NBER WORKING PAPER SERIES

SELECTION ON MORAL HAZARD IN HEALTH INSURANCE

Liran Einav
Amy Finkelstein
Stephen P. Ryan
Paul Schrimpf
Mark R. Cullen

Working Paper 16969
http://www.nber.org/papers/w16969

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2011

We are grateful to Felicia Bayer, Brenda Barlek, Chance Cassidy, Fran Fripovits, Frank Patrick, and Mike Williams for innumerable conversations explaining the institutional environment of Alcoa, to Colleen Barry, Susan Busch, Linda Cantley, Deron Galusha, James Hill, Sally Vegso, and especially Marty Slade for providing and explaining the data, to Tatyana Deryugina, Sean Klein, Michael Powell, Iuliana Pascu, and James Wang for outstanding research assistance, and to Ben Handel, Justine Hastings, Jim Heckman, Igal Hendel, Nathan Hendren, Kate Ho, Pat Kline, Jon Levin, Matt Notowidigdo, Phil Reny, Rob Townsend, and numerous seminar participants for helpful comments and suggestions. The data were provided as part of an ongoing service and research agreement between Alcoa, Inc. and Stanford, under which Stanford faculty, in collaboration with faculty and staff at Yale University, perform jointly agreed-upon ongoing and ad hoc research projects on workers' health, injury, disability, and health care, and Mark Cullen serves as Senior Medical Advisor for Alcoa, Inc. We gratefully acknowledge support from the NIA (R01 AG032449), the National Science Foundation Grant SES-0643037 (Einav), the Alfred P. Sloan Foundation (Finkelstein), the John D. and Catherine T. MacArthur Foundation Network on Socioeconomic Status and Health, and Alcoa, Inc. (Cullen), and the U.S. Social Security Administration through grant #5 RRC08098400-03-00 to the National Bureau of Economic Research as part of the SSA Retirement Research Consortium. Einav also acknowledges the hospitality of the Center for Advanced Study in the Behavioral Sciences at Stanford. The findings and conclusions expressed are solely those of the authors and do not represent the views of SSA, any agency of the Federal Government, or the NBER.

Selection on Moral Hazard in Health Insurance
Liran Einav, Amy Finkelstein, Stephen P. Ryan, Paul Schrimpf, and Mark R. Cullen
NBER Working Paper No. 16969
April 2011
JEL No. D12,D82,G22

ABSTRACT

In this paper we explore the possibility that individuals may select insurance coverage in part based on their anticipated behavioral response to the insurance contract. Such "selection on moral hazard" can have important implications for attempts to combat either selection or moral hazard. We explore these issues using individual-level panel data from a single firm, which contain information about health insurance options, choices, and subsequent claims. To identify the behavioral response to health insurance coverage and the heterogeneity in it, we take advantage of a change in the health insurance options offered to some, but not all of the firm's employees. We begin with descriptive evidence that is suggestive of both heterogeneous moral hazard as well as selection on it, with individuals who select more coverage also appearing to exhibit greater behavioral response to that coverage. To formalize this analysis and explore its implications, we develop and estimate a model of plan choice and medical utilization. The results from the modeling exercise echo the descriptive evidence, and allow for further explorations of the interaction between selection and moral hazard. For example, one implication of our estimates is that abstracting from selection on moral hazard could lead one to substantially over-estimate the spending reduction associated with introducing a high deductible health insurance option.

Liran Einav
Stanford University
Department of Economics
579 Serra Mall
Stanford, CA 94305-6072
and NBER
leinav@stanford.edu

Amy Finkelstein
Department of Economics
MIT E52-274C
50 Memorial Drive
Cambridge, MA 02142
and NBER
afink@mit.edu

Stephen P. Ryan
MIT Department of Economics
E52-262C
50 Memorial Drive
Cambridge, MA 02142
and NBER
sryan@mit.edu

Paul Schrimpf
Department of Economics
MIT
50 Memorial Drive
Cambridge, MA 02142
paul_s@mit.edu

Mark R. Cullen
Stanford School of Medicine
251 Campus Drive, MSOB x338
Stanford, CA 94306
and NBER
mrcullen@stanford.edu

1 Introduction

Economic analysis of market failure in insurance markets tends to analyze selection and moral hazard as distinct phenomena. In this paper, we explore the potential for selection on moral hazard in insurance markets. By this we mean the possibility that moral hazard effects are heterogeneous across individuals, and that individuals' selection of insurance coverage is affected by their anticipated behavioral response to coverage – their "moral hazard type." We examine these issues empirically in the context of employer-provided health insurance in the United States. Specifically, in addition to "traditional" selection based on one's health risk, we also examine selection on the expected incremental medical spending due to insurance.

Such selection on moral hazard has implications for the standard analysis of both selection and moral hazard. For example, a standard – and ubiquitous – approach to mitigating selection in insurance markets is risk adjustment, i.e. pricing on observable characteristics that predict one's insurance claims. However, the potential for selection on moral hazard suggests that monitoring techniques that are usually thought of as reducing moral hazard – such as cost sharing that varies across categories of claims with differential scope for moral hazard – may also have important benefits in combatting adverse selection. In contrast, a standard approach to mitigating moral hazard is to offer plans with higher consumer cost sharing. But if individuals' anticipated behavioral response to coverage affects their propensity to select such plans, the magnitude of the behavioral response could be much lower (or much higher) from what would be achieved if plan choices were unrelated to the behavioral response. As we discuss in more detail below, not only the existence of selection on moral hazard but also the sign of any relationship between anticipated behavioral response and demand for higher coverage is ex ante ambiguous. Ultimately, these are empirical questions. To our knowledge, however, there is no empirical work on selection on moral hazard in insurance markets.

Health insurance provides a particularly interesting setting in which to explore these issues. Both selection and moral hazard have been well-documented in the employer-provided health insurance market in the United States. Moreover, given the extensive government involvement in health insurance, as well as the concern about the size and rapid growth of the health care sector, there is considerable academic and public policy interest in how to mitigate both selection and moral hazard in this market.

Recognition of the possibility of selection on moral hazard, however, highlights potentially important limitations of analyzing these problems in isolation. For example, the sizable empirical literature on the likely spending reductions that could be achieved through higher consumer cost sharing has intentionally focused on isolating and exploring exogenous changes in cost sharing – such as those induced by the famous Rand experiment (Manning et al., 1987; Newhouse, 1993). Yet, the very same feature that solves the causal inference problem – namely randomization (or attempts to approximate it in the subsequent quasi-experimental literature on this topic) – removes the endogenous choice element. It thus abstracts, by design, from any selection on moral hazard, which could have important implications for the spending reductions achieved through offering

plans with higher consumer cost sharing, especially since substantial plan choice is now the norm not only in private health insurance but also increasingly in public health insurance programs, such as Medicare Part D.

We explore these issues using data on the U.S. workers at Alcoa Inc., a large multinational producer of aluminum and related products. We observe individual-level data on the health insurance options, choices, and subsequent medical utilization of employees (and their dependents); we also observe relatively rich demographic information, including health risk scores. Crucially for identifying and estimating moral hazard, we observe variation in the health insurance options offered to different groups of workers. In an effort to control health spending, Alcoa began introducing a new set of health insurance options in 2004, designed to encourage employees to move into plans with substantially higher consumer cost sharing. We calculate that, if there were no change in behavior, the move from the original options to the new options would have increased the average share of spending paid out of pocket from 13 to 28 percent. We exploit the fact that, for unionized employees, the introduction of the new health insurance options was phased in gradually, as the new health insurance options could only be introduced when existing union contracts expired.

We begin by providing descriptive and motivating evidence on moral hazard in our setting. Difference-in-differences estimates suggest that the new options are associated with an average reduction in total medical spending of about $600 (11 percent) per employee. We find evidence consistent with heterogeneity in this moral hazard effect, such as larger spending reductions for older relative to younger employees, and for sicker relative to healthier employees. We also present suggestive evidence of selection on moral hazard, with those who initially selected more coverage appearing to have a greater behavioral response to a change in coverage.

In order to formalize the analysis of selection on moral hazard and to explore some of its implications, we develop a utility-maximizing model of individual health insurance plan choices and claims. This allows us to precisely define "moral hazard" (a term whose usage is far from standardized in the literature) and, within the context of our model, identify selection on it. The model draws heavily on a relatively standard two-period framework for modeling health insurance demand and subsequent medical care utilization (as in, e.g., Cardon and Hendel, 2001). In the first period, a risk-averse expected-utility-maximizing individual makes optimal coverage choices based on his risk aversion, health expectations, and anticipated behavioral response to the contract choice. In the second period, health is realized and individuals make optimal medical expenditure decisions based on their realized health as well as on their chosen coverage. It is this last effect which generates what we term moral hazard, with a larger responsiveness corresponding to a higher "moral hazard type." We allow for unobserved heterogeneity along three dimensions: health expectations, risk aversion, and moral hazard, and for flexible correlation across these three.

An individual's optimal health insurance choice involves a trade-off of higher up-front premiums in exchange for lower ex-post out-of-pocket spending. All else equal, willingness to pay for coverage is increasing in the individual's health expectation and his risk aversion; these are standard results. In addition, in our model, all else equal, willingness to pay for coverage is increasing in the individual's moral hazard type: individuals with a greater behavioral response to coverage benefit more

from more coverage, since they will consume more care as a result. This is the "selection on moral hazard" comparative static that is the focus of our paper. Empirically, however, the sign (let alone the magnitude) of any selection on moral hazard is ambiguous and depends on the heterogeneity in moral hazard as well as the correlation between moral hazard type and the other primitives that affect health insurance choice, expected health and risk aversion.

We use this model, together with the data on individual plan options, plan choice, and subsequent medical spending, to recover the joint distribution of individuals' (unobserved) health type, risk aversion, and moral hazard type. The econometric model and its identification share many properties with some of our earlier work on insurance (Cohen and Einav, 2007; Einav, Finkelstein, and Schrimpf, 2010). The inclusion of moral hazard and heterogeneity in it is new. The panel structure of the data and the staggered timing of the introduction of the new coverage options are key in allowing us to identify this new element. The model is estimated using Markov Chain Monte Carlo Gibbs sampler, and its fit appears reasonable.

Qualitatively, the model's results are consistent with the descriptive evidence of selection on moral hazard. We find that individuals who exhibit a greater behavioral response to coverage are more likely to choose higher coverage plans. Quantitatively, we estimate substantial heterogeneity in moral hazard and selection on it. We focus on the counterfactual of moving from the most comprehensive to the least comprehensive of the new options – essentially moving them from a no deductible plan to a high ($3,000 for family coverage) deductible plan. In terms of heterogeneity in moral hazard, we find that the standard deviation across individuals of the spending reduction that would be achieved by this change in plans is more than twice the average. In terms of selection on moral hazard, we find that for determining the choice between these two plans, selection on moral hazard is roughly as important as "traditional" selection on health risk, and considerably more important than selection on risk aversion.

We use the model to examine some of the implications of the selection on moral hazard we detect for spending and for welfare. In terms of spending, our results suggest that if we were to introduce the high deductible plan in a setting where previously there was only the no deductible option, and price it so that 10 percent of the population chooses the high deductible plan, spending for those who choose the high deductible plan would fall by approximately $130 per person. By contrast, were we to ignore selection on moral hazard and assume that the 10 percent who chose the high deductible plan were randomly drawn, we would have estimated a spending reduction for those moved to the high deductible plan more than 2.5 times as large, at about $350 per person. In terms of welfare, we estimate that about two-thirds of the welfare gain that can be achieved in our setting by perfect risk adjustment that eliminates adverse selection could be achieved if better monitoring technologies eliminated selection on moral hazard. While our quantitative estimates are specific to our setting and our modeling choices, they nonetheless provide an interesting example of the potential for selection on moral hazard to play a non-trivial role in the analysis of both selection and moral hazard.

Our paper is related to several distinct literatures. As previously noted, our modeling approach is closely related to that of Cardon and Hendel (2001), which is also the approach taken by Bajari

et al. (2010), Carlin and Town (2010), and Handel (2010) in modeling health insurance plan choice. Like us, all of these papers have allowed for selection based on expected health risk. Our paper differs in our focus on identifying and estimating moral hazard – and in particular heterogeneous moral hazard – and in examining the relationship between moral hazard type and plan choice. From a methodological perspective, we also differ from these and many other discrete choice models in that we do not allow for a choice-specific, i.i.d. error term, which does not seem appealing given the vertically rankable nature of our choices.

Our analysis of the spending reduction associated with changes in cost sharing is related to a sizable experimental and quasi-experimental literature in health economics analyzing the impact of higher consumer cost sharing on spending. The difference-in-differences exercises with which we begin our analysis is very much in the spirit of this literature, which searches for identifying variation in consumer health plans to isolate the causal impact of consumer cost sharing on health spending. Our central difference-in-differences estimate translates into an implied arc elasticity of medical spending with respect to the average out-of-pocket cost share of about -0.14. This is broadly similar to the findings of the existing experimental and quasi-experimental literature, which tends to produce arc elasticities in the range of -0.1 to -0.4, with the "central" Rand elasticity estimate of -0.2 (see Chandra, Gruber, and McKnight (2010) for a recent review). However, our subsequent exploration of heterogeneity in this average moral hazard effect and selection on it suggests the need for caution in using such estimates, which do not account for endogenous plan selection, for forecasting the likely spending effects of introducing the option of plans with higher consumer cost sharing. It also suggests that one can embed the basic identification approach of the difference-in-differences framework in a model that allows for and investigates such endogenous selection.

Our examination of selection on moral hazard is motivated in part by the growing empirical literature demonstrating that selection in insurance markets often occurs on dimensions other than risk. This literature has tended to abstract from moral hazard, and focused on selection on preferences, such as risk aversion (Finkelstein and McGarry, 2006; Cohen and Einav, 2007), cognition (Fang, Keane, and Silverman, 2008), or desire for wealth after death (Einav, Finkelstein, and Schrimpf, 2010). Our exploration of selection on moral hazard highlights another potential dimension of selection and one that, we believe, has particularly interesting implications for contract design (in contexts where moral hazard is important). For many questions the extent to which selection occurs on the basis of expected health type or risk aversion does not matter (see, e.g., Einav, Finkelstein, and Cullen, 2010). However, as we illustrate in this paper, for questions regarding the design of contracts to reduce selection and the implications of contract design for spending, the extent to which selection is based on moral hazard can be important.

Despite its potential importance, we are not aware of any empirical work attempting to identify and analyze selection on moral hazard in insurance markets.[1] The basic idea of selection on moral

[1]Karlan and Zinman (2009) observe that selection in a credit market may be on unobserved risk and/or on anticipated effort, although they do not empirically distinguish between the two.

hazard, however, is not unique to us. Similar ideas have appeared in several other contexts. For example, in the context of appliance choices and phone plan choices, respectively, Dubin and McFadden (1984) and Miravete (2003) estimate models in which the choice is allowed to depend on subsequent utilization, which in turn may respond to the utilization price. One general way to think about the concept of selection on moral hazard is in the context of estimating a treatment effect of insurance coverage on medical expenditure. Within such a framework, selection on health risk would be equivalent to heterogeneity in (and selection on) the level (or constant term in a regression of medical spending on insurance coverage), while selection on moral hazard can be thought of as heterogeneity in (and selection on) the slope coefficient. Indeed, Heckman, Urzua and Vytlacil (2006) present an econometric examination of the properties of IV estimators when individuals select into treatment in part based on their anticipated response to the treatment, a phenomenon they refer to as "essential heterogeneity." They subsequently apply these ideas in the context of the returns to education in Carneiro, Heckman, and Vytlacil (2010).

The rest of the paper proceeds as follows. Section 2 describes the data and Section 3 presents descriptive evidence of moral hazard, heterogeneity in moral hazard, and selection on it in our data. Section 4 develops a two-period model of an individual's health insurance plan choice and spending decisions. Building on this model, Section 5 presents the econometric specification and describes its identification and estimation, and Section 6 presents our results, as well as illustrates some of their implications for spending and welfare. The last section concludes.

2 Setting and Data

We study health insurance choices and medical care utilization of the U.S.-based workers (and their dependents) at Alcoa, Inc., a large multinational producer of aluminum and related products. Our main analysis is based on data from 2003 and 2004, although for some of the analyses we extend the sample through 2006.

In 2004, in an effort to control health care spending by encouraging employees to move into plans with substantially higher consumer cost sharing, Alcoa introduced a new set of health insurance PPO options. The new options were introduced gradually to different employees based on their union affiliation, since new benefits could only be introduced when an existing union contract expired. The staggered timing in the transition from one set of insurance options to another provides a plausibly exogenous source of variation that can help us identify the impact of health insurance on medical care utilization, which is what we mean throughout by the term "moral hazard."

Our data contain the menu of health insurance options available to each employee, the employee's coverage choices, and detailed, claim-level information on his (and any covered dependents') medical care utilization and expenditures for the year.[2] The data also contain relatively rich demographic information (compared to typical claims data), including the employee's union

[2] Health insurance choices are made in November, during the open enrollment period, and apply for the subsequent calendar year. They can be changed during the year only if the employee has a qualifying event, which is not common.

affiliation, employment type (hourly or salary), age, race, gender, annual earnings, job tenure at the company, and the number and ages of other insured family members. In addition, we obtained a summary proxy of an individual's health based on software that predicts future medical spending on the basis of previous years' detailed medical diagnoses and claims, as well as basic demographics (age and gender); importantly for our purposes, this generated "health risk score" is not a function of the individual's coverage choice.[3]

Sample definition and demographics Alcoa has about 45,000 active employees per year. We exclude about 15 percent of the sample whose data are not suited to our analytical framework.[4] Given the source of variation used to identify moral hazard, we concentrate on the approximately one third of Alcoa workers who are unionized.[5] We further exclude the approximately two thirds of unionized workers that are covered by the Master Steel Workers' agreement. These workers faced only one PPO option which was left unchanged over our sample period. Finally, we exclude the approximately 10 percent of unionized employees who choose HMOs or who opt out of Alcoa-provided insurance, thus limiting our sample to employees enrolled in one of Alcoa's PPO plans.[6]

Our baseline sample therefore consists of the approximately 4,000 unionized workers (each year) not covered by the Master agreement. These workers belong to one of 28 different unions. Table 1 (top row) provides some descriptive statistics on the demographic characteristics of our baseline sample in 2003. Our sample is 72 percent white, 84 percent male, with an average age of 41,

[3] This is a relatively sophisticated way of predicting medical spending as it takes into account the differential persistence of different types of medical claims (e.g., diabetes vs. car accident) in addition to overall utilization, demographics, and a rich set of interactions among these measures. The particular software we use is a risk adjustment tool called DXCG risk solution which was developed by *Verisk Health* and is used by, among other organizations, the Center for Medicare and Medicaid services in determining reimbursement rates in Medicare Advantage. See Bundorf, Levin, and Mahoney (2009), Carlin and Town (2010), and Handel (2010) for other examples of academic uses of this type of predictive diagnostic software.

[4] The biggest reduction in sample size comes from excluding workers who are not at the company for the entire year (for whom we do not observe complete annual medical expenditures). In addition, we exclude employees who are outside the traditional benefit structure of the company (for example because they were working for a recently acquired company with a different (grandfathered) benefit structure); for such employees we do not have detailed information on their insurance options and choices. We also exclude a small number of employees because of missing data or data discrepancies.

[5] Approximately 70 percent of Alcoa workers are hourly employees, and approximately half of these are unionized. Salaried workers are not unionized.

[6] As is typical in claims data sets, we lack information for employees who choose an HMO or who opt out of employer coverage on both the details of their insurance coverage and their medical care utilization. Of course, this raises potential sample selection concerns. Reassuringly, as we show in Appendix A, the change in PPO health insurance options does not appear to be associated with a statistically or economically significant change in the fraction of employees who choose one of these excluded options.

average annual income of about $31,000, and an average tenure of about 10 years at the company. Approximately one quarter of the sample has single (employee only) coverage, while the rest also cover additional dependents. The health risk score is calibrated to be interpreted as predicted medical spending relative to a randomly drawn person under 65 in the nationally representative population; Table 1 indicates that, on average, individuals in our sample have predicted medical spending that is about 5 percent lower than this benchmark.

The remaining rows of Table 1 show summary statistics for four different groups of employees based on when they were switched to the new benefit options (i.e. four different treatment groups); we discuss this comparison when we present our difference-in-differences strategy and results below.

As noted, our main analysis is based on the 2003 and 2004 data (7,570 employee-years and 4,477 unique employees). We exclude the 2005 and 2006 data from our primary analysis because it introduces two challenges for estimation of our plan choice model. First, the relative price of comprehensive coverage on the new options was raised substantially in 2005 and raised further in 2006, yet remarkably few employees already in the new option set changed their plans. This is consistent with substantial evidence of inertial behavior in health insurance plan choices (Handel, 2010; Carlin and Town, 2010). Rather than modeling this behavior (e.g., as switching costs), we prefer instead to restrict the data to a time period where they are less central to understanding plan choices. Of course, plan choice for individuals under the old options may also reflect inertial factors (indeed, as we will show in Table 3 below, plan switching is extremely rare (about 1 percent) for employees whose options did not change in 2004), but the pricing under the old options is not changing during our sample period, making any such inertia less central for trying to understand current choices. Second, the pricing in 2006 is such that it is hard to rationalize some of the plan choices in which there is considerable mass, without extending the model to include some combination of switching costs, additional plan features, and/or biased expectations; again, we prefer to avoid these issues in the context of our primary question of interest.

The main drawback to limiting the data to 2003 and 2004 is that less than one-fifth of our sample were offered the new benefits starting in 2004, while another half of the sample was transitioned to the new benefits in 2005 and 2006 (Table 1, column (1)). Therefore, for some of the descriptive evidence we report in this section (which does not require an explicit model of plan choice) we use data from 2003-2006. This sample produces qualitatively similar descriptive results to the 2003-2004 sample, but the larger sample size allows for greater precision (and hence probing) in our descriptive exercises.

Medical spending We have detailed, claim-level information on medical expenditures and utilization. Our primary use of these data is to construct annual total medical spending for each employee (and his covered dependents). In Appendix A, we also use these data in a less aggregated way to break out spending by category (i.e., doctor's office, outpatient, inpatient, and other).

Figure 1 graphs the distribution of medical spending for our sample. We show the distribution separately for the approximately three-quarters of our sample with non-single coverage and the remainder with single employee coverage; not surprisingly, average spending is substantially higher

in the former group. Across all employees, the average annual spending (on themselves and their covered dependents) is about $5,200.[7] As is typical, medical expenditures are extremely skewed. For example, for non-single coverage, average spending ($6,100) is about 2.5 times greater than the median spending ($2,400), about 4 percent of our baseline sample has no spending, while each of the employees in the top decile spends over $13,000.

Health insurance options and choices An attractive feature of our setting is that the PPO plans in both the original and new regimes differ (within and across regimes) only in their consumer cost sharing requirements. They are identical on all non-cost sharing features, such as the network definition. Table 2 summarizes the original and new plan options and the fraction of employees who choose each option in our baseline sample. Employees may choose from up to four coverage tiers: single (employee only) coverage, or one of three non-single coverage tiers (employee plus spouse, employee plus children, or family). In our analysis we take coverage tier as given, assuming that it is primarily driven by family structure.[8]

There were three PPO options under the old benefits and five entirely different PPO options under the new benefits. Because there was no option of "staying in your existing plan" – the five new options were all distinct from the three old options in both their name and their design – individuals did not have the option of passively being defaulted into their existing coverage. We show in Table 3 below that plan choices for those who are switched to the new options are also consistent with the notion of "active" choices. As a result, we suspect that defaults did not play an important role in the choice of new benefits. Indeed, although option 4 was the default coverage option, it was not the most common choice (Table 2).[9] In the robustness section we provide additional analysis that suggests that the importance of defaults for our analysis is negligible.

The primary change from the old to the new benefits was to offer plans with higher deductibles and to increase the lowest out-of-pocket maximum.[10] As shown in the table, under the new options there was a shift to plans with higher consumer cost sharing. Under the old options virtually all

[7] A little over one quarter of total spending is in doctor offices, about one third is for inpatient hospitalizations, and about one third is for outpatient services. About half of the remaining four percent of spending is accounted for by emergency room visits.

[8] Employee premiums vary across the four coverage tiers according to fixed ratios. Cost sharing provisions differ only between single and non-single coverage. Specifically, for a given PPO, deductibles and out-of-pocket maxima are twice as great for any non-single coverage tier as they are for single coverage. As shown in Table 1, about one quarter of the sample chooses single coverage. Within non-single coverage, slightly over half choose family coverage, 30 percent choose employee plus spouse, and about 16 percent choose employee plus children (not shown).

[9] Also consistent with a large amount of "active" choices, although the old option 2 and the new option 5 are identical in all the aspects we model, only about half the employees who chose the old option 2 choose the new option 5, presumably reflecting the change in choice set (including relative pricing).

[10] At a point in time, prices within a coverage tier vary slightly across employees (in the range of several hundred dollars) under either the old or new options, depending on the employee's affiliation (see Einav, Finkelstein, and Cullen (2010) for more detail). Premiums were constant over time under the old options; as mentioned, under the

employees faced no deductible. Looking at employees with non-single coverage in Panel B (patterns for single coverage employees are similar), about two fifths faced a $2,000 out-of-pocket maximum while three-fifths faced a $5,000 out-of-pocket maximum. By contrast, under the new options, about a third of the employees faced a deductible, and all of them faced a high out-of-pocket maximum of at least $5,000 for non-single coverage.[11]

As one way to summarize the differences in consumer cost sharing under the different plans, we used the plan rules to simulate the average share of medical spending that would be paid out of pocket (counterfactually for most individuals) under different plans for all 2003 employees and their realized medical claims.[12] Less generous plans correspond to those with higher consumer cost sharing. The results are summarized in the third row of each panel of Table 2. Combining the information on average enrollment shares of the different plans with our calculation of the average cost sharing in the different plans, we estimate that, holding spending behavior constant, the change from the original options to the new options on average would have more than doubled the share of spending paid out of pocket, from about 13 to 28 percent.[13]

The plan descriptions in Table 2, and the subsequent parameterization of our model in Section 5, abstract from some additional details. First, while we model all plans as having a 10 percent in-network consumer coinsurance after the plan deductible is reached for all care, under the old options doctor visits and ER visits had in fact co-pays rather than coinsurance.[14] Second, we have summarized (and modeled) the in-network features only. All of the plans have higher (less generous) consumer cost sharing for care consumed out of network rather than in network. We

new options, premiums were increased substantially (and cross-employee differences were removed) in 2005 and 2006 (not shown).

[11] A $5,000 ($2,500) out-of-pocket maximum for non-single (single) coverage is rarely binding. With no deductible and a 10 percent consumer cost sharing, the employee must have $50,000 ($25,000) in total annual medical expenditures to hit this out-of-pocket maximum. Using the realized claims, we calculate that only about one percent of the employees would hit the out-of-pocket maximum in a given year. By contrast, under the old options the lowest out-of-pocket maximum was $2,000 ($1,000) for non-single (single) coverage, corresponding to total annual spending of $20,000 ($10,000). Using the same realized claims distribution, we calculate that about 5.5 percent of employees would hit this out-of-pocket maximum.

[12] By constructing (counterfactually) the share of a given (constant) set of medical expenditures that would be covered by different plans, we are able to construct a measure of the relative comprehensiveness of different plans that is purged of the confounding factors of selection and moral hazard that influence the actual out-of-pocket share of medical expenditures covered by each plan.

[13] These numbers are based on the average out of pocket shares by plan calculated in Table 2 and the plan shares for the 2003-2006 sample (not shown). Using the 2003-2004 sample's plan shares (shown in Table 2) we estimate that the move to the new options would on average raise the average out of pocket share from 12 to 25 percent.

[14] Specifically they had doctor and ER co-pays of $15 and $75 respectively, or $10 and $50 depending on the plan. In practice, given the average costs of a doctor visit ($115) and an ER visit ($730) in our data, the switch from the co-pay to coinsurance did not make much difference for predicted out-of-pocket spending.

choose to model only the in-network rules (where more than 95% of spending occurs) in order to avoid having to model the decision to go in or out of network. Third, while in general the new options were designed to have higher consumer cost sharing, a wider set of preventive care services (including regular physicals, screenings, and well baby care) were covered with no consumer cost sharing under the new options; these preventive services account for less than 2 percent of medical spending in our sample. Finally, the least comprehensive of the new options (option 1) includes a health reimbursement account (HRA) into which the employer makes tax-free contributions that the employee can draw on to pay for out-of-pocket medical expenses, or roll over for subsequent years. In the robustness section we explore alternative models that try to account for these distinctive features of this option.

Table 3 shows plan transitions for employees who were in the old options in both 2003 and 2004 and for employees who were switched from the old to the new options in 2004. Two main features emerge. First, almost all employees (almost 99 percent) under the old options in both years maintain the same coverage, which is to be expected given that the options and their prices did not change (but could also be driven by inertia in plan choices). Second, for those who get switched to the new options in 2004, there is far from a perfect correlation in the rank ordering of their choices under the old and new options. Over 40 percent of individuals move from the highest possible coverage under the old option to something other than the highest possible coverage under the new options, or vice versa. This is consistent with individuals making more "active" choices under the new options, as suggested earlier.

3 Descriptive Evidence of Moral Hazard

We start by presenting some basic descriptive evidence of moral hazard in our setting. The analysis provides a feel for the basic identification strategy for moral hazard. It also provides suggestive evidence of heterogeneity in moral hazard and selection on it. At the same time, our descriptive exercise points to the difficulty in identifying heterogeneity in moral hazard and selection on it without a formal model of moral hazard. The suggestive evidence as well as its important limitations together motivate our subsequent modeling exercise, which we turn to in the next section.

Descriptive estimates of moral hazard We start with the (easier) empirical task of documenting the existence of some form of asymmetric information in our data. Table 4 reports realized medical spending as a function of insurance coverage in our baseline sample. The analysis – which is in the spirit of Chiappori and Salanie's (2000) "positive correlation test" – shows that under either the old or new options individuals who choose more comprehensive coverage have systematically higher (contemporaneous) spending. This is consistent with the presence of adverse selection and/or moral hazard in our data.

To identify moral hazard in the data separately from adverse selection, we take advantage of the variation in the option set faced by different groups of employees. Table 5 presents this basic

difference-in-differences evidence of moral hazard for our baseline sample. Specifically, we show various moments of the spending distribution in 2003 and in 2004 for the control group (employees who are covered by the old options in both years) and the treatment group (employees who are switched to the new options in 2004). The results show a strikingly consistent pattern across all the various moments of the spending distribution: spending falls for the treatment group, and tends to increase slightly for the control group.

The results in Table 5 also suggest slight differences in 2003 spending for the treatment group relative to the control group, although these cross-sectional differences are, for the most part, small relative to the changes over time within the treatment group. More generally, the bottom four rows of Table 1 indicate differences in demographics as well as initial spending across all four of the treatment groups. In Appendix A we therefore explore in depth the sensitivity of our difference-in-differences estimates to controlling for observable differences across employees. We also investigate in the Appendix the validity of the underlying identifying assumption behind the difference-in-differences estimates, namely that absent the changes in health insurance benefits these different groups would have experienced similar trends in health spending. We find these results generally quite reassuring.

Table 6 summarizes our central difference-in-differences estimates (which we then explore in more detail in Appendix A). Columns (1)-(3) show the results for our baseline 2003-2004 sample. The first column shows the difference-in-differences estimate when the dependent variable is measured in dollars. Such a specification assumes that the moral hazard effects of insurance occurs in levels. This is consistent with the model we write down in the next section. However, both because it is possible that the moral hazard effect is in fact proportional to spending, and because one may be concerned about the results being driven by a few outliers with extremely high spending, in columns (2) and (3) we investigate specifications that give rise to a proportional moral hazard effect. Given the large fraction of employees with zero spending, we cannot estimate the model in simple logs. Instead, in column (2) we report estimates from a specification in which spending, m, is measured by $\log(1+m)$,[15] and column (3) reports a quasi-maximum likelihood Poisson model.[16] The results suggest that the move to the new options is associated with an economically significant decline in spending.

An important concern about the results in columns (1)-(3) is that they are not very precise. This is reflected in the large standard errors of the estimate, and in the relatively large differences in the quantitative implications of the different specifications. This lack of precision is driven by the fact that only about one-fifth of the employees in our sample are switched to the new benefits in 2004 (Table 1, column (1)). Therefore, in columns (4)-(6) we report analogous estimates from

[15] Given that almost all individuals spend at least several hundred dollars (Figure 1), the results are not sensitive to the choice of 1 relative to some other small numbers. For the same reason, the estimated coefficients can be approximately interpreted as elasticities.

[16] The QMLE-Poisson model requires only that the conditional mean be correctly specified for the estimates to be consistent. See, e.g., Wooldridge (2002, Chapter 19) for more discussion.

the 2003-2006 sample, during which more than half of the employees switched to the new benefits. As expected, the standard error of our estimates decreases substantially, and the quantitative implications of the results become much more stable across specifications. The estimated spending reduction is now statistically significant at the 5 percent level, with the point estimates suggesting a reduction of spending of about $600 (column (4)) or 11-17% (columns (5) and (6)). In Appendix A we show that the reduction in spending appears to arise entirely through reduced doctor and outpatient spending, with no evidence of a discernible effect on inpatient spending.[17]

Following common practice in this literature, we can compute a back-of-the-envelope elasticity of health spending with respect to the out-of-pocket cost sharing by combining these estimates of the spending reduction with the estimates in Table 2 of the average cost sharing of different plans (holding behavior constant). Given the distribution of employees across the different plans, the numbers in Table 2 suggest that the change from the old options to the new options should increase the average share of out-of-pocket spending from 12.6 percent to 28.4 percent in the 2003-2006 sample. Combining the point estimate of a $591 reduction in spending (Table 6, column (4)) with our calculation of the increase in cost sharing, our estimates imply an arc elasticity of medical spending with respect to out-of-pocket cost sharing of about -0.14.[18] This is broadly similar to the widely used Rand experiment arc-elasticity of medical spending of -0.2 (Manning et al., 1987; Keeler and Rolph, 1988). Subsequent studies that have used quasi-experimental variation in health insurance plans have tended to estimate elasticities of medical spending in the range of -0.1 to -0.4.[19]

Heterogeneity in moral hazard A necessary (but not sufficient) condition for selection on moral hazard is that there is heterogeneity in individuals' responsiveness to consumer cost sharing. To our knowledge, the experimental and quasi-experimental literature in health economics analyzing the impact of higher consumer cost sharing on spending has focused on average effects and largely ignored potential heterogeneity. This may in part reflect the fact that because health realizations are, by their nature, partially random, testing for heterogeneity in moral hazard is not trivial. It is particularly challenging without an explicit model of the nature of moral hazard which can, for example, provide guidance as to whether the effect of consumer cost sharing is ad-

[17] The reduction in outpatient spending appears to occur entirely on the intensive margin, while the reduction in doctor spending may occur entirely through a reduction in doctor visits.

[18] We compute an arc elasticity, in which the proportional change in spending (and in consumer cost sharing) is calculated relative to the average observed across the old and new options, so that our results are more directly comparable with the existing literature. The arc elasticity is calculated as $\frac{(q_2-q_1)/(q_1+q_2)/2}{(p_2-p_1)/(p_1+p_2)/2}$ where p denotes the average consumer cost sharing rate. For the 2003-2006 sample, the proportional change in spending and cost sharing is 11% and 77%, respectively.

[19] See Chandra, Gruber, and McKnight (2010), who provide a recent review of some of this literature as well as one of the estimated elasticities.

ditive or multiplicative.[20] In addition, because changes in health insurance change the consumer's (non-linear) budget set and individuals will vary as to where on the budget set they are, a careful examination of heterogeneity in moral hazard involves modeling this heterogeneity in the "treatment" associated with a change in health insurance plan; see Einav and Finkelstein (2011) for an exploration of related issues. In our specific context, a further subtlety is that it is the menu of plan options that varies in a quasi-experimental fashion, rather than the plan itself, making the actual individual coverage endogenous. All of these considerations motivate our formal modeling of moral hazard and of plan choice in the next section.

We begin, however, by first presenting some suggestive evidence in the data of what might plausibly be heterogeneity in moral hazard. One approach is to look at the distribution of spending changes across individuals. In the context of a model with an additive separable moral hazard effect (such as the one we develop in the next section), homogeneous moral hazard would imply a constant (additive) change in spending for all individuals. The results in Table 5 showing the difference-in-differences estimates at different quantiles of the distribution indicate that the change in spending associated with the change in insurance options is higher at higher quantiles. Due to censoring at zero this is mechanically true (and therefore not particularly informative) at the lower spending quantiles, but even comparing quantiles above the median shows a marked pattern of larger effects at larger quantiles.[21] Of course, since individuals may move quantiles with the change in options, this is not evidence of heterogeneity per se, but it is nonetheless suggestive.

Table 7 presents additional suggestive evidence of heterogeneous (level or proportional) moral hazard effects by reporting the difference-in-differences estimates separately for observably different groups of workers. Specifically, we show the estimated reduction in spending associated with the change from the old to the new options separately for workers above and below the median age (panel A), male vs. female workers (panel B), workers above and below the median income (panel C), and workers of above and below median health risk score (panel D). We discuss the final panel (panel E) later.

A difficulty with trying to infer heterogeneity in moral hazard from heterogeneous changes in spending across demographic groups is that differential changes in spending may reflect either heterogeneous treatment effects (the object of interest) or heterogeneous treatments (i.e., greater

[20] Without such a model, a nonparametric test for whether there is heterogeneity in moral hazard effects is possible to construct when there is no choice in health insurance and an exogenous change in health insurance coverage. In this case, a nonparametric test can be developed by relying on the panel nature of the data and comparing the joint distribution (before and after the introduction of a new benefit) of the quantiles of medical spending for the treatment group relative to the control group; the change in individual's spending rank (i.e. the joint distribution of the quantiles of spending) in the control group provides an estimate of the variation in ranking across individuals in their spending to expect simply from the random nature of health realizations. However, when an endogenous plan choice is present (as in our setting), a nonparametric test for heterogeneity in moral hazard is more challenging.

[21] Kowalski (2010) finds similar patterns in her quantile treatment estimates using a different identification strategy in a different firm.

changes in cost sharing for some groups than for others, given their endogenous plan choices). Separating these two requires a more explicit model of plan choices as well as how the cost sharing features of the plan affect the spending decision. Again, we do this formally in the context of the model we develop below. However, to get a loose sense of the variation in the change in cost sharing across groups, in columns (5) and (6) we report the average out of pocket share for each demographic group under the old and new options; column (7) reports the increase in the average out of pocket share associated with the change in options, which provides a metric by which to measure the treatment.

The estimates in Table 7 – while generally not precise – are suggestive of heterogenous moral hazard. The top two rows show that the reduction in spending associated with the new options is an order of magnitude higher for older workers than for younger workers, despite what appears to be a somewhat larger increase in the average out of pocket share for the younger workers (column (7)). Panel B indicates similar point estimates for male and female workers, despite the fact that males experience a larger increase in the out of pocket share. Similarly, panel C indicates similar point estimates for higher and lower income workers, but a somewhat larger increase in the out of pocket share for higher income workers. Finally, panel D indicates that the less healthy experience a substantial decline in spending while the more healthy experience no statistically detectable decline in spending, despite a larger increase in the out of pocket share for the more healthy.

While many of the estimates are quite imprecise, the results are suggestive of larger behavioral responses to consumer cost sharing for older workers than younger workers and for sicker workers than healthier workers, and perhaps also for female workers relative to male workers and for lower income workers relative to higher income workers. While suggestive, this type of exercise also points to the limitations of inferring heterogeneity in moral hazard across individuals from such simple descriptive evidence. For example, the parameterization of the "treatment" effects by the average out of pocket share obscures both the endogenous plan choice from within the menu of options as well as the different expected (end of year) marginal price faced by different individuals in the same plan based on their health status, which in principle should guide their utilization decisions.

Selection on moral hazard As discussed in the introduction, the pure comparative static of selection on moral hazard (holding all other factors that determine plan choice constant) is that individuals with a greater behavioral response to coverage (i.e., a larger moral hazard effect) will choose greater coverage. We therefore look for descriptive evidence of the relationship between an individual's behavioral responsiveness to coverage and their coverage choice. Some suggestive evidence of selection on moral hazard comes from the fact that older workers and sicker workers – whom we saw in Panel A may have larger moral hazard effects than younger workers and healthier workers respectively – also choose more comprehensive insurance under both the new and original plan options (not shown). Of course, older and sicker workers also have higher medical spending so that it is difficult to know from this evidence alone whether their insurance choice is driven by their expected health or their anticipated behavioral response to coverage.

Slightly more direct evidence of selection on moral hazard comes from comparing the estimated

behavioral response (estimated by examining the change in spending with the change from the original to the new options) between those who chose more vs. less coverage under the original options. The last panel of Table 7 presents the estimated treatment effect of the move from the original to the new options separately for individuals who chose more coverage under the original options in 2003 compared to those who chose less coverage under the original options in 2003.[22] Consistent with selection on moral hazard, we estimate a reduction in spending associated with the move from the old options to the new options that is more than twice as large for those who originally had more coverage than those who originally had less coverage, even though the reduction in cost sharing associated with the change in options (i.e., the treatment) is substantially larger for those who had less coverage. We do not have enough precision, however, to reject the null that estimated spending reductions are the same across the two groups. Moreover, we are once again confronted with the need to model the endogenous plan choice from among the new option as well as the variation in expected end of year marginal price induced by variation in health status.

Overall, we view the findings as suggestive descriptive evidence of selection on moral hazard of the expected sign. The rest of the paper now investigates this phenomenon more formally by developing and estimating a model of individual coverage choice and health care utilization. The model allows us to formalize more precisely the notion of "moral hazard," and aids in the identification of heterogeneity in moral hazard and selection on it. It also allows us to quantify selection on moral hazard and explore its implications through various counterfactual exercises.

4 A model of coverage choice and utilization

We now present a stylized model of individual coverage choice and health care utilization which we will then use as the main ingredient in our econometric specification and counterfactual exercises. The model is designed to allow us to isolate and examine separately three different potential determinants of insurance coverage choice: health expectations, risk aversion, and "moral hazard type."

We consider a two period model. In the first period, a risk-averse expected-utility maximizing individual makes an optimal health insurance coverage choice, using his available information to form his expectation regarding his subsequent health realization. In the second period, the individual observes his realized health and makes an optimal health care utilization decision, which depends on the realized health as well as on his coverage. It is this last effect which leads to what we call moral hazard. This general modeling framework is similar to the one used in existing empirical

[22] Specifically, we compare individuals who picked option 3 ("more coverage") under the original options to those who picked option 2 ("less coverage") under the original options. To do this analysis we need to limit the sample to the approximately 85 percent of the sample who was already employed at the firm by 2003 and in one of these two options. The estimated change in spending associated with the move from the old to the new options for this subsample is -859 (standard error 245), compared to -592 (standard error 264) in the full 2003-2006 sample (Table 5, column (4)).

models of demand for health insurance and medical spending (Cardon and Hendel, 2001; Bajari et al., 2010; Carlin and Town, 2010; Handel, 2010).

We begin with notation. This is a model of individual behavior, so we omit i subscripts to simplify notation; in the next section, where we take the model to the data, we describe how individuals may vary. At the time of his utilization choice (period 2), an individual is characterized by two objects: his health realization λ, and his "moral hazard type" ω. The health realization λ captures the uncertain aspect of demand for healthcare, with individuals with higher λ being sicker and demanding greater healthcare consumption. The moral hazard type ω determines how responsive health care utilization decisions are to insurance coverage. In other words, ω affects the individual's price elasticity of demand for healthcare with respect to its (out of pocket) price, with individuals with higher ω being more price elastic and therefore increasing their utilization more sharply in response to greater insurance coverage.

At the time of coverage choice (period 1), an individual is characterized by three objects: $F_\lambda(\cdot)$, ω, and ψ. The first, $F_\lambda(\cdot)$, represents the individual's expectation about his subsequent health risk λ. It is precisely the (natural) assumption that individuals do not know λ with certainty at the time of coverage choice, which leads them to demand insurance. The second object that enters the individual's coverage choice is his moral hazard type ω, which determines his period 2 price elasticity of demand for health care. Because individuals are forward looking, they anticipate that their price sensitivity will subsequently affect their utilization choices, and this in turn affects their utility from different coverages. It is this channel that creates the potential for selection on moral hazard, which is the main focus of our paper. Finally, the third object is ψ, which captures the individual's coefficient of absolute risk aversion. Importantly, unlike ω and $F_\lambda(\cdot)$, which enter the coverage choice but also affect (deterministically and stochastically, respectively) utilization decisions, risk preferences affect coverage choice but play no direct role in utilization decisions.

Utilization choice In the second period, insurance coverage, denoted by j, is taken as given. We assume that the individual's health care utilization decision is made in order to maximize a tradeoff between health and money, with higher ω individuals putting greater weight on health. Specifically, we assume that the individual's second period utility is separable in health and money and can be written as $u(m;\lambda,\omega) = h(m-\lambda;\omega) + y(m)$, where $m \geq 0$ is the monetized utilization choice, λ is the monetized health realization, and $y(m)$ is the residual income. Naturally, $y(m)$ is decreasing in m at a rate that depends on coverage. In contrast, we assume that $h(m-\lambda;\omega)$ is concave in its first argument, so that it is increasing for low levels of utilization (when treatment presumably improves health) and is decreasing eventually (when there is no further health benefit from treatment and time costs dominate). Thus, we assume that the marginal benefit from incremental utilization is decreasing. Using this formulation, we think of λ, the underlying health realization, as shifting the level of optimal utilization m^*. Finally, we assume that $h(m-\lambda;\omega)$ is increasing in its second argument, but this is purely a normalization which (as we will see below) allows us to interpret individuals with higher ω as those who are more elastic with respect to the price of medical utilization.

We parametrize further so that the second-period utility function is given by

$$u(m;\lambda,\omega,j) = \underbrace{\left[(m-\lambda) - \frac{1}{2\omega}(m-\lambda)^2\right]}_{h(m-\lambda;\omega)} + \underbrace{[y - c_j(m) - p_j]}_{y(m)}. \tag{1}$$

That is, we assume that $h(m-\lambda;\omega)$ is quadratic in its first argument, with ω affecting its curvature. We also explicitly write the residual income as the initial income y minus the premium p_j associated with coverage j and the out-of-pocket expenditure $c_j(m)$ associated with utilization m under coverage j. Because y and p_j are taken as given (at the time of utilization choice), it will be convenient to define

$$\widetilde{u}(m;\lambda,\omega,j) = \left[(m-\lambda) - \frac{1}{2\omega}(m-\lambda)^2\right] - c_j(m), \tag{2}$$

so that $u(m;\lambda,\omega,j) = \widetilde{u}(m;\lambda,\omega,j) + y - p_j$.

Given this parameterization, the optimal utilization is given by

$$m^*(\lambda,\omega,j) = \arg\max_{m\geq 0} u(m;\lambda,\omega,j). \tag{3}$$

It will also be convenient to denote $u^*(\lambda,\omega,j) \equiv u(m^*(\lambda,\omega,j);\lambda,\omega,j)$ and $\widetilde{u}^*(\lambda,\omega,j) \equiv \widetilde{u}(m^*(\lambda,\omega,j);\lambda,\omega,j)$.

To facilitate intuition, we consider here optimal utilization for the case of a linear (i.e., constant coinsurance) coverage contract, so that $c_j(m) = c \cdot m$ where $c \in [0,1]$. Full insurance is therefore given by $c = 0$ and no insurance is given by $c = 1$. The first order condition implied by the optimization problem in equation (3) is therefore given by $1 - \frac{1}{\omega}(m-\lambda) - c = 0$, or

$$m^*(\lambda,\omega,c) = \max\left[0, \lambda + \omega(1-c)\right]. \tag{4}$$

Thus, abstracting from the potential truncation of utilization at zero, the individual will spend $m^* = \lambda$ with no insurance (i.e., $c = 1$) and $m^* = \lambda + \omega$ with full insurance (i.e., $c = 0$). Thus, the utilization response to the change in coverage from full to no insurance is ω; utilization responds more to changes in coverage for individuals of greater moral hazard type (i.e., higher ω). One way to think about this model of moral hazard, therefore, is that λ represents non-discretionary health care shocks that all individuals will pay to treat, regardless of insurance. There is also discretionary health care utilization (such as whether to go to the doctor when confronted with a minor pain or irritation, for example) which, without insurance will not be undertaken. With insurance, some amount of this discretionary care will be consumed, with individuals with a higher ω consuming more of this discretionary care when they are insured.[23]

[23] We have written the model as if it is the individual who makes all the utilization decisions. In practice, many of the decisions are also affected by physicians. To the extent that physicians also respond to the individual's coverage (and they are likely to), our interpretation of moral hazard should be thought of as some combination of both the individual's and the physician's responses.

Coverage choice In the first period, the individual faces a fairly standard insurance coverage choice. As mentioned, we assume that the individual is an expected-utility maximizer, with a coefficient of absolute risk aversion of ψ. We further assume that the individual's von Neumann Morgenstern (vNM) utility function is of the constant absolute risk aversion (CARA) form, $w(x) = -\exp(-\psi x)$. In a typical insurance setting $w(x)$ is defined solely over financial outcomes. However, because moral hazard is present, individuals trade off income and health and therefore $w(x)$ is defined over the realized second-period utility $u^*(\lambda, \omega, j)$. We note that income enters $u^*(\lambda, \omega, j)$ additively with a coefficient of one, so $u^*(\lambda, \omega, j)$ is monetized and can still be thought of in dollars, as in the regular case.

Consider now a set of coverage options J, with each option $j \in J$ defined by its premium p_j and coverage function $c_j(m)$. Following the above assumptions, the individual will then evaluate his expected utility from each option,

$$v_j(F_\lambda(\cdot), \omega, \psi) = -\int \exp(-\psi u^*(\lambda, \omega, j)) dF_\lambda(\lambda), \tag{5}$$

with his optimal coverage choice given by

$$j^*(F_\lambda(\cdot), \omega, \psi) = \arg\max_{j \in J} v_j(F_\lambda(\cdot), \omega, \psi). \tag{6}$$

Measuring welfare and efficient contracts Our standard measure of consumer welfare in this context will be the notion of certainty equivalent. That is, for an individual defined by $(F_\lambda(\cdot), \omega, \psi)$, we denote the certainty equivalent to a contract j by the scalar e_j that solves $-\exp(-\psi e_j) = v_j(F_\lambda(\cdot), \omega, \psi)$, or

$$e_j(F_\lambda(\cdot), \omega, \psi) \equiv -\frac{1}{\psi} \ln \left[\int \exp(-\psi u^*(\lambda, \omega, j)) dF_\lambda(\lambda) \right]. \tag{7}$$

Our assumption of CARA utility over (additively separable) income and health implies no income effects. To see the implications of no income effects, we can substitute $u^*(\lambda, \omega, j) = \widetilde{u}^*(\lambda, \omega, j) + y - p_j$ into equation (7) and reorganize to obtain

$$\begin{aligned} e_j(F_\lambda(\cdot), \omega, \psi) &\equiv \widetilde{e}_j(F_\lambda(\cdot), \omega, \psi) + y - p_j \equiv \\ &\equiv -\frac{1}{\psi} \ln \left[\int \exp(-\psi \widetilde{u}^*(\lambda, \omega, j)) dF_\lambda(\lambda) \right] + y - p_j, \end{aligned} \tag{8}$$

so that $\widetilde{e}_j(F_\lambda(\cdot), \omega, \psi)$ captures the welfare from coverage, and residual income enters additively. Using this notation, differences in $\widetilde{e}(\cdot)$ across contracts with different coverages capture the willingness to pay for coverage. For example, an individual defined by $(F_\lambda(\cdot), \omega, \psi)$ is willing to pay at most $\widetilde{e}_k(F_\lambda(\cdot), \omega, \psi) - \widetilde{e}_j(F_\lambda(\cdot), \omega, \psi)$ in order to increase his coverage from j to k.

Equation (8) can also be used to characterize the comparative statics of willingness to pay for more coverage with respect to the model's primitives. In general, willingness to pay for more coverage is increasing in risk aversion ψ and in risk $F_\lambda(\cdot)$ (in a first order stochastic dominance

sense).[24] Given our specific parametrization, willingness to pay for more coverage is also increasing in moral hazard type ω.[25]

We assume that insurance providers are risk neutral, so that the provider's welfare is given by his expected profits, or

$$\pi_j(F_\lambda(\cdot), \omega) \equiv p_j - \int \left[m^*(\lambda, \omega, j) - c_j\left(m^*(\lambda, \omega, j)\right)\right] dF_\lambda(\lambda), \tag{9}$$

where the integrand captures the share of the utilization covered by the provider under contract j. Total surplus s_j is then given by

$$s_j(F_\lambda(\cdot), \omega, \psi) = e_j(F_\lambda(\cdot), \omega, \psi) + \pi_j(F_\lambda(\cdot), \omega) = \widetilde{e}_j(F_\lambda(\cdot), \omega, \psi) + y - \int \left[m^*(\lambda, \omega, j) - c_j\left(m^*(\lambda, \omega, j)\right)\right] dF_\lambda(\lambda). \tag{10}$$

That is, total surplus is simply certainty equivalent minus expected cost.

Finally, it may be useful to characterize the nature of the efficient contract in this setting. Because of our CARA assumptions, premiums are a transfer which do not affect total surplus. Therefore, the efficient contract can be characterized by the efficient coverage function $c^*(\cdot)$ that maximizes total surplus (as given by equation (10)) over the set of possible coverage functions. Such optimal contracts would trade off two offsetting forces. On one hand, an individual is risk averse while the provider is risk natural, so optimal risk sharing implies full coverage, under which the individual is not exposed to risk. On the other hand, the presence of moral hazard makes an insured individual's privately optimal utilization choice socially inefficient; any positive insurance coverage makes the individual face a healthcare price which is lower than the social cost of healthcare, leading to excessive utilization. Efficient contracts will therefore resolve this tradeoff by some form of partial coverage (Arrow, 1971; Holmstrom, 1979). For example, it is easy to see that no insurance ($c^*(m) = m$) is efficient if individuals are risk neutral or face no risk ($F_\lambda(\cdot)$ is degenerate), and that full insurance ($c^*(m) = 0$) is efficient when moral hazard is not present ($\omega = 0$). In all other situations, the efficient contract is some form of partial insurance.

Discussion Before turning to estimation, a few of our specific modeling choices above merit some further discussion.

[24] These comparative statics do not always hold. The model has unappealing properties when a significant portion of the distribution of λ is over the negative range, in which case the individual is exposed to a somewhat artificial uninsurable (background) risk (since spending is truncated at zero). We are not particularly concerned about this feature, however, as our estimated parameters do not give rise to it, and because we have experimented with a (non-elegant) modification to the model that does not have this feature, and the overall results were similar.

[25] In a more general model, ω is associated with two effects. One is the increased utilization, which increases willingness to pay. The second effect is the increased flexibility to adjust utilization as a function of the realized uncertainty (λ), which in turn reduces risk exposure and reduces willingness to pay for insurance. Our specific parameterization was designed to have spending under no insurance unaffected by ω; this eliminates this latter effect, and therefore makes the comparative statics unambiguous.

Terminology. The key conceptual distinction we are interested in is the possibility that selection is not only driven by "traditional" selection, on the expected level of medical expenditures ($F_\lambda(\cdot)$), but also by selection on the basis of the incremental medical expenditure with respect to increased coverage (ω). We refer to this latter effect as "moral hazard."

The use of the term "moral hazard" to refer to the responsiveness of medical care utilization to insurance coverage dates back at least to Arrow (1963). Consistent with the notion of "hidden action" – as is typically associated with the term "moral hazard" – it has been conjectured that health insurance may induce individuals to exert less (unobserved) effort in maintaining their health. However, in the context of health insurance the term "moral hazard" is more typically used to refer to the price elasticity of demand for health care, conditional on underlying health status (Pauly, 1968; Cutler and Zeckhauser, 2000). We thus follow this abuse of terminology, and use the term in a similar way. In other words, our model, like most in this literature, does not consider the potential impact of insurance on underlying health λ.

As a result, the asymmetric information problem that we associate with "moral hazard" is arguably more accurately described as one of hidden information (rather than of hidden action). The individual's actions (utilization) are observed and contractible, but his underlying health λ is hidden information which, if contractible, would be the efficient object of reimbursement. For our purposes, whether the problem is one of hidden information or hidden action is simply an issue of appropriate usage of terminology, and here we simply follow convention.[26]

Additive effect of moral hazard. We made the strong choice to model moral hazard (ω) as a level shift in spending that is (except due to the truncation of spending at zero) independent of one's health (λ) (see, e.g., equation (4)). This is primarily for analytical tractability. Our choice of the utility function in equation (1) is designed to achieve a straightforward economic interpretation of the key parameters of interest in the first order condition (4). In particular, it is designed so that λ (health status) is the monetized health spending without insurance (i.e., one's "nondiscretionary" spending), and ω (moral hazard) captures incremental, "discretionary" spending as individuals are moved from no insurance to full insurance. This allows us to straightforwardly measure and compare the magnitude of (and heterogeneity in) health risk λ and moral hazard ω. In alternative models that would give rise to a model of utilization that is non-separable in λ and ω, the monetized moral hazard effect would depend on both parameters; in this case it would be much more difficult to define (and analyze) the choices and behavior of high vs. low moral hazard types distinctly from high vs. low health status types.

We should note that this analytical tractability does not come at the obvious expense of realism. In other words, it is not a priori obvious whether or not moral hazard affects individuals in a manner

[26] There are two potential justifications given in the literature for why the impact of insurance on medical expenditures, conditional on health status, may constitute hidden action. First, patients and physicians may take less effort to shop around for better prices when they are insured (Arrow, 1963). Second, if insurance affects the quantity of care consumed, Cutler and Zeckhauser (2000) argue that this still constitutes hidden action since "though the action itself (seeking medical care) is not hidden, the motivation behind it is."

that is additively separable from their health. It does not strikes us as unreasonable to assume that whether or not one chooses to seek care for some minor skin irritation may not be affected by one's overall severity of illness. Of course, it is also not unreasonable to imagine that the responsiveness of medical care utilization to insurance coverage could depend on one's underlying health (λ) for a variety of reasons; for example, individuals who are sicker arguably have more occasions to make medical decisions and therefore to "exercise" moral hazard. Our finding that all of the spending reduction associated with the move to the new options seems to come from reductions in doctor and outpatient utilization and not in (the much more expensive) inpatient utilization (see Appendix A) suggests that the "right" model of moral hazard may not be one where the effects are multiplicative in underlying health. Still, moral hazard effects may also not be completely independent of health. Importantly, our set-up does not preclude this. Although we do not explicitly model this complementarity "within" an individual, our empirical specification below will allow for this in the aggregate, by modeling a cross-sectional distribution that allows for an arbitrary correlation between individual's moral hazard type and health risk.

Source of moral hazard. We do not explicitly model the underlying "source" of moral hazard (ω) and potential heterogeneity in it. The level of an individual's ω presumably derives from some combination of the individual's value of time (income), his disutility of doctor visits, his underlying health conditions (e.g., how discretionary they are), and so on. It may also relate to one's risk aversion regarding future health conditions. We have modeled the second period utility in a static way, with no uncertainty. As a result, moral hazard is not directly determined by risk aversion. Nonetheless, one can well imagine that more risk averse people might be less sensitive to price in making their medical care consumption decisions, making them have a lower ω in the context of our model. This is not inconsistent with our model. Our empirical specification below will allow for an arbitrary correlation between moral hazard type (ω) and risk aversion (ψ). We will also allow ω to vary with various observable characteristics, that may provide guidance on its sources.

Welfare. Finally, we note that our model assumes that any moral-hazard induced expenditure represents pure waste from a societal perspective. In other words, we assume that individuals would consume the socially optimal amount of medical care if they were uninsured. We view this as a natural benchmark rather than a normative statement about the healthcare industry. In practice, in the absence of insurance medical expenditures may be too high or too low relative to efficient levels. For example, in the absence of subsidies, liquidity constrained and/or myopic consumers may under-consume medical care that has no immediate payoff, particularly preventive care. Absent any clear guidance as to the nature and magnitude of any such frictions, we abstract from them in our model. In our particular setting, we also suspect that the induced reduction in medical spending with the move to the new, less generous health insurance options is not likely to have involved a socially sub-optimal reduction in consumption of some aspects of medical care since the price of preventive care actually decreased with the move to the new options.[27]

[27] Busch et al. (2006) and Cabral (2010) estimate that the move to the new options had no effect on the use of preventive care, perhaps because at the same time that the price of preventive care was lowered, the price of physician

5 Econometric specification

5.1 Parameterization

We now turn to specify a more complete econometric model that is based on the economic model of individual coverage choice and utilization developed in the last section. This will allow us to jointly estimate coverage choices and utilization, relate the estimated parameters of the model to underlying economic objects of interest, and quantify how spending and welfare may be affected under various counterfactuals. The additional modeling assumptions in this section are of two different natures. First, we will need to specify more parametrically some of the objects introduced earlier (e.g., individuals' beliefs $F_\lambda(\cdot)$). Second, we need to specify what form of heterogeneity we allow across individuals, and for a given individual over time.

Our unit of observation is an employee i, in a given year t. We abstract from the specifics of the timing and nature of claims, and, as we have done so far, simply code utilization m_{it} as the total medical spending (in dollars) for the entire year. The individual faces the choice set of either the original plan options or the new plan options (as described in Table 2), depending on the year and the employee's union affiliation, which dictates whether and when he was switched to the new benefits options.

Using the model of Section 4, recall that individuals are defined by three objects: their beliefs about their subsequent health status $F_\lambda(\cdot)$, their moral hazard parameter ω, and their risk aversion ψ. We assume that ω_i and ψ_i may vary across employees, but are constant for a given employee over time. It is the potential heterogeneity in ω_i which is the focus of the paper. We also assume that $F_\lambda(\cdot)$ is a (shifted) lognormal distribution with parameters $\mu_{\lambda,it}$, $\sigma_{\lambda,i}$, with support $(\kappa_{\lambda,i}, \infty)$, as explained below. That is, beliefs about health also vary across employees, and we allow $\mu_{\lambda,it}$ to be time varying to reflect the possibility that information about one's health evolves with time.

At the time of coverage choice individuals believe that

$$\log\left(\lambda_{it} - \kappa_{\lambda,i}\right) \sim N(\mu_{\lambda,it}, \sigma^2_{\lambda,i}), \tag{11}$$

and these beliefs are correct. Assuming a lognormal distribution for λ is natural, as the distribution of annual health expenditures is highly skewed (see Figure 1). The additional parameter $\kappa_{\lambda,i}$ is used in order to capture the significant fraction of individuals who have no spending over an entire year. When $\kappa_{\lambda,i}$ is negative, the support of the implied distribution of λ_{it} is expanded, allowing for λ_{it} to obtain negative values, which in turn implies (when ω_i is not too large) zero spending. The parameter $\sigma_{\lambda,i}$ indicates the precision of the individual's information about his subsequent health.

It is the heterogeneity in $\mu_{\lambda,it}$, $\sigma_{\lambda,i}$, and $\kappa_{\lambda,i}$ that gives rise to the traditional form of adverse selection on the basis of expected health, i.e. on the basis of expected λ (denoted $\overline{\lambda}$) which is given by

$$\overline{\lambda}\left(\mu_\lambda, \sigma_\lambda, \kappa_\lambda\right) = \exp\left(\mu_\lambda + \frac{1}{2}\sigma^2_\lambda\right) + \kappa_\lambda. \tag{12}$$

visits (which are likely complements to the use of preventive care) was raised. As mentioned, preventive care is about 2 percent of overall spending in our sample.

That is, higher $\mu_{\lambda,it}$, $\sigma_{\lambda,i}$, or $\kappa_{\lambda,i}$ are all associated with higher expected λ, which all else equal leads to greater expected medical spending and greater cost by the insurance provider.[28] All else equal, individuals with higher $\mu_{\lambda,it}$, $\sigma_{\lambda,i}$, or $\kappa_{\lambda,i}$ also prefer to choose greater coverage, thus giving rise to adverse selection.

Let x_{it} denote a vector of observables which are taken as given, and let $\overline{x_i}$ denote their within-individual average. In order to link the latent variables to observables, we make several parametric assumptions. First, we assume that $\log \omega_i$, $\log \psi_i$, and $\overline{\mu_{\lambda,i}}$ (which denotes the average (over time) of $\mu_{\lambda,it}$ for a given individual i) are drawn from a jointly normal distribution, such that[29]

$$\begin{pmatrix} \overline{\mu_{\lambda,i}} \\ \log \omega_i \\ \log \psi_i \end{pmatrix} \sim N\left(\begin{pmatrix} \overline{x_i}\beta_\lambda \\ \overline{x_i}\beta_\omega \\ \overline{x_i}\beta_\psi \end{pmatrix}, \begin{pmatrix} \sigma^2_{\bar{\mu}} & \sigma_{\mu,\omega} & \sigma_{\mu,\psi} \\ \sigma_{\mu,\omega} & \sigma^2_{\omega} & \sigma_{\omega,\psi} \\ \sigma_{\mu,\psi} & \sigma_{\omega,\psi} & \sigma^2_{\psi} \end{pmatrix} \right). \tag{13}$$

We then assume a random effects structure on μ_{it}, so that μ_{it} varies over time, but is correlated within an employee, such that

$$\mu_{\lambda,it} = \overline{\mu_{\lambda,i}} + (x_{it} - \overline{x_i})\,\beta_\lambda + \epsilon_{\lambda,it}, \tag{14}$$

where $\epsilon_{\lambda,it}$ is an i.i.d. normally distributed error term, with variance σ^2_ϵ. The variance of $\mu_{\lambda,it}$ is then $\sigma^2_\mu = \sigma^2_{\bar{\mu}} + \sigma^2_\epsilon$. Finally, we assume that

$$\sigma^{-2}_{\lambda,i} \sim \Gamma(\gamma_1, \gamma_2)\mathbf{1}\{\sigma^2_{\lambda,i} \leq \bar{\sigma}^2\} \tag{15}$$

and that

$$\kappa_{\lambda,i} \sim N\left(\overline{x_i}\beta_\kappa, \sigma^2_\kappa\right). \tag{16}$$

That is, $\sigma^2_{\lambda,i}$ is drawn from a right truncated inverse gamma distribution,[30] and $\kappa_{\lambda,i}$ is drawn from a normal distribution, and both are drawn independently from the other latent variables.

Thus, overall we estimate four vectors of mean shifters (β_λ, β_ω, β_ψ, β_κ), eight variance and covariance parameters (σ_μ, σ_ε,σ_ω,σ_ψ,σ_κ,$\sigma_{\mu,\omega}$,$\sigma_{\mu,\psi}$,$\sigma_{\omega,\psi}$), and two additional parameters (γ_1, γ_2) that determine the distribution of $\sigma^{-2}_{\lambda,i}$. Of course, an important decision is what observables $\overline{x_i}$ shift which primitive, and whether we would like any observables to be excluded from one or more of the (four) equations. To pay particular attention to the underlying variation emphasized in Section 2, in all the specifications we experiment with, we include in $\overline{x_i}$ treatment group fixed effects for each of the four treatment groups (see Table 1), as well as a year fixed effect on $\mu_{\lambda,it}$, the only time varying latent variable. We also include coverage tier fixed effects since both the choice sets and

[28] Note that expected medical spending of an individual is closely related but not identical to $\overline{\lambda}$, since both moral hazard and the restriction that spending be non negative create a wedge between expected medical spending and expected health (see, e.g., equation (4)).

[29] For notational simplicity we consider $\overline{x_i}$ to be the super-set of covariates, and implicitly assume some coefficient restrictions if we allow for different mean shifters for different latent variables.

[30] We truncate the distribution of $\sigma^{-2}_{\lambda,i}$ because the non-truncated distribution causes the unconditional distribution of λ_{it} to have no moments.

spending varies substantially by coverage tier (see Table 2 and Figure 1, respectively), and a rich set of demographics, specifically age, gender, job tenure, income, and health risk scores.

5.2 Estimation

We estimate the model using Markov Chain Monte Carlo (MCMC) Gibbs sampling. The multi-dimensional unobserved heterogeneity naturally lends itself to such methods, as the iterative sampling allows us to avoid evaluating multi-dimensional integrals numerically, which is computationally cumbersome. The key observation is that the model we developed is sufficiently flexible so that we can augment the latent variables into the model and formulate a hierarchical statistical model. To see this, let $\theta_1 = \left\{\beta_\lambda, \beta_\omega, \beta_\psi, \beta_\kappa; \sigma_\mu, \sigma_\varepsilon, \sigma_\omega, \sigma_\psi, \sigma_\kappa, \sigma_{\mu,\omega}, \sigma_{\mu,\psi}, \sigma_{\omega,\psi}; \gamma_1, \gamma_2\right\}$ be the set of parameters we are interested in, and let $\theta_2 = \left\{\lambda_{it}, \mu_{\lambda,it}, \sigma_{\lambda,i}, \kappa_{\lambda,i}, \omega_i, \psi_i\right\}_{i=1,t=2003}^{i=N,t=2004}$ be the set of employee-year latent variables. The model is set up so that, even conditional on θ_1, we can always rationalize the observed data – namely, plan choice and medical utilization – by appropriately finding a set of latent variables for each individual, θ_2.

Thus, the iterative procedure is straightforward. We can first sample from the distribution of θ_1 conditional on θ_2. Because, conditional on θ_2, there is no additional information in the data about θ_1 this part of the sampling is simple and quite standard. Then, we can sample from the distribution of θ_2 conditional on θ_1 and the information available in the data. This latter step is of course more customized toward our specific model, but does not introduce any conceptual difficulties. The full sampling procedure, the specific prior distributions we impose, and the resultant posteriors are described in detail in Appendix B. We verified using Monte Carlo simulations that the procedure seems to work quite effectively, and is pretty robust to initial values. For our baseline results, the estimation seems to converge after about 5,000 iterations of the Gibbs sampler, so we drop the first 10,000 draws and use the last 10,000 draws of each variable to report our results. The results we report are based on the posterior mean and posterior standard deviation from these 10,000 draws.

One important difficulty that our model introduces is related to our decision to not allow for an additive separable plan-specific error term. It is extremely common in applications of discrete choice (such as ours) to add such error terms, and often to assume that they are distributed i.i.d. across plans and individuals. Such error terms serve two important roles. First, they allow the researcher to rationalize any choice observed in the data through a large enough error term. Second, their independence makes the objective function of any M-estimator smooth, which is computationally attractive for numerical optimization. In the context of our application, however, we view such error terms as economically unappealing. The options from which individuals in our sample choose are financially rankable and are identical in their non-financial features. This makes one wonder what such error terms would capture that is outside of our model. The clear ranking of the options also makes the i.i.d. nature of the error terms not very appealing. Instead, we introduce a fair amount of heterogeneity along the other dimensions of our model. Some of this heterogeneity (e.g., the heterogeneity in $\sigma_{\lambda,i}$ and $\kappa_{\lambda,i}$) is richer than the minimum required to capture the key economic forces we would like to capture, but this richness is what allows us to rationalize all observed choices

in the data. This still leads to a model which is not very attractive for numerical optimization, which is one important reason why we use Gibbs sampling.[31]

5.3 Identification

We now discuss the identification of the model. Conditional on the individual-behavior model described in Section 4, the object of interest that we seek to identify is the joint distribution of $F_\lambda(\cdot)$, ω, and ψ. We have data on individuals' health insurance options, choices, and medical spending. Throughout the paper we make the strong assumption that individual beliefs about their subsequent health status ($F_\lambda(\cdot)$) are correct.[32] The model and its identification share many properties with some of our earlier work on insurance (Cohen and Einav, 2007; Einav, Finkelstein, and Schrimpf, 2010). The key novel element is that we now allow for moral hazard, and heterogeneity in it. The panel structure of the data and the staggered timing of the introduction of the new options are key in allowing us to identify this new element. We organize our discussion of identification in two steps. We first consider nonparametric identification of our model with ideal data, and then discuss the ways in which our actual data is different from the ideal, thus requiring us to make additional parametric assumptions that aid in identification.

Identification with ideal data The two features of our data set that are instrumental for identification are the panel structure of the data and the exogenous change in the health insurance options available to employees. In the ideal setting, we consider a case in which we observe individuals for a sufficiently long period before and a sufficiently long period after the change in coverage. Moreover, we assume that the choice set from which employees can choose coverage is continuous (for example, one can imagine a continuous coinsurance rate, and an increasing and differentiable mapping from coinsurance rate to premium).

In such a setting, our model is non-parametrically identified. To see this, note that such data provide us with two medical expenditure distributions, $G_i^{before}(m)$ and $G_i^{after}(m)$, for each individual i. Using the realized utility model (during the second period of the model), these two distributions allow us to recover for each individual $F_{i,\lambda}(\cdot)$ and ω_i. To see this, recall that abstracting from the truncation of medical spending at zero, our model implies that medical expenditure m_{it} is equal to $\lambda_{it} + \omega_i(1 - c_t)$. If $F_{i,\lambda}(\cdot)$ is stable over time,[33] one can regress (for each employee i

[31] In addition to our previous work (Cohen and Einav, 2007; Einav, Finkelstein, and Schrimpf, 2010), several other papers have estimated a discrete choice model without an i.i.d. error, for similar reasons. These include Keane and Moffitt (1998), Berry and Pakes (2007), and Goettler and Clay (forthcoming).

[32] While it is reasonable to question this assumption, absent direct data on beliefs some assumption about beliefs is essential for identification. Otherwise, it is not possible to distinguish beliefs from other preferences that only affect choices, such as risk aversion (see Einav, Finkelstein and Schrimpf (2010) for a more detailed discussion of this point). While we could instead assume some other (pre-specified) form of biased beliefs, correct beliefs seem like a natural starting point.

[33] If $F_{i,\lambda}(\cdot)$ changes over time, one could parameterize, identify, and estimate the autocorrelation structure with a

separately) m_{it} on a dummy variable that is equal to 1 after the change. The estimated coefficient on the dummy variable would be then an estimate of $\omega_i \left(c_{after} - c_{before}\right)$, providing an estimate of ω_i. The distribution of λ_{it} can then be recovered by observing that $\lambda_{it} = m_{it} - \omega_i(1 - c_t)$, which is known.

Conditional on $F_{i,\lambda}(\cdot)$ and ω_i, individual i's choice from a continuous set of options provides a unique mapping from choices to his coefficient of absolute risk aversion since – conditional on $F_{i,\lambda}(\cdot)$ and ω_i – the coefficient of risk aversion is the only unknown primitive that may shift employees' choices, and it does so monotonically. Thus, using information about $F_{i,\lambda}(\cdot)$ and ω_i and individual i's choice from the continuous option set,[34] we can recover ψ_i. Since we recovered $F_{i,\lambda}(\cdot)$, ω_i, and ψ_i for each employee, we can now combine these estimates for our entire sample, and obtain the joint distributions of $F_\lambda(\cdot)$, ω, and ψ.

Identification with our specific data Our actual data depart from the ideal data described above in two main ways. First, although we have a panel structure, we only observe individuals for two periods in the baseline sample (that is limited to 2003 and 2004). Second, the choice set is highly discrete (including three to five options) rather than continuous. We thus make additional parametric assumptions to aid us in identification. This implies that our identification in the actual estimation cannot rely anymore on identifying the individual-specific parameters employee-by-employee. Rather, we observe a distribution of medical expenditures before the change and a distribution for medical expenditure after the change. We then identify the model by comparing the distribution after with the distribution before.

We can now think first about the identification of moral hazard. A comparison of spending distributions before and after a change in health insurance options may be contaminated by other confounders that change over time. Therefore, analogously to the difference-in-differences strategy of the reduced form (Section 3), we use the majority of the sample for which the options did not change during our sample period as a control group. We can therefore conceptually think of identification in our baseline sample as if we follow a *stable* population before and after a treatment, using the control population to adjust for any time-varying effects.

To gain intuition for our identification of moral hazard, consider a set of individuals who chose the same sequence of plans in 2003 and 2004. Of course, this is a selected subset of the population, a point that we will return to below. Without moral hazard, the distribution of medical expenditures for this group, before and after the change, would have remained the same. With moral hazard, spending under the new, say, lower coverage plan is lower. Loosely, and abstracting from truncation of spending at zero, the overall difference in the level of spending identifies the average moral hazard effect. Since our model implies that the moral hazard parameter affects spending additively, the

sufficiently long panel (the health risk score variable, which varies over time for a given individual, is quite useful in this regard). We therefore treat $F_{i,\lambda}(\cdot)$ as stable over time throughout this section.

[34] This can be done using either the options set before the change or after. In fact, the ideal data leads to over identification, so could allow us to test or enrich the model.

extent of heterogeneity in moral hazard is identified by the difference in the distributions, quantile-by-quantile.

Once the distribution of moral hazard, ω_i, is known, the remaining identification challenge is very similar to our earlier work (Cohen and Einav, 2007; Einav, Finkelstein, and Schrimpf, 2010). Conditional on the distribution of ω_i, our data provide information about coverage choices and subsequent realizations. By assuming that $F_{i,\lambda}(\cdot)$ follows a lognormal distribution, we can map the data on choices and spending to the remaining primitives of risk aversion ψ_i and risk types $F_{i,\lambda}(\cdot)$. Intuition for this is perhaps most easily seen in two steps (although in practice it is more efficient to estimate all parameters simultaneously, as we do). The observed distribution of medical spending (net of the known moral hazard) provides information on the distribution of health risk $F_{i,\lambda}(\cdot)$; conditional on health risk and moral hazard, the choice of insurance identifies risk aversion (ψ_i). We assume a three-dimensional heterogeneity in $F_{i,\lambda}(\cdot)$ – in mean μ, variance σ, and offset κ. Loosely, the distribution of the mean is primarily identified by the first moment of the spending distribution, the distribution of the variance by the second moment, and the distribution of κ is primarily driven by the extent of zero spending across different choices.

Two difficulties still remain. First, the choice set is discrete, so choices can only map to intervals of risk aversion. Second, while the distribution of ω_i across individuals is known, the specific value of ω_i is not known for each individual. Here, the parametric assumption regarding the joint normal distribution of $\log \omega_i$, $\overline{\mu_{\lambda,i}}$, and $\log \psi_i$ is useful, as it allows us to integrate over all possible values within each such choice interval. The final step is to repeat a similar argument for each observed sequence of choices, which together aggregate to the joint distribution of the population as a whole, which is the object we wish to identify.

6 Results

6.1 Parameter estimates

Table 8(a) presents the estimated parameters from estimating the model on the baseline sample of 7,570 employee-years. The top panel presents the estimated coefficients on the mean shifters of the four latent variables: $\mu_{\lambda,it}$ and $\kappa_{\lambda,i}$ which affect expected health risk ($E(\lambda_{it})$), ω_i which affects moral hazard, and ψ_i which captures risk aversion. The middle panel report the estimated variance-covariance matrix and the bottom panel reports the estimates of the rest of the parameters. In Table 8(b) we report some implied quantities of interest that are derived from the estimates. The latter may be more easy to interpret, so we focus much of the discussion on them.

Overall, as shown in the top panel of Table 8(b), the estimates imply an average health risk ($E(\lambda)$) of about \$4,340 per employee-year. We estimate an average moral hazard parameter (ω) that is about 30 percent of the average health risk, or about \$1,330 dollar; by way of context, recall that ω is approximately the size of the spending effect as we move individuals from no insurance

to full insurance (see equation (4)).[35]

We estimate statistically significant and economically large heterogeneity in each one of the components: health, moral hazard, and risk aversion. One way to gauge the magnitude of this heterogeneity is in the top panel of Table 8(b). Our estimates indicate a standard deviation for expected health risk ($E(\lambda)$) of about \$5,100, or a coefficient of variation of about 1.2; the standard deviation of realized health (λ) is, not surprisingly, much larger at \$25,000 (not shown). Moral hazard (ω) is also estimated to be highly heterogenous, with a standard deviation across employees of about \$3,200, or a coefficient of variation that is greater than 2. Finally, we estimate a coefficient of variation for absolute risk aversion (ψ) that is about one.

The unconditional correlations (Table 8(b), middle panel) are all statistically significant, and their signs seem reasonable. We estimate that the unconditional correlation between moral hazard (ω) and expected health risk ($E(\lambda)$) is positive and reasonably important (0.24).This likely reflects the fact that in our model moral hazard type (ω) is measured in absolute (dollar) terms rather than relative to health, so individuals with higher $E(\lambda)$ (i.e., worse health) have more opportunities to exercise moral hazard. The correlation between risk aversion and health risk (and moral hazard) is negative, perhaps reflecting the fact that individuals who are more risk averse are also those who take better care of their health. A similar pattern was documented by Finkelstein and McGarry (2006) in the context of long-term care insurance. Finally, as may be expected, we estimate a strong correlation in $\mu_{\lambda,it}$ over time, of 0.5 (not shown), suggesting that much of an individual's health risk is persistent over time, for example due to chronic conditions.

The signs of the covariates seem generally sensible.[36] The bottom panel of Table 8(b) summarizes the effects of covariates on $E(\lambda)$ by combining their separate effects on μ_λ and κ_λ. As could be expected, the health risk scores are an important predictor of expected health risk $E(\lambda)$, shifting it by thousands of dollars in the expected direction. We also estimate that female employees and employees with non-single coverage are associated with worse expected health (higher $E(\lambda)$). One should interpret these latter effects cautiously, however, as health risk scores are partialled out and are highly correlated with these other variables. This may also explain why the residual effect of income and tenure on expected health appears negligible.

Our estimates also imply (top panel of Table 8(a)) that employees with higher (i.e., worse) health risk scores are associated with greater moral hazard and lower risk aversion. Again, this likely reflects our choice to model moral hazard in absolute terms rather than relative to health. Conditional on health risk scores, employees with single coverage appear to be associated with

[35] We estimate an average coefficient of absolute risk aversion of about 0.0019, but caution against trying to compare this to existing estimates. In our model, realized utility is a function of both health risk and financial risk, while in other papers that estimate risk aversion from insurance choices (e.g., Cohen and Einav, 2007; Handel, 2010) realized utility is only over financial risk. Thus, the estimated "level" of risk aversion is not directly comparable; indeed, one could add a separable health related component to utility that is affected only by λ to change the risk aversion estimates, without altering anything else in the model.

[36] The covariates appear to explain about 55 to 60 percent of the variation in each of $E(\lambda)$, ω, and ψ.

greater moral hazard as well as with greater risk aversion. This may represent different process of decision making regarding health coverage and health care utilization when regarding one self vs. his family members.

6.2 Model fit

In Table 9 we report the actual and predicted plan choice probabilities. We fit the choices of employees who are choosing from the original plan options remarkably well. The fit of the choices from the new options is also reasonable, but not as good as the fit for the original options. This is likely because there are many fewer employees in the baseline sample who are subject to the new options. Thus, to the extent that the same model attempts to rationalize the choices from both the old and new options, it is natural that more weight is given to trying to fit choices from the old menu, leading to slightly worse fit for those choosing from the new menu.

Figure 2 reports the actual and predicted distributions of medical expenditure. The top panel reports the fit for the individuals facing the old options, and the bottom panel reports the fit for the individuals facing the new options. Overall, the fit is quite reasonable. For example, the predicted average spending is within 10 percent of actual average spending under both the original and new options, and the medians also fit quite well. We tend to over predict the fraction of individuals who have no spending under the new options, but this again is likely driven by the relatively small number of employees who are switched to the new options in our estimation sample.[37]

Finally, we note that if we simulate data based on our parameter estimates and then run the difference-in-differences analysis we report on in Section 3, we predict about an 8 percent reduction in spending associated with moving from the old option set to the new option set. This is broadly similar to the difference-in-differences estimates we obtained for the actual data (Table 6, columns (1)-(3)). However, given how imprecise our difference-in-differences estimates are, both in the actual data and in the simulated data, we caution against making too much of any comparison. The lack of statistical significance of the difference-in-differences estimate in the simulated data, relative to the reasonably precise estimates of the model parameters, suggests that a more complete model of unobservable heterogeneity and endogenous plan choice is important in increasing precision.

6.3 Moral hazard estimates

The parameter ω_i captures moral hazard in our model . Recall that, abstracting from the truncation of spending at zero, employee i would spend λ_{it} in year t if he had no insurance, and with full insurance would spend $\lambda_{it}+ \omega_i$. Thus, ω_i can be thought of as the scope for moral hazard. As discussed, the top panel of Table 8(b) reports that the estimated average of ω_i is about 1,330 dollars, or about 30% of the estimated health risk (the average of λ_{it}).

[37] To conserve on space, Figure 2 pools individuals across coverage tiers, but the fit within singles or non-singles looks similar to the results pooled by coverage tier, and the predicted differences in spending between singles and non-singles are similar to the observed ones.

Table 10 reports an alternative way one could quantify moral hazard. In the top row of the table, we calculate each employee's expected decline in medical expenditure as we move him from the highest to the lowest coverage in the new options. We will feature the move (or choice) between these two options in all of our subsequent counterfactual exercises. Recall that, as we have modeled these options, moving from the highest to the lowest coverage primarily entails moving someone from a plan with no deductible to a plan with a high deductible, specifically a $3,000 deductible for non-single coverage, or $1,500 for single coverage (Table 2). We estimate that the average spending effect from this move is $348. The second row reports a similar exercise, but considers moving individuals from full insurance to no insurance. We estimate an average spending reduction of $1,273; this is slightly lower than the average ω_i of $1,330 reported earlier (see Table 8(b)) precisely because of the truncation of spending at zero.

These economically meaningful estimates of moral hazard satisfy one necessary condition for selection on moral hazard – the focus of our paper – to be important. A second necessary condition is that moral hazard be heterogeneous. Indeed, we find important heterogeneity in our moral hazard estimates across individuals. For example, the estimated variance of $log(\omega)$ is about one, and highly statistically significant (Table 8(a)), implying that an employee who is one standard deviation above the mean is associated with a moral hazard parameter that is almost three times greater than the mean, and an employee who is one standard deviation below the mean has a moral hazard parameter that is less than a half of the mean. As shown in the top of Table 8(b), across individuals, the standard deviation of ω_i is almost $3,200, and the coefficient of variation of ω is more than 2.

Again, Table 10 reports more empirically-motivated measures of heterogeneity in moral hazard. The top row shows that the spending decline as we move individuals from the no deductible plan to the high deductible plan has a standard deviation of $749, compared to the mean of $348. The median spending reduction is only $48, while the 90th percentile exhibits a spending reduction of more than a thousand dollars. Similarly, as we move individuals from full insurance to no insurance, we estimate that the median reduction in spending is $310, but the 90th percentile of the spending reduction distribution is greater than $3,000.

We briefly explored the extent of heterogeneity in moral hazard implied by our results by the same observable characteristics we explored in the descriptive evidence in Table 7. We found generally similar results. For example, we estimate the average ω to be larger for older vs. younger individuals ($1,590 vs. $1,080, respectively). We also estimate a larger average ω for those who chose more vs. less coverage in 2003, which is consistent with selection on moral hazard. We now turn to a more systematic examination of selection on moral hazard.

6.4 Selection on moral hazard

The fact that individuals are heterogeneous in their moral hazard response to coverage does not of course mean that they select on it in any quantitatively meaningful way. That is, it is conceivable that heterogeneity in other factors is more important in determining plan choice. As one way to

gauge the quantitative importance of selection on moral hazard, we examine how the choice of coverage varies with the quantiles of the marginal distribution of moral hazard ω, and compare this to how the choice of coverage varies with the quantiles of the marginal distribution of risk aversion ψ, and of expected health risk $E(\lambda)$. Once again, we focus on the choice between the highest coverage and lowest coverage plan in the new options (see Table 2). Loosely, our exercise resembles the introduction of a high deductible health insurance plan into a setting where previously there was only a no deductible plan. We set the premiums so that, on average, 10 percent of our sample chooses the high deductible plan.

Figure 3 reports the results. It shows the fraction of individuals choosing the high deductible coverage, conditional on the individual being in each quantile of the marginal distribution of moral hazard ω, of risk aversion ψ, and of expected health risk $E(\lambda)$. We present two different sets of results. The top panel presents the pattern while taking as given the underlying correlation structure among these objects. This panel can be thought of as giving the empirical answer to the question of how much selection there is, on net, on each of the latent primitives that we model. Given the flexible correlations we allow for, these patterns are a-priori of ambiguous sign. The bottom panel repeats the same exercise but "shuts down" the effect of the correlation structure. To do so, we compute the marginal distributions (unconditional on observables) of each of the three latent variables that affect plan choice (ω, ψ, and $E(\lambda)$), and draw values for the other two latent variables independently of the value of the variable for which the graph is drawn. This panel can be thought of as giving the answer to the conceptual comparative static exercise of how much selection there is on one latent factor, holding the other factors constant. As discussed previously, demand for higher coverage generally increases in expected health risk, in risk aversion, and in moral hazard. Our purpose here is to assess the relative magnitudes. Taken together, the two panels help inform not only whether empirically there is selection on moral hazard and of what sign (top panel) but also the extent to which any such selection is primarily "direct" selection based on moral hazard rather than "indirect" selection arising from the correlation structure between moral hazard and other factors which may be driving plan choice.

The results in the top panel indicate that empirically there is selection on moral hazard of the expected sign, with higher moral hazard types (higher ω) less likely to choose the high deductible plan. In terms of the substantive importance of this selection, both panels reveal a similar qualitative pattern: selection on moral hazard is substantially larger than selection on risk aversion and of similar magnitude to selection on health risk. For example, the top panel indicates that moving from the 10th percentile to the 90th percentile of the moral hazard distribution is associated with about a 23 percentage point decline in the demand for the high deductible plan, while moving from the 10th to the 90th percentile of the expected health risk distribution is associated with about a 24 percentage point decline in the demand for the high deductible plan. While some of this reflects the underlying correlation structure, the "pure" comparative static shown in the bottom panel produces quite comparable magnitudes. This suggests that much of this selection on moral hazard is "direct" selection. In other words, in making plan choices, individuals select not only based on their expected level of spending that they would incur with no insurance, but also on

their expected slope, or incremental spending due to insurance.

By contrast, we find selection on risk aversion considerably less important than selection on either moral hazard or expected health. In our data (see Figure 3(a)) there is very little variation in demand for the high deductible plan across the centiles of the risk aversion distribution (reflecting various correlations), and even the "pure" comparative static (Figure 3(b)) suggests only about a 15 percentage point range between the 10th and 90th percentile.

6.5 Implications for spending

We investigate the implications of the selection on moral hazard that we detect for attempts to combat moral hazard through higher consumer cost sharing. To this end, we perform counterfactual analyses of the spending reduction associated with introducing a lower coverage option. Given our finding that higher "moral hazard types" prefer greater coverage, accounting for this selection on moral hazard suggests that introducing plans with greater consumer cost sharing will produce less of a spending reduction than would be estimated if selection on moral hazard were ignored, and it were assumed that those who select the lower coverage option are drawn at random from the "moral hazard type" distribution.

In the health care sector, the impact of consumer cost sharing on moral hazard is an issue of considerable policy as well as academic interest. The size and rapid growth of the health care sector, and the pressure this places on public sector budgets, has created great interest among both academics and policymakers in possible approaches to reducing health care spending. Encouraging individuals to enroll in plans with higher consumer cost sharing, such as the tax-advantaged Health Savings Accounts (HSAs) designed to increase enrollment in high deductible plans, is seen as one potentially promising approach to reducing health spending.

To examine the implications of selection on moral hazard for analysis of such efforts, Figure 4 engages in the same exercise as in Figure 3 of giving employees in our sample a choice between the no deductible and high deductible health insurance plans in the new options. In Figure 3 we fixed the price of each option and reported the fraction of each quantile of a latent variable who choose each plan. In Figure 4 we instead gradually increase the (relative) price of the higher coverage (no deductible) option, and ask how selected is the group of employees who endogenously select the lower coverage (high deductible) option at each given price. To show the extent of selection, the figure reports the average per employee decline in annual spending for those employees who endogenously select the high deductible plan at each price.

The figure illustrates strong selection on moral hazard, especially when the share of the high deductible plan is small. For example, when the price of the no deductible coverage is low enough so that only 10 percent of the employees select the high deductible coverage, the average (per employee) spending decline for those who select the high deductible plan instead of the no deductible plan is just over $130. By contrast, were all employees to choose the high deductible plan instead of the no deductible plan, we estimate the per employee spending decline would be about $350. As noted in the introduction, the common practice in the literature on health insurance and moral

hazard is to look for experimental variation that randomly moves individuals across plans. Such variation would recover the unconditional average effect of coverage (which is $348 in our context); this does not account for selection on moral hazard and will therefore substantially over-estimate the spending reduction associated with the introduction of the high deductible plan when only a small share of individuals select it.

This selection reflects the earlier observation that, all else equal, individuals that are associated with higher moral hazard (higher ω_i) have higher willingness to pay for insurance, and are therefore the ones that would be the last to switch to the lowest coverage, as we gradually increase the price of highest coverage. It is somewhat interesting that in our setting the selection on moral hazard becomes less important (i.e., the slope of the line in Figure 4 becomes less steep) at higher levels of prices for the no deductible plans (which leads to greater fractions choosing the high deductible plan). The same underlying forces are still in play, but are offset by the correlation structure with other primitives.

6.6 Implications for welfare

Our findings of selection on moral hazard also have implications for policies aimed at reducing selection. Analysis of how to mitigate selection often focuses on risk adjustment – whereby individual's insurance premiums are adjusted on the basis of individual covariates (such as age, gender, and prior health conditions) that are predictive of expected medical spending. From this perspective, the potential for selection on moral hazard suggests that investments in better monitoring technologies – such as coinsurance that varies across diagnoses (e.g., heart attack vs. headache) or types of healthcare (e.g., prescription drugs vs. inpatient services) with different behavioral responsiveness to insurance – may also be effective at ameliorating adverse selection.

Our final set of counterfactual analyses considers these issues of contract design by using our model to go further out of sample to analyze the impact of alternative contract designs on social welfare. Table 11 reports our results. Once again we restrict our attention to a choice between the no deductible and high deductible plans under the new options (Table 2, options 5 and 1 respectively). Throughout this section we make the simplifying assumption of perfect competition for the incremental coverage among providers of the no deducible plan, so that the incremental price of the no deductible plan breaks even for those who provide it: incremental price is equal to incremental cost.[38] We report the implications of various counterfactual contracts for the equilibrium (incremental) premium for the no deductible plan, the share choosing this plan, expected spending per employee, and total welfare (or surplus) per employee. Our primary focus is on the consequence of different contract designs for total welfare (i.e., the sum of consumer welfare and producer welfare) which in our context is the certainty equivalent minus expected costs (see equation (10)).

The first row presents the "status quo" benchmark contract with no (additional) screening or monitoring. As with the observed contracts in our data, individuals are offered a "uniform" price

[38] We normalize the price of the lower coverage option to zero. Given our assumptions of CARA utility and a realized utility that is additively separable in income, the price level does not affect plan choice or welfare.

that only varies by coverage tier, and insurance companies reimburse medical spending, regardless of its origin, based on their contract rules. We estimate that the competitive, average incremental price for the no deductible plan (relative to the high deductible plan) is about $1,570, and that at this competitive price 90% of the employees would select the no deductible plan. We normalize total welfare per employee in this "status quo" benchmark to be zero, so that we can more easily compare the welfare gains from alternative contract designs.

The second row presents our "perfect screening" counterfactual, which eliminates adverse selection. Specifically, we assume that insurers can observe and price on all the determinants of health care utilization that the individual knows at the time of his insurance choice – i.e., all of the components of $F(\lambda)$ as well as ω. We solve for the incremental price of the no deductible plan that breaks even for each employee individually, thereby eliminating the adverse selection that arises from uniform pricing. The results indicate that, as expected, the elimination of adverse selection leads to a lower (average) incremental premium for the no deductible plan, increased coverage (i.e., greater fraction choosing the no deductible plan), and higher welfare. It also leads to lower expected spending since the risk-based pricing disproportionately shifts higher moral hazard (ω) individuals into lower coverage. We estimate the welfare gain per employee from eliminating adverse selection to be about $52.[39]

Of particular interest is the contribution of eliminating selection on moral hazard to the welfare gain from eliminating selection. Row 3 explores this by reporting the welfare gain from eliminating only selection on moral hazard (ω) but continuing to allow selection on health risk ($F(\lambda)$). Specifically, we allow insurers to observe ω and price on it, but not on $F(\lambda)$. This is of course not a very sensible scenario, since presumably if insurers could observe ω they could also refuse to reimburse on it, and thus eliminate moral hazard entirely (not just selection on moral hazard). But it is a conceptually useful way to examine the welfare cost of different sources of selection. The results in row 3 suggest that the welfare cost of selection on moral hazard is $34, or about 65 percent of the $52 total welfare cost of selection from row 2.

In an analogous fashion, we can investigate the contribution of eliminating selection on moral hazard to the total welfare gain from eliminating moral hazard. In our setting, the welfare gain from eliminating moral hazard stems from two sources: removing the allocative inefficiency that arises from selection on moral hazard and eliminating the "traditional" moral hazard distortion that comes through socially inefficient over-utilization of health care. We show the results from eliminating moral hazard in the fourth row, which presents our "perfect monitoring" counterfactual. Here we assume that insurance coverage only applies to "λ-related" spending, which in the context of our model means that instead of reimbursing based on actual spending (i.e., reimbursing $m - c_j(m)$), the

[39] By way of perspective, we calculate the total surplus from perfect screening relative to everyone being in the high deductible plan to be $1,084, so that mispricing due to adverse selection appears to reduce welfare by only a small amount relative to the total surplus at stake. Although not the focus of our paper, this finding is consistent with other recent empirical papers on the welfare costs arising from inefficient pricing due to adverse selection; see Einav, Finkelstein, and Levin (2010) for a discussion of some of this recent literature.

contracts reimburse $\max\{\lambda, 0\} - c_j(\max\{\lambda, 0\})$ regardless of what the actual spending is. In such situations, optimizing individuals would spend $\max\{\lambda, 0\}$, which would be the socially efficient level of spending. Row 4 of Table 11 indicates that, relative to the status quo (row 1), this elimination of moral hazard reduces spending by more than \$1,100 per employee (column 3) and increases welfare by about \$490 per employee, which is an order of magnitude greater relative to the welfare gain associated with eliminating adverse selection through perfect screening (row 2).

To examine the relative contribution of selection on moral hazard to this welfare cost, in row 5 we again consider an artificial counterfactual. Specifically, we assume that individuals make their contract choices in the first period as if they are faced with the "perfect monitoring" contracts (row 4), but then in the second period make their spending decision faced with the observed contracts that reimburse in the same manner as the actual contracts (i.e., reimburse based on m rather than based on λ). This allows us to isolate the welfare gain from eliminating solely selection on moral hazard, while preserving the distortion in second period consumption caused by moral hazard. The results suggest that eliminating selection on moral hazard can achieve welfare gains of \$25 per employee, or only about 5 percent of the total welfare cost of moral hazard (row 4).

Overall, these results suggest that, in our setting, selection on moral hazard contributes non-trivially to the total welfare cost of selection, but contributes much less relative to the total welfare cost of moral hazard. At a broad level, our findings suggest that in thinking about contract design, traditional approaches to combatting moral hazard may well aid in combatting selection, and possibly vice versa. Of course, our quantitative estimates undoubtedly depend on our specific setting (contracts and population) and on our modeling assumptions. While there is not much we can do about the former (at least in the current paper), we investigate the latter in the next section.

6.7 Robustness

Table 12 briefly explores the robustness of some of our main findings. Overall, we find that the main results are quite stable across alternative specifications. All the alternative specifications we explore give rise to quantitatively similar estimates of average moral hazard (column (1)), heterogeneity in moral hazard (column (2)), selection on moral hazard (column (4)), the implications of accounting for selection on moral hazard for the spending reduction that can be achieved by offering a high deductible plan (column (5) vs. column (1)), and the contribution of selection on moral hazard to the overall welfare cost of adverse selection (columns (7) relative to column (6)).

The first row replicates our baseline findings reported earlier. The next two rows explore the sensitivity of our findings to trying to account for various institutional features that our baseline specification abstracted from. Row 2 explores the sensitivity of our findings to trying to account for the fact that the lowest coverage option under the new options (option 1) has a health reimbursement account (HRA) component (see Section 2 for details) which we abstracted from in our econometric specification. To do so, we simply drop from the sample the 2004 observations associated with employees who chose option 1 when offered the new choice set (roughly 6 percent of those offered the new choice set).

Row 3 provides one way of gauging the potential importance of "passive choices" for our results. As noted earlier, an attraction of our setting is that for employees who are offered the new choice set in 2004, there is no option of staying with their existing plan. However, there were defaults for those who did not make an "active" choice under the new options. To account for – and exclude – a set of potentially passive choosers, we identified all individuals whose coverage choices under the new benefit options for each of five different insurance options (health, drug, dental, short-term disability, and long-term disability) are consistent with the defaults for those five options.[40] Row 3 shows the results of excluding the 2004 observations for the approximately 12 percent of individuals offered the new options for whom all of their coverage decisions are consistent with the default options.

The remaining rows of the table investigate the sensitivity of our findings to some alternative natural parameterizations of the model. In row 4 we remove all of the demographic covariates from the model (i.e., age, gender, job tenure, income, and health risk score) leaving only indicator variables for year and treatment group (to capture the quasi-experimental variation in the option set) and coverage tier dummies (because the prices of the options depend on coverage tier). In row 5 we allow for heteroskedastic errors, by letting all the parameters in the variance-covariance matrix (see equation (13)) depend on all the covariates. In row 6, instead of assuming that $\log \omega_i$, $\log \psi_i$, and $\overline{\mu_{\lambda,i}}$ are drawn from a joint normal distribution, we assume that they are drawn from a mixture of two normals**.**

While there is, of course, a potentially limitless set of alternative specifications one could investigate, we found the stability of the core results to the natural ones we tried reassuring about the stability of our model estimates within our context. As noted previously, whether or not the results would generalize – quantitatively or even qualitatively – to other option sets, populations, or different models of coverage choice and utilization – is of course an open question.

7 Conclusions

This paper takes a first step toward marrying empirical analysis of selection with that of moral hazard. The active (and growing) empirical literature on insurance demand has focused almost exclusively on selection on risk type or risk preferences, and largely abstracted from moral hazard.[41] The large and venerable literature on moral hazard in insurance has largely focused on average moral hazard effects, abstracting from potential heterogeneity as well as potential selection on that heterogeneity. In this paper we introduced the (to our knowledge) previously overlooked potential for selection on moral hazard, or in other words, the possibility that individuals' anticipated behavioral response to insurance contracts affects their contract choice.

[40] Employees make their choices for each insurance domain all at the same time, on the same benefit worksheet during open enrollment period. Einav, Finkelstein, Pascu, and Cullen (2010) provide more detail and discussion of these other benefits options and choices.

[41] See Einav, Finkelstein, and Levin (2010) for a recent discussion of this literature.

We explored the existence, nature, and implications of selection on moral hazard empirically in the context of the employer-provided market for health insurance in the United States. We estimate substantial heterogeneity in moral hazard and selection on it, with individuals who have a greater behavioral response to the contract (i.e., greater "moral hazard type") demanding more coverage. We estimate that "moral hazard type" is roughly as important as health expectations in determining whether to buy a higher or lower deductible. In other words, selection based on the expected slope of spending (i.e., incremental spending due to insurance) appears about as quantitatively important in our setting as "traditional" selection based on the expected level of spending (i.e., health risk type). Such selection on moral hazard can have important implications for traditional analysis of either selection or moral hazard. For example, we estimate that if we ignored selection on moral hazard, we could estimate a spending reduction associated with introducing a high deductible plan that is substantially larger than what we estimate when we account for the fact that those who select the high deductible plan have a disproportionately low behavioral response to such cost sharing.

Needless to say, our quantitative estimates are highly specific to our particular population and our particular counterfactual analyses. Nonetheless, at a broad level, they illustrate the potential importance of selection on moral hazard for understanding the welfare consequences of both selection and moral hazard. They also illustrate some of the potential implications of selection on moral hazard for policies designed to ameliorate these welfare costs. They suggest, for example, that efforts to reduce health spending by introducing health insurance options with high consumer cost sharing – such as the high deductible plans available through Health Savings Accounts – may produce substantially smaller spending reductions than would have been expected based on the existing estimates of moral hazard effects in health insurance which has ignored selection on moral hazard. They also suggest that improvements in monitoring technology – traditionally thought of as a way to reduce moral hazard – may have the ancillary benefit of ameliorating some of the welfare costs of selection.

Given the importance of the topic, we hope that future work will explore selection on moral hazard in other contexts and in other ways. As noted, we know of very little work that even examines heterogeneity in moral hazard effects, let alone selection of insurance on this heterogeneity. Both the approaches taken in this paper and those suggested (but not explored) by Einav, Finkelstein and Cullen (2010, Section III.D) for estimating heterogeneity in moral hazard effects and its correlation with demand should be fruitful to apply in other settings. In addition, our analysis has focused exclusively on the spending and welfare implications of selection on moral hazard for a given set of contracts; it would be interesting to consider, both theoretically and empirically, the implications of selection on moral hazard for richer analyses of contract designs.

References

Arrow, Kenneth (1963). "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review* 53(5), 941-973.

Arrow, Kenneth (1971). *Essays in the Theory of Risk-Bearing.* Markham: Chicago, IL.

Bajari, Patrick, Han Hong, Ahmed Khwaja, and Christina Marsh (2010). "Moral Hazard, Adverse Selection and Health Expenditures: A Semiparametric Analysis." Manuscript, University of Minnesota.

Berry, Steve, and Ariel Pakes (2007). "The Pure Characteristics Model of Demand." *International Economic Review* 48(4), 1193-1225.

Bundorf, Kate M., Jonathan Levin, and Neale Mahoney (2009). " Pricing and Welfare in Health Plan Choice." Mimeo, Stanford University.

Busch, Susan H., Colleen L. Barry, Sally J. Vegso, Jody L. Sindelar, and Mark R. Cullen (2006). "Effects Of A Cost-Sharing Exemption On Use Of Preventive Services At One Large Employer." *Health Affairs* 25(6), 1529-1536.

Cabral, Marika (2009). "How Does Preventive Care Usage Change When Patient Prices Decrease?" Manuscript, Stanford University.

Cameron, A. Colin, Jonah Gelbach, and Douglas Miller (2011). "Robust Inference with Multi-way Clustering." *Journal of Business and Economic Statistics* 29(2), 238-249.

Cardon, James H., and Igal Hendel (2001). "Asymmetric Information in Health Insurance: Evidence from The National Medical Expenditure Survey." *Rand Journal of Economics* 32, 408-427.

Carneiro, Pedro, James J. Heckman, and Edward Vytlacil (2010). "Estimating Marginal returns to Education." NBER Working Paper No. 16474.

Carlin, Caroline, and Robert Town (2010). "Adverse Selection: The Dog that Didn't Bite." Manuscript, University of Minnesota.

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

Chiappori, Pierre-André, and Bernard Salanié (2000). "Testing for Asymmetric Information in Insurance Markets." *Journal of Political Economy* 108, 56-78.

Cohen, Alma, and Liran Einav (2007). "Estimating Risk Preferences from Deductible Choice." *American Economic Review* 97(3), 745-788.

Cutler, David and Richard Zeckhauser. (2000). "The Anatomy of Health Insurance." In A.J. Culyer and J.P. Newhouse (eds) *Handbook of Health Economics* Volume 1A. Amsterdam: North-Halland.

Dubin, Jeffrey A., and Daniel L. McFadden (1984). "An Econometric Analysis of Residential Electric Appliance Holdings and Consumption." *Econometrica* 52(2), 345-362.

Einav, Liran and Amy Finkelstein (2011). "Moral hazard in health insurance: how important is forward looking behavior?" Manuscript in progress, MIT and Stanford University.

Einav, Liran, Amy Finkelstein, and Mark R. Cullen (2010). "Estimating Welfare in Insurance Markets using Variation in Prices." *Quarterly Journal of Economics* 125(3), 877-921.

Einav, Liran, Amy Finkelstein, and Jonathan Levin (2010) "Beyond Testing: Empirical Models of Insurance Markets." *Annual Review of Economics*, 2, September 2010, 311-336.

Einav, Liran, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen (2010). "How General are Risk Preferences? Choices Under Uncertainty in Different Domains." NBER Working Paper 15686.

Einav, Liran, Amy Finkelstein, and Paul Schrimpf (2010). "Optimal Mandates and The Welfare Cost of Asymmetric Information: Evidence from The U.K. Annuity Market." *Econometrica* 78(3), 1031-1092.

Fang, Hanming, Michael Keane, and Dan Silverman (2008). "Sources of Advantageous Selection: Evidence from the Medigap Insurance Market." *Journal of Political Economy* 116, 303-350.

Finkelstein, Amy, and Kathleen McGarry (2006). "Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market." *American Economic Review* 96(4), 938-958.

Gilks, W. R., N.G. Best, and K.K.C. Tan (1995). "Adaptive Rejection Metropolis Sampling." *Applied Statistics* 44, 455-472.

Goettler, Ronald L., and Karen Clay (forthcoming). "Tariff Choice with Consumer Learning and Switching Costs." *Journal of Marketing Research*, forthcoming.

Handel, Benjamin (2010). "Adverse Selection and Switching Costs in Health Insurance Markets: When Nudging Hurts." Manuscript, UC Berkeley.

Heckman, James J., Sergio Urzua, and Edward Vytlacil (2006). "Understanding instrumental variables in models with essential heterogeneity." *Review of Economics and Statistics*, 88, 389-432.

Holmstrom, Bengt (1979). "Moral hazard and observability." *Bell Journal of Economics* 10(1), 74-91.

Karlan, Dean, and Jonathan Zinman (2009). "Observing Unoberservables: Identifying Information Asymmetries with a Consumer Credit Field Experiment." *Econometrica,* 77(6), 1993-2008.

Keane, Michael, and Robert Moffitt (1998)."A Structural Model of Multiple Welfare Program Participation and Labor Supply." *International Economic Review* 39(3), 553-589.

Keeler, Emmett B., and John E. Rolph (1988). "The Demand for Episodes of Treatment in the Health Insurance Experiment." *Journal of Health Economics* 7, 337-367.

Kowalski, Amanda E. (2010). "Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care." NBER Working Paper No. 15085.

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

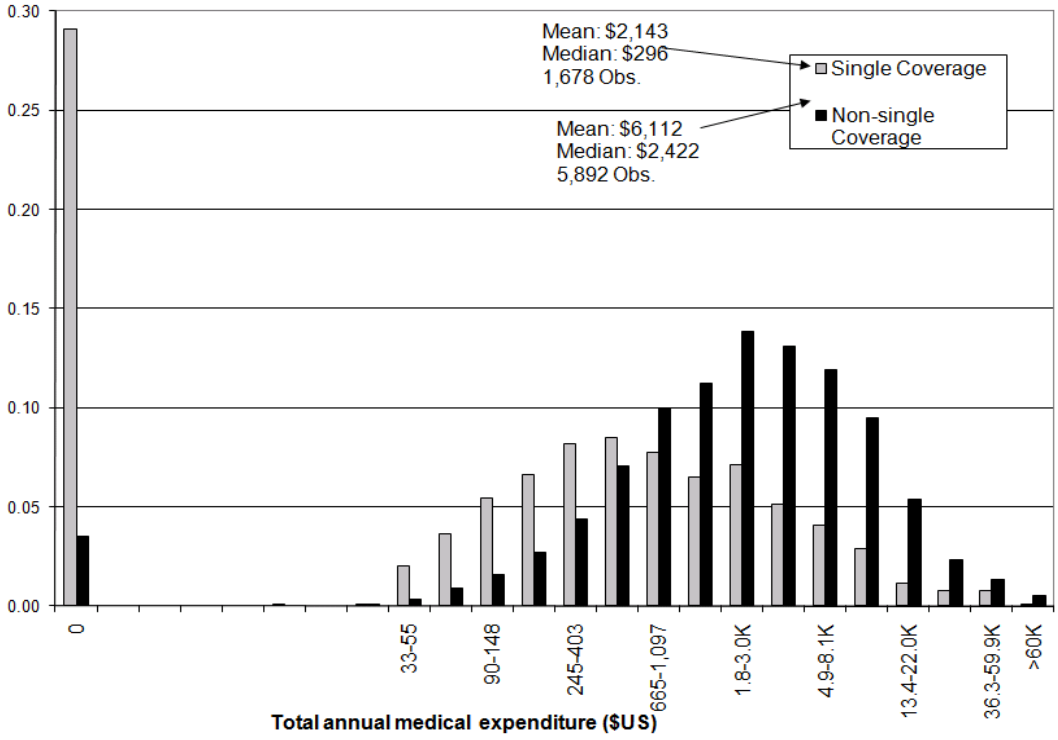
Miravete, Eugenio J. (2003). "Choosing the Wrong Calling Plan? Ignorance and Learning." *American Economic Review* 93(1), 297-310.

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

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

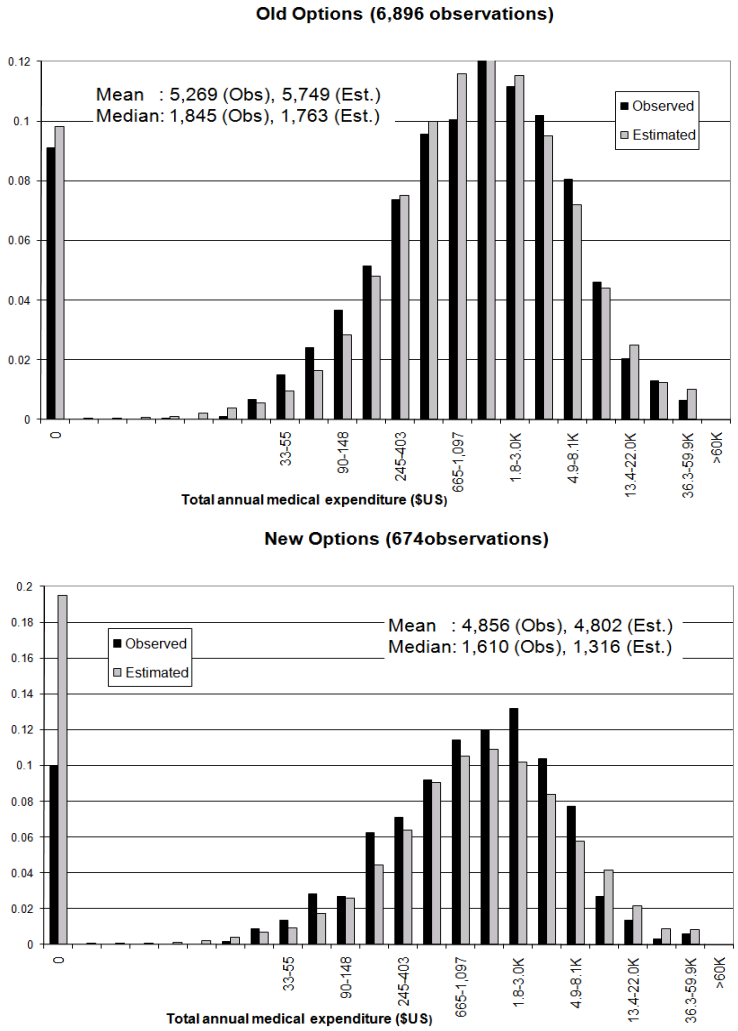
Wooldridge, Jeffrey M. (2002). *Econometric Analysis of Cross Section and Panel Data.* MIT Press: Cambridge, MA.

Figure 1: The cross-sectional distribution of medical expenditure



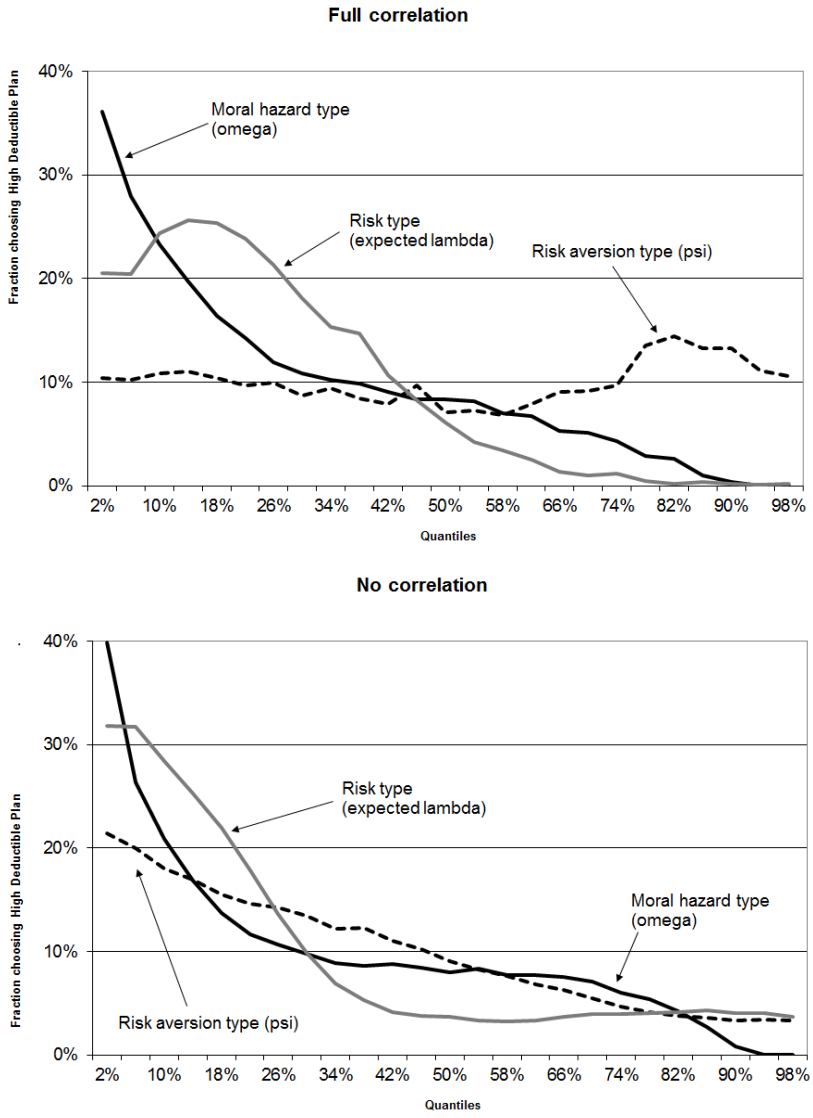
The figure presents the distribution of total annual medical expenditure for each employee (and any covered dependents) in our baseline sample. The graph uses a log scale, such that the second bin covers expenditure lower than exp(0.5), the next covers expenditures between exp(0.5) and exp(1), and so on; the x-axis labels show the corresponding dollar amounts of selected bins. An observation is an employee-year, pooling data from 2003 and 2004. The grey bars correspond to employees with a single coverage, while the black bars correspond to employees who also covered additional dependents (spouse, children, or both).

Figure 2: Model fit – medical spending distributions



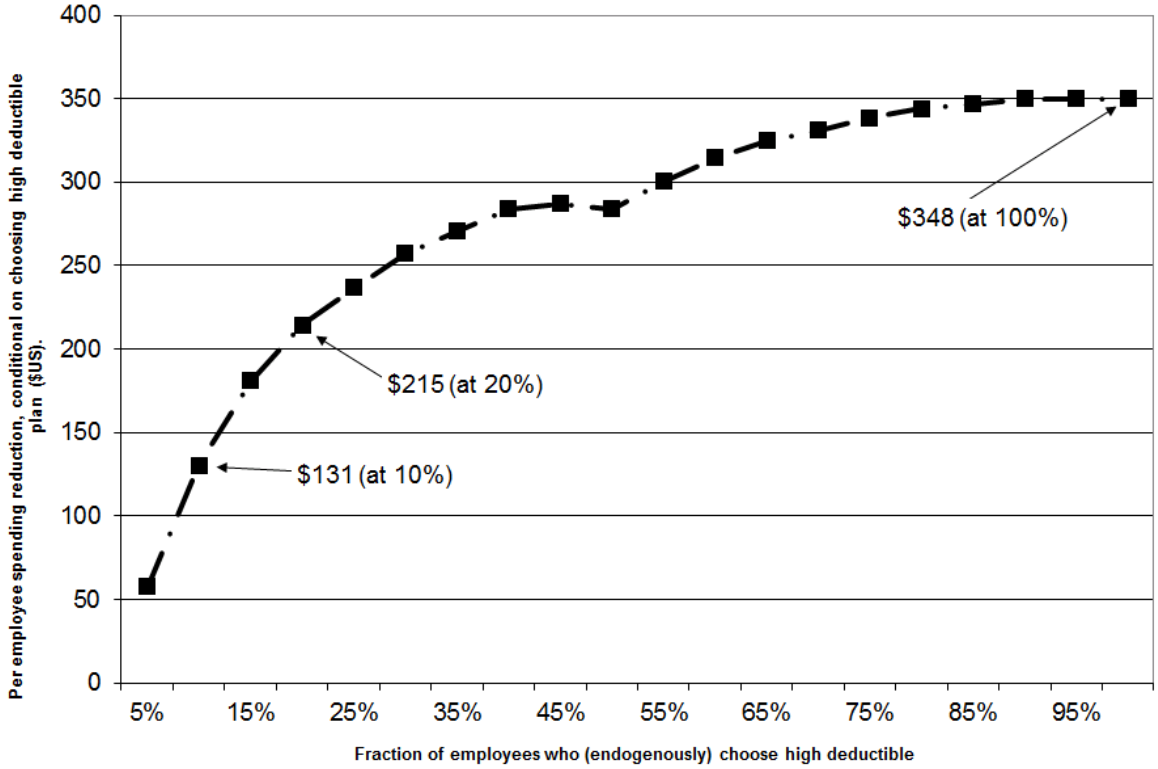
The figure presents the distribution of total annual medical expenditure, in the data and in model simulations based on the estimated parameters. The graph uses a log scale, such that the second bin covers expenditure lower than exp(0.5), the next covers expenditures between exp(0.5) and exp(1), and so on; the x-axis labels show the corresponding dollar amounts of selected bins. The top panel compares spending of individuals who faced the original options, and the bottom panel compares the spending distribution of individuals who faced the new options.

Figure 3: Selection on moral hazard relative to other sources of selection



The figure illustrates the relative importance of the three different sources of selection that we model. We consider an individual's choice between two available options: the no deductible and high deductible plans among the new set of options (see Table 2, options 5 and 1 respectively). We assume the observed (averaged within each coverage tier) premiums for these two options. Each point in the figure indicates the fraction of individuals choosing the high deductible (i.e. low coverage) option relative to the no deductible (high coverage) option. We consider three sources of selection: $E(\lambda)$ (risk), ω (moral hazard), and ψ (risk aversion). For each of them, we compute the fraction choosing the high deductible at different quantiles of the distribution. In the top panel, we take into account the correlation between each component and the others, while in the bottom panel we repeat the same exercise but draw the other components of the model randomly from their marginal distribution (that is, assuming no correlation).

Figure 4: Spending implications of selection on moral hazard



The figure illustrates the potential spending implications arising from selection on moral hazard. To construct the table, we use an exercise similar to the one used for Figure 3. For each individual, we use the model estimates to compute his decline in expected annual expenditure as we move him from the highest coverage (no deductible) to the lowest coverage (high deductible) in the new benefits options (see Table 2, options 5 and 1 respectively). We then vary the relative price of the highest coverage, allowing employees to endogenously choose between the two options, and report the per-employee expected decrease in spending for the group of individuals who chooses the lowest coverage at each price. Without selection on moral hazard, the curve would have been flat. Selection on moral hazard implies that those with the lowest moral hazard effects of insurance are those who have the lowest willingness to pay for incremental coverage and are therefore the first (as the price of coverage increases) to switch from higher to lower coverage. Ceteris paribus, therefore, selection on moral hazard generates an upward sloping curve; this can be offset through the correlation between moral hazard and other components of demand (such as risk aversion or health risk).

Table 1: Summary statistics for 2003 sample

	Obs.	Average Age	Average Annual Income	Average Tenure with Alcoa	Fraction Male	Fraction White	Fraction Single Coverage	Average Health Risk Score	Avg number of insured family members (if non-single coverage)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Baseline sample	3,995	41.3	31,292	10.2	0.84	0.72	0.23	0.95	2.8
Switched in 2004	682	44.5	39,715	15.5	0.96	0.85	0.21	1.06	2.7
Switched in 2005	974	39.7	25,532	8.2	0.73	0.44	0.25	0.91	2.8
Switched in 2006	1,075	38.3	29,952	5.7	0.86	0.82	0.23	0.86	2.9
Switched after 2006	1,264	43.3	32,316	12.7	0.85	0.79	0.22	1.01	2.6

Top row presents statistics based on the 2003 data for our baseline sample, which covers all hourly union workers not covered by the Master Steelworker's Agreement (except those that get dropped in the process of the data cleaning described in the text). The subsequent rows ("Switched in 2004," "Switched in 2005," and so on) partition our baseline sample based on the year in which employees were switched to the new set of health insurance options. Average health risk score in column (8) gives predicted spending for our population (based on demographics and medical conditions) relative to a nationally representative population under age 65; see text for more details. Total annual medical spending in column (10) is for employees and any covered dependents.

[a] Health risk score is normalized, so that 1 indicates the expected medical expenditure for a random draw from a nationally representative under-65 population. To construct this table, we assign each employee the average risk score of all covered family members.

Table 2: Old and new health plans

Panel A: Single coverage (N=1,679)

	Original Plan Options			New Plan Options				
	Option 1	Option 2	Option 3	Option 1[a]	Option 2	Option 3	Option 4	Option 5
Plan features:								
Deductible	1,000	0	0	1,500	750	500	250	0
Out of Pocket Maximum	5,000	2,500	1,000	4,500	3,750	3,500	2,750	2,500
Average Share of Spending Paid Out of Pocket[b]	0.580	0.150	0.111	0.819	0.724	0.660	0.535	0.112
Employee Premium[c]	0	351	1,222	0	132	224	336	496
Fraction choosing each option[d]	3.3%	63.5%	33.2%	14.1%	0.0%	2.2%	37.8%	45.9%

Panel B: Non-single coverage (N=5,895)

	Original Plan Options			New Plan Options				
	Option 1	Option 2	Option 3	Option 1[a]	Option 2	Option 3	Option 4	Option 5
Plan features:								
Deductible	2,000	0	0	3,000	1,500	1,000	500	0
Out of Pocket Maximum	10,000	5,000	2,000	9,000	7,500	7,000	5,500	5,000
Average Share of Spending Paid Out of Pocket[b]	0.495	0.130	0.098	0.732	0.600	0.520	0.387	0.111
Employee Premium[c]	0	354	1,297	0	364	620	914	1,306
Fraction choosing each option[d]	0.6%	56.1%	43.3%	3.9%	0.6%	1.8%	24.4%	69.3%

The table summarizes the key features of the original and new health insurance coverage options. The features shown apply to in-network spending. Not shown are coinsurance rates (applied to those who reached the deductible but have yet to reach the out-of-pocket maximum) which are 10% in all plans (old and new). There are some other small differences between the original and new options that are associated with out-of-network spending, preventive care, and certain treatments associated with co-pays rather than coinsurance in the original set of options. See text for further details.

[a] The New Option 1 includes a health reimbursement arrangement (HRA). Every year the employer sets aside $750 (for single; $1,250 for non-single coverage) that the employee can use (tax free) to pay for a variety of expenses such as deductibles and coinsurance payments. Unused HRA funds roll over to future years and, eventually, can be used during retirement to finance health insurance, provided through the company or through COBRA. Our baseline model abstracts from the HRA component of New Option 1.

[b] To compute the average share of spending out of pocket, we use the 2003 claims from all individuals in the baseline sample and apply each option's coverage details to this (common) sample. We then compute, for each option, the ratio of the resultant out-of-pocket expenses to the total claim mounts, and report the average (across employees in 2003) for each option. As a result, our computed average share of spending out of pocket abstracts from any differential behavioral effect of each contract. It does, however, account for the unmodeled small differences between the new and original options described above and in the text.

[c] Premiums are normalized so that the lowest coverage is free for all employees. This is true in both the original and new options, up to small variation of several hundred dollars across employees. We report the average premium for employees in the baseline sample, pooling 2003 and 2004. Premiums vary by coverage tier; there is also some additional variation (across employees within coverage tier) in the incremental premiums associated with greater coverage options. The variation is based on the business unit to which each employees belongs (see Einav, Finkelstein, and Cullen, 2010).

[d] Statistics are based on all employee in the baseline sample, pooling 2003 and 2004.

Table 3: Plan transitions

Old options in 2004

		Highest coverage	All Other coverages
2003	Highest coverage	40.0%	0.5%
	All Other coverages	0.6%	58.9%

New options in 2004

		Highest coverage	All other coverages
2003	Highest coverage	32.0%	15.8%
	All other coverages	27.8%	24.5%

The table shows transition matrices across plan options for those in the old options in both 2003 and 2004 (top panel) and those who are switched to the new options in 2004 (bottom panel). Under the original options, the highest coverage is option 3. Under the new options, the highest coverage in option 5. See Table 2 for coverage details. The sample is limited to the 6186 employees (82% of the baseline sample) who are in the data in both 2003 and 2004.

Table 4: Spending patterns by coverage level

	Single Coverage			Non-Single Coverage		
	Count	Mean	Median	Count	Mean	Median
Original Plan Options						
Highest coverage	512	3,130	557	2,318	6,634	2,670
All other coverages	1,031	1,795	233	3,035	5,768	2,288
New Plan Options						
Highest coverage	62	1,650	447	375	6,858	2,630
All other coverages	73	560	52	164	3,405	1,481

The table shows (contemporaneous) spending by coverage choice. Under the original options, the highest coverage is option 3. Under the new options, the highest coverage is option 5. See Table 2 for coverage details.

Table 5: Basic difference-in-differences in baseline sample

	Obs.	Mean	Fraction with zero spending	10th	25th	50th	75th	90th
Control (Switched after 2004)								
2003 spending	3,313	5,300	0.09	52	426	1,775	5,178	11,984
2004 spending	2,901	5,250	0.09	55	517	1,889	5,589	12,253
Treated (Switched in 2004)								
2003 spending	682	5,201	0.08	79	581	1,957	5,048	12,644
2004 spending	674	4,856	0.10	22	447	1,610	4,622	9,467
Treated-Control Differences (levels)								
2003 spending		-99	-0.01	27	155	182	-130	660
2004 spending		-394	0.01	-33	-70	-279	-967	-2,786
2004-2003 Difference (levels)								
Control (switched after 2004)		-50	0.00	3	91	114	411	269
Treated (Switched in 2004)		-345	0.02	-57	-134	-347	-426	-3,177
Difference in differences (levels)		-295	0.02	-60	-225	-461	-837	-3,446
Difference (percentages)								
Control (switched after 2004)		-0.9%	0.0%	5.8%	21.4%	6.4%	7.9%	2.2%
Treated (Switched in 2004)		-6.6%	22.0%	-72.2%	-23.1%	-17.7%	-8.4%	-25.1%
Diff. in differences		-5.7%	22.0%	-77.9%	-44.4%	-24.2%	-16.4%	-27.4%

Table 6: Difference-in-differences estimates of impact of change in health insurance options on annual medical spending

	2003-2004 sample			2003-2006 sample		
	OLS in levels (1)	OLS in logs (2)	QMLE-Poisson (3)	OLS in levels (4)	OLS in logs (5)	QMLE-Poisson (6)
Estimated Treatment effect	-297.2	-0.35	-0.06	-591.8	-0.175	-0.114
	(753.7)	(0.19)	(0.15)	(264.2)	(0.12)	(0.048)
	[0.70]	[0.08]	[0.69]	[0.034]	[0.17]	[0.018]
Mean Dependent Variable	5,232	6.91	5,232	5,392	6.9	5,392
N	7,570	7,570	7,570	14,638	14,638	14,638

The table shows the difference-in-difference estimate of the spending reduction associated with moving from the old options to the new options. The unit of observation is an employee-year. Dependent variable is the total annual medical spending for each employee and any covered dependents (or log of 1 + total spending in column (2) and column (5)). The coefficient shown is the coefficient on an indicator variable that is equal to 1 if the employee's treatment group is offered the new health insurance options that year, and 0 otherwise. All regressions include year and treatment group fixed effects. We classify employees into one of four possible treatment groups - switched in 2004, switched in 2005, switched in 2006, or switched later - based on his union affiliation which determines the year in which he is switched to the new health insurance options. Estimation is either by OLS or QMLE Poisson as indicated in the column headings. Standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions; p-values are in [square brackets]. Columns (1)-(3) show estimates for the 2003-2004 sample; Columns (3)-(6) expand the sample to include 2003-2006.

Table 7: Suggestive evidence of heterogeneous moral hazard and of selection on moral hazard

		Obs.	Mean spending	Estimated change in spending associated with change in options (levels)		Avg Out-of-Pocket Share (Old Options)	Avg Out-of-Pocket Share (New Options)	Increase in Out-of-Pocket Share
				Coeff.	Std. Err.			
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
(A)	Above median age (of 43)	6,972	6,263	-1,302	(799)	12.4	27.8	15.4
	Below or equal to median age (of 43)	7,666	4,600	-85.8	(483)	12.9	29.5	16.6
(B)	Male	12,373	5,442	-604	(293)	12.6	29.1	16.5
	Female	2,265	5,120	-579	(693)	12.9	25.8	12.9
(C)	Above median income (of $31,000)	7,322	5,669	-364	(602)	12.2	29.1	16.9
	Below median income (of $31,000)	7,316	5,116	-301	(397)	13	28.1	15.1
(D)	Above median health	7,320	3,321	488	(330)	15.3	31	15.7
	Below median health	7,318	7,462	-1525	(540)	14.7	25.9	11.2
(E)	Less coverage in 2003	6,997	5,003	-621	(513)	13.4	32	18.6
	More coverage in 2003	5,229	6,296	-1,336	(596)	10.1	23.5	13.4

The table shows results for different groups of workers (shown in different rows) in the 2003-2006 sample. Column (1) reports the number of employee-years in the sample, and column (2) reports their mean annual medical spending over the sample period. Columns (3) and (4) report, respectively, the coefficient and standard error of the estimated change in spending associated with moving from the old to the new options. This is based on a difference-in-differences regression on the 2003-2006 sample; we report in columns (3) and (4) the coefficient and standard error on an indicator variable that is equal to 1 if the employee's treatment group is offered the new health insurance options that year, and 0 otherwise. The dependent variable is always total annual medical spending for each employee and any covered dependents. All regressions include year and treatment group fixed effects. Standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions. Columns (5) and (6) show the average out of pocket share within each group under the old and new options respectively. These are calculated based on the share of employees within each group in each plan, and the plan specific out of pocket shares shown in Table 2 (which are computed on a common sample of workers across plans). Column (7) reports the increase in the average out of pocket share for each group associated with moving from the old options to the new options. In panel (D), the sample is split into above and below median health based on the employee's health risk score, which is a prediction of future medical spending on the basis of prior year detailed medical diagnoses and claims, as well as demographics. In panel (E), the sample is limited to employees who are employed at the firm in 2003 and who choose either "more coverage" (option 3 from Table 2) or "less coverage" (option 2 from Table 2) in 2003.

Table 8(a): Parameter estimates

Mean Shifters				
	μ_λ (Health risk)	κ_λ (Health risk)	$ln(\omega)$ (Moral hazard)	$ln(\psi)$ (Risk aversion)
Constant	6.11 (0.14)	-389 (73)	5.31 (0.24)	-5.57 (0.10)
Coverage tier				
Single	(omitted)	(omitted)	(omitted)	(omitted)
Family	0.19 (0.08)	57 (51)	-0.58 (0.18)	-0.88 (0.07)
Emp+Spouse	0.27 (0.09)	44 (53)	-0.66 (0.22)	-0.95 (0.07)
Emp+Children	0.24 (0.08)	185 (47)	-0.28 (0.21)	-0.91 (0.06)
Treatment group				
Switch 2004	-0.01 (0.07)	-278 (43)	-0.24 (0.11)	-0.31 (0.05)
Switch 2005	-0.10 (0.06)	-78 (38)	0.07 (0.12)	-0.23 (0.05)
Switch 2006	0.12 (0.07)	-94 (37)	0.01 (0.12)	-0.07 (0.05)
Switch later	(omitted)	(omitted)	(omitted)	(omitted)
Demographics				
Age	-0.01 (0.003)	-5 (1.8)	-0.01 (0.006)	0.01 (0.002)
Female	0.18 (0.08)	94 (39)	-0.08 (0.13)	-0.07 (0.06)
Job Tenure	0.002 (0.003)	-2.3 (1.6)	0.002 (0.004)	0.003 (0.002)
Income	0.003 (0.002)	6 (0.9)	0.001 (0.003)	-0.0003 (0.001)
Health risk score				
1st quartile (< 1.119)	(omitted)	(omitted)	(omitted)	(omitted)
2nd quartile (1.119 to 1.863)	0.91 (0.07)	305 (59)	0.13 (0.29)	-0.41 (0.06)
3rd quartile (1.863 to 2.834)	1.48 (0.08)	242 (81)	1.79 (0.27)	-0.66 (0.06)
4th quartile (> 2.834)	2.05 (0.09)	-416 (120)	3.38 (0.22)	-0.89 (0.07)
2004 Time dummy	-0.12 (0.02)	--	--	--
Variance-covariance matrix				
	μ_λ_bar	$ln(\omega)$	$ln(\psi)$	
μ_λ_bar	0.20 (0.03)	-0.03 (0.04)	-0.12 (0.02)	
$ln(\omega)$	--	0.98 (0.08)	-0.01 (0.03)	
$ln(\psi)$	--	--	0.25 (0.02)	
Additional parameters				
σ_μ	0.33 (0.03)			
σ_κ	290 (12)			
γ_1	0.04 (0.004)			
γ_2	15 (1.2)			

The table presents our baseline parameter estimates based on our baseline sample of 7,570 employees. As described in the text, the estimates are based on a Gibbs sampler; the table reports the posterior mean and the posterior standard deviations in parentheses. All time varying demographics are set to their mean over the two years, except for health risk score which we allow to be a time varying shifter of μ_λ. "Treatment group" dummies refer to indicator variables based on the year your union's benefits were switched to the new benefits. Higher risk scores correspond to worse predicted health.

Table 8(b): Implied quantities

Unconditional statistics			
	$E(\lambda)$	ω	ψ
Average	4,340 (200)	1,330 (59)	0.0019 (0.00002)
Std. Deviation	5,130 (343)	3,190 (320)	0.0020 (0.00007)
Unconditional correlations			
	$E(\lambda)$	ω	ψ
$E(\lambda)$	1.00	0.24 (0.03)	-0.36 (0.01)
ω	--	1.00	-0.15 (0.01)
ψ	--	--	1.00
Marginal Effects on $E(\lambda)$			
Coverage tier			
Single	(omitted)		
Family	360 (75)		
Emp+Spouse	700 (85)		
Emp+Children	300 (69)		
Treatment group			
Switch 2004	260 (71)		
Switch 2005	-320 (67)		
Switch 2006	640 (70)		
Switch later	(omitted)		
Demographics			
Age	-42 (3.6)		
Female	720 (71)		
Job Tenure	5.9 (3.1)		
Income	13 (2)		
Health risk score			
1st quartile (< 1.119)	(omitted)		
2nd quartile (1.119 to 1.863)	1,600 (96)		
3rd quartile (1.863 to 2.834)	4,000 (200)		
4th quartile (> 2.834)	8,500 (370)		
2004 Time dummy	-590 (29)		

The table reports some implied quantities of interest that are derived from the estimated parameters in Table 8(a). Posterior standard deviations are in parentheses.

Table 9: Model fit – choice probabilities

Original options (N = 6,896)		
Plan	*Data*	*Model*
Option 1	1.2%	2.0%
Option 2	58%	57%
Option 3	41%	41%
New options (N = 674)		
Plan	*Data*	*Model*
Option 1	5.9%	5.0%
Option 2	0.5%	5.0%
Option 3	1.9%	1.0%
Option 4	27%	11%
Option 5	65%	76%

The table reports the actual and predicted choice probabilities of each plan. Plans are numbered from lowest to highest coverage. For plan details see Table 2.

Table 10: Spending implications of moral hazard estimates

	Mean	Std. Dev.	10th	25th	50th	75th	90th
Spending difference as we move from no to high deductible plan	348	749	0	0	48	316	1,028
Spending difference as we move from full to no insurance	1,273	3,181	0	86	310	1,126	3,236

The table reports the implied spending implications if we move different employees across plans. For each employee, we use the model estimates to compute his decline in expected annual expenditure as we change his insurance plan. In the top row, we move each employee from the highest coverage option under the new benefit options (option 5) to the lowest coverage option under the new benefit options (option 1); roughly speaking, this entails moving from a plan with no deducible to a plan with a high deductible; see Table 2 for more details. In the bottom row, we move each employee from full to no insurance. The table then summarizes the cross-sectional distribution of the spending effects. The estimates are primarily driven by the estimated distribution of ω, but they take into account the truncation of spending at zero by integrating over the conditional (on ω) distribution of λ.

Table 11: Spending and welfare effects of asymmetric information

		Average equilibrium (incremental) premium	No deductible plan share	Expected spending per employee	Total welfare per employee
(1)	"Status quo": no screening or monitoring	1,568	0.90	5,318	normalized to 0
(2)	"Perfect screening": premiums depend on F(lambda) and omega	1,491	0.91	5,248	52
(3)	"Imperfect screening": premiums depend on omega (but not on F(lambda))	1,523	0.88	5,265	34
(4)	"Perfect monitoring": contracts reimburse only "lambda-related" spending	1,139	0.94	4,185	490
(5)	"Imperfect monitoring": perfect monitoring assumed for choice (but not for utilization)	1,139	0.94	5,327	25

The table reports the spending and welfare effects from a set of counterfactual contracts described in the text. All exercises are applied to a setting in which the only two options available are the no deductible plan and the high deductible plan under the new benefit options (i.e. option 5 and option 1, respectively; see Table 2). Equilibrium premiums are computed as the incremental (relative) premium for the no deductible plan that equals the expected incremental costs associated with providing the no deductible plan to those who choose it. The no deductible plan share is calculated based on the choice probabilities as a function of equilibrium premiums. Expected spending and total welfare are computed based on these choices. Row 1 assumes the "status quo" asymmetric information contracts, which a "uniform" price that varies only by coverage tier. Row 2 assumes "perfect screening", so that contracts are priced based on ω_i and all components of $F_i(\lambda)$ and adverse selection is eliminated. Row 3 assumes "imperfect screening", in which contracts are priced based only on ω_i.Row 4 assumes "perfect monitoring" so that moral hazard is eliminated. Specifically we assume the insurance provider can counterfactually observe (and not reimburse) spending that is associated with moral hazard; spending associated with health – realization of λ – are reimbursed according to the observed contracts. Row 5 assumes "imperfect monitoring" in which, ex ante individuals choose contracts under the assumption that there will be perfect monitoring (i.e. spending associated with moral hazard will not be reimbursed), but ex-post (after they choose their contract but before they make their spending decision) the contracts are changed to be the standard contracts that reimburse all medical spending regardless of its origin.

Table 12: Robustness

	Average moral hazard effect	Moral hazard: coefficient of variation	90-10 difference in probability choose high deductible plan		Average moral hazard effect for "selected" group	Welfare effect of "Perfect Screening"	Welfare effect of "Imperfect Screening"
			Quantiles of risk	Quantiles of moral hazard			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(1) Baseline specification	348	2.15	24.2%	22.9%	131	52	34
(2) Omitting new option 1	424	1.98	23.3%	23.0%	160	61	43
(3) Omitting potentially passive choosers	336	2.18	26.9%	19.8%	153	56	36
(4) No demographic covariates	270	2.11	27.4%	21.4%	114	33	16
(5) Allowing heteroskedasticity	277	2.18	20.1%	18.0%	112	58	28
(6) Mixture of two normals	355	2.19	28.6%	20.7%	144	60	38

The table reports summary results from a variety of specifications. Column 1 ("average moral hazard effect") reports the average (per employee) reduction in spending associated with moving everyone from the no deductible plan to the high deductible plan under the new benefit options (i.e. option 5 and option 1 respectively) and column 2 ("moral hazard: coefficient of variation") reports the standard deviation of this effect relative to the mean; the baseline numbers are shown in Table 10, row 1. Columns 3 and 4 report the difference in the probability an individual chooses the high deductible plan compared to the no deductible plan (if it is priced so that on average 10 percent of the population chooses the high deductible plan) by the quantiles of the marginal distribution of risk type ($E(\lambda)$) and the quantiles of the marginal distribution of moral hazard (ω), respectively; the baseline estimates are shown in Figure 3a. Column 5 reports the average (per employee) reduction in spending for those who choose the high deductible plan when, starting from the no deductible plan, the price of the high deductible plan is set so that only 10 percent of employees select the high deductible plan (see Figure 4 for the baseline estimate). Columns 6 and 7 show, respectively, the welfare gain from "perfect screening" – i.e. contracts are priced based on ω_i and all components of $F_i(\lambda)$ and adverse selection is eliminated – and the welfare gain from "imperfect screening" – i.e. contracts are priced based only on ω_i; the baseline results were shown in Table 11, rows 2 and 3 respectively.

Each row reports the results from a different specification . Row 1 replicates the baseline specification. All other rows show a single deviation from the baseline as specified. Row 2 shows the results omitting the employees who chose the new option 1 in 2004. Row 3 shows the results omitting individuals who may potentially be "passive choosers" in 2004. Row 4 omits all of the demographic covariates from the baseline specification (age, gender, job tenure, income, and health risk score), leaving only dummies for coverage tier, year your benefits were switched, and whether it is 2004. Row 5 allows the variance-covariance matrix (see equation (13)) to depend on the covariates. Row 6 allows the joint distribution of the latent variables (in equation (13)) to be more flexible by allowing it to follow a mixture of two normal distributions.

Appendix A: Additional descriptive results on moral hazard

In this appendix we report in more detail on the results of our difference-in-differences analysis of the impact of the change in health insurance options on healthcare spending and utilization. Specifically, we estimate the impact of the change in coverage separately for different types of healthcare utilization, investigate the validity of our identifying assumption, and explore a number of other additional potential concerns with the analysis. All of the results shown are for the 2003-2006 sample.

Econometric framework The basic difference-in-differences specification (which we used in Tables 5 and 6) is:

$$y_{ijt} = \alpha_j + \delta_t + \beta \cdot Treat_{jt} + \mathbf{x}'_{ijt}\phi + \varepsilon_{ijt}, \tag{17}$$

where y_{ijt} is the outcome variable of interest for employee i in treatment group j at time t. We classify each employee i into one of four possible treatment groups – "switched in 2004," "switched in 2005," "switched in 2006," and "switched later" – based on his union affiliation which determines the year in which he is switched to the new set of health insurance options. The coefficients α_j represent a full set of treatment group fixed effects; these control for any fixed differences across treatment groups. The vector of δ_t 's represents a full set of year fixed effects; these control (flexibly) for any common secular year-to-year changes across all treatment groups.[42] The vector x denotes a set of employee demographic covariates that are included in some of our specifications; there are no such covariates in our baseline specification. We adjust the standard errors to allow for an arbitrary variance-covariance matrix within each of the 28 different unions in our sample.[43]

The main coefficient of interest is β, the coefficient on the variable $Treat_{jt}$. The variable $Treat_{jt}$ is an indicator variable that is equal to 1 if group j is offered the new health insurance options in year t, and 0 otherwise. For example, for the group "switched in 2004" $Treat_{jt}$ is 0 in 2003, and 1 in 2004 and subsequent years, while for the "switched later" group the variable $Treat_{jt}$ is 0 in all years.

Impact on types of medical spending and care utilization Appendix Table A1 examines the impact of the change in health insurance options on the various components of health care spending and health care utilization. We can break out health care spending into doctor visits (approximately 25 percent of the total), outpatient spending (approximately 35 percent of the total), inpatient spending (approximately 35 percent of the total), and other (which accounts for about 4 percent of spending, about half of which

[42] An annual measure is a natural unit of time since it is both the unit of time during which the set of health insurance incentives apply (i.e., cost sharing requirements reset at the beginning of the year) and the time over which the choice of health insurance contract is made. In some additional analysis below we also report results at the quarterly level, which allows for a finer examination of pre- and post-period dynamics.

[43] Ideally, we would allow for an arbitrary variance-covariance matrix within each of the four treatment groups, but we are concerned about small sample biases with such few clusters (Cameron, Gelbach, and Miller, 2011). Below we report alternative results aggregated to the treatment group level in which we estimate the model by Generalized Least Squares (GLS) and allow for both heterosketasticity as well as treatment-group specific auto-correlation parameters. These tend to produce similar point estimates and smaller standard errors relative to our baseline specification.

is due to emergency room visits). Column (1) shows our baseline results for 2003-2006 for total spending (i.e., Table 6, column (4)). It indicates that the change from the old health insurance options to the new health insurance options was associated with, on average, a $591 (11 percent) reduction in annual medical spending.

Columns (2) through (5) show estimates separately for spending on doctor visits, spending on outpatient visits, spending on inpatient visits and other spending. We detect a statistically significant decline in annual doctor spending of $220 (15 percent) and in annual outpatient spending of $310 (16 percent). The point estimates for inpatient spending suggest a statistically insignificant decline in inpatient spending of $117 (6 percent).

In addition to spending, we are able to measure utilization on the extensive margin. We define doctor visits as the total number of doctor visits by anyone in the household covered by the insurance (limited to a maximum of one per day). On average, an employee has 12 doctor visits for covered members in a given year. Outpatient visits are defined in an identical manner, where the average is 3 outpatient visits per year. We also code an indicator variable for whether there are any inpatient hospitalizations for anyone insured over the year; on average 14 percent of the employees have an inpatient hospitalization in a given year.

Columns (6) through (8) show the estimated effects on these measures of utilization. We estimate that the change in health insurance options is associated with a statistically and economically significant decline in the average number of annual doctor visits 1.9 (16 percent). Given the average cost of a doctor visit in our data of about $115, it is possible that the decline in spending on doctor visits comes entirely on the extensive margin. There is no evidence of an economically or statistically significant impact of the change in health insurance options on outpatient visits or inpatient hospitalization. The estimated decline in outpatient spending therefore presumably reflects a decrease in the intensity of treatment (i.e., spending conditional on the visit).

Validity of identifying assumption The identifying assumption in interpreting the difference-in-differences β coefficient from equation (17) as the causal impact of the change in health insurance options on the outcome of interest is that absent the change in health insurance options, employees in the different treatment groups would have otherwise experienced similar changes in their healthcare utilization or spending. Employees who are switched at different times differ in some of their demographics as well as in their 2003 (pre period) spending (see Table 1).Such observable differences across the treatment groups is not a problem per se for our difference-in-differences analysis which uses group fixed effects and therefore controls for any time-invariant differences across the treatment group. It naturally, however, raises concerns about the validity of our identifying assumption.

We undertake two types of analysis designed to help shed light on the likely validity of the identifying assumption. First, as our most direct investigations, we examine whether outcomes were trending similarly across the different groups in the periods prior to the change in health insurance options. These results are quite reassuring; there is no evidence of any substantively or statistically significant declines in spending in the several quarters prior to the change in health insurance options. Second, as a more indirect investigation, we also examine the sensitivity of our baseline results to controlling for observable characteristics of the employees. Again, it is quite reassuring that the basic OLS estimate in the 2003-2006 sample is not

particularly sensitive to controlling for observable worker characteristics.

Dynamics. To compare pre-period trends across the treatment groups we disaggregated the data from the annual to the quarterly level (so that t now denotes quarters rather than years) and estimate:

$$y_{ijt} = \alpha_j + \delta_t + \beta \cdot Treat_{jt} + \phi \cdot Treat_{jt,0} + \varepsilon_{ijt} \tag{18}$$

where $Treat_{jt,0}$ is an indicator variable for whether it is the quarter before group j is switched to the new health insurance options. The variable $Treat_{jt,0}$ acts as a pre-specification test; it will be informative of whether there are any differential trends in the outcome variables of interest across different treatment groups *before* the change in health insurance options. We estimate equation (18) at the quarterly rather than annual level primarily because at the annual level we would not be able to estimate pre period trends for the first treatment group (who is switched in 2004) which is roughly one-fifth of our sample, as there is only one year (2003) of pre data for this group. Another advantage of the quarterly specification is that it allows us to test for anticipation effects which presumably are most likely to occur immediately prior to the switch.[44]

Appendix Table A2 reports the results from estimating equation (18). In the interest of brevity, we report results for total spending only; results from components of spending (or utilization) are broadly similar (not shown). Column (1) reports the results from estimating equation (18) without the pre-period specification variable $Treat_{jt,0}$. It is therefore the exact analog of equation (17) but at the quarterly level rather than annual level. Correspondingly, therefore, the estimated coefficient on $Treat_{jt}$ is one-quarter the level of what we estimated in column (4) of Table 6. Column (2) of Table A2 shows the results when the pre-period variable $Treat_{jt,0}$ is included in the regression. The estimated main effect (the coefficient on $Treat_{jt}$) is virtually unaffected by the inclusion of this additional variable, although the standard error increases noticeably. More importantly, the coefficient on the pre-period specification test variable $Treat_{jt,0}$ is the opposite sign, statistically insignificant, and less than one-third the magnitude of the main effect. This goes some way toward assuaging concerns that the estimated effect is just picking up differential trends across groups.

A potential concern with quarterly level data is that results may be much more sensitive to outliers. To investigate this concern, in columns (3) and (4) we repeat the analysis in columns (1) and (2) but censor the dependent variable at the 99th percentile. Comparing columns (1) and (3), we see very similar point estimates on the estimated treatment effect (-148 in the uncensored estimate in column (1) and -157 in the censored estimate in column (3)) but a substantially lower standard error (65.76 vs. 43.62); this comparison is consistent with little or no economic incentive effect at the 99th percentile and therefore the introduction of noise from including the estimates above this point.[45] The pre-specification test on the censored data in

[44] In specifications at the quarterly level the δ_t represent a full set of quarter-of-year fixed effects rather than year fixed effects.

[45] The 99th percentile of the spending distribution is $57,500 for non-single coverage and $29,600 for single coverage. This level exceeds the out-of-pocket maximum on all plans with any non trivial mass except for the lowest coverage option (option 1) under the new plan options (see Table 2). Censoring the data at a spending level above the out of pocket maximum of the lowest coverage plan is conceptually valid since any spending above this amount cannot

column (4) shows a virtually identical main effect to the censored estimate in column (3), however now the pre period effect is not only statistically insignificant but substantively trivial (with a coefficient of -0.3.31 (standard error = 69) it is about two orders of magnitude smaller the main effect with a coefficient of -157). Finally, in column (5), as a further check on the validity of the identifying assumption, we re-estimate equation (18) with the addition of treatment-group specific linear trends; this allows each treatment group to be on a different (linear) trend over the 2003-2006 period and investigates whether the switch in health insurance options is associated with a change in spending for the treatment group relative to its average trend, relative to the changes in spending experienced at the same calendar time by other treatment groups relative to their own trends. The fact that the main estimate remains quite similar in magnitude is consistent with the evidence that these groups are not in fact on very different trends which are driving the estimated effect of the change in health insurance.

To more thoroughly examine the full range of pre-period dynamics, as well as to examine the dynamics in the timing of the post-period in any impact of the change in health insurance regime on the outcomes of interest, we also estimate a more flexible version of this quarterly specification that includes a full set of dummies for the number of quarters it has been since (or until) the switch. Specifically, we estimate

$$y_{ijt} = \alpha_j + \delta_t + \sum_{k=-11}^{12} \lambda_k Switch_{ijt,k} + \varepsilon_{ijt}, \tag{19}$$

where $Switch_{ijt,k}$ is an indicator variable for whether individual i is in a group j which at time t is k quarters away from the switch in health insurance options. The period $k = 1$ corresponds to the first quarter in which the group is under the new health insurance options, while $k = 0$ corresponds to the quarter right before the switch to the new health insurance options, etc. Thus, for example, for the "Switched in 2004" group, $Switch_{ijt,1}$is turned on (equal to 1) in the first quarter of 2004, while $Switch_{ijt,-3}$ is turned on the first quarter of 2003, and $Switch_{ijt,12}$ is turned on in the last quarter of 2006; for the "Switched later" group, all $Switch_{ijt,k}$ variables are set to 0. We examine periods from $k = -11$ (i.e., 12 quarters or 3 years before the switch) through $k = 12$ (i.e., 12 quarters or 3 years after the switch) although of course not all treatment groups can be used in identifying each of these periods (a point we return to below).

The coefficients of interest are the time pattern on the $\lambda_k's$, the coefficients on the $Switch_{ijt,k}$ indicators. Column (6) of Table A2 shows the coefficients on the λ_k's from estimating equation (19) on the outcome variable of total spending. We show (and focus our attention on) only the four quarters before and four quarters after the switch, since these are all identified off of the full sample; by contrast, coefficients further removed from $k = 0$ are identified off of only some of the groups; as a result, the time pattern at longer intervals potentially conflates the true time pattern with heterogeneous treatment effects across the groups identifying different coefficients.[46] We observe two interesting (and reassuring) features of the time pattern.

be affected by the cost-sharing features of the plan, except via income effects. To the extent that our censoring level is lower than the highest out of pocket maximum, censoring the dependent variable should bias downward our estimated effect of increased cost sharing. In practice, the results in Appendix Table A2 do not suggest any substantive downward bias.

[46]For example, employees in the "Switched in 2006" group do not contribute to the identification of the parameter

First, we can see that the decline in spending after the switch to the new regime happens pretty much instantaneously. This is reassuring as the timing of the effect suggests that we are estimating the effect of the change in plans, rather than some confounding factor. Second, there is no systematic trend in spending in the quarters before the switch for select relative to other groups with other timing; while the pattern is admittedly quite noisy it is relatively flat. This is re-assuring in further supporting the likely validity of the identifying assumption that absent this change in plans, the different groups would have been on similar trends in spending.

Sensitivity to covariates. An alternative way to shed light on the likely validity of the identifying assumption is to explore the sensitivity of the results to the inclusion of covariates. Appendix Table A3 explores these issues. This analysis is all done at an annual level. Column (1) replicates the baseline results from Table 6, column (4). Column (2) of Table A3 shows the results with the addition of controls for coverage tier. Column (3) adds controls for a wider set of employee demographic characteristics: in addition to whether they have single coverage, we control for their age, gender, risk score, the number of dependents insured on the policy, whether they are white, the number of years they have been at Alcoa, and their annual salary; this specification is shown to mimic the one we used in our baseline modeling approach below. The results in columns (1) through (3) indicate the results are not sensitive – in either magnitude or precision – to controlling for employee demographics; the baseline estimate of a $591 decline in spending associated with the move to the new PPO options changes to a $523 or $537 when the controls are added. As a stronger set of controls, we can include individual fixed effects for employees in the sample for more than one year. Column (4) shows the baseline results limited to the approximately half of employees who are in our data in all four years**.** The point estimate of the decline in spending associated with the move to the new PPO options is noticeably larger ($966) in this subsample, presumably reflecting heterogeneity in treatment effects and/or the treatment (i.e., plan selection) itself. More interestingly for our purposes, column (5) shows that the point estimate is unaffected ($966) by the inclusion of individual fixed effects in this subsample. Overall, we view the robustness of our results to various inclusions of covariates as reassuring with respect to the validity of the identifying assumption.

Additional sensitivity analyses Finally, Appendix Table A4 explores a variety of additional concerns and sensitivity analysis. One concern, noted earlier, is with sample selection. Specifically, we excluded from our analysis the 11 percent of employees who choose to opt out of insurance or choose the HMO option (available in all years and to all our employees) rather than one of the PPO options we study. To the extent that the new PPO options were more or less attractive to employees – in either their benefit design and/or their pricing – this raises concerns that our treatment variable (the offering of the new PPO options) could affect selection out of our sample and thus bias our estimates. To investigate this, we added back in the excluded individuals and re-estimated equation (17) for the binary dependent variable of whether the employee chose a non PPO option (i.e., is excluded from our baseline sample). The results indicate that the new options are associated with a statistically insignificant and economically small 2.1 percentage point decline in the

estimates beyond the third quarter under the new policy, while individuals in the "Switched in 2004" group do not contribute to the identification of the parameter estimates beyond the third quarter prior to the policy.

probability of an employee choosing a non PPO option. We suspect this reflects the fact that the excluded options are sufficiently horizontally differentiated from the PPO options that they are largely determined by other factors (outside insurance options, taste for HMO plan, etc.) and thus not that sensitive on the margin to redesigns of the PPO options; consistent with this, in Einav, Finkelstein, and Cullen (2010) we find that variation in the relative prices of the five new PPO options also does not have an economically or statistically significant association with the decision to choose one of these non PPO options. This is also consistent with Handel (2010)'s finding – in the context of a different employer provided health insurance setting – that individuals in a PPO are unlikely to subsequently choose an HMO when the set of HMO and PPO options change.

Another concern noted above was the treatment of the standard errors. Our baseline specification adjusts for an arbitrary variance-covariance matrix within each of the 28 unions (whose contracts determine which of the four treatment groups the employee is in). To investigate the sensitivity of our estimates to this approach, we follow the estimation approach pursued by Chandra, Gruber and McKnight (2010) in a similar context. Specifically, we aggregate our employee-level data to the treatment group level and estimate the treatment group by quarter data using Generalized Least Squares (GLS), with a treatment-group specific auto correlation parameter and variance**.** Column (3) of Table A4 reports the results of this estimation; for comparison purposes, column (2) reproduces the results of the quarterly OLS estimation of the employee-level regression, with clustering at the union level (see Table A2, column (1)). We are reassured that these two specifications yield not only similar point estimates (-$147.8 in column (2) and -$164.4 in column (3)) but also very similar standard errors; indeed, the standard errors are slightly smaller in the GLS specification than in our baseline OLS specification.

Appendix B: Sampling algorithm

Throughout, we will let Y denote the data. $\Theta = (\theta_1, \theta_2)$ is the set of parameters. We will write $\Theta_{-\beta}$ for all the parameters except β. We will use the following notation for the variance of the latent variables:

$$V\begin{pmatrix} \omega \\ \psi \\ \mu_{\lambda,i,2003} \\ \mu_{\lambda,i,2004} \end{pmatrix} = \Sigma = \begin{pmatrix} \sigma^2_\omega & \sigma_{\omega,\psi} & \sigma_{\mu,\omega} & \sigma_{\mu,\omega} \\ \sigma_{\omega,\psi} & \sigma^2_\psi & \sigma_{\mu,\psi} & \sigma_{\mu,\psi} \\ \sigma_{\mu,\omega} & \sigma_{\mu,\psi} & \sigma^2_\mu & \sigma_{\mu_{03},\mu_{04}} \\ \sigma_{\mu,\omega} & \sigma_{\mu,\psi} & \sigma_{\mu_{03},\mu_{04}} & \sigma^2_\mu \end{pmatrix}. \tag{20}$$

Suppose now that we have some initial draws of the parameters. We sample each parameter conditional on the others and the data as follows.

- Draw $\beta = (\beta_\omega, \beta_\psi, \beta_\lambda, \beta_\kappa)|\Theta_{-\beta}, \omega_i, \psi_i, \lambda_{it}, \mu_{it}, \sigma_i, Y$. Given $\omega_i, \psi_i, \lambda_{it}, \mu_{it}, \sigma_i, \kappa_i$, the vector β does not enter the density of the data. Spending depends only on (λ_{it}, ω_i) and plan choices depend only on $(\mu_{\lambda,it}, \sigma_i, \kappa_i, \omega_i, \psi_i)$. Therefore, the distribution of $\beta|\Theta_{-\beta}, \omega_i, \psi_i, \lambda_{it}, \mu_{\lambda,it}, , \sigma_i, \kappa_i, Y$ does not depend

on Y. Leaving out the prior for now, the posterior of β is:

$$\begin{aligned} f(\beta|\Theta_{-\beta},\omega_i,\psi_i,\lambda_{it},\mu_{\lambda,it},\sigma_i,\kappa_i) &\propto \prod_{i=1}^{N} f(\lambda_{it}|\mu_{\lambda,it},\sigma_i,\omega_i,\psi_i,\theta_{-\beta},\beta) f(\mu_{\lambda,it},\sigma_i,\omega_i,\psi_i|\theta_{-\beta},\beta) \\ &\propto \prod_{i=1}^{N} e^{-\frac{1}{2}\left(\frac{\log(\lambda_{it}-\kappa_i)-\mu_i}{\sigma_i}\right)^2} \exp\left[-\frac{1}{2}(u_i - x_i\beta)'^{-1}(u_i - x_i\beta)\right] f(\sigma_i|k,\theta) \\ &\propto \exp\left(-\frac{1}{2}(\beta-\hat{\beta})'\left(X'^{-1}\otimes I_N)X\right)(\beta-\hat{\beta})\right) \end{aligned} \tag{21}$$

where

$$u_i = (\log\omega_i, \log\psi_i, \mu_{i,2003}, \mu_{i,2004}, \kappa_i) \qquad \underbrace{U}_{5N\times 1} = \begin{pmatrix} \log\omega \\ \log\psi \\ \mu \end{pmatrix} \tag{22}$$

$$\underbrace{\beta}_{k_\omega+k_\psi+k_\lambda+k_\kappa\times 1} = [\beta_\omega, \beta_\psi, \beta_\lambda, \beta_\kappa] \qquad \underbrace{X}_{5N\times k_\omega+k_\psi+k_\lambda+k_\kappa} = \text{diag}\left(x^\omega, x^\psi, \begin{pmatrix} x^\lambda_{2003} \\ x^\lambda_{2004} \end{pmatrix}, x^\kappa\right)$$

and

$$\hat{\beta} = \left(X'^{-1}\otimes I_N)X\right)^{-1}\left(X'^{-1}\otimes I_N)U\right) \tag{23}$$

Hence, with a diffuse prior, the posterior of β is simply

$$N(\hat{\beta}, \left(X'^{-1}\otimes I_N)X\right)^{-1}) \tag{24}$$

With a $N(\beta_0, V_0)$ prior, the posterior of β would be

$$N(\bar{\beta}, \left(X'^{-1}\otimes I_N)X + V_0^{-1}\right)^{-1}) \tag{25}$$

with

$$\bar{\beta} = \left(X'^{-1}\otimes I_N)X + V_0^{-1}\right)^{-1}\left(X'^{-1}\otimes I_N)U\hat{\beta} + V_0^{-1}\beta_0\right) \tag{26}$$

- Draw $\Sigma|\Theta_{-\Sigma}, Y$. In order to impose the restrictions on Σ above (for example, that $cov(\mu_{\lambda,2003},\omega) = cov(\mu_{\lambda,2004},\omega)$ and $cov(\mu_{\lambda,2003},\psi) = cov(\mu_{\lambda,2004},\psi)$), we sample Σ in various pieces. To do this, it is useful to define α as the coefficient from regressing $\mu_{\lambda,it} - x^\lambda_{it}\beta_\lambda$ on $\log\omega - x^\omega\beta_\omega$ and $\log\psi - x^\psi\beta_\psi$. That is,

$$\alpha = \begin{pmatrix} \alpha_\omega \\ \alpha_\psi \end{pmatrix} = \begin{pmatrix} \sigma^2_\omega & \sigma_{\omega,\psi} \\ \sigma_{\omega,\psi} & \sigma^2_\psi \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{\omega,\mu} \\ \sigma_{\psi,\mu} \end{pmatrix} \tag{27}$$

Using this notation, we can write

$$\mu_{\lambda,it} - x^\lambda_{it}\beta_\lambda = \alpha_\omega(\log\omega_i - x^\omega_i\beta_\omega) + \alpha_\psi(\log\psi_i - x^\psi_i\beta_\psi) + \epsilon_{it} \tag{28}$$

Where ϵ_{it} is normally distributed and independent of $\log\omega - x^{\omega}\beta_{\omega}$ and $\log\psi - x^{\psi}\beta_{\psi}$. We parameterize the variance of $(\epsilon_{i,2003}, \epsilon_{i,2004})$ as

$$V\begin{pmatrix}\epsilon_{i,2003}\\ \epsilon_{i,2004}\end{pmatrix} = \frac{\sigma_\epsilon^2}{1-\rho^2}\begin{pmatrix}1 & \rho\\ \rho & 1\end{pmatrix} \tag{29}$$

That is, we think of ϵ as coming from an AR(1) process. Note that for $T = 2$, as in our baseline model, specifying that ϵ follows an AR(1) process carries no restriction – we could just as well simply say that ϵ has some variance matrix. However, our sampling algorithm and code are written for generic T, and for $T \geq 3$, the AR(1) assumption is a meaningful restriction.

– Draw $\Sigma_{\omega,\psi} = \begin{pmatrix}\sigma_\omega^2 & \sigma_{\omega,\psi}\\ \sigma_{\omega,\psi} & \sigma_\psi^2\end{pmatrix} | \Theta_{-\Sigma}, Y$. As above, the posterior of $\Sigma_{\omega,\psi}$ given the latent variables and the data does not depend on the data. Standard calculations show that if the prior for $\Sigma_{\omega,\psi}$ is $IW(A, m)$ then its posterior is $IW\left(n\hat{\Sigma}_{\omega,\psi} + A, n + m\right)$ where

$$\hat{\Sigma}_{\omega,\psi} = \frac{1}{n}\sum \begin{pmatrix}\log\omega_i - x_i^{\omega}\beta_{\omega}\\ \log\psi_i - x_i^{\psi}\beta_{\psi}\end{pmatrix}\begin{pmatrix}\log\omega_i - x_i^{\omega}\beta_{\omega}\\ \log\psi_i - x_i^{\psi}\beta_{\psi}\end{pmatrix}' \tag{30}$$

– Draw $\alpha | \Theta_{-\alpha}, \omega_i, \psi_i, \lambda_{it}, \mu_i, \sigma_i, Y$. As above, the posterior of α given the latent variables and the data does not depend on the data. Ignoring any prior for now, the posterior is

$$\begin{aligned} f(\alpha)|\Theta_{-\alpha}, \omega_i, \psi_i, \lambda_{it}, \mu_i, \sigma_i) \propto & \prod_{i=1}^{N}\prod_{t=1}^{T} f(\lambda_{it}|\mu_i, \sigma_i, \omega_i, \psi_i, \Theta_{-\alpha}, \alpha_{\omega,\psi}) \times \\ & \times f(\mu_i|\omega_i, \psi_i, \Theta_{-\alpha_{\omega,\psi}}, \alpha) f(\sigma_i|k, \theta) f(\omega_i, \psi_i|\Theta) \\ \propto & \prod_{i=1}^{N}\prod_{t=1}^{T} \exp\left(-\frac{1}{2}\left(\frac{\tilde{y}_{it} - \tilde{x}_i\alpha}{\alpha_\epsilon/\sqrt{1-\rho^2}}\right)^2\right) \end{aligned} \tag{31}$$

where $\tilde{y}_{it} = (\mu_{it} - x_{it}^{\lambda}\beta_{\lambda})$ and $\tilde{x}_i = \begin{pmatrix}\log\omega_i - x_i^{\omega}\beta_{\omega}\\ \log\psi_i - x_i^{\psi}\beta_{\psi}\end{pmatrix}$. The usual calculations would show that if the prior for α is $N(b_0, V_0)$, then the posterior is:

$$N\left(\left((1-\rho^2)\alpha_\epsilon^{-2}X'X + V_0\right)^{-1}\left((1-\rho^2)\alpha_\epsilon^{-2}X'Y + V_0 b_0\right), \left((1-\rho^2)\alpha_\epsilon^{-2}X'X + V_0\right)^{-1}\right) \tag{32}$$

where X is $(\tilde{x}_1, ..., \tilde{x}_N)'$ repeated twice, and Y is $(\tilde{y}_{1,2003}, .., \tilde{y}_{N,2003}, \tilde{y}_{1,2004}, ..., \tilde{y}_{N,2004})'$.

– Draw
$sigma_\epsilon^2 | \Theta_{-\sigma_\epsilon^2}, \omega_i, \psi_i, \lambda_{it}, \mu_i, \sigma_i, Y$. The same reasoning as for α shows that with a $\Gamma(a_1, a_2)$ prior, the posterior of σ_ϵ^{-2} is $\Gamma\left(N + a_1, 1/\left(\frac{1-\rho^2}{2}\sum_{it}(\tilde{y}_{it} - \tilde{x}_i\alpha_{\omega,\psi})^2 + 1/a_2\right)\right)$.

– Draw $\rho | \Theta_{-\rho}, \mu_{it}\omega_i, \psi_i, \lambda_{it}, \sigma_i, Y$. As above, the posterior of ρ given the latent variables and the data does not depend on the data. The distribution of ρ given the latent variables is proportional

to

$$f(\rho|\mu_{it}\omega_i,\psi_i,\lambda_{it},\sigma_i,\Theta_{-\rho}) \propto \prod_{i,t} f(\mu_{it}|\rho,\omega_i,\psi_i,\Theta_{-\rho}) \tag{33}$$
$$\propto \prod_{i=1}^{N} \sqrt{1-\rho^2}\exp(-\frac{1}{2}(1-\rho^2)\epsilon_{i1}^2)\prod_{t=2}^{T}\exp\left[-\frac{1}{2}(\epsilon_{it}-\rho\epsilon_{i,t-1})^2\right]$$
$$\propto (1-\rho^2)^{N/2}\exp\left[-\frac{1}{2}(\rho-\hat{\rho})'\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\epsilon_{i,t-1}^2\right)(\rho-\hat{\rho})\right]$$

where $\hat{\rho} = \frac{\sum_{i=1}^{N} \epsilon i1^2 + 2\sum_{t=2}^{T}\epsilon_{it}\epsilon_{it-1}}{2\sum_{i=1}^{N}\sum_{t=2}^{T}\epsilon_{it-1}^2}$, so ρ has the density of a normal truncated to $[-1,1]$ and scaled by $(1-\rho^2)^{N/2}$. [47] We sample from it using a metropolis sampler with candidate density,

$$N\left(\rho_{\text{current}}, N^{-1/2}\right) \tag{34}$$

This leads to an acceptance rate between 0.3 and 0.5 for a wide range of sample sizes.

- Draw $\lambda_{it},\omega_i|\Theta_{-\lambda,-\omega},Y$. This means drawing λ,ω from the region that rationalizes the observed choices and spending. The likelihood of the latent variables given spending m and choice j is:

$$f(\lambda_i,\omega_i|\Theta_{-\lambda,-\omega}) \propto \prod_{t=1}^{T} e^{\frac{-1}{2}(\frac{\log\lambda_{it}-\mu_{it}}{\sigma_i})^2} e^{-\frac{-1}{2}(\frac{\log\omega_i - m_i^o}{s^o})^2}\mathbf{1}(j^*(\omega,\psi,\mu,\kappa,\sigma)=j)\mathbf{1}(m^*(\lambda,\omega)=m) \tag{35}$$

where $m_i^o = x_i^\omega\beta_\omega + (\mu_i - x_i^\lambda\beta_\lambda)\delta_\mu + (\log\psi - x_i^\psi\beta_\psi)\delta_\psi$ and $s^o = \sqrt{\sigma_\omega^2 - S_{\omega,(\mu,\psi)}\Sigma_{\mu,\psi}^{-1}S_{(\mu,\psi),\omega}}$ with $\delta = \Sigma_{\mu,\psi}^{-1}S_{(\mu,\psi),\omega}$ and $S_{(\mu,\psi),\omega}$ the vector of covariances between ω and (μ,ψ) and $\Sigma_{\mu,\psi}$ the variance of (μ,ψ). We can do accept-reject sampling to sample from the region where $j^*(\omega,\psi,\mu,\sigma,\kappa) = J$. However, the area where $m^*(\lambda,\omega) = m$ has measure zero, so accept-reject sampling will not work. Instead, we have to more carefully characterize spending(λ,ω) to sample from the appropriate area. Let d be the chosen plan's deductible, x the maximum out of pocket sending, and c the copayment rate. A person chooses m to maximize utility:

$$\max_m (m-\lambda) - \frac{1}{2\omega}(m-\lambda)^2 \begin{cases} m & m < d \\ d + c(m-d) & m \geq d \,\&\, d + c(m-d) < x \\ x & d + c(m-d) \geq x \end{cases} \tag{36}$$

There are four possible solutions for m: 0, λ, $\lambda + (1-c)\omega$, and $\lambda + \omega$. We check whether each of these satisfy the constraints in (36) and compare the utilities of the ones that do.

We sample from the distribution of the latent variables subject to $m^*(\lambda,\omega) = m$ using a Metropolis-

[47] We tried to sample from this density using rejection sampling. We drew $\rho^* \sim TN(\hat{\rho}, v_\rho, -1, 1)$ and accepted with probability $(1-\rho^2)^{N/2}$, unfortunately this leads to unacceptably low acceptance rates.

Hastings sampler. The density of ω_i given m_{it} is

$$f(\omega_i|\{m_{it}\},m_i^o,s^o)proptoe^{-\frac{-1}{2}(\frac{\log\omega_i-m_i^o}{s^o})^2}\prod_{t=1}^{T}\begin{pmatrix}\mathbf{1}\{m=0\}P\left(m=0|\omega\right)\\+\mathbf{1}\{0<m<d\}P(m=m_{it}|\omega)\\+\mathbf{1}\{d<m<x\}\frac{1}{m_{it}-(1-c)\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-(1-c)\omega_i)-\mu_{\lambda,it}}{\sigma_i}\right)^2}\\+\mathbf{1}\{m>x\}+\mathbf{1}\{x<m\}\frac{1}{m_{it}-\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-\omega_i)-\mu_{\lambda,it}}{\sigma_i}\right)^2}\end{pmatrix} \tag{37}$$

We sample from this density by:

1. Sample

$$\omega\sim\tilde{f}(\omega|...)\propto e^{-\frac{-1}{2}(\frac{\log\omega_i-m_i^o}{s^o})^2}\prod_{t=1}^{T}\begin{bmatrix}\mathbf{1}\{d<m<x\}\frac{1}{m_{it}-(1-c)\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-(1-c)\omega_i)-\mu_{\lambda,it}}{\sigma_{\lambda,i}}\right)^2}\\+\mathbf{1}\{x<m\}\frac{1}{m_{it}-\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-\omega_i)-\mu_{\lambda,it}}{\sigma_{\lambda,i}}\right)^2}\end{bmatrix} \tag{38}$$

   We sample from this density using the Metropolis-Hastings algorithm with a normal candidate density for $\log\omega$. For each draw of ω_i, we run five metropolis iterations.

2. If $m_{it}=0$ for any t, draw $\log\lambda_{it}\sim N(\mu_{\lambda,it},\sigma_{\lambda,i})$.
3. If $0<m_{it}<d$, set $\lambda_{it}=m_{it}$
4. Accept ω_i if the observed m_{it} is the solution to (36) and $j_{it}=j^*(\omega_i,\psi_i,\mu_{\lambda,it},\sigma_{\lambda,i},\kappa_i)$ for all t, else repeat.

- For $t=2003,2004$, draw $\mu_{it}|\Theta_{-\mu},Y$. The posterior is a normal distribution truncated to the region where the choices implied by the model match the choices in the data. We repeatedly draw from this normal distribution until the choices match. The joint distribution of $\log\psi_i,\log\omega_i,\{\mu_{is}\},\log(\lambda_{it}-\kappa_i)$ is normal with mean $(x_i^\psi\beta_\psi,x_\omega^i\beta_\omega,\{x_{is}^\lambda\beta_\lambda\},x_{it}^\lambda\beta_\lambda)$ and variance

$$V_i=\begin{pmatrix}&\Sigma&&\begin{pmatrix}\sigma_{\omega,\mu}\\\sigma_{\psi,\mu}\\\sigma_\mu^2\\\sigma_\mu^2\end{pmatrix}\\\begin{pmatrix}\sigma_{\omega,\mu}&\sigma_{\psi,\mu}&\sigma_\mu^2&\sigma_\mu^2\end{pmatrix}&&\sigma_i^2+\sigma_\mu^2\end{pmatrix} \tag{39}$$

Note that we do not need to condition on $\log\lambda_{is}$ for $s\neq t$, because conditional on μ_{is}, μ_{it} and $\log\lambda_{is}$ are independent. Let $C_{\mu_t,(\omega,\psi,\mu_s,\lambda_t)}$ be the vector of covariances between μ_{it} and the other latent variables, $V_{-\mu_t;i}$ be V_i with the row and column for μ_{it} deleted, and

$$e_i=\begin{pmatrix}\log\omega_i\\\log\psi_i\\\mu_{is}\\\lambda_{it}\end{pmatrix}-\begin{pmatrix}x_\omega^i\beta_\omega\\x_i^\psi\beta_\psi\\x_\lambda^{is}\beta_\lambda\\x_\lambda^{it}\beta_\lambda\end{pmatrix} \tag{40}$$

The posterior mean of μ_{it} is then $e_i\delta_i$ with $\delta_i = C_{\mu_t,(\omega,\psi,\mu_s,\lambda)}V^{-1}_{-\mu_t;i}$, and the variance is $\sigma^2_\mu - C_{\mu_t,(\omega,\psi,\mu_s,\lambda)}V^{-1}_{-\mu_t;i}C'_{\mu_t,(\omega,\psi,\mu_s,\lambda)}$.

- Draw $\psi_i|\Theta_{-\psi}, Y$. As with μ_{it}, the posterior will be a normal distribution truncated to the region where the choices implied by the model match the choices in the data. We repeatedly draw from this normal distribution until the choices match. Define e_i as when sampling μ_{it}, but leave out λ_{it}. Also, let $C_{\psi,(\omega,\mu)}$ be the vector of covariances of ψ and (ω,μ) and $\Sigma_{-\psi}$ be Σ with the row and column for ψ removed. Then, the posterior distribution of ψ is

$$N\left(e_i\Sigma^{-1}_{-\psi}C_{\psi,(\omega,\mu)'}, \sigma^2_\psi - C_{\psi,(\omega,\mu)}\Sigma^{-1}_{-\psi}C'_{\psi,(\omega,\mu)}\right) \tag{41}$$

- Draw $\sigma_i|\Theta_{-\sigma_i}, Y$.

$$\begin{aligned} f(\sigma_i|\log\lambda_{it},\mu_i,\theta,k) \propto & \mathbf{1}\{\sigma^2<\bar{\sigma}^2\}\frac{1}{\sigma^{2(k-1)}}e^{-\sigma^{-2}/\theta}\prod_t\frac{1}{\sigma}e^{-\frac{1}{2}\left(\frac{\log\lambda_{it}-\mu}{\sigma_i}\right)^2} \\ \propto & \mathbf{1}\{\sigma^2<\bar{\sigma}^2\}\frac{1}{\sigma^{2(k-1+T/2)}}e^{-\sigma^{-2}(1/\theta+\frac{1}{2}\sum_t(\log\lambda_{it}-\mu_i)^2)} \end{aligned} \tag{42}$$

So the posterior of σ_i^{-2} is $\Gamma(k+T/2, \frac{2\theta}{2+\theta\sum_t(\log\lambda_{it}-\mu_i)^2})\mathbf{1}\{\sigma_i^2<\bar{\sigma}^2$, a truncated Gamma distribution.

- Draw $\gamma_1|\Theta_{-\gamma_1}, Y, ...$

$$\begin{aligned} f(\gamma_1|\sigma_i,k,Y,...) \propto & \prod f(\sigma_i|\gamma_1,k)p(\gamma_1) \\ \propto & \prod\sigma_i^{-2(k-1)}\frac{e^{-\sigma_i^{-2}/\gamma_1}}{\gamma_1^k\Gamma(k)}\frac{1}{1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1)}p(\gamma_1) \\ \propto & (1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1))^{-N}(1/\gamma_1)^{Nk}e^{-(1/\gamma_1)\sum\sigma_i^{-2}}p(\gamma_1) \\ \propto & (1/\gamma_1)^{Nk+k_0-1}e^{-(1/\gamma_1)\frac{\gamma_{1,0}\sum\sigma_i^{-2}+1}{\gamma_{1,0}}} \times (1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1))^{-N} \end{aligned} \tag{43}$$

where the prior for $1/\gamma_1$ is $\Gamma(k_0,\gamma_{1,0})$. This is a gamma distribution times some weighting function. Therefore, we use a metropolis sampler with candidate density for $1/\gamma_1$ a $\Gamma(Nk+k_0, \frac{\gamma_{1,0}}{\gamma_{1,0}\sum\sigma_i^{-2}+1})$. Given the current estimates, $1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1)$ is very close to one, so this metropolis sampler accepts nearly all draws.

- Draw $\gamma_2|\Theta_{-\gamma_2}, Y$.

$$\begin{aligned} f(k|\sigma_i,\theta,...) \propto & \prod\sigma_i^{-2(k-1)}\frac{e^{-\sigma_i^{-2}/\theta}}{\theta^k\Gamma(k)}p(k)(1-F_\Gamma(\bar{\sigma}^{-2};k,\theta))^{-N} \\ \propto & \frac{e^{k\sum\log\sigma_i^{-2}+\log\theta^{-N}}}{\Gamma(k)^N}p(k)(1-F_\Gamma(\bar{\sigma}^{-2};k,\theta))^{-N} \end{aligned} \tag{44}$$

which is a nonstandard distribution. We use the adaptive rejection metropolis sampling (ARMS) method of Gilks, Best, and Tan (1995) to sample from it. This is a hybrid accept-reject and metropolis sampling scheme. It is designed to sample from log-concave and nearly log-concave densities efficiently. Without the $(1-F_\Gamma(\bar{\sigma}^{-2};k,\theta))^{-N}$ term, this density would be log-concave (it may be log-concave anyway), and ARMS can sample from it very efficiently.

Appendix Table A1: Impact of change in health insurance options on components of health spending and utilization

	Spending					Utilization		
	Total Spending	Spending on Doctor Visits	Spending on Outpatient Visits	Spending on Inpatient Visits	Remaining Spending	Number of Doctor Visits	Number of Outpatient Visits	Any Inpatient Visits
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimated treatment effect	-591.81	-220.37	-310.32	-116.69	55.91	-1.94	-0.0005	-0.017
	(264.26)	(69.32)	(137.89)	(246.17)	(69.34)	(0.37)	(0.27)	(0.011)
	[0.034]	[0.004]	[0.033]	[0.639]	[0.427]	[0.000]	[0.999]	[0.155]
Mean Dep. Var.	5392	1475	1922	1804	191	12.2	3	0.14

The table shows the difference-in-difference estimate of the impact of the move from the old to the new options on various components of health care spending and utilization. All columns show the coefficient on $TREAT$ from estimating equation 17 by OLS for the dependent variable given in the column heading. Unit of observation is an employee-year. All regressions include year and treatment group fixed effects. We classify employees into one of four possible treatment groups - switched in 2004, switched in 2005, switched in 2006, or switched later - based on his union affiliation which determines the year in which he is switched to the new health insurance options. Standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions; p-values are in [square brackets].Sample is 2003-2006. N = 14,638.

Appendix Table A2: Impact of change in health insurance options on spending (quarterly data)

	Total Spending		Total Spending, Censored at 99th percentile			
	Baseline	Pre-specification test	Baseline	Pre-specification test	Col (4) w treatment-group-specific linear trend	More dynamics
	(1)	(2)	(3)	(4)	(5)	(6)
$TREAT_{jt}$	-147.87	-139.44	-156.85	-157.54	-185.65	
	(66.04)	(85.22)	(43.60)	(50.49)	(74.82)	
	[0.034]	[0.113]	[0.001]	[0.004]	[0.020]	
$TREAT_{jt,0}$		40.78		-3.31	-5.69	
		(158.49)		(69.21)	(76.00)	
		[0.799]		[0.962]	[0.941]	
$TREAT_{jt,-3}$						58.59
						(60.91)
$TREAT_{jt,-2}$						-2.46
						(90.69)
$TREAT_{jt,-1}$						-42.03
						(69.75)
$TREAT_{jt,0}$						0 (reference period)
$TREAT_{jt,1}$						-121.79
						(53.47)
$TREAT_{jt,2}$						-187.06
						(77.21)
$TREAT_{jt,3}$						-118.35
						(65.90)
$TREAT_{jt,4}$						-197.82
						(61.78)
Mean dep. Var.	1348			1125		

The table shows the difference-in-difference estimate of the impact of the move from the old to the new options. Specifically, columns 1 through 5 show the results from estimating equation 18 (and column 6 shows results from estimating equation 19) by OLS for the dependent variable total quarterly health spending. Unit of observation is an employee-quarter. The variable $TREAT_{jt}$ is an indicator variable for whether treatment group j is offered the new health insurance options in quarter t. The variable $Treat_{jt,0}$ is an indicator variable for whether it is the quarter before group j is switched to the new health insurance options. The variable $TREAT_{jt,k}$ is an indicator variable for whether it is k quarters since quarter 0 (i.e. the quarter before the switch). All regressions include quarter and treatment group fixed effects; column 5 also includes a treatment group-specific linear trend. We classify employees into one of four possible treatment groups - switched in 2004, switched in 2005, switched in 2006, or switched later - based on his union affiliation which determines the year in which he is switched to the new health insurance options. Standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions; p-values are in [square brackets].Sample is 2003-2006. N = 58,552.

Appendix Table A3: Sensitivity of annual difference-in-differences estimates to controlling for observables

	Baseline (no covariates)	Adding control for coverage tier	Adding additional demographic controls	At Alcoa all four years	At Alcoa all four years, w individual fixed effects.
	(3)	(2)	(3)	(4)	(5)
$TREAT_{jt}$	-591.81	-522.74	-537.96	-965.92	-965.92
	(264.26)	(267.29)	(264.33)	(302.33)	(349.04)
	[0.034]	[0.061]	[0.052]	[0.004]	[0.012]
Mean Dep. Var.		5392		5438	
N		14,638		7,580	

The table examines the sensitivity of the annual difference-in-differences estimates of the impact of the move from the old to the new options on total annual medical spending. All columns show the coefficient on $TREAT$ from estimating equation 17 by OLS for the dependent variable total annual medical spending. Unit of observation is an employee-year. All regressions include quarter and treatment group fixed effects. We classify employees into one of four possible treatment groups - switched in 2004, switched in 2005, switched in 2006, or switched later - based on his union affiliation which determines the year in which he is switched to the new health insurance options. Standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions; p-values are in [square brackets].Sample is 2003-2006. Column 1 replicates the baseline results (from Table 6, column 4). In column 2 we control for coverage tier. In column 3 we control for coverage tier, employee age, risk score, employee gender, number of dependents insured on the policy, whether the employee is white, the number of years the employee has been at Alcoa, and the employee's annual salary. Column 4 limits the sample to employees who are at Alcoa (and in our data) for all four years. Column 5 adds employee fixed effects to the sample in column 4.

Appendix Table A4: Additional sensitivity analysis

	Dependent variable: choose a non-PPO option	Dependent variable: total spending	
		Baseline quarterly specification (OLS)	GLS estimation at Treatment group - quartelry level
	(1)	(2)	(3)
$TREAT_{jt}$	-0.021	-147.87	-166.43
	(0.024)	(66.04)	(61.22)
	[0.376]	[0.034]	[0.007]
Mean dep var	0.106	1348	1364
N	16366	58,552	64

The table examines some additional sensitivity of the annual difference-in-differences estimates of the impact of the move from the old to the new options on total annual medical spending. All regressions include year and treatment group fixed effects. Column 1 shows the coefficient on $TREAT$ from estimating equation 17 by OLS on the baseline 2003-2006 sample, plus the employees who choose a non-PPO option; the dependent variable is an indicator variable for whether the employee chose a non PPO option; unit of observation is an employee-year. In columns 2 and 3 the dependent variable is total spending. Column 2 shows the coefficient on $TREAT$ from estimating equation 18 by OLS at the employee-quarter level. Column 3 shows the coefficient on $TREAT$ from estimating equation 18 by GLS with a panel-specific auto correlation parameter and variance at the treatment group - quarter level. In columns 1 and 2 standard errors (in parentheses) are adjusted for an arbitrary variance-covariance matrix within each of the 28 unions; p-values are in [square brackets].