al. used claims from a single insurer with 11,075 enrollees in the treatment arm.

2.3. Conceptual framework

There are several mechanisms through which a deductible could bring about the hypothesized reductions in spending. To understand how, we start with a baseline plan with no cost-sharing and no other forms of managed-care which could restrict a patient's consumption. With such a plan, a patient would face no consequences for consuming either too great a quantity of services or services that are too expensive, as the entire cost would be borne by the insurance provider. According to the canonical Grossman model, under this scenario the only factor that would limit the healthcare consumption of the patient would be the time that it takes time to receive the medical treatment when faced with competing interests pressuring their time constraint (Grossman, 1972). When adding in a deductible, however, a patient now bears the cost of services up until they reach their annual deductible limit, which puts pressure on their budget constraint. Now, as with other goods, the patient must decide whether the value of a given medical service as measured by the production of health is worth the out-of-pocket spot-price of that medical service. This could change the patient's behavior in three distinct ways: (1) the patient could forgo medical care entirely if they believe there is no care that they could receive that would be worth the out-of-pocket price; (2) the patient could shop around for better prices on a particular service until they find a price at which it makes sense to receive the service; or (3) they could substitute a particular service for another service that is either cheaper, more effective, or of generally greater value. A thorough investigation into these mechanisms can be found in Brot-Goldberg et al. (2017), which provides evidence that forgoing care across the board appears to be the dominant mechanism through which a deductible acts. What is currently unknown is whether patients are capable of assessing the value of various procedures. Since the low-value services in the present analysis have been identified as those that provide no evidence-based clinical benefit, we would expect to find quantity reductions in these services.

3. Methods

3.1. Data

In this analysis, we use data from the Truven Health MarketScan® Research Databases, which is sourced from large, self-insured US employers. The Truven data has been shown to be broadly representative of the United States employer-based health insurance market (Haviland et al., 2016; Eisenberg et al., 2017). We constructed measures of low-value service use and spending using the Outpatient Services Table, Inpatient Services Table, and the Annual Enrollment Summary Table, which include individual-level enrollment data and privately paid medical claims.

3.1.1. Defining deductible levels

Our empirical strategy purports to look at the effect of a firm's decision to offer a plan with a high deductible on enrollee spending in that firm. However, one challenge in this approach is that our data does not have specific plan benefit designs, including information on deductible levels. In lieu of this, MarketScan provides a categorical plan-type variable, which takes on one of nine values corresponding to the type of plan that an enrollee is in (e.g. Preferred Provider Organization, Health Maintenance Organization). To circumvent the issue of not knowing the true deductible level of a plan a prior, we estimate deductible levels for each plan-type in a firm in a given year using empirical spending data from the enrollees, in accordance with related literature (Zhang et al., 2018).

To estimate deductibles, we summed the deductible spending amount on each claim in a calendar year for enrollees with coverage for the full calendar year. We identified enrollees who reached their deductible limit as those who had at least three consecutive claims without deductible spending. Within each firm-year-plan-type, we calculated the distribution of annual deductible spending for those we flagged as having reached their deductible limit, and used the 80^{th} percentile of spending within these groupings as the estimated deductible level for that firm-year-plan-type. While this measure is an imperfect estimate of the true plan deductible level, we believe Fig. 1 provides an intuitive test of accuracy. In this figure, we can see that plan-types where one would expect a low or no deductible (such as Health Maintenance Organizations) overwhelmingly have deductibles below $500, whereas plan types where one would expect higher deductibles (High-Deductible Health Plans and Consumer-Directed Health Plans) present with a large proportion of their distribution at higher deductible levels.

3.2. Study sample

The study used data from 30 firms that reported continuously from 2008 to 2013. We characterized our sample inclusions criteria at two levels: the firm level and the individual level. At the firm-level, the treatment arm was comprised of 7 firms that had 2 full years of data in which no plan with an estimated deductible greater than $500 was offered, followed by a post-period where they offered at least one plan with an estimated deductible greater than $500 for three consecutive calendar years. The control arm was comprised of 23 firms that never offered a plan with a deductible over $500 between 2008 and 2013. Firm offer of a high-deductible plan was defined as a firm for which at least 3%[1] of employees were enrolled in a high-deductible plan as of January 1^{st} of that calendar year, whereas control firms had no enrollees in these plans. Enrollees were identified as being in a high-deductible plan based on the estimated deductible level of that plan, as empirically derived for each firm-year-plan-type. Finally, we excluded firms that had abnormal fluctuations in their levels of employment, defined as an increase or decrease from the prior year of more than 50%. This criterion served to prevent undue influence from firms that underwent major restructuring (e.g. large branch openings or closings), which could exacerbate the threat of changes in enrollee composition and undermine our empirical strategy.

At the individual level, we first identified individuals aged 18–64 who were continuously enrolled for at least two full calendar years (24 months). The first 12 months of data was not included in the analytic file but was used to calculate baseline comorbidity scores and to provide a suitable window to observe patient clinical history necessary to identify our low-value services (see Appendix Table A1 for details). By using baseline data to calculate comorbidity

[1] 3% was chosen arbitrarily as a number that would be large enough to signal that a non-trivial number of enrollees actually had the opportunity to enroll, but low enough to not exclude firms with low uptake.

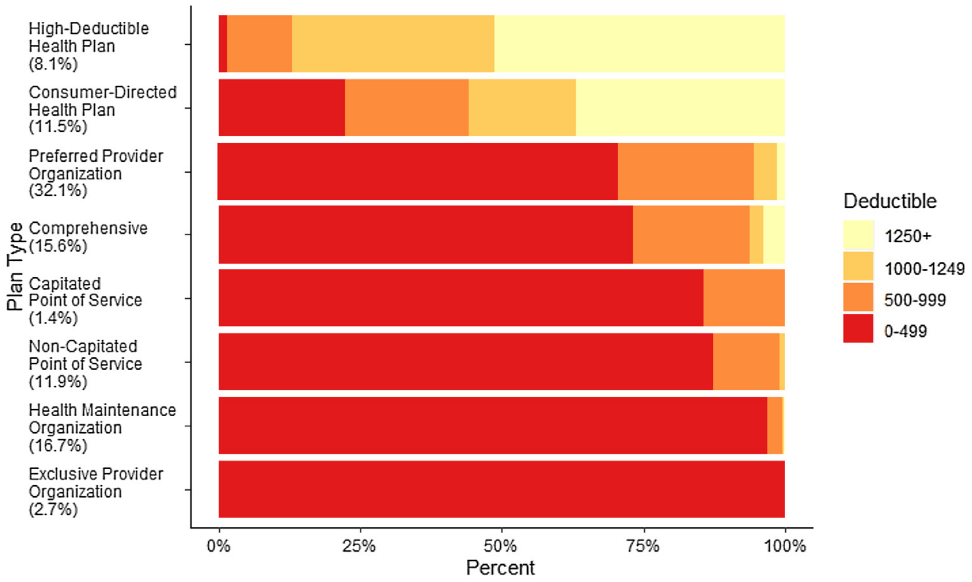

Fig. 1. Distribution of Estimated Deductibles by Plan-type.
Notes: Fig. 1 shows the distribution of estimated deductible levels for each plan-type variable associated with an enrollee in the MarketScan eligibility file. The percentages in parentheses on the y-axis represent the share of offered plans from 2008 to 2013 for each plan-type (e.g. 8.1% of plans offered by all firms from 2008 to 2013 were high-deductible health plans). The x-axis shows the percent of firm-year-plans in a plan-type with the given estimated deductible level. Deductibles were estimated by summing the deductible payment on all claims for an enrollee in a calendar year, and identifying enrollees who reached their deductible limit. An enrollee who has reached the deductible limit is defined as one who has three consecutive claims with no deductible spending. A distribution of annual deductible spending from these enrollees is then taken at the firm-year-plan-type level. The 80th percentile value of all enrollees who reached their limit in each group is assumed to be the deductible level for this firm-year-plan-type.

scores we sought to minimize the threat of unobservable correlations between this measure and high-deductible plan offer. Beyond these calculations, this baseline year did not contribute to our estimates. The second continuous calendar year of data formed our first year of analytic data for each individual. If the individual was continuously enrolled for longer than 12 months, we kept data for patients enrolled in 12 month intervals (12, 24, 36, etc.) only since our analysis was conducted at the patient-year level to avoid any potential seasonality bias.

3.3. Outcome measures

Our main dependent variables were total annual outpatient spending and annual spending on 24 low-value services performed in an outpatient setting. The rationale for focusing on outpatient spending stems from the prior cost-sharing literature which has found that demand for outpatient spending is more elastic than spending in other settings, such as inpatient stays (Haviland et al., 2016). In addition to general outpatient spending, subcategories for outpatient imaging and outpatient laboratory spending were also constructed to provide further detail about the effect of high-deductible plans on specific types of spending. Overall outpatient procedures were defined as all those found in the Outpatient Services Table that did not specify "Inpatient Hospital" as a place of service. Outpatient imaging and laboratory procedures included all outpatient services with a Current Procedural Terminology code in the range of 70000–79999 and 80000–89999, respectively.

We constructed 24 measures of low-value services and assessed spending on these services. These measures have been have been used in several studies of low-value service use and spending (Reid et al., 2016, 2017; Schwartz et al., 2015, 2014; Schwartz et al., 2018) and are informed by the recommendations of the ABIM Foundation's Choosing Wisely campaign, the US Preventive Services Task Force, the National Institute for Health and Care Excellence, and the Canadian Agency for Drugs and Technologies in Health ("About the Health Technology Assessment Service", 2019; "Clinician Lists," 2021; "NICE Do not Do prompts," 2013; "Recommendations for Primary Care Practice," 2019). The set of 24 low-value services in the present study were selected specifically to capture services for which our non-elderly, commercially insured population would be eligible. The full list of our procedures and their definitions can be found in Appendix Table A1. While our selected measures are not meant to be an exhaustive list of low-value services, it is expected that this cross section of services can be used as a proxy for a broader array of low-value treatments and help both in characterizing patterns in low-value utilization as well as measuring the impact of incentive-based mechanisms designed to reduce overtreatment more generally.

Spending for low-value services was captured in three ways depending on the complexity of the service. First, for many simpler low-value services, we used the payment amount listed on the claim to determine spending for that service. These services included vitamin D testing, adnexal cyst imaging, back scan within six weeks of back pain without a warranting cause, bone density testing, carotid artery screening for syncope, carotid scanning for asymptomatic patients, CT for rhinosinusitis, electroencephalogram for headache, head imaging for uncomplicated headache, head imaging for syncope, HPV testing in women younger than 30, imaging for plantar fasciitis, preoperative echocardiography, and T3 measurements in hypothyroidism. Alternatively, for some more complex low-value services there is a set of related co-occurring procedures that are typically performed during the same visit. To more accurately capture the total cost associated with these services, spending was defined as the sum of the payment for the low-value service and the payments from all of the expected co-occurring procedures taking place on the same day. The services that this cost-aggregating methodology was applied to were homocysteine testing in cardiovascular disease, parathyroid hormone testing for stage I-III chronic kidney disease, hypercoagulability testing for venous thromboembolism, preoperative chest radiography, preoperative pulmonary function testing, stress testing in stable coronary artery disease, and inferior vena cava filters to prevent pulmonary embolism. For our most complex low-value services it is likely that there are several co-occurring procedures and services that are too numerous and variable to be specifically enumerated. To capture spending for these complex low-value services, we took the sum of all payments from all outpatient claims occurring on the same day as the low-value service. The services that this procedure applied to were renal artery angioplasty or stent, and spinal injection for lower back pain. Spending for these individual procedures were then aggregated to the person-year level, and summed together to form a single outcome for overall outpatient low-value spending. We then created subcategories of low-value spending where we grouped our low-value services into low-value outpatient imaging, and low-value outpatient laboratory spending measures. A more detailed definition of each of the 24 low-value services, the co-occurring procedures that potentially contributed to their spending totals, and the grouping of the services into our two subcategories (outpatient imaging and outpatient laboratory) can be found in Appendix A1. For our main analysis, each of our person-year spending outcomes was winsorized at the 5^{th} and 95^{th} percentiles to remove the influence of extreme outliers, with the exception of those with \$0 annual spending, which were left as \$0.

3.4. Independent variable

Our independent variable for the main analysis was a binary treatment and post-year interaction variable (treatXpost) representing observations for enrollees in treatment firms in any one of the years when a high-deductible plan was offered. Our main analysis used this single post-period interaction variable, though we also tested an event-study specification with three binary independent variables – treatXpost1, treatXpost2, and treatXpost3 – where each represented enrollee presence in a treatment firm in the first, second, or third year since the firm began offering a high-deductible plan (Appendix Table A2).

3.5. Covariates

Patient characteristics included in these analyses were age, sex, and comorbidity as assessed by the sum of the number of conditions contributing to the Charlson Comorbidity Index (Charlson et al., 1987) in the first calendar year in which the enrollee was observed.

3.6. Intent-to-Treat analysis

The goal of this analysis is to estimate the causal effect of high-deductible plans on enrollee spending. To accomplish this, we took two main approaches. The first used an intent-to-treat design to analyze the effect of a firm *offering* a high-deductible plan. We employed a generalized difference-in-differences design with two-way fixed effects for firm and calendar year using the following baseline specification:

$$Y_{ift} = f(\beta_0 + \beta_1 treatXpost_{ft} + X_{it}\beta_2 + firm_f + year_t) + \varepsilon_{ift} \quad (1)$$

Where Y_{ift} is annual spending for enrollee i in firm f in year t, $treatXpost_{ft}$ is an interaction variable equal to one if the f is a treatment firm and t is a post year, $firm_f$ is a vector of firm fixed effects, $year_t$ is a vector of calendar year fixed effects, and X_{it} is a vector of characteristics of enrollee i in year t, including age, sex, and the Charlson comorbidity count from the baseline year.

To account for the large mass of enrollees with \$0 in annual spending, as well as the heavily right-skewed distribution of healthcare costs, we estimated the mean spending using a two-part model with a probit first stage to determine the probability of having any spending for an outcome, and a generalized linear model with a log-link and gamma family to estimate the spending in the second stage, as recommended in the literature on modelling healthcare spending (Deb and Norton, 2018; Farewell et al., 2017). We also tested specifications using a logit first stage and a Gaussian family in the second stage in sensitivity analyses (Appendix Table A3). Standard errors were clustered at the firm level as our main independent variable varies at the firm level and we may expect correlations in the error terms between enrollees at the same firm.

β_1 was our parameter of interest, representing the difference-in-differences estimate in mean annual enrollee spending between our treatment and control enrollees over the three years following a high-deductible plan offer. We calculated these results two ways: first as the average marginal effect for each outcome (using estimates from both the first and second stage of the two-part model), which can be interpreted as the difference in mean annual spending per enrollee in treatment firms after offering a high-deductible plan. Second, we estimated the percent change associated with being in a treatment firm in the post period by calculating the semi-elasticity of spending, represented by $\frac{\delta(\ln(y))}{\delta treatXpost}$.

We estimated this model for each of our six outcomes of interest – annual spending on outpatient, imaging, laboratory, low-value outpatient, low-value imaging, and low-value laboratory services.

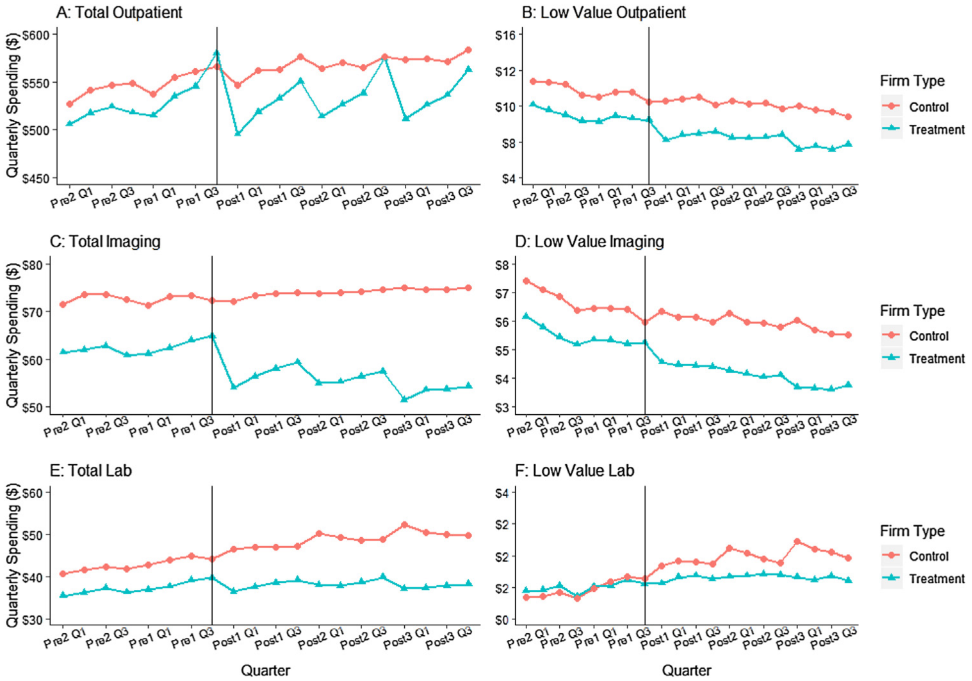

Fig. 2. Mean Quarterly Spending Per Enrollee.

Notes: Mean quarterly low-value spending per enrollee for treatment firms and control firms. Total spending graphs for our 3 categories are on the left (Panels A, C, and E), while low-value spending graphs are on the right (Panels B, D, and F). Total spending per person for each category defined as the sum of all payments on any claim in the Annual Outpatient Table without an explicit service location of 'inpatient hospital' for Panel A, claims with a CPT code in the range 70000–79999 for Panel C, and claims with a CPT code in the range 80000–89999 for Panel E. Panels B, D, and F are the sum of payments on claims for low-value services grouped into each category as described in Appendix Table A1. The vertical line is placed at Q4 of Pre-Year 1, immediately before a high-deductible offer. The control firm pre and post-periods are calculated as a weighted average of calendar year mean spending per enrollee matching the enrollee-level distribution of calendar years in each period for the treatment firms.

3.7. Identification

The key identifying assumption of our model was that changes in mean spending would have been the same between treatment and control firms had the treatment firms not offered a high-deductible plan. This assumption is ultimately untestable, but researchers typically gain confidence that the assumption is reasonable by comparing trends in outcome variables in the periods before an event and confirming that the trendlines of the treatment and control outcomes are parallel until the event occurs. To analyze the validity of this assumption, we graphically present the unadjusted spending for each outcome by quarter in Fig. 2. The graphs show that the treatment and control spending trends for all outcomes are approximately parallel in the pre-period. Further, a formal test of pre-trends is conducted in Appendix Table A4, where the coefficient represents a relative difference in quarterly spending in pre-year2 relative to the immediate pre-year, excluding Q4 of the immediate pre-year.[2] Results of the formal pre-trends analysis indicate that there are no significant differences in the trends in five of the six outcomes. For the sixth outcome – low-value laboratory spending – there is a significant downward trend in spending in the treatment firms relative to the control firms. However, Fig. 2F is provides compelling visual evidence that this significant difference is not economically meaningful.

3.8. Local average treatment effect

We used the firm-level offer of a high-deductible plan as an instrument for individual enrollment in a high-deductible plan, and used a two-stage least squares model to estimate the local average treatment effect of switching into a high-deductible plan on spending. This approach has been used in similar analyses to estimate the endogenous effect of enrollment on continuous outcomes (Eisenberg et al., 2017; Finkelstein et al., 2012). The empirical model is as follows:

$$Enroll_{ift} = \pi_0 + \pi_1 treatXpost_{ft} + X_{it}\pi_2 + firm_f + year_t + \mu_{ift} \tag{2}$$

$$Y_{ift} = \delta_0 + \delta_1 \widehat{Enroll}_{ift} + X_{it}\delta_2 + firm_f + year_t + \lambda_{ift} \tag{3}$$

Where $Enroll_{ift}$ is the binary outcome for whether an employee i of firm f enrolled in a high-deductible plan in

[2] Q4 is omitted from this particular analysis as there is evidence in the literature that if an enrollee anticipates switching into a high-deductible plan, it could lead to a stockpiling of medical services in the period immediately before the switch (Eisenberg et al., 2017; Reid et al., 2017). This supports an argument that the spending in pre-period Q4 is directly attributable to the high-deductible plan offer – a possibility that we explored further in a sensitivity analysis (Table 5)

year t after firm offer, and $\widehat{Enroll}_{ift}$ is the conditional probability that an enrollee opted into a high-deductible plan estimated in the first-stage.

3.9. Gradient effect of deductible offers

While our main independent variable in the above analyses is an indicator for whether a firm offered a plan with a deductible of at least \$500, we also test for a gradient effect for even higher deductibles. To test for heterogeneity in the effects of being offered plans with different deductible levels, we allow for firm-year indicators for offering plans of increasingly high deductibles:

$$Y_{ift} = f(\gamma_0 + \gamma_1 treatXpostX500_{ft} + \gamma_2 treatXpostX1000_{ft} + \gamma_3 treatXpostX1250_{ft} + X_{it}\gamma_4 + firm_f + year_t) + \sigma_{ift} \quad (4)$$

Where $treatXpostX500_{ft}$ is an indicator for firms whose highest-deductible plan has a deductible between \$500-\$999 in year t, $treatXpostX1000_{ft}$ for firms whose highest level plan has a deductible between \$1000-\$1249 in year t, and $treatXpostX1250_{ft}$ for firms who offer a plan with a deductible over \$1250. While each indicator represents the plan with the highest deductible offered, it is possible that firms offer multiple plans with a range of deductibles, though the indicators are mutually exclusive at the firm-year level.

3.10. Spending reduction decomposition

As highlighted previously, there are broadly three mechanisms that could results in a spending reduction – reducing the price per service through price shopping, or reducing quantity of services through forgoing care or substituting for different services. Prior work on this topic has found that the majority of spending reductions come from outright reductions in utilization as opposed to substitutions or price shopping (Brot-Goldberg et al., 2017). Our contribution in this space comes from decomposing the spending reductions specifically for low-value services to see if they follow the same patterns identified in prior literature. To do so, we employ the same empirical strategy from (1) using new outcome variables for each individual low-value services. To test for changes in the quantity of services received, we take as outcome variables the annual count of each low-value service for each enrollee. To test for price shopping, we calculate the average payment per low-value service per enrollee per year, conditional on the enrollee having had received that service in the given calendar year. These results can be found in Table 6.

3.11. Robustness checks

3.11.1. Continuous enrollment restriction

One potential threat to the validity of our strategy is that the composition of our firms could be changing over time in a way that is correlated with high-deductible plan offer and the use of low-value services, thus biasing our results. For example, the potential for lower premiums associated with a high-deductible plan could have attracted new, healthier enrollees to firms that offer these plans, or unobserved changes could be forcing out less healthy workers, which would result in lower mean annual spending that is not attributable to our proposed mechanism of changing enrollee healthcare consumption behaviors. To test for this, we restricted our sample to enrollees who were employed at a firm for the entire observation window and model our outcomes on this continuously enrolled sample (Table 2 Panel B). In addition, we re-estimated our models with individual fixed-effects rather than firm fixed effects to account for unobserved, time-invariant differences in the enrollees (Appendix Table A5)

3.11.2. Anticipation period

As alluded to above, one potential effect of firm offer of a high-deductible plan is anticipatory stockpiling of services in the period immediately preceding the offer. Given our main model specification, this effect would bias our results towards finding a greater reduction in mean annual spending since it would appear as though treatment firm pre-period spending is higher than it would have been without the offer, and the decreased spending in the first quarter of the post-period would be larger than can be causally attributed to high-deductible plan offer since it would include the result of both the cost-sharing and a satiation in the demand for healthcare services brought on by the stockpiling. To account for this, we performed a sensitivity analysis where we included pre-period Q4 spending in our post-period, which allowed us to formally attribute the stockpiling effect to the high-deductible plan offer (Table 5).

3.11.3. Checking outlier influence

While our main analysis used winsorized measures to estimate differences with more generalizability and accuracy than using the raw values, it is plausible that the effects of firm high-deductible plan offer could be nullified or even reversed if the outliers are consistently affected in the opposite way. We believe this is unlikely as we hypothesize that, if anything, healthier people would be more likely to remain in our treatment firms as a result of high-deductible plan offer, whereas less healthy, more expensive employees would be more likely to leave, which should lead to fewer extreme spending outliers in the right tail of our treatment firms relative to control firms. By winsorizing our results, we expect to bias our results against a finding that a high-deductible plan offer reduces mean enrollee spending as the process is most likely to disproportionately lower control firm post-period spending relative to treatment firm spending. Nonetheless, we re-estimated our models using unwinsorized outcomes and present our results in Appendix Table A6.

3.11.4. Exploring alternative mechanisms

Another potential threat to the internal validity of our design is that firms may be employing several measures to combat overspending. For example, firms may be offering Preferred Provider Organizations that have increasingly narrow physician networks with low-cost providers. Unfortunately, our lack of plan benefit design does not allow us to investigate these possibilities in great depth. However, using the available information, we take several steps to assuage concerns that co-occurring interventions are biasing our results. First, to test whether enrollees in alternative plan types are also experiencing spending reductions, we re-estimate our models excluding enrollees who are in each of the alternative plan-types[3] piecewise by plan-type. The rationale is that if enrollees in these other plans had non-deductible reasons to reduce spending, the magnitude of their spending reductions would be captured in our original ITT. Excluding these enrollees would therefore reduce the magnitude of our ITT results if they were indeed contributing meaningfully to the firm-level reductions, but would have no effect on the ITT results (or increase them) if these enrollees were not reducing (or increasing) spending after the firm offer of high-deductible plans.

Another concern is that our empirical strategy is based on using a binary indicator for a firm beginning to offer a high-deductible plan, where our interpretation of the results relies on the assumption that the deductible level substantially increased from the pre to the post-period in our treatment firms relative to control firms. However, it is in theory possible that the change in deductible level is actually greater in control firms even though they never offer a plan with a deductible over $500, if, in an extreme example, these firms went from deductible offerings of $0 in the pre-period to $499 in the post, whereas treatment firms went from $499 to $500. Appendix Table A7 provides descriptive evidence that this possible threat is unlikely to be present, as both the estimated deductible level and deductible spending for enrollees in control firms (as well as low-deductible enrollees in treatment firms) stay relatively constant between pre and post-period when compared to those switching into high-deductible plans.

3.12. Limitations

There are several limitations with our study. First, claims data may not provide a sufficiently rich set of observable characteristics to allow us to perfectly ascribe low-value status to any given procedure, as it is possible that a procedure was appropriate for a given enrollee based on their medical condition which was not fully conveyed in claims. Additionally, we could not observe services that an enrollee received that were paid for completely out of pocket (i.e. cash payments not processed by the insurance provider). Our chosen subset of low-value services, while in line with the recommendations of respected clinical sources and the recent literature, is another limiting factor of our analysis. Our 24 procedures are only a small number of the total number of low-value services that can be provided and were chosen primarily because of the consensus of their low-value status as opposed to their representativeness or impact on our population. We assumed that these services could be proxies for all low-value services, though to our knowledge this assumption has yet to be tested. As a result of using this cross-section of low-value services, our absolute magnitudes for low-value spending are quite small relative to what may be expected from the literature on overuse.[4] Finally, while we were able to test for the effects of alternative mechanisms that low-deductible plan types may have been providing, our approach in unable to test if plans that increased the deductible level also simultaneously implemented other mechanisms to reduce spending such as price-transparency tools (Brot-Goldberg et al., 2017).

4. Results

4.1. Main results

As shown in Table 1, our study included 23 control firms that never offered a high-deductible plan during the study period, representing 8,821,184 person-years, and 7 treatment firms representing 3,168,199 person-years, for a total of 11,989,383 person-years in the main analytic sample. Treatment firms had enrollees that were slightly younger (average age of 43.40 vs 45.15) and healthier (Charlson comorbidity score of 0.14 vs 0.17). They also had higher pre-period annual deductible spending ($92.16 compared to $79.47), higher plan deductible levels ($135.37 compared to $96.56), and lower levels of spending in all categories except for low-value laboratory spending. It is expected that the enrollees in the treatment firms, being younger, healthier, and having slightly higher deductibles in the pre-period, have lower levels of spending in this period, as can be seen in Fig. 2. Going from the pre-period to the post-period, treatment firms had a far greater increase in deductible spending relative to control firms (a 182% increase compared to a 17% increase), verifying that there is indeed a substantial increase in cost-sharing in treatment firms. In the post-period, 62.26% of enrollees in treatment firms are enrolled in a plan with a deductible of at least $500. Descriptive statistics in Table 1 also show that the largest compositional shifts in plan-types for enrollees in treatment firms came primarily from losing members in non-capitated point-of-service plans and large gains in enrollment in High-Deductible Health Plans and Consumer-Directed Health Plans (both known for their high deductibles), with a small increase in enrollment in Preferred Provider Organizations. Control firms see a large shift towards more PPOs, with small losses in all other plan-types, and no enrollment in HDHPs or CDHPs.

Fig. 2 shows graphs of spending over time for each of our six categories, with a vertical line overlaid on the final quarter before the offer a high-deductible plan in the treatment firms. As can be seen, spending trends appear to be parallel prior to the offer for all categories (formally tested in Appendix Table A4), with a stark drop in treatment firm spending in the first quarter when a high-deductible plan was offered. This level change appears to persist in the following three years. There is interesting cyclical behavior

[3] We exclude the enrollee in all years if they are enrolled in a given plan type at any point in the post-period and are in a treatment firm. Further, if the enrollee is in a given plan type but also has a deductible of at least $500, we retain them in the sample.

[4] Mean annual spending for our subset of low-value outpatient, imaging, and laboratory services is $39.51, $23.24, and $6.94, respectively.

Table 1
Pre-Post Descriptive Summary.

	Pre		Post	
	Treatment	Control	Treatment	Control
Individual Characteristics				
Female	52.47%	53.39%	52.49%	53.10%
Age	43.40	45.15	43.20	44.07
Charlson Comorbidity Score	0.14	0.17	0.12	0.13
Expenditures				
Deductible Spending	$92.16	$79.47	$260.41	$93.21
Outpatient Spending	$2121.26	$2191.63	$2130.16	$2277.01
Imaging Spending	$249.71	$290.52	$221.81	$296.38
Laboratory Spending	$149.32	$171.21	$152.34	$195.78
Low-Value Outpatient Spending	$37.23	$44.26	$30.64	$40.25
Low-Value Imaging Spending	$21.85	$26.50	$16.43	$23.79
Low-Value Laboratory Spending	$5.99	$5.91	$6.61	$7.95
Deductible Amount				
Estimated Deductible Level	$135.37	$96.56	$522.75	$132.18
Low Deductible (<$500)	99.68%	100.00%	37.74%	100.00%
$500 -$999	0.01%	0.00%	50.88%	0.00%
$1000-$1249	0.02%	0.00%	5.62%	0.00%
$1250+	0.30%	0.00%	5.77%	0.00%
Enrollee Plan Types				
Comprehensive	5.01%	7.01%	2.04%	3.50%
EPO	1.85%	0.00%	0.02%	0.00%
HMO	11.48%	20.81%	8.47%	16.50%
Non-cap POS	34.84%	17.95%	16.55%	13.94%
PPO	40.12%	54.23%	44.71%	66.06%
Cap POS	6.49%	0.00%	1.83%	0.00%
CDHP	0.06%	0.00%	16.79%	0.00%
HDHP	0.15%	0.00%	9.59%	0.00%
Annual Enrollees per Firm	96755	65365	86363	63741
Firms	7	23	7	23
Person-Years	1,354,574	3,590,716	1,813,625	5,230,468

Notes: Descriptive statistics presented over the pre and post-period for treatment (high-deductible offering) and control (no high-deductible offer) firms. The pre-period for the treatment firms are the 2 years prior to the offer of a plan with a deductible greater than or equal to $500, and the post-period is the three consecutive years following this offer where such an offer exists in all 3 years. The pre-period for control firms is a weighted average of calendar year 2008 through 2010 matching the enrollee-level distribution of calendar years in the pre-period for the treatment firms, and the post-period is similarly a weighted average of 2010–2013.

that is most noticeable in Fig. 2A and C, where spending dips most dramatically in Q1 of every post-period year, and then quickly rises, peaking in Q4 of each post-year. This is often observed in studies in which deductibles are influencing behavior, as individuals appear to be highly myopic and affected by spot-prices as opposed to expectations of annual spending, even for those who expect to exceed their deductible by year's end (Brot-Goldberg et al., 2017). That is, in the beginning of the calendar year when the annual deductible resets, enrollees are less likely to consume services, whereas by the late year many enrollees have reached their deductible limit and no longer face the same levels of cost-sharing, resulting in increased spending. These patterns, clearly visible in Panels A and C, and suggestively present in Panels B, D, and E, provide further evidence that the increased deductible is contributing to the differences in spending behavior in our treatment firms. Similar trends can be seen when disaggregating the treatment firms into those offering at least one plan with a $500, $1000, and $1250 deductible, as can be seen in Appendix Fig. A1. The unadjusted trends seem to indicate a gradient effect, with the subset of firms that offer plans with higher deductibles appearing to experience greater reductions in spending after the offer.

Table 2 shows the results of our main ITT analysis. Panel A has values for our main analytic sample where enrollees were only required to be enrolled for at least one calendar year to be included in the sample. The average marginal effect of firm offer is statistically significant for all categories, with absolute reductions in low-value (total) spending of $5.23 ($105.77) for outpatient services, $4.81 ($47.32) for imaging services, and $1.41 ($22.74) for laboratory services.

Since the magnitude in dollars of low-value and total spending for a category are so different, it is helpful to estimate a percent change of the respective spending category to enable us to compare the effects of the reductions in low-value spending in relation to total spending. Here we saw significant low-value (total) reductions of 13.7% (5.2%) in outpatient spending, 21.2% (17.7%) in imaging spending, and 22.2% (13.6%) in laboratory spending. To compare whether spending reductions come disproportionately from low-value spending, we compare the magnitudes of these low-value and total spending coefficients in each category with a Z-statistic developed to test for differences in regression coefficients computed from different models using the same sample that has been used in prior literature (Clogg et al., 1995; Greenland et al., 1999;

Table 2
Marginal Impact of HDHP Offer on Spending.

	Outpatient		Imaging		Laboratory	
Panel A: 12 month continuous enrollment	Low-Value	Total	Low-Value	Total	Low-Value	Total
Dif-in-Dif (Δ\$)	−5.23***	−105.77***	−4.81***	−47.32***	−1.41***	−22.74***
	(1.36)	(34.72)	(0.81)	(14.64)	(0.19)	(6.90)
Dif-in-Dif (Δ%)	−0.137***	−0.052***	−0.212***	−0.177***	−0.222***	−0.136***
	(0.035)	(0.016)	(0.035)	(0.053)	(0.027)	(0.039)
Mean Costs	39.51	2211.15	23.24	278.34	6.94	176.96
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Panel B: Continuously enrolled the entire observation period	Low-Value	Total	Low-Value	Total	Low-Value	Total
Dif-in-Dif (Δ\$)	−5.98***	−136.51***	−4.92***	−51.62***	−1.55***	−24.81***
	(1.34)	(33.26)	(0.81)	(13.54)	(0.22)	(7.12)
Dif-in-Dif (Δ%)	−0.166***	−0.067***	−0.227***	−0.196***	−0.242***	−0.147***
	(0.036)	(0.015)	(0.036)	(0.049)	(0.032)	(0.040)
Mean Costs	37.31	2184.42	22.19	273.28	6.96	176.73
Observations	6891863	6891863	6891863	6891863	6891863	6891863
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Panel A: Model used is a two-part model with a probit first stage and GLM second stage with gamma family and log link using Stata's 'twopm' command (Belotti et al., 2015). Average marginal effects using the 'dydx' option of the 'margins' command for the row labelled (Δ\$) and 'eydx' for the row labelled for the row labelled (Δ%) are presented. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed. Standard errors are clustered at the firm level. Panel B: The same as Panel A, but subset by those who are continuously enrolled in a firm from 2008 to 2013

Paternoster et al., 1998; Stafford et al., 2007). Results of this comparison between low-value and total spending coefficients within each category returns Z-scores of -2.21 for outpatient spending, −0.55 for imaging spending, and -1.81 for laboratory spending. The interpretation of these Z-scores using a conservative two-tailed Z-test is that there is a significant difference between low-value and total outpatient reductions ($p < 0.05$), no measurable difference in reductions in low-value and total imaging spending, and a borderline significant difference in low-value and total laboratory reductions ($p < 0.10$).

Table 2 Panel B presents the same estimates after restricting the sample to those who were continuously enrolled for the entire observation period to control for the effects of changing firm composition in response to the high-deductible plan offer. The results from these models have the same sign and significance, and are qualitatively similar to those using the full sample, as are the results of the Z-test for significant spending differences.

Table 3 shows the results of our LATE analysis, using firm offer as an instrument for enrollment in a high-deductible plan to estimate the local average treatment effect of switching into a high-deductible plan on spending. The first stage is the same across all models where high-deductible plan offer is significantly associated with high-deductible plan enrollment ($F = 37.49$). In Panel B, we show the second stage estimates of predicted high-deductible plan enrollment on spending. We estimate highly significant reductions in low-value (total) spending of \$6.40 (\$193.72) for outpatient spending, \$5.70 (\$67.27) for imaging spending, and \$2.56 (\$40.62) for laboratory spending.

Table 4 shows the effects of increasingly large deductible offers on spending. Table 4 Panel A show the ITT effect of a firm whose highest-level plan has an estimated deductible that falls within the range of \$500-\$999 in the first tranche, \$1000-\$1249 in the second tranche, and exceeds \$1250 in the third tranche. The first row within each tranche is the ITT estimate in absolute dollars, the second row is the standard error, and the third row is the share of enrollees in a firm who are enrolled in a plan with a deductible of at least \$500. The rationale for including the share of enrollees in a plan of at least \$500 instead of the share of those in the highest-level plan is that firms often offer plans with a range of deductibles, and the ITT effect captures all of these plans, not simply the highest-leveled one.

One might expect to see a gradient effect in the results as the level of the deductible offer increases, which does not always hold in Panel A. However, this would not be expected if the share of those in high-deductible plans decreased as the deductible increases, as is the case in our data. In order to interpret the effects in light of the uptake rate, we performed a back-of-the-envelope calculation in Table 4 Panel B in which we divide the ITT coefficient by the rate of uptake, approximating the effect on those who actually enrolled in a high-deductible plan in the spirit of a LATE analysis. While significance testing is not appropriate for these back-of-the-envelope estimates, the results return the expected increasing gradient in effect size as the level of the deductible offer increases.

Table 5 shows the decomposition of the spending reductions into reductions in the quantity of services received and the average price per service. Here, the rows represent the dependent variable in each model and the columns represent whether the dependent variable is referring to annual counts per enrollee, or average price per service per enrollee conditional on having at least one of the services indicated. Looking first at the aggregate measures, we see a 15.55% (11.93%) decrease in the number of low-value (total) imaging procedures, a 15.39% (10.59%) decrease in low-value (total) laboratory procedures, and 12.79% (5.61%) decrease in the quantity of all outpatient procedures, all of which are statistically significant. For price, we observe an 8.65% (10.77%) decrease in low-value (total)

Table 3
Local Average Treatment Effect of Switching into a HDHP on Annual Spending.

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
Panel A: First Stage						
HDHP Offer	0.614***	0.614***	0.614***	0.614***	0.614***	0.614***
	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
F-Stat	37.49	37.49	37.49	37.49	37.49	37.49
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Panel B: Second Stage						
HDHP Enrollment	−6.40***	−193.72***	−5.70***	−67.27***	−2.56***	−40.62***
	(2.45)	(62.68)	(2.17)	(25.46)	(0.87)	(13.09)
Mean Costs	39.51	2211.15	23.24	278.34	6.94	176.96
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Model used is two-stage least squares using Stata's 'ivreg2' command. The first stage is a linear probability model predicting enrollment in a plan with a deductible of at least $500 and the second stage is the effect of predicted enrollment in such a plan on spending for each outcome. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed. Standard errors clustered at the firm level are displayed. Coefficients represent the change (in $) in annual spending as a result of switching into a plan with a deductible of at least $500.

Table 4
Marginal Impact of HDHP Offer by Maximum Deductible Offered.

	Outpatient		Imaging		Laboratory	
Panel A: ITT by maximum deductible offered	Low-Value	Total	Low-Value	Total	Low-Value	Total
ITT (Δ$) Firm Offer $500-$999	−6.65***	−41.83	−4.18***	−29.46***	−1.91***	−24.38***
SE	(0.84)	(60.32)	(0.62)	(9.91)	(0.15)	(5.01)
Uptake	0.86	0.86	0.86	0.86	0.86	0.86
ITT (Δ$) Firm Offer $1000-$1249	−4.00***	−136.67***	−4.23***	−44.74***	−0.90***	−16.50**
SE	(0.79)	(31.17)	(0.79)	(14.36)	(0.2)	(7.88)
Uptake	0.58	0.58	0.58	0.58	0.58	0.58
ITT (Δ$) Firm Offer $1250+	−5.92***	−124.39***	−6.47***	−69.86***	−1.54***	−30.51***
SE	(2.06)	(34.72)	(1.46)	(23.69)	(0.31)	(10.91)
Uptake	0.44	0.44	0.44	0.44	0.44	0.44
Panel B: Estimated Treatment on the Treated						
Dif-in-Dif (Δ$) Firm Offer 500	−7.75	−48.74	−4.87	−34.33	−2.23	−28.41
Dif-in-Dif (Δ$) Firm Offer 1000	−6.95	−237.38	−7.35	−77.71	−1.56	−28.66
Dif-in-Dif (Δ$) Firm Offer 1250	−13.34	−280.33	−14.58	−157.44	−3.47	−68.76
Mean	39.51	2211.15	23.24	278.34	6.94	176.96
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Standard errors clustered at the firm level are displayed. In Panel A, each tranche represents the effects of a high-deductible offer for firms whose maximum offered deductible was (1) between $500-$999, (2) between $1000-$1249, and (3) above $1250. Note that a firm may have offered multiple plans with deductible levels below these amounts, but must have offered at least one plan with an estimated deductible level in the respective range. Additionally, a firm may be in one tranche in one year, but in another tranche in another year if they increased or decreased the deductible level of their highest-level plan. The first row in each tranche is the coefficient on the indicator variable for that deductible level (treatXdeductible500_999). The second row is the standard error of the coefficient. The third row represents the percent of enrollees in these firms enrolled in a high-deductible plan (at least $500). Panel B uses the coefficient and uptake rates from Panel A to crudely calculate the effects on those who enrolled in a high-deductible plan in each firm (coefficient/uptake).

outpatient services and a 5.26% decrease in total lab procedures, with the other measures not being measurably different from 0. In general, every individual low-value procedure has a downward trend in average quantity received, and 11 out of 24 of these were statistically significant. Compared to price reductions, we observe a mostly downward trend with some exceptions (e.g. renal stenting), though far fewer procedures attain statistical significance, with only 5 of 24 significantly decreasing and renal stenting increasing significantly. A scan of the table indicates the point estimates of the quantity reductions are generally of a larger magnitude than those of price reductions (in 22 of 24 individual procedures and in all aggregate groupings of procedures). Overall, we provide some evidence to suggest that our spending reductions are mainly due to quantity reductions, though there does appear to be some price-shopping as well.

4.2. Robustness check results

To test whether our results were robust to fourth quarter anticipatory stockpiling behavior, we re-estimated our models shifting the post-period back to include Q4 of the immediate pre-period to account for the effect of antici-

Table 5
Impact of Deductible Offer on Annual Quantity and Price per Service for each Low Value Service.

	Quantity		Price	
Procedure	Dif-in-Dif (Δ%)	SE	Dif-in-Dif (Δ%)	SE
All Imaging Procedures	−0.1193***	(0.0459)	−0.1077**	(0.0452)
Low Value Imaging Procedures	−0.1555***	(0.0562)	−0.0865***	(0.0216)
Preoperative Stress Testing	−0.1009	(0.0617)	−0.0507	(0.0384)
Stress Testing for Coronary Disease	−0.1636	(0.1376)	−0.1176**	(0.0440)
Carotid Scanning (Asymptomatic)	−0.0964	(0.0633)	−0.0151	(0.0471)
Carotid Artery Screen for Syncope	−0.0468	(0.0812)	0.0352	(0.0927)
Adnexal Cyst Imaging	−0.1687**	(0.0745)	−0.0432	(0.0495)
CT for Rhinosinusitus	−0.2072***	(0.0665)	−0.0517	(0.0372)
Head Imaging for Syncope	−0.1892**	(0.0764)	−0.2166**	(0.0866)
Head Imaging for Headache	−0.1858***	(0.0454)	−0.1342***	(0.0403)
Back Scan	−0.1447**	(0.0658)	−0.0985***	(0.0289)
Imaging for Plantar Fasciitis	−0.1608**	(0.0758)	−0.0142	(0.0636)
Bone Density Testing	−0.1534	(0.0997)	−0.0263	(0.0593)
Preoperative Radiography	−0.1389*	(0.0714)	−0.1327*	(0.0686)
All Laboratory Procedures	−0.1059**	(0.0459)	−0.0526*	(0.0269)
Low Value Laboratory Procedures	−0.1539*	(0.0786)	−0.0651	(0.0430)
Vitamin D Screening	−0.1356	(0.1082)	−0.0557	(0.0529)
Homocysteine Testing	−0.0992	(0.1483)	0.0989	(0.1182)
Hypercoagulability in DVT	−0.3138***	(0.1136)	−0.5129	(2.9913)
HPV Testing in Women < 30	−0.0376	(0.0724)	0.0000	(0.0000)
PTH Testing in CKD	−0.0696	(0.0718)	−0.0200	(0.0479)
T3 Measurements in Hypothyroidism	−0.1500**	(0.0596)	−0.0383	(0.0461)
All Outpatient Procedures	−0.0561***	(0.0213)	0.0277	(0.0183)
(Other) Low Value Outpatient Procedures	−0.1279**	(0.0501)	−0.0443	(0.0470)
EEG for Headache	−0.0375	(0.0867)	−0.0667	(0.0987)
IVC Filter Placement	−0.0841	(0.2489)	−0.0528	(0.1086)
Preoperative Echocardiography	−0.1579**	(0.0803)	−0.0005	(0.0485)
Preoperative PFT	−0.2078***	(0.0629)	−0.1043	(0.1388)
Renal Stenting	−0.2992	(0.3968)	0.4540*	(0.2482)
Spinal Injections for Low Back Pain	−0.0021	(0.0331)	−0.0104	(0.0859)

*** $p < 0.01$, **$p < 0.05$, *$p < 0.10$. Unit of observation is person-year. Standard errors clustered at the firm level are displayed. The first column represents the dependent variable in each regression, where each one is one of 24 low-value services or an aggregation of these services. Column 2 represents the percent change in the mean number of each service received per person, per year. Model used is a negative binomial model using Stata's 'nbreg' command, with average marginal effects taken using the 'eydx' option of the 'margins' command. The third column is the standard error for this regression. Column 4 is the percent change in the price per service for each of these services, conditional on the enrollee receiving at least one of the services in a given calendar year. The model used is OLS, with average marginal effects taken using the 'eydx' option of the 'margins' command. The final column is the standard errors from this regression.

Table 6
Effect of HDHP Offer and Enrollment on Quarterly Spending Treating the Fourth Quarter of the Pre-Year as the Post-Period.

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
Panel A: ITT						
Dif-in-Dif (Δ$)	−1.15***	−14.82***	−1.00***	−9.79***	−0.34***	−4.68***
	(0.32)	(5.63)	(0.18)	(3.08)	(0.04)	(1.40)
Dif-in-Dif (Δ%)	−0.12***	−0.03***	−0.17***	−0.14***	−0.21***	−0.11***
	(0.03)	(0.01)	(0.03)	(0.04)	(0.03)	(0.03)
Panel B: LATE						
HDHP Enrollment (Δ$)	−1.45**	−35.27**	−1.27**	−15.89**	−0.70***	−9.90***
	(0.66)	(13.83)	(0.59)	(6.42)	(0.24)	(3.47)
Mean Costs	9.88	552.79	5.81	69.58	1.73	44.24
Observations	47957532	47957532	47957532	47957532	47957532	47957532
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** $p < 0.01$, **$p < 0.05$, *$p < 0.10$. Unit of observation is person-quarter. The pre-period is defined as the first seven quarters in the 2 calendar years before the offer of a high-deductible plan. The post-period consists of Q4 of the calendar immediately year before the firm offer plus 3 calendar years following offer. Standard errors are clustered at the firm level. Panel A: Model used is a two-part model with a probit first stage and GLM second stage with gamma family and log link using Stata's 'twopm' command. Average marginal effects using the 'dydx' option of the 'margins' command for the row labelled (Δ$) and 'eydx' for the row labelled for the row labelled (Δ%) are presented. Panel B: Model used is two-stage least squares using Stata's 'ivreg2′ command. The first stage is a linear probability model predicting enrollment in a plan with a deductible of at least $500 and the second stage is the effect of predicted enrollment in such a plan on spending for each category. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed.

patory stockpiling of services. These results can be found in Table 6. In the ITT analysis (Panel A), the proportional reductions in spending are smaller than in other specifications, though their directionality and significance are the same as in our main specification and other sensitivity analyses. The LATE analysis also retains overall directionality and significance in all models.

Appendix Table A8 shows the results of analysis when dropping enrollees in our treatment firms who take up alternative, low-deductible plan types in the post-period. The threat that this table seeks to address is if treatment firms are encouraging reductions in spending through unobserved changes to plan benefits beyond the deductible, which would bias our result towards finding larger ITT and LATE results than can be causally attributed to high-deductible plans. While this problem cannot be completely accounted for without detailed information on plan design that is not available in our data, we attempted to address this concern by systematically excluding one low-deductible plan-type at a time to see if it changes the magnitude of our ITT estimates. If unobserved characteristics of other plan types (e.g. narrower physician networks in PPOs) were contributing to spending reductions, we would expect the magnitude of our ITT results to decrease relative to our main findings, and if those unobserved characteristics increased spending, then our main results would represent a lower-bound for the effect of the high-deductible offer. As can be seen in Appendix Table A8, when comparing estimates in tranche 2 through 6 with the top tranche (which is our main result from Table 2), the ITT estimates actually increase in every model, suggesting that enrollees in these low-deductible plans may actually be biasing us against finding reductions. The seventh tranche (the effects when excluding capitated POS plans) has some estimates that are slightly less than our main results, but no deviations change the result in a qualitatively significant way.[5] The final tranche acts as a placebo test and shows the results when dropping those in high-deductible plans. Here we see that in four of six models, there is no significant decrease in spending if we do not include those who enroll in a high-deductible plan. However, both low-value imaging and low-value lab spending show significant reductions, and all other outcomes have a downward trend in spending. This suggests that other unobservable plan characteristics besides deductibles may be contributing to our observed reductions in spending, though these contributions do not appear to be the primary mechanism through which we spending reductions occur, as evidenced by most of these models showing no measurable difference in spending.

To test whether our results are robust to spending outliers, we performed the analyses using unwinsorized spending outcomes (Appendix Table A6). Including these spending outliers shows results that are qualitatively similar to our main specification, though, unsurprisingly, with increased variance around the estimates. The sign of our results is the same in all models, and the significance and magnitude are the same for all models except total outpatient spending, which is no longer significantly different from zero in the ITT analysis (Panel A), though the point estimate is still negative and qualitatively large.

While our tests for parallel pre-trends gives us confidence that there were no spending-related abnormalities that may have prompted our treatment firms to offer high-deductible plans, there is a worry that the firm offer was not completely exogenous. To test for this, we regressed the decision to offer a high-deductible on observable firm characteristics, including share of males, average enrollee age, average baseline Charlson Comorbidity score of enrollees, and average number of enrollees per year. As shown in Appendix Table A9, we find no significant association between these characteristics and the decision to offer a high-deductible plan, adding to our confidence that the plan offer was as good as exogenous.

Finally, in accordance with the recent literature on differences-in-differences designs with time-varying treatments, we performed a treatment effect decomposition to determine what type of comparisons were driving our results (Goodman-Bacon, 2019). We find in Appendix Figs. A2–A7 that our results are overwhelmingly driven by comparisons of our treatment firms with our 'untreated' firms, with weights on the comparison of 'early' high-deductible adopters to 'late' adopters being nearly zero.[6] We also note that the weights on the difference-in-differences estimators between early and late adopters are negative, indicating that there is heterogeneity in the treatment effects over time, which may constitute a violation of the assumptions of our time-varying difference-in-differences design according to the Goodman-Bacon time-varying difference-in-differences decomposition theorem. Goodman-Bacon's recommendation in this case is to re-estimate the model using an event-study design, which we provide in Appendix Table A2. The results from the event-study model are very similar to those found in all of our other model specifications, leading us to believe that the violations of the assumption of treatment homogeneity over time is not producing a meaningful bias in our overall results.

5. Discussion

While there is a large and growing literature on both the effects of high-deductible plans and low-value services, ours is the first study to our knowledge whose design allows for a rigorous estimation of the causal effect of high-deductible plans on measures of low-value spending. Our

[5] The largest change by proportion is total laboratory spending, which goes from a reduction of $22.74 to $22.35, representing a 1.7% decrease. Low-value lab spending goes from a reduction of $1.41 to $1.39 (−1.4%), total imaging goes from $47.32 to $46.90 (−0.8%) and low-value outpatient goes from $5.23 to $5.21 (−0.3%).

[6] In the parlance of Goodman-Bacon (2019), untreated firms are those that never receive the treatment, and this is distinct from control firms, since according to his decomposition theorem, when there is difference in the time of treatment, early treatment firms will eventually act as controls for later treatment firms, and vice-versa. In this paper, we have simply referred to our untreated firms as control firms, which is admittedly less precise, though more intuitive given that the nature of this paper is not methodological.

ITT analysis found that those that were offered a high-deductible plan spent $5.23 less on low-value outpatient services, representing a 13.7% decrease. Our results also provide insight into the more nuanced question of the ability of high-deductible plans to induce efficient spending, as opposed to simply causing indiscriminate reductions in spending. Since it is reasonable to expect that encouraging efficient spending would disproportionately reduce low-value spending, our findings suggest that high-deductible plans are a potentially viable way to encourage value-based decision making.

While our findings are broadly consistent with the most closely related study, Reid et al. (2017), in that total spending for each of our three categories decreases, our present findings depart from this earlier study in that we find suggestive evidence that high-deductible plans reduce low-value spending disproportionately relative to overall spending. We have identified two potential reasons for the discrepancies. First, our analysis has a much larger sample size (3,168,199 treatment person-years compared to 11,075 treatment person-years), which allows us to detect smaller differences. The prior 2017 study found signs on their point estimates that implied a trend towards disproportionate low-value spending reductions, but were not able to achieve conventional levels of statistical significance to draw such conclusions. Additionally, the empirical strategy employed in this study is robust to time-invariant unobservable characteristics of the enrollees who opt into high-deductible plans, whereas the Reid et al. study relied on a matching-on-observables strategy that may have allowed unobserved differences in the treatment and control groups to bias the results.

The findings warrant further investigation given their departure from such benchmark studies as the RAND Health Insurance Experiment (HIE), which concluded that deductibles resulted in across-the-board spending reductions as opposed to value-based ones (Manning et al., 1987). One potential reason for the differences is the sample. Whereas the RAND HIE used a random sample of the population, our study focuses exclusively on those who are employed, and therefore less likely to be of lower socioeconomic status than those in the HIE. This SES discrepancy between modern observational studies and the HIE has been acknowledged in a more recent reflection on the classic experiment, which suggests that those of low SES may be particularly susceptible to indiscriminate spending reductions (Chernew and Newhouse, 2008). Additionally, as evidenced by the very existence of the Choosing Wisely campaign from which our measures were derived, practitioners have been becoming increasingly aware of the adverse effects of remaining value-agnostic. It has been found in other recent studies that physicians take the financial burden of treatment on patients into account when prescribing drugs when information on minimizing costs is available (Carrera et al., 2018). It is possible that this increased awareness could prompt providers to help patients make value-based decisions in instances where the patient is more exposed to the financial burden of treatment.

Results of our decomposition analysis, while hindered by the relative rarity of the set of low-value services used in our analysis, imply that these low-value spending reductions are coming disproportionately from reductions in utilization of these services, either because patients are avoiding these services entirely or are substituting more valuable treatments. This further bolsters our interpretation that high-deductible plans incentivize value-based decision making, as patients are not simply looking for better prices on care that is widely considered to be useless.

It should be noted that there was no significant difference in the reductions in low-value and total spending for imaging services. While this may just be caused by our inability to detect the difference given the variability in spending for these services, it warrants further exploration. As alluded to above, the presence of asymmetrical information in medical care markets makes it plausible that it is a combination of a patient's financial incentive and a provider's participation in steering patients away from low-value services that contribute to the reductions. If this were the case, the lack of value-based reductions may be the result of less physician participation in the utilization decision. Providers may be less likely to make this distinction for imaging services than for others services due to the financial incentives surrounding them. Prior studies have found that physician behavior is particularly responsive to the financial incentives regarding imaging services, perhaps due to high costs of acquiring the advanced technology required to perform the services, or the favorable reimbursement rates for these services (Baker, 2010; Clemens and Gottlieb, 2014). Another explanation is that physicians may be less aware of the relative value of various imaging services relative to laboratory or other outpatient services. Ultimately though, spending did decrease substantially for both low-value and total imaging, regardless of the relative value of these reductions.

Our findings have relevant policy implications. This study indicates that deductibles may be considered one of the most useful tools in our toolkit for incentivizing value-based decision making in a time when much attention is being given to incorporating value-based mechanisms into our healthcare system. Given these findings, managers and policymakers should consider encouraging greater use of plan designs with sizeable deductibles. However, while our study has optimistic results, it should be noted that this study did not directly test for adverse health outcomes, and was unable to explore individual-level heterogeneity. Of particular concern is that those of lower socioeconomic status may be especially sensitive to high deductibles and therefore more prone to eliminating valuable services, promulgating existing health inequities. As has been demonstrated in the case of low-income diabetes patients, high-deductible plans can lead to a rationing of services that result not only in harmful health effects, but more expensive emergency department care in the long-run (Wharam et al., 2017). Care should be taken in the application of our findings to real-world settings, with particular attention paid to potential negative effects on vulnerable populations. Further study is needed to determine the generalizability of our results to different subpopulations.

Our study has provided robust results indicating that high-deductible plans are effective at reducing low-value spending, and we present mixed evidence that they do so by promoting value-based decision making. This is cause for optimism among proponents of patient-side cost-sharing. Future research should look at the changes in health outcomes of patients enrolling in these plans, as well as explore potential heterogeneity in their effects. Overall, we conclude that high-deductibles plans offer an effective and promising way to bend the healthcare cost-curve, and do so by encouraging value-based decision making.

Funding source

Agency for Healthcare Research and Quality and NIHCM Foundation.

This project was supported by grant number T32HS000029 from the Agency for Healthcare Research and Quality, grant number T32MH122357 from the National Institute of Mental Health and by a grant from the NIHCM foundation. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality.

CRediT authorship contribution statement

Brendan Rabideau: Conceptualization, Methodology, Software, Validation, Formal analysis, Data curation, Writing - original draft, Writing - review & editing, Visualization, Project administration. **Matthew D. Eisenberg:** Conceptualization, Methodology, Writing - review & editing, Supervision, Project administration. **Rachel Reid:** Conceptualization, Methodology, Writing - review & editing. **Neeraj Sood:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision, Project administration, Funding acquisition.

Acknowledgments

The authors wish to thank Jean Roth and NBER for their helpful assistance in obtaining the Truven Marketscan Data.

Appendix A.

Table A1
Specifications for Low-Value Health Care Service Measures.

Procedure	Definition	Criteria	Applicable Codes
Imaging Procedures			
	Carotid imaging with syncope diagnosis for patients without stroke or TIA history, and without diagnosis of stroke, TIA, or focal neurologic symptoms in claim	Identifying CPTs	36222, 36223, 36224, 70498, 70547, 70548, 70549, 93880, 93882, 3100F
	Inclusion	1-year look-back, ICD9:7802, 9921	
	Exclusion	Stroke/TIA by CCW, ICD9: 430, 431, 43301, 43311, 43321, 43331,43381, 43391, 43400, 43401, 43410, 43411,43490, 43491, 4350, 4351, 4353, 4358, 4359, 436, 99702, V1254, 3623, 36284, 781xx, 7820, 78451, 78452, 78459, 781xx	
	Additional Costs		
	Carotid imaging not associated with inpatient or emergency care for patients without a history of stroke or TIA and without a diagnosis of stroke, TIA, or focal neurologic symptoms in claim	Identifying CPTs	36222, 36223, 36224, 70498, 70547, 70548, 70549, 93880, 93882, 3100F
	Inclusion	1-year look-back	
	Exclusion	Stroke/TIA by CCW, Hospitalization associated with ED or ED up to 14 days before procedure, ICD9: 430, 431, 43301, 43311, 43321, 43331, 43381, 43391, 43400, 43401, 43410, 43411, 43490, 43491, 4350, 4351, 4353, 4358, 4359, 436, 99702, V1254, 3623, 36284, 7802, 781xx, 7820, 78451, 78452, 78459, 781xx	
	Additional Costs		
	Stress test not associated with inpatient or ED care for patients with an established diagnosis of acute myocardial infarction (≥3 mo before)	Identifying CPTs	93015, 93016, 93017, 93018, 93350, 93351, 78451, 78452, 78453, 78454, 78460, 78461, 78464, 78465, 78472, 78473, 78481, 78483, 78491, 78492
	Inclusion	AMI by CCW > 3 months before procedure, 1-year look-back	
	Exclusion	Hospitalization associated with ED, or ED up to 14 days before procedure	
	Additional Costs	Any of the following: CPT: 93303–93352, 93000–93042, 78414–78499, 75552–75564, 75571–75574, A9500-A9700, J0150, J0152, J0280, J1245, J1250, J2785 on the same day	
	Two or more echography procedures within 60d of primary diagnosis of adnexal cyst	Identifying CPTs	76857, 76830
	Inclusion	1-year look-back, Prior cyst testing within 60 days, ICD9: 6200, 6201, 6202	

Table A1 (*Continued*)

Procedure Imaging Procedures	Definition	Criteria	Applicable Codes
	Exclusion		
	Additional Costs		
	Head CT or MR imaging with syncope diagnosis and no diagnoses in claim warranting imaging	Identifying CPTs	70450, 70460, 70470, 70551, 70552, 70553
	Inclusion	ICD9: 7802, 9921	
	Exclusion	ICD9: 78097, 7820, V1254, 345xx, 800xx, 801xx, 802xx, 803xx, 804xx, 850xx, 851,852xx, 853xx, 854xx, 870xx, 871xx, 872xx, 873xx, 910xx, 920xx, 921xx, 781xx, V10xx, 7803x, 7845x, 9590x, 43xxx	
	Additional Costs		
	Brain CT or MR imaging with non-posttraumatic, non-thunderclap headache diagnosis, and no diagnoses in claim warranting imaging	Identifying CPTs	70450, 70460, 70470, 70551, 70552, 70553
	Inclusion	ICD9: 30781, 7840, 339xx, 346x	
	Exclusion	ICD9: 33920, 33921, 33922, 33943, 4465, 78097, V1254, 345xx, 800xx, 801xx, 802xx, 803xx, 804xx, 850xx, 851xx, 852xx, 853xx, 854xx, 870xx, 871xx, 872xx, 873xx, 781xx, V10xx, 3463x, 3466x, 7803x, 7845x, 9590x, 43xxx, 140xx-208xx, 230xx-239xx	
	Additional Costs		
	Maxillofacial CT with sinusitis diagnosis and no sinusitis complications, immune deficiencies, nasal polyps, or head/face trauma in claim and no sinusitis diagnosis 30–365 d before imaging	Identifying CPTs	70486, 70487, 70488
	Inclusion	1-year look-back, 461xx, 473xx	
	Exclusion	No chronic sinusitis (previous sinusitis procedure occurring between 30days and 1 year before the current claim), ICD9: 07953, 37600, 2770x, 9590x, 471xx, 373xx, 800xx, 801xx, 802xx, 803xx, 804xx, 850xx, 851xx, 852xx, 853xx, 854xx, 870xx, 871xx, 872xx, 873xx, 910xx, 920xx, 921xx, 042xx, 279xx	
	Additional Costs		
	Bone density test within 2y of prior bone density test, with established osteoporosis diagnosis	Identifying CPTs	76977, 77078, 77079, 77080, 77083, 78350, 78351
	Inclusion	Prior bone density testing within 2 years, osteoporosis diagnosis within the last year (73300, 73301, 73302, 73303, 73309), 2-year look-back	

Table A1 (*Continued*)

Procedure Imaging Procedures	Definition	Criteria	Applicable Codes
	Exclusion		
	Additional Costs		
	Back imaging with low back pain diagnosis occurring within 6 wk of initial back pain diagnosis and no diagnoses in claim warranting imaging	Identifying CPTs	72010, 72020, 72052, 72100, 72110, 72114, 72120, 72200, 72202, 72220, 72131, 72132, 72133, 72141, 72142, 72146, 72147, 72148, 72149, 72156, 72157, 72158
	Inclusion	Within 6 weeks of first diagnosis of back pain, 1-year look-back, ICD9: 7213, 72190, 72210, 72252, 7226, 72293, 72402, 7242, 7243, 7244, 7245, 7246, 72470, 72471, 72479, 7385, 7393, 7394, 8460, 8461, 8462, 8463, 8468, 8469, 8472	
	Exclusion	No chronic history of back pain (former diagnosis > 6 weeks prior), ICD9: 92611, 92612, 304460, 4210, 4211, 4219, 78079, 01xxx, 86xxx, 952xx, 958xx, 959xx, 038xx, 730xx, 929xx, 7292x, 7830x, 7832x, 7808x, 2859x, 140xx-208xx, 230xx-239xx, 850xx-854xx, 800xx-839xx, 905xx-909xx, 3054x-3057x, 3040x-3042x	
	Additional Costs		
	Radiographic or MR imaging with plantar fasciitis diagnosis within 2 w of initial diagnosis	Identifying CPTs	73620, 73630, 73650, 73718, 73719, 73720, 76880, 76881, 76882
	Inclusion	ICD9: 72871, 7294	
	Exclusion		
	Additional Costs		
	Chest radiograph not associated with inpatient or ED care, $\leq$30d before low/intermediate risk non-cardiothoracic surgery	Identifying CPTs	71010, 71015, 71020, 71021, 71022, 71023, 71030, 71034, 71035
	Inclusion	Non-Cardiothoracic surgery happening up to 30 days in the future (CPT: 19120, 19125, 47562, 47563, 49560, 58558; BETOS: P1x, P3D, P4A, P4B, P4C, P5C, P5D, P8A, P8G)	
	Exclusion	Inpatient or Emergency Setting	
	Additional Costs	CPT: 93303–93352 on the same day	
	Stress EKG, echocardiogram, or nuclear imaging, not associated with inpatient or ED care, $\leq$30d before low/intermediate-risk non-cardiothoracic surgery	Identifying CPTs	78451, 78452, 78453, 78454, 78460, 78461, 78464, 78465, 78472, 78473, 78481, 78483, 78491, 78492, 93015, 93016, 93017, 93018, 93350, 93351
	Inclusion	Non-Cardiothoracic surgery happening up to 30 days in the future (CPT: 19120, 19125, 47562, 47563, 49560, 58558; BETOS: P1x, P3D, P4A, P4B, P4C, P5C, P5D, P8A, P8G)	

Table A1 (*Continued*)

Procedure Imaging Procedures	Definition	Criteria	Applicable Codes
	Exclusion	Inpatient or Emergency Setting	
	Additional Costs		
Laboratory Procedures			
	Calcitriol test without hypercalcemia, secondary hyperparathyroidism, or other hypercalcemia condition (sarcoidosis, TB, or selected neoplasms) in claim, or CKD history; no hypercalcemia diagnosis in past 30d	Identifying CPTs	82652
	Inclusion	1-year look-back	
	Exclusion	CKD by CCW, ICD9: 27542, 58881, 1890, 1891, 1830, 135xx, 173xx, 174xx, 175xx, 188xx, 200xx, 201xx, 202xx, 203xx, 204xx, 205xx, 206xx, 207xx, 208xx, 01xxx	
	Additional Costs		
	Homocysteine test with no diagnoses of folate or vitamin B12 deficiencies in claim and no folate or vitamin B12 test in prior claims	Identifying CPTs	83090
	Inclusion	1-year look-back	
	Exclusion	History of B12 or Folate Disorders (2662, 2704, 2810, 2811, 2812, 2859)	
	Additional Costs	CPT: 36415 on the same day	
	HPV test in female patients younger than age 30	Identifying CPTs	87622, 87620, 90649, 87621, 90650
	Inclusion	Female, <30 years old	
	Exclusion		
	Additional Costs		
	Hypercoagulable state laboratory test within 30d after lower extremity DVT or PE diagnosis; no evidence of recurrent thrombosis (i.e., DVT or PE diagnosis >90 d before claim)	Identifying CPTs	83090, 85300, 85303, 85306, 85613, 86147
	Inclusion	Deep Vein Thrombosis/Pulmonary Embolism diagnosis within 30 days (4151, 4510, 45111, 45119, 4512, 45181, 4519, 4534, V1251), 1-year look-back	
	Exclusion	Recurrent Deep Vein Thrombosis/Pulmonary Embolism (defined as a DVT/PE diagnosis >90 days before the current diagnosis)	
	Additional Costs	CPT: 83890–83914 on the same day	
	PTH test for CKD; no dialysis services before or $\leq$30 d after test, no hypercalcemia diagnosis during year	Identifying CPTs	83970

Table A1 (*Continued*)

Procedure	Definition	Criteria	Applicable Codes
Imaging Procedures			
	Inclusion	CKD by CCW, 1-year look-back	
	Exclusion	No prior dialysis, no upcoming dialysis within 30 days, no hypercalcemia in 2009	
	Additional Costs	36415 on the same day	
		Identifying CPTs	84480, 84481
	Total or free T3 measurement in patient with hypothyroidism diagnosis during year		
	Inclusion	Hypothyroidism within 1 year (244xx), 1-year look-back	
	Exclusion		
	Additional Costs		
Other Outpatient Procedures			
		Identifying CPTs	Before 2012: 75940 In and After 2012: 37191
	Any IVC filter placement		
	Inclusion		
	Exclusion		
	Additional Costs	Any of the following: CPT: 36010, 37620, 75825, 76937 on the same day	
		Identifying CPTs	35471, 35450, 37205, 37207, 75966, 75960
	Renal/visceral angioplasty or stent placement with renal atherosclerosis or renovascular hypertension diagnosis in procedure claim		

Table A1 (*Continued*)

Procedure Imaging Procedures	Definition	Criteria	Applicable Codes
	Inclusion	ICD9: 4401, 40501, 40511, 40591	
	Exclusion		
	Additional Costs	All procedures occurring on the same day	
	EEG with headache diagnosis in claim, and no epilepsy or convulsions in current or prior claims	Identifying CPTs	95812, 95813, 95816, 95819, 95822, 95827, 95830, 95957
	Inclusion	1-year look-back, ICD9: 30781, 7840, 339xx, 346x	
	Exclusion	History of epilepsy (7803x, 7810x)	
	Additional Costs		
	Outpatient epidural, facet, or trigger point injections for low back pain, excluding etanercept; no radiculopathy diagnoses in claim	Identifying CPTs	62311, 64483, 20552, 20553, 64493, 64475
	Inclusion	Must be Outpatient or Office visit, ICD9: 7213, 72190, 72210, 7222, 72252, 7226, 72280, 72283, 72293, 72400, 72402, 72403, 7242, 7245,7246, 72470, 72471, 72479, 7384, 7385,7393,7384,7385, 7393, 7394, 75612, 8460, 8461, 8462,8463, 8468, 8469,8472	
	Exclusion	ICD9: 72142, 72191, 72270, 72273, 7243, 7244	
	Additional Costs	All procedures occurring on the same day	
	Echocardiogram not associated with inpatient or ED care, ≤30d before low/ intermediate-risk non-cardiothoracic surgery	Identifying CPTs	93303, 93304, 93306, 93307, 93308, 93312, 93315, 93318
	Inclusion	Non-Cardiothoracic surgery happening up to 30 days in the future (CPT: 19120, 19125, 47562, 47563, 49560, 58558; BETOS: P1x, P3D, P4A, P4B, P4C, P5C, P5D, P8A, P8G)	
	Exclusion	Inpatient or Emergency Setting	
	Additional Costs		
	PFT not associated with inpatient or ED care, ≤30d before low/intermediate-risk surgery	Identifying CPTs	94010
	Inclusion	Specified surgery happening up to 30 days in the future (BETOS: P1x, P2x, P3D, P4A, P4B, P4C, P5C, P5D, P8A, P8G)	
	Exclusion	Inpatient or Emergency Setting	
	Additional Costs	CPT: 94010–94799, 93720–93722 on the same day	

Table A2
Effect of Firm-level HDHP Offer on Annual Spending (by Post-year).

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
Dif-in-Dif (Δ$) (Post Year 1)	−4.85***	−112.60***	−3.58***	−34.51***	−1.05***	−15.07***
	(0.94)	(42.47)	(0.56)	(11.65)	(0.16)	(5.42)
Dif-in-Dif (Δ$) (Post Year 2)	−4.42***	−81.63**	−4.47***	−43.18***	−1.30***	−20.40**
	(1.54)	(36.06)	(0.97)	(15.32)	(0.23)	(8.30)
Dif-in-Dif (Δ$) (Post Year 3)	−6.55***	−121.55***	−6.88***	−67.87***	−1.95***	−34.55***
	(1.87)	(40.19)	(1.27)	(19.40)	(0.24)	(8.18)
Dif-in-Dif (Δ%) (Post Year 1)	−0.13***	−0.05***	−0.16***	−0.13***	−0.17***	−0.09***
	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)	(0.03)
Dif-in-Dif (Δ%) (Post Year 2)	−0.12***	−0.04**	−0.20***	−0.16***	−0.21***	−0.12***
	(0.04)	(0.02)	(0.04)	(0.05)	(0.03)	(0.05)
Dif-in-Dif (Δ%) (Post Year 3)	−0.17***	−0.06***	−0.30***	−0.25***	−0.30***	−0.21***
	(0.05)	(0.02)	(0.05)	(0.07)	(0.04)	(0.05)
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Pre-period is the calendar year before high-deductible offer, Post Year 1 is the DD estimator in the first year of the offer, Post Year 2 is the following year, and Post Year 3 is third year. Standard errors are clustered at the firm level. Model used is a GLM with gamma family and log link and included covariates and firm and calendar year fixed effects. Average marginal effects using the 'dydx' option of the 'margins' command for the rows labelled (Δ$) and 'eydx' for the row labelled for the rows labelled (Δ%) are presented. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed.

Table A3
Marginal Impact of HDHP Offer on Spending Under Different Model Specifications.

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
Panel A: Logit First Stage, Gamma Family						
Dif-in-Dif (Δ$)	−5.36***	−104.61***	−4.93***	−47.22***	−1.40***	−22.64***
	(1.36)	(34.73)	(0.81)	(14.62)	(0.19)	(6.91)
Dif-in-Dif (Δ%)	−0.14***	−0.05***	−0.21***	−0.18***	−0.21***	−0.14***
	(0.03)	(0.02)	(0.03)	(0.05)	(0.03)	(0.04)
Panel B: Probit First Stage, Gaussian Family						
Dif-in-Dif (Δ$)	−4.56***	−104.47***	−4.15***	−41.45***	−1.48***	−23.38***
	(1.36)	(30.61)	(0.71)	(11.33)	(0.20)	(6.20)
Dif-in-Dif (Δ%)	−0.12***	−0.05***	−0.18***	−0.16***	−0.23***	−0.14***
	(0.03)	(0.01)	(0.03)	(0.04)	(0.03)	(0.04)
Mean Costs	39.51	2211.15	23.24	278.34	6.94	176.96
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Standard errors clustered at the firm level are displayed. Panel A: Model used is a two-part model with a logit first stage and GLM second stage with gamma family and log link. Panel B: Model used is a two-part model with a probit first stage and GLM second stage with gaussian family and log link. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed. For both Panels, average marginal effects using the 'dydx' option of the 'margins' command for the row labelled (Δ$) and 'eydx' for the row labelled for the row labelled (Δ%) are presented.

Table A4
Quarterly Spending Prior to High-Deductible Offer.

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
PreYear2 X Treat	−0.01	−3.75	−0.11	−0.62	0.21***	0.95
(relative to immediate pre-year)	(0.23)	(3.10)	(0.10)	(1.34)	(0.04)	(1.08)
Observations	47276429	47276429	47276429	47276429	47276429	47276429
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-quarter. The coefficient represents the interaction of the treatment indicator with the 'PreYear2' indicator, and is the additional level change for enrollees in treatment firms in relative to the reference period, 'PreYear1'. Q4 of PreYear1 is omitted due to concerns of anticipatory stockpiling of services, as explored in Table 5. Standard errors are clustered at the firm level. Model used is a two-part model with a probit first stage and GLM second stage with gamma family and log link using Stata's 'twopm' command. Average marginal effects using the 'dydx' option of the 'margins' command. Standard errors are clustered at the firm level. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed.

Table A5
Marginal Impact of HDHP Offer on Spending Using Individual Fixed Effects.

	Outpatient		Imaging		Laboratory	
Sample: Continuously enrolled the entire observation period	Low-Value	Total	Low-Value	Total	Low-Value	Total
Dif-in-Dif (Δ$)	−5.93***	−141.40***	−4.35**	−50.11***	−1.80**	−28.38***
	(1.88)	(41.38)	(1.62)	(16.73)	(0.70)	(9.65)
Dif-in-Dif (Δ%)	−0.16***	−0.07***	−0.20**	−0.19***	−0.29**	−0.17***
	(0.05)	(0.02)	(0.08)	(0.06)	(0.12)	(0.06)
Mean Costs	37.31	2184.42	22.19	273.28	6.96	176.73
Observations	6891863	6891863	6891863	6891863	6891863	6891863
Individual Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Model used is a fixed effects regression using Stata's 'areg' command. Average marginal effects using the 'dydx' option of the 'margins' command for the row labelled (Δ$) and 'eydx' for the row labelled for the row labelled (Δ%) are presented. Individual fixed effects are used instead of firm fixed effects. Standard errors are clustered at the firm level.

Table A6
Effect of HDHP Offer and Enrollment on Annual, Unwinsorized Spending.

	Outpatient		Imaging		Laboratory	
	Low-Value	Total	Low-Value	Total	Low-Value	Total
Panel A: ITT						
Dif-in-Dif (Δ$)	−5.71***	−80.15	−5.61***	−74.09***	−1.56***	−35.87***
	(1.89)	(64.54)	(1.00)	(22.98)	(0.23)	(11.37)
Dif-in-Dif (Δ%)	−0.14***	−0.03	−0.23***	−0.21***	−0.23***	−0.16***
	(0.04)	(0.02)	(0.04)	(0.06)	(0.03)	(0.05)
Panel B: LATE						
HDHP Enrollment (Δ$)	−7.22**	−182.02**	−6.81***	−107.33***	−2.76***	−65.82***
	(2.82)	(89.46)	(2.38)	(38.72)	(0.92)	(19.14)
Mean Costs	43.67	3214.08	25.36	373.40	7.38	228.15
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Firm Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Covariates	Yes	Yes	Yes	Yes	Yes	Yes

*** p < 0.01, **p < 0.05, *p < 0.10. Unit of observation is person-year. Panel A: Model used is a two-part model with a probit first stage and GLM second stage with gamma family and log link using Stata's 'twopm' command. Average marginal effects using the 'dydx' option of the 'margins' command for the row labelled (Δ$) and 'eydx' for the row labelled for the row labelled (Δ%) are presented. Panel B: Model used is two-stage least squares using Stata's 'ivreg2′ command. The first stage is a linear probability model predicting enrollment in a plan with a deductible of at least $500 and the second stage is the effect of predicted enrollment in such a plan on spending for each category. Covariates include age, sex, and Charlson Comorbidity Score in the first year that a patient was observed. Standard errors are clustered at the firm level.

Table A7
Pre-Post Descriptive Summary by HDHP Uptake.

	Pre			Post		
	Treatment		Control	Treatment		Control
	Never HDHP	Future HDHP		Non-HDHP	HDHP	
Individual Characteristics						
Female	51.30%	53.52%	53.39%	50.85%	53.18%	53.10%
Age	43.76	43.08	45.15	44.67	42.59	44.07
Charlson Comorbidity Score	0.16	0.12	0.17	0.15	0.11	0.13
Expenditures						
Deductible Spending	$76.67	$106.04	$79.47	$71.17	$340.52	$93.21
Outpatient Spending	$2168.08	$2079.29	$2191.63	$2319.07	$2050.20	$2277.01
Imaging Spending	$231.96	$265.61	$290.52	$206.23	$228.40	$296.38
Laboratory Spending	$142.86	$155.11	$171.21	$154.30	$151.51	$195.78
Low-Value Outpatient Spending	$38.01	$36.53	$44.26	$33.02	$29.63	$40.25
Low-Value Imaging Spending	$20.95	$22.65	$26.50	$16.07	$16.57	$23.79
Low-Value Laboratory Spending	$5.50	$6.43	$5.91	$6.30	$6.75	$7.95
Deductible Amount						
Estimated Deductible Level	$119.36	$149.72	$96.56	$126.34	$690.55	$132.18
Low Deductible (<$500)	99.78%	99.59%	100.00%	100.00%	11.38%	100.00%
$500 -$ 999	0.00%	0.02%	0.00%	0.00%	72.42%	0.00%
$1000-$1249	0.01%	0.02%	0.00%	0.00%	8.00%	0.00%
$1250+	0.21%	0.37%	0.00%	0.00%	8.21%	0.00%

Table A7 (*Continued*)

	Pre			Post		
	Treatment		Control	Treatment		Control
	Never HDHP	Future HDHP		Non-HDHP	HDHP	
Firm Plan Types						
Comprehensive	6.19%	3.96%	7.01%	4.64%	0.93%	3.50%
EPO	2.70%	1.09%	0.00%	0.07%	0.00%	0.00%
HMO	16.16%	7.28%	20.81%	27.37%	0.46%	16.50%
Non-Capitated POS	37.35%	32.58%	17.95%	35.68%	8.46%	13.94%
PPO	30.96%	48.32%	54.23%	27.60%	51.95%	66.06%
Cap POS	6.55%	6.44%	0.00%	4.63%	0.64%	0.00%
CDHP	0.03%	0.09%	0.00%	0.00%	23.90%	0.00%
HDHP	0.06%	0.23%	0.00%	0.00%	13.65%	0.00%
Annual Enrollees per Firm	96755	96755	65365	86363	86363	63741
Firms	7	7	23	7	7	23
Person-Years	640,246	714,328	3,590,716	539,388	1,274,237	5,230,468

Notes: Descriptive statistics presented over the pre and post period for treatment (high-deductible offering) and control (no high-deductible offer) firms. The pre-period for the treatment firms are the 2 years prior to the offer of a deductible greater than or equal to $500, and the post-period is the three consecutive years following this offer where such an offer exists in all 3 years. The pre-period for control firms is a weighted average of calendar year 2008 through 2010 matching the enrollee-level distribution of calendar years in the pre-period for the treatment firms, and the post-period is a weighted average of 2010-2013. Never HDHP takers in the pre-period treatment firms are those who will not be enrolled in deductible plan > $500 in the post-period, whereas future HDHP takers are those that will enroll in one after the offer, and is presented to show heterogeneity for enrollees within the treatment firms.

Table A8
Effects of HDHP Offer Excluding Alternative Plan Designs.

	Outpatient		Imaging		Laboratory	
	Low Value	Total	Low Value	Total	Low Value	Total
Full Sample	−5.23***	−105.77***	−4.81***	−47.32***	−1.41***	−22.74***
	(1.36)	(34.72)	(0.81)	(14.64)	(0.19)	(6.90)
Observations	11989383	11989383	11989383	11989383	11989383	11989383
Drop Comprehensive Plans	−5.37***	−110.34***	−4.86***	−47.66***	−1.42***	−23.00***
	(1.35)	(36.41)	(0.83)	(14.83)	(0.19)	(6.94)
Observations	11925117	11925117	11925117	11925117	11925117	11925117
Drop EPOs	−5.24***	−105.80***	−4.82***	−47.35***	−1.41***	−22.76***
	(1.36)	(34.71)	(0.81)	(14.63)	(0.19)	(6.89)
Observations	11988975	11988975	11988975	11988975	11988975	11988975
Drop HMOs	−5.59***	−106.44***	−5.23***	−53.58***	−1.47***	−25.63***
	(1.53)	(36.50)	(0.78)	(14.10)	(0.18)	(6.77)
Observations	11750545	11750545	11750545	11750545	11750545	11750545
Drop Non-Capitated POS Plans	−6.93***	−150.35***	−5.48***	−52.53***	−1.51***	−25.46***
	(1.04)	(56.80)	(1.01)	(18.19)	(0.20)	(7.39)
Observations	11374013	11374013	11374013	11374013	11374013	11374013
Drop PPOs	−5.77***	−124.52***	−5.34***	−52.87***	−1.58***	−25.93***
	(1.45)	(34.28)	(0.83)	(14.15)	(0.18)	(6.23)
Observations	11713242	11713242	11713242	11713242	11713242	11713242
Drop Capitated POS Plans	−5.21***	−108.51***	−4.84***	−46.90***	−1.39***	−22.35***
	(1.29)	(36.19)	(0.83)	(14.90)	(0.18)	(6.79)
Observations	11896706	11896706	11896706	11896706	11896706	11896706
Drop > $500 Deductibles	−2.01	−30.09	−2.89***	−28.58	−1.32***	−13.38
	(2.77)	(66.27)	(0.89)	(15.41)	(0.32)	(8.38)
Observations	10000818	10000818	10000818	10000818	10000818	10000818

*** p < 0.001, **p < 0.01, *p < 0.05. Unit of observation is person-year. Model used is a two-part model with a probit first stage and GLM second stage with gamma family and log link using Stata's 'twopm' command. Average marginal effects using the 'dydx' option of the 'margins' command are presented. The first panel is identical to Panel A from Table 2 and serves as a baseline for comparison. Subsequent Panels exclude all instances for enrollees who were in the displayed plan-type in the post-period, in a treatment firm. All Panels save the final one keep enrollees in the stated plan type if they also had a deductible of at least $500 (i.e. in Panel 6, if an enrollee is both in a PPO and has a deductible of $500, they are not excluded).

Table A9
Probability of Firm Offer of a High-Deductible Plan Based on Observable Firm Characteristics.

Firm Characteristics	Probability of Deductible Offer
Percent Male	−0.41
	(0.90)
Average Comorbidity Score	−2.52
	(2.17)
Average Age	−0.01
	(0.02)
Number of Enrollees	0.00
	(0.00)
Observations	172

*$p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Unit of observation is the firm-year. Model run is a probit model using Stata's 'probit' command, and coefficients presented are the marginal effects returned using the 'margins' command's 'dydy' option. Standard errors are clustered at the firm level.

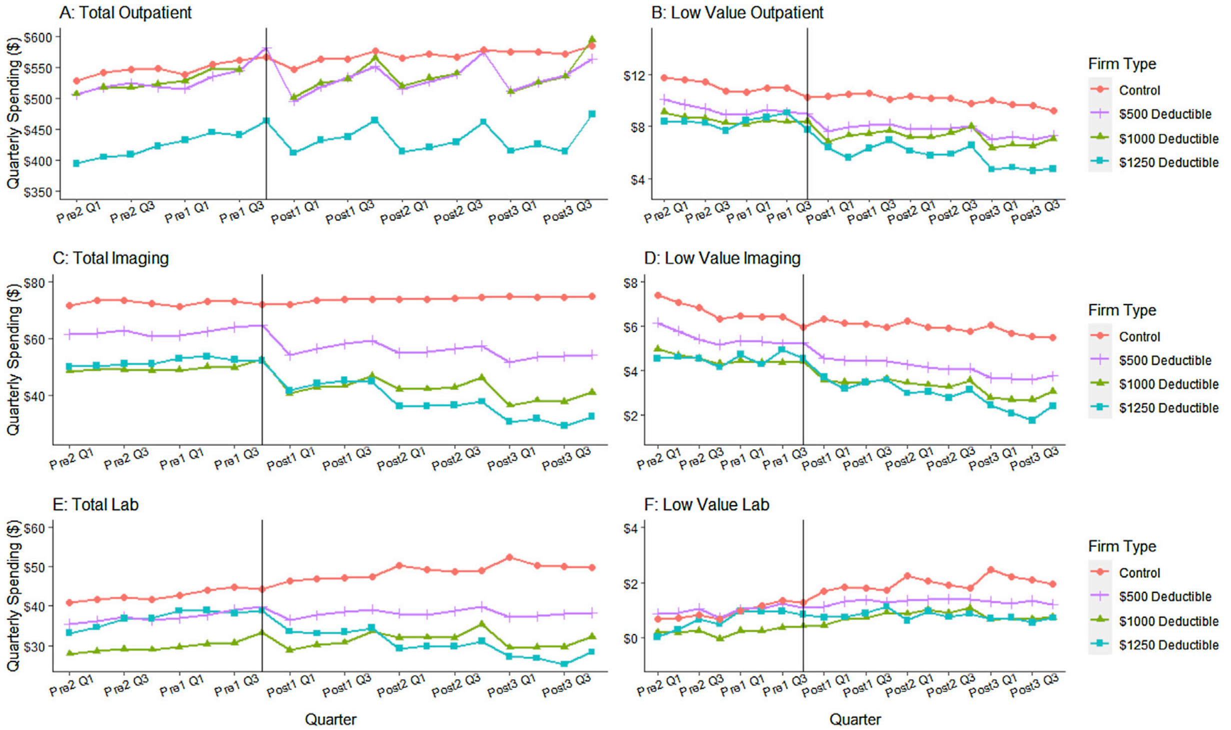

Fig. A1. Mean Quarterly Spending Per Enrollee by Firm Deductible Level.
Notes: Mean quarterly low-value spending per enrollee for treatment firms and control firms. Total spending graphs for our 3 categories are on the left (Panels A, C, and E), while low-value spending graphs are on the right (Panels B, D, and F). Total spending per person for each category defined as the sum of all payments on any claim in the Annual Outpatient Table without an explicit service location of 'inpatient hospital' for Panel A, claims with a CPT code in the range 70000–79999 for Panel C, and claims with a CPT code in the range 80000–89999 for Panel E. Panels B, D, and F are the sum of payments on claims for low-value services grouped into each category as described in Appendix Table A1. The vertical line is placed at Q4 of Pre-Year 1, immediately before a high-deductible offer. The control firm pre and post periods are calculated as a weighted average of calendar year mean spending per enrollee matching the enrollee-level distribution of calendar years in each period for the treatment firms. Deductible groups are determined by the deductible level of the plan with the highest deductible offered by a firm (e.g. $1250 Deductible firms offer at least 1 plan with a deductible of at least $1250).

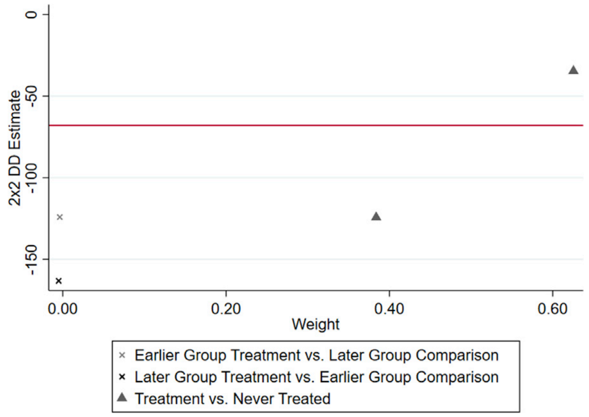

Fig. A2. Goodman-Bacon Decomposition for Total Outpatient Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked 'Δ' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata (Goodman-Bacon et al., 2019).

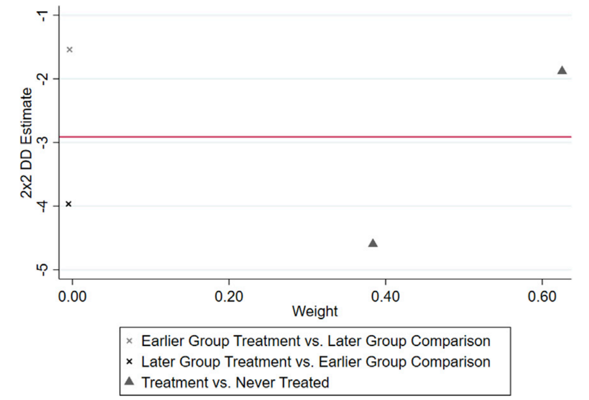

Fig. A3. Goodman-Bacon Decomposition for Low-Value Outpatient Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked 'Δ' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata.

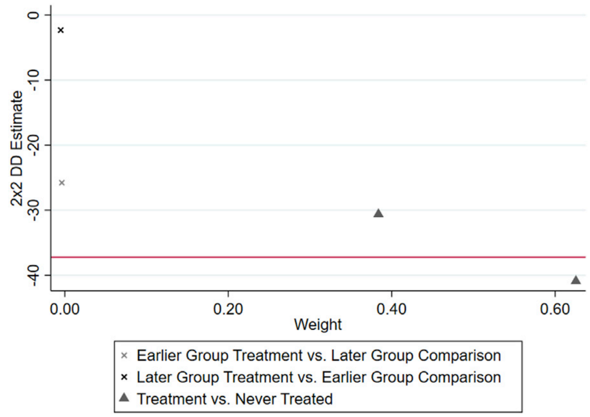

Fig. A4. Goodman-Bacon Decomposition for Total Imaging Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked 'Δ' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata.

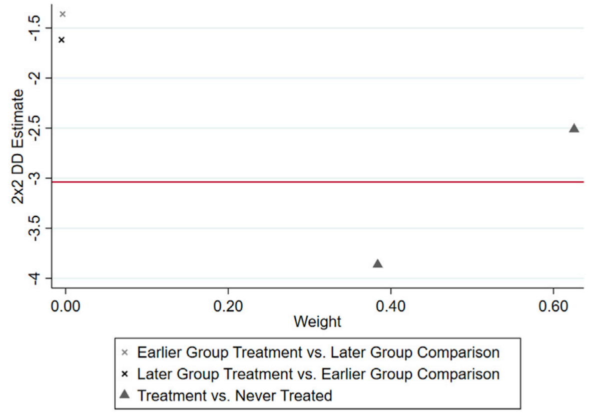

Fig. A5. Goodman-Bacon Decomposition for Low-Value Imaging Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked 'Δ' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata.

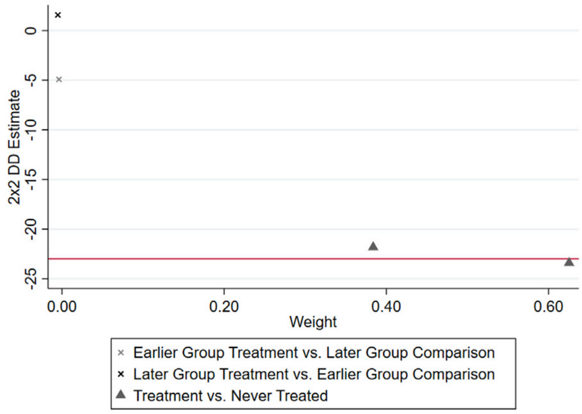

Fig. A6. Goodman-Bacon Decomposition for Total Lab Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked '∆' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata.

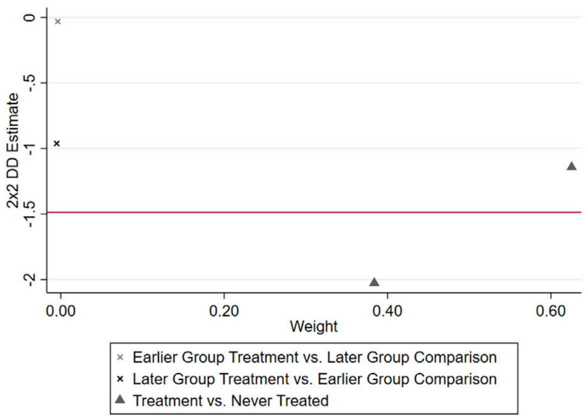

Fig. A7. Goodman-Bacon Decomposition for Low Value Lab Spending.
Notes: Results from the Goodman-Bacon decomposition, which decomposes the single DD estimator from a design using time-varying treatments into each of its component 2 × 2 DD estimators. Some treatment firms begin offering a high-deductible plan in 2010, others begin offering one in 2011. Symbols marked 'X' represent the DD coefficient when using late-offer treatment firms as a control for early treatment firms in 2010, and early treatment firms as a control for late treatment firms after 2011. Symbols marked '∆' represent DD coefficients when comparing early-treatment firms against the untreated firms, and late-treatment firms against the untreated firms. Command used is 'ddtiming' from Stata.

References

About the Health Technology Assessment Service | CADTH.ca. (n.d.). Retrieved February 5, 2019, from https://www.cadth.ca/about-cadth/what-we-do/products-services/hta.

Baker, L.C., 2010. Acquisition of MRI equipment by doctors drives up imaging use and spending. Health Aff. 29 (12), 2252–2259, http://dx.doi.org/10.1377/hlthaff.2009.1099.

Barnett, M.L., Linder, J.A., Clark, C.R., Sommers, B.D., 2017. Low-value medical services in the safety-net population. JAMA Intern. Med., http://dx.doi.org/10.1001/jamainternmed.2017.0401.

Beaudin-Seiler, B., Ciarametaro, M., Dubois, R., Lee, J., Fendrick, A.M., 2016. Reducing Low-Value Care., http://dx.doi.org/10.1377/hblog20160920.056666.

Beeuwkes Buntin, M., Haviland, A.M., McDevitt, R., Sood, N., 2011. Healthcare spending and preventive care in high-deductible and consumer-directed health plans. Am. J. Manage. Care 17 (3), 222–230, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21504258.

Belotti, F., Deb, P., Manning, W.G., Norton, E.C., 2015. Twopm: two-part models. Stata J. 15 (1), 3–20, http://dx.doi.org/10.1177/1536867x1501500102.

Berwick, D.M., Hackbarth, A.D., 2012. Eliminating waste in US health care. JAMA – J. Am. Med. Assoc., http://dx.doi.org/10.1001/jama.2012.362.

Brot-Goldberg, Z.C., Chandra, A., Handel, B.R., Kolstad, J.T., 2017. What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics*. Q. J. Econ. 132 (3), 1261–1318, http://dx.doi.org/10.1093/qje/qjx013.

Carrera, M., Goldman, D.P., Joyce, G., Sood, N., 2018. Do physicians respond to the costs and cost-sensitivity of their patients? Am. Econ. J. Econ. Policy 10 (1), 113–152, http://dx.doi.org/10.1257/pol.20150523.

Carter, E.A., Morin, P.E., Lind, K.D., 2017. Costs and trends in utilization of low-value services among older adults with commercial insurance or medicare advantage. Med. Care, http://dx.doi.org/10.1097/MLR.0000000000000809.

Cassel, C.K., Guest, J.A., 2012. Choosing wisely: helping physicians and patients make smart decisions about their care. JAMA – J. Am. Med. Assoc., http://dx.doi.org/10.1001/jama.2012.476.

Charlesworth, C.J., Meath, T.H.A., Schwartz, A.L., McConnell, K.J., 2016. Comparison of low-value care in medicaid vs commercially insured populations. JAMA Intern. Med. 176 (7), 998, http://dx.doi.org/10.1001/jamainternmed.2016.2086.

Charlson, M.E., Pompei, P., Ales, K.L., MacKenzie, C.R., 1987. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. J. Chronic Dis. 40 (5), 373–383, http://dx.doi.org/10.1016/0021-9681(87)90171-8.

Chernew, M., Newhouse, J., 2008. What does the RAND health insurance experiment tell us about the impact of patient cost sharing on health outcomes? Am. J. Manage. Care, Retrieved from https://www.ajmc.com/journals/issue/2008/2008-07-vol14-n7/jul08-3414p412-414/.

Claxton, G., Rae, M., Long, M., Damico, A., Whitmore, H., Foster, G., 2016. Health benefits in 2016: family premiums rose modestly, and offer rates remained stable. Health Aff. 35 (10), 1908–1917, http://dx.doi.org/10.1377/hlthaff.2016.0951.

Claxton, G., Rae, M., Long, M., Damico, A., Whitmore, H., Foster, G., 2017. Health benefits in 2017: stable coverage, workers faced considerable variation in costs. Health Aff. 36 (10), 1838–1847, http://dx.doi.org/10.1377/hlthaff.2017.0919.

Clemens, J., Gottlieb, J.D., 2014. Do physicians' financial incentives affect medical treatment and patient health? Am. Econ. Rev. 104 (4), 1320–1349, http://dx.doi.org/10.1257/aer.104.4.1320.

Clinician Lists. (n.d.). Retrieved February 5, 2019, from http://www.choosingwisely.org/clinician-lists/.

Clogg, C.C., Petkova, E., Haritou, A., 1995. Statistical methods for comparing regression coefficients between models. Am. J. Sociol. 100 (5), 1261–1293, http://dx.doi.org/10.1086/230638.

Deb, P., Norton, E.C., 2018. Modeling health care expenditures and use. Annu. Rev. Public Health 39, 489–505, http://dx.doi.org/10.1146/annurev-publhealth.

Dolan, R., 2016. High-deductible health plans. Health Aff., http://dx.doi.org/10.1377/hpb20160204.950878.

Eisenberg, M.D., Haviland, A.M., Mehrotra, A., Huckfeldt, P.J., Sood, N., 2017. The long term effects of "Consumer-Directed" health plans on preventive care use. J. Health Econ. 55, 61–75, http://dx.doi.org/10.1016/j.jhealeco.2017.06.008.

Farewell, V.T., Long, D.L., Tom, B.D.M., Yiu, S., Su, L., 2017. Two-part and related regression models for longitudinal data. Annu. Rev. Stat. Appl. 4 (1), 283–315, http://dx.doi.org/10.1146/annurev-statistics-060116-054131.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., et al., 2012. The Oregon health insurance experiment: evidence from the first year. Q. J. Econ. 127 (3), 1057–1106, Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/23293397.

Gawande, A.A., Colla, C.H., Halpern, S.D., Landon, B.E., 2014. Avoiding low-value care. N. Engl. J. Med. 370 (14), e21, http://dx.doi.org/10.1056/NEJMp1401245.
Goodman-Bacon, A., 2019. Difference-in-Differences with variation in treatment timing. National Bureau Econ. Res., http://dx.doi.org/10.3386/w25018.
Goodman-Bacon, A., Goldring, T., & Nichols, A. (2019). tgoldring.com. Retrieved April 13, 2020, from https://tgoldring.com/projects/ddtiming.html.
Greene, J., Hibbard, J.H., Dixon, A., Tusler, M., 2006. Which consumers are ready for consumer-directed health plans? J. Consum. Policy 29 (3), 247–262, http://dx.doi.org/10.1007/s10603-006-9008-z.
Greenland, S., Robins, J.M., Pearl, J., 1999. Confounding and collapsibility in causal inference. Source: Statistical Sci. 14.
Grossman, M., 1972. On the concept of health capital and the demand for health. J. Polit. Econ. 80 (2), 223–255, http://dx.doi.org/10.1086/259880.
Hartman, M., Martin, A.B., Espinosa, N., Catlin, A., The National Health Expenditure Acc, T. N. H. E. A, 2018. National health care spending in 2016: spending and enrollment growth slow after initial coverage expansions. Health Aff. 37 (1), 150–160, http://dx.doi.org/10.1377/hlthaff.2017.1299.
Haviland, A.M., Sood, N., McDevitt, R., Marquis, M.S., 2011. How do consumer-directed health plans affect vulnerable populations? Forum Health Econ. Policy 14 (2), http://dx.doi.org/10.2202/1558-9544.1248.
Haviland, A.M., Marquis, M.S., McDevitt, R.D., Sood, N., 2012. Growth of consumer-directed health plans to one-half of all employer-sponsored insurance could save $57 billion annually. Health Aff. 31 (5), 1009–1015, http://dx.doi.org/10.1377/hlthaff.2011.0369.
Haviland, A.M., Eisenberg, M.D., Mehrotra, A., Huckfeldt, P.J., Sood, N., 2016. Do "Consumer-Directed" health plans bend the cost curve over time? J. Health Econ. 46, 33–51, http://dx.doi.org/10.1016/J.JHEALECO.2016.01.001.
Manning, W.G., Newhouse, J.P., Duan, N., Keeler, E.B., Leibowitz, A., 1987. Health insurance and the demand for medical care: evidence from a randomized experiment. Am. Econ. Rev. 77 (3), 251–277, Retrieved from https://econpapers.repec.org/article/aeaaecrev/v_3a77_3ay_3a1987_3ai_3a3_3ap_3a251-77.htm.
McDevitt, R.D., Haviland, A.M., Lore, R., Laudenberger, L., Eisenberg, M., Sood, N., 2014. Risk selection into consumer-directed health plans: an analysis of family choices within large employers. Health Serv. Res. 49 (2), 609–627, http://dx.doi.org/10.1111/1475-6773.12121.
NICE Do not Do prompts | NICE (2013). Retrieved February 5, 2019, from https://www.nice.org.uk/sharedlearning/nice-do-not-do-prompts.
PATERNOSTER, R., BRAME, R., MAZEROLLE, P., PIQUERO, A., 1998. Using the correct statistical test for the equality of Regression Coefficients. Criminology 36 (4), 859–866, http://dx.doi.org/10.1111/j.1745-9125.1998.tb01268.x.
Recommendations for Primary Care Practice - US Preventive Services Task Force. (n.d.). Retrieved February 5, 2019, from https://www.uspreventiveservicestaskforce.org/Page/Name/recommendations.
Reid, R.O., Rabideau, B., Sood, N., 2016. Low-value health care services in a commercially insured population. JAMA Intern. Med. 176 (10), http://dx.doi.org/10.1001/jamainternmed.2016.5031.
Reid, Rachel O., Rabideau, B., Sood, N., 2017. Impact of consumer-directed health plans on low-value healthcare. Am. J. Manage. Care.
Rosenthal, M.B., Colla, C.H., Morden, N.E., Sequist, T.D., Mainor, A.J., Li, Z., Nguyen, K.H., 2018. Overuse and insurance plan type in a privately insured population HHS public access. Am. J. Manage. Care 24.
Schwartz, A.L., Landon, B.E., Elshaug, A.G., Chernew, M.E., McWilliams, J.M., 2014. Measuring low-value care in medicare. JAMA Intern. Med., http://dx.doi.org/10.1001/jamainternmed.2014.1541.
Schwartz, A.L., Chernew, M.E., Landon, B.E., McWilliams, J.M., 2015. Changes in low-value services in year 1 of the medicare pioneer accountable care organization program. JAMA Intern. Med. 175 (11), 1815, http://dx.doi.org/10.1001/jamainternmed.2015.4525.
Schwartz, A.L., Zaslavsky, A.M., Landon, B.E., Chernew, M.E., McWilliams, J.M., 2018. Low-value service use in provider organizations. Health Serv. Res., http://dx.doi.org/10.1111/1475-6773.12597.
Sinaiko, A.D., Mehrotra, A., Sood, N., 2016. Cost-sharing obligations, high-deductible health plan growth, and shopping for health care. JAMA Intern. Med. 176 (3), 395, http://dx.doi.org/10.1001/jamainternmed.2015.7554.
Sood, N., Wagner, Z., Huckfeldt, P., Haviland, A.M., 2013. Price shopping in consumer-directed health plans. Forum Health Econ. Policy 16 (1), 35–53, http://dx.doi.org/10.1515/fhep-2012-0028.
Spinal imaging for acute low-back pain | Choosing Wisely. (n.d.). Retrieved February 3, 2019, from http://www.choosingwisely.org/clinician-lists/aca-spinal-imaging-for-acute-low-back-pain/.
Stafford, M., Chandola, T., Marmot, M., 2007. Association between fear of crime and mental health and physical functioning. Am. J. Public Health 97 (11), 2076–2081, http://dx.doi.org/10.2105/AJPH.2006.097154.
The Healthcare Imperative. (2010). Washington, D.C.: National Academies Press. https://doi.org/10.17226/12750.
Tollen, L.A., Ross, M.N., Poor, S., 2004. Risk segmentation related to the offering of a consumer-directed health plan: a case study of humana inc. Health Serv. Res., http://dx.doi.org/10.1111/j.1475-6773.2004.00281.x.
Umscheid, C.A., 2013. Snapshots of low-value medical care. JAMA Intern. Med. 173 (3), 186, http://dx.doi.org/10.1001/jamainternmed.2013.1532.
Wharam, James Frank, Graves, A.J., Landon, B.E., Zhang, F., Soumerai, S.B., Ross-Degnan, D., 2011. Two-year trends in colorectal Cancer Screening after switch to a high-deductible health plan. Med. Care 49 (9), 865–871, http://dx.doi.org/10.1097/MLR.0b013e31821b35d8.
Wharam, J.Frank, Zhang, F., Eggleston, E.M., Lu, C.Y., Soumerai, S., Ross-Degnan, D., 2017. Diabetes outpatient care and acute complications before and after high-deductible insurance enrollment. JAMA Intern. Med. 177 (3), 358, http://dx.doi.org/10.1001/jamainternmed.2016.8411.
Zhang, X., Haviland, A., Mehrotra, A., Huckfeldt, P., Wagner, Z., Sood, N., 2018. Does Enrollment in High-Deductible Health Plans Encourage Price Shopping? Health Serv. Res. 53, 2718–2734, http://dx.doi.org/10.1111/1475-6773.12784.