Online Appendix

"Selection on Moral Hazard in Health Insurance"
by Einav, Finkelstein, Ryan, Schrimpf, and Cullen

Appendix A: Construction of the baseline sample.

Alcoa has about 45,000 active employees per year. We start by excluding about 15% of the sample whose data are not suited to our analytical framework. 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.

Given the source of variation used to identify moral hazard, we concentrate on the approximately one third of Alcoa workers who are unionized; approximately 70% of Alcoa workers are hourly employees, and approximately half of these are unionized (salaried workers are not unionized). 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% 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.[1]

Appendix B: 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:

$$y_{ijt} = \alpha_j + \delta_t + \beta \cdot Treat_{jt} + \mathbf{x}'_{ijt}\phi + \varepsilon_{ijt}, \quad (1)$$

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

[1] 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 B below, 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.

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.[2] 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.[3]

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% of the total), outpatient spending (approximately 35% of the total), inpatient spending (approximately 35% of the total), and other (which accounts for about 4% of spending, about half of which 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%) 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%) and in annual outpatient spending of $310 (16%). The point estimates for inpatient spending suggest a statistically insignificant decline in inpatient spending of $117 (6%).

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% 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

[2] 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.

[3] 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.

in the average number of annual doctor visits 1.9 (16%). 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 (1) 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:

$$y_{ijt} = \alpha_j + \delta_t + \beta \cdot Treat_{jt} + \phi \cdot Treat_{jt,0} + \varepsilon_{ijt} \tag{2}$$

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 (2) 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.[4]

Appendix Table A2 reports the results from estimating equation (2). In the interest of brevity, we report results for total spending only; results from components of spending (or utilization) are broadly similar (not

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

shown). Column (1) reports the results from estimating equation (2) without the pre-period specification variable $Treat_{jt,0}$. It is therefore the exact analog of equation (1) 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.[5] The pre-specification test on the censored data in 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 (2) 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

[5] 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 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.

$$y_{ijt} = \alpha_j + \delta_t + \sum_{k=-11}^{12} \lambda_k Switch_{ijt,k} + \varepsilon_{ijt}, \quad (3)$$

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 (3) 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.[6] We observe two interesting (and reassuring) features of the time pattern. 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

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

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% 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 (1) 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 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 (2011)'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 C: Suggestive evidence of heterogeneity in and selection on moral hazard.

Heterogeneity in moral hazard. We begin by 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 developed in Section I), 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.[7] 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.

Appendix Table A5 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 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 Appendix Table A5 – 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

[7] Kowalski (2010) finds similar patterns in her quantile treatment estimates using a different identification strategy in a different firm. We should also point out that the frequency of reaching the out-of-pocket maximum is less than 5% even under the most generous plan in the data, so a "zero" marginal price is unlikely to affect spending at the 90th percentiles (the highest quantile presented in the table).

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 expected 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 A5 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.[8] 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

[8] 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% 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)).

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.

Appendix D: 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:

$$V \begin{pmatrix} \omega \\ \psi \\ \mu_{\lambda,i,2003} \\ \mu_{\lambda,i,2004} \end{pmatrix} = \Sigma = \begin{pmatrix} \sigma_\omega^2 & \sigma_{\omega,\psi} & \sigma_{\mu,\omega} & \sigma_{\mu,\omega} \\ \sigma_{\omega,\psi} & \sigma_\psi^2 & \sigma_{\mu,\psi} & \sigma_{\mu,\psi} \\ \sigma_{\mu,\omega} & \sigma_{\mu,\psi} & \sigma_\mu^2 & \sigma_{\mu_{03},\mu_{04}} \\ \sigma_{\mu,\omega} & \sigma_{\mu,\psi} & \sigma_{\mu_{03},\mu_{04}} & \sigma_\mu^2 \end{pmatrix}. \tag{4}$$

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:

$$\begin{aligned} f(\beta|\Theta_{-\beta}, \omega_i, \psi_i, \lambda_{it}, \mu_{\lambda,it}, \sigma_i, \kappa_i) &\propto \prod_{i=1}^N f(\lambda_{it}|\mu_{\lambda,it}, \sigma_i, \omega_i, \psi_i, \theta_{-\beta}, \beta) f(\mu_{\lambda,it}, \sigma_i, \omega_i, \psi_i|\theta_{-\beta}, \beta) \\ &\propto \prod_{i=1}^N e^{-\frac{1}{2}\left(\frac{\log(\lambda_{it}-\kappa_i)-\mu_i}{\sigma_i}\right)^2} \exp\left[-\frac{1}{2}(u_i - x_i\beta)'^{-1}(u_i - x_i\beta)\right] f(\sigma_i|k,\theta) \\ &\propto \exp\left(-\frac{1}{2}(\beta-\hat{\beta})'\left(X'^{-1} \otimes I_N)X\right)(\beta-\hat{\beta})\right) \end{aligned} \tag{5}$$

where

$$u_i = (\log\omega_i, \log\psi_i, \mu_{i,2003}, \mu_{i,2004}, \kappa_i) \qquad \underbrace{U}_{5N\times 1} = \begin{pmatrix} \log\omega \\ \log\psi \\ \mu \end{pmatrix} \tag{6}$$

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

and

$$\hat{\beta} = \left(X'^{-1} \otimes I_N)X\right)^{-1}\left(X'^{-1} \otimes I_N)U\right) \tag{7}$$

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

$$N(\hat{\beta}, \left(X'^{-1} \otimes I_N)X\right)^{-1}) \tag{8}$$

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

$$N(\bar{\beta}, \left(X'^{-1} \otimes I_N)X + V_0^{-1}\right)^{-1}) \tag{9}$$

with

$$\bar{\beta} = \left(X'^{-1} \otimes I_N)X + V_0^{-1}\right)^{-1}\left(X'^{-1} \otimes I_N)U\hat{\beta} + V_0^{-1}\beta_0\right) \tag{10}$$

- 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 $\mu_{\lambda,it} - x_{it}^{\lambda}\beta_{\lambda}$ on $\log\omega - x^{\omega}\beta_{\omega}$ and $\log\psi - x^{\psi}\beta_{\psi}$. That is,

$$\alpha = \begin{pmatrix} \alpha_\omega \\ \alpha_\psi \end{pmatrix} = \begin{pmatrix} \sigma_\omega^2 & \sigma_{\omega,\psi} \\ \sigma_{\omega,\psi} & \sigma_\psi^2 \end{pmatrix}^{-1} \begin{pmatrix} \sigma_{\omega,\mu} \\ \sigma_{\psi,\mu} \end{pmatrix} \tag{11}$$

Using this notation, we can write

$$\mu_{\lambda,it} - x_{it}^{\lambda}\beta_{\lambda} = \alpha_\omega(\log\omega_i - x_i^{\omega}\beta_\omega) + \alpha_\psi(\log\psi_i - x_i^{\psi}\beta_\psi) + \epsilon_{it} \tag{12}$$

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

$$V\begin{pmatrix} \epsilon_{i,2003} \\ \epsilon_{i,2004} \end{pmatrix} = \frac{\sigma_\epsilon^2}{1-\rho^2}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \tag{13}$$

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

$$\hat{\Sigma}_{\omega,\psi} = \frac{1}{n}\sum \begin{pmatrix} \log\omega_i - x_i^{\omega}\beta_\omega \\ \log\psi_i - x_i^{\psi}\beta_\psi \end{pmatrix}\begin{pmatrix} \log\omega_i - x_i^{\omega}\beta_\omega \\ \log\psi_i - x_i^{\psi}\beta_\psi \end{pmatrix}' \tag{14}$$

  - 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

$$\begin{aligned} f(\alpha)|\Theta_{-\alpha}, \omega_i, \psi_i, \lambda_{it}, \mu_i, \sigma_i) \propto & \prod_{i=1}^{N}\prod_{t=1}^{T} f(\lambda_{it}|\mu_i, \sigma_i, \omega_i, \psi_i, \Theta_{-\alpha}, \alpha_{\omega,\psi})\times \\ & \times f(\mu_i|\omega_i, \psi_i, \Theta_{-\alpha_{\omega,\psi}}, \alpha) f(\sigma_i|k,\theta) f(\omega_i, \psi_i|\Theta) \\ \propto & \prod_{i=1}^{N}\prod_{t=1}^{T} \exp\left(-\frac{1}{2}\left(\frac{\tilde{y}_{it} - \tilde{x}_i\alpha}{\alpha_\epsilon/\sqrt{1-\rho^2}}\right)^2\right) \end{aligned} \tag{15}$$

  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:

$$N\left(\left((1-\rho^2)\alpha_\epsilon^{-2}X'X + V_0\right)^{-1}\left((1-\rho^2)\alpha_\epsilon^{-2}X'Y + V_0 b_0\right), \left((1-\rho^2)\alpha_\epsilon^{-2}X'X + V_0\right)^{-1}\right) \tag{16}$$

  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

$$
\begin{aligned}
f(\rho|\mu_{it}\omega_i, \psi_i, \lambda_{it}, \sigma_i, \Theta_{-\rho}) \propto & \prod_{i,t} f(\mu_{it}|\rho, \omega_i, \psi_i, \Theta_{-\rho}) \\
\propto & \prod_{i=1}^{N} \sqrt{1-\rho^2}\exp(-\frac{1}{2}(1-\rho^2)\epsilon_{i1}^2)\prod_{t=2}^{T}\exp\left[-\frac{1}{2}(\epsilon_{it} - \rho\epsilon_{i,t-1})^2\right] \\
\propto & (1-\rho^2)^{N/2}\exp\left[-\frac{1}{2}(\rho-\hat{\rho})'\left(\sum_{i=1}^{N}\sum_{t=2}^{T}\epsilon_{i,t-1}^2\right)(\rho-\hat{\rho})\right]
\end{aligned} \tag{17}
$$

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 $(1-\rho^2)^{N/2}$. [9] We sample from it using a metropolis sampler with candidate density,

$$
N\left(\rho_{\text{current}}, N^{-1/2}\right) \tag{18}
$$

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:

$$
f(\lambda_i, \omega_i|\Theta_{-\lambda,-\omega}) \propto \prod_{t=1}^{T} e^{\frac{-1}{2}(\frac{\log\lambda_{it}-\mu_{it}}{\sigma_i})^2} e^{-\frac{-1}{2}(\frac{\log\omega_i - m_i^o}{s^o})^2}\mathbf{1}(j^*(\omega,\psi,\mu,\kappa,\sigma) = j)\mathbf{1}(m^*(\lambda,\omega) = m) \tag{19}
$$

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:

$$
\max_m(m-\lambda) - \frac{1}{2\omega}(m-\lambda)^2 \begin{cases} m & m < d \\ d + c(m-d) & m \geq d \,\&\, d + c(m-d) < x \\ x & d + c(m-d) \geq x \end{cases} \tag{20}
$$

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 (20) 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-

[9] 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

$$f(\omega_i|\{m_{it}\}, m_i^o, s^o) proptoe^{-\frac{-1}{2}(\frac{\log\omega_i - m_i^o}{s^o})^2} \prod_{t=1}^{T} \begin{pmatrix} \mathbf{1}\{m=0\}P\,(m=0|\omega) \\ +\mathbf{1}\{0<m<d\}P(m=m_{it}|\omega) \\ +\mathbf{1}\{d<m<x\}\frac{1}{m_{it}-(1-c)\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-(1-c)\omega_i)-\mu_{\lambda,it}}{\sigma_i}\right)^2} \\ +\mathbf{1}\{m>x\}+\mathbf{1}\{x<m\}\frac{1}{m_{it}-\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-\omega_i)-\mu_{\lambda,it}}{\sigma_i}\right)^2} \end{pmatrix} \tag{21}$$

We sample from this density by:

1. Sample

$$\omega \sim \tilde{f}(\omega|...) \propto e^{-\frac{-1}{2}(\frac{\log\omega_i - m_i^o}{s^o})^2} \prod_{t=1}^{T} \begin{bmatrix} \mathbf{1}\{d<m<x\}\frac{1}{m_{it}-(1-c)\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-(1-c)\omega_i)-\mu_{\lambda,it}}{\sigma_{\lambda,i}}\right)^2} \\ +\mathbf{1}\{x<m\}\frac{1}{m_{it}-\omega}e^{\frac{-1}{2}\left(\frac{\log(m_{it}-\omega_i)-\mu_{\lambda,it}}{\sigma_{\lambda,i}}\right)^2} \end{bmatrix} \tag{22}$$

   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 (20) 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

$$V_i = \begin{pmatrix} \Sigma & \begin{pmatrix} \sigma_{\omega,\mu} \\ \sigma_{\psi,\mu} \\ \sigma_{\mu}^2 \\ \sigma_{\mu}^2 \end{pmatrix} \\ \begin{pmatrix} \sigma_{\omega,\mu} & \sigma_{\psi,\mu} & \sigma_{\mu}^2 & \sigma_{\mu}^2 \end{pmatrix} & \sigma_i^2 + \sigma_{\mu}^2 \end{pmatrix} \tag{23}$$

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

$$e_i = \begin{pmatrix} \log\omega_i \\ \log\psi_i \\ \mu_{is} \\ \lambda_{it} \end{pmatrix} - \begin{pmatrix} x_{\omega}^{i}\beta_{\omega} \\ x_i^{\psi}\beta_{\psi} \\ x_{\lambda}^{is}\beta_{\lambda} \\ x_{\lambda}^{it}\beta_{\lambda} \end{pmatrix} \tag{24}$$

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

- 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

$$N\left(e_i\Sigma_{-\psi}^{-1}C_{\psi,(\omega,\mu)'},\sigma_\psi^2-C_{\psi,(\omega,\mu)}\Sigma_{-\psi}^{-1}C'_{\psi,(\omega,\mu)}\right) \tag{25}$$

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

$$\begin{aligned} f(\sigma_i|\log\lambda_{it},\mu_i,\theta,k) &\propto \mathbf{1}\{\sigma^2<\bar{\sigma}^2\}\frac{1}{\sigma^{2(k-1)}}e^{-\sigma^{-2}/\theta}\prod_t\frac{1}{\sigma}e^{-\frac{1}{2}\left(\frac{\log\lambda_{it}-\mu}{\sigma_i}\right)^2} \\ &\propto \mathbf{1}\{\sigma^2<\bar{\sigma}^2\}\frac{1}{\sigma^{2(k-1+T/2)}}e^{-\sigma^{-2}(1/\theta+\frac{1}{2}\sum_t(\log\lambda_{it}-\mu_i)^2)} \end{aligned} \tag{26}$$

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

$$\begin{aligned} f(\gamma_1|\sigma_i,k,Y,...) &\propto \prod f(\sigma_i|\gamma_1,k)p(\gamma_1) \\ &\propto \prod \sigma_i^{-2(k-1)}\frac{e^{-\sigma_i^{-2}/\gamma_1}}{\gamma_1^k\Gamma(k)}\frac{1}{1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1)}p(\gamma_1) \\ &\propto (1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1))^{-N}(1/\gamma_1)^{Nk}e^{-(1/\gamma_1)\sum\sigma_i^{-2}}p(\gamma_1) \\ &\propto (1/\gamma_1)^{Nk+k_0-1}e^{-(1/\gamma_1)\frac{\gamma_{1,0}\sum\sigma_i^{-2}+1}{\gamma_{1,0}}}\times(1-F_\Gamma(\bar{\sigma}^{-2};k,\gamma_1))^{-N} \end{aligned} \tag{27}$$

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

$$\begin{aligned} f(k|\sigma_i,\theta,...) &\propto \prod\sigma_i^{-2(k-1)}\frac{e^{-\sigma_i^{-2}/\theta}}{\theta^k\Gamma(k)}p(k)(1-F_\Gamma(\bar{\sigma}^{-2};k,\theta))^{-N} \\ &\propto \frac{e^{k\sum\log\sigma_i^{-2}+\log\theta^{-N}}}{\Gamma(k)^N}p(k)(1-F_\Gamma(\bar{\sigma}^{-2};k,\theta))^{-N} \end{aligned} \tag{28}$$

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 E: Heterogeneity in moral hazard in a "multiplicative model."

To explore whether our findings of substantial heterogeneity in moral hazard are simply an artifact of the additive way that moral hazard affect utilization in the model of Section I, we estimated a slightly modified model, in which moral hazard enters multiplicatively. Specifically, we use the same model and econometric specification, except that we replace equation (1) in the main text with the following expression:

$$u(m;\lambda,\omega,j) = \left[(m-\lambda) - \frac{1}{2\omega\lambda}(m-\lambda)^2\right] + \left[y - c_j(m) - p_j\right]. \tag{29}$$

That is, we keep the utility function specification the same, except that we add λ to the denominator of the second component. One can check that this small modification implies, in the context of a linear contract, that optimal utilization is given by

$$m^*(\lambda,\omega,c) = \max\left[0, \lambda(1+\omega(1-c))\right]. \tag{30}$$

That is, ω now affects the optimal spending multiplicatively, rather than additively as in equation (3) in the main text. Note that in this alternative model, "moral hazard" – i.e. the difference in spending between no insurance ($c = 1$) and full insurance ($c = 0$) – is now $\lambda\omega$ rather than ω as in the original model; as a result, when choosing insurance one's moral hazard type is uncertain. The rest of the model specification remains the same.

Appendix Tables A6 and A7 report the results from the estimation of this multiplicative model. The tables correspond to Table 7(a) and Table 9 in the main text. As one can observe, the qualitative features of the results remain similar. For example, the heterogeneity in ω is still substantial, with a coefficient of variation of about 2.5 (bottom of Appendix Table A6), and the qualitative pattern reported in Appendix Table A7 is quite similar, although slightly smaller, to the pattern shown in Table 9 of the main text.

Appendix F: Robustness checks of the main, model-based findings.

Appendix Table A8 briefly explores the robustness of some of our main findings to alternative econometric specifications of the baseline model. 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 II 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% 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.[10] Row 3 shows the results of excluding the 2004 observations for the approximately 12% 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 (8) of the main text) 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.

References

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.

Einav, Liran, Amy Finkelstein, Iuliana Pascu, and Mark R. Cullen. 2012. "How General are Risk Preferences? Choices Under Uncertainty in Different Domains." *American Economic Review* 102(6): 2606-2638.

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

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

[10] 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 (2012) provide more detail and discussion of these other benefits options and choices.

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 (1) 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 2 (and column 6 shows results from estimating equation 3) 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 1 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 1 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 2 by OLS at the employee-quarter level. Column 3 shows the coefficient on $TREAT$ from estimating equation 2 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].

Appendix Table A5: 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) Coeff.	Std. Err.	Avg Out-of-Pocket Share (Old Options)	Avg Out-of-Pocket Share (New Options)	Increase in Out-of-Pocket Share
		(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.

Appendix Table A6: Parameter estimates from a multiplicative model

Mean Shifters				
	μ_λ (Health risk)	κ_λ (Health risk)	$ln(\omega)$ (Moral hazard)	$ln(\psi)$ (Risk aversion)
Constant	6.47 (0.12)	355 (62.8)	-2.06 (0.22)	-1.76 (0.69)
Coverage tier				
Single	(omitted)	(omitted)	(omitted)	(omitted)
Family	-0.396 (0.07)	-47.1 (41.1)	0.56 (0.14)	-0.66 (0.34)
Emp+Spouse	-0.262 (0.07)	-24 (35.9)	0.71 (0.14)	-0.42 (0.31)
Emp+Children	-0.022 (0.08)	-157 (31.2)	0.46 (0.12)	-0.65 (0.32)
Treatment group				
Switch 2004	0.138 (0.07)	370 (38.8)	-0.05 (0.13)	-1.39 (0.36)
Switch 2005	-0.054 (0.07)	53.8 (30.3)	0.70 (0.17)	1.70 (0.45)
Switch 2006	0.116 (0.05)	70.8 (25.7)	0.41 (0.14)	1.53 (0.61)
Switch later	(omitted)	(omitted)	(omitted)	(omitted)
Demographics				
Age	-0.016 (0.002)	4.78 (1.30)	-0.0015 (0.005)	-0.006 (0.011)
Female	0.043 (0.060)	-104 (31.4)	-0.39 (0.12)	-1.18 (0.38)
Job Tenure	0.003 (0.002)	0.73 (1.35)	-0.025 (0.005)	-0.021 (0.01)
Income	0.0004 (0.0014)	-7.99 (0.89)	-0.002 (0.003)	0.001 (0.007)
Health risk score				
1st quartile (< 1.119)	(omitted)	(omitted)	(omitted)	(omitted)
2nd quartile (1.119 to 1.863)	1.17 (0.068)	-254 (36.2)	-0.43 (0.13)	-0.40 (0.32)
3rd quartile (1.863 to 2.834)	2.05 (0.073)	-424 (42.5)	-0.46 (0.14)	-0.57 (0.36)
4th quartile (> 2.834)	3.11 (0.079)	-409 (56.9)	-0.51 (0.14)	-1.09 (0.39)
2004 Time dummy	-0.10 (0.028)	--	--	--
Variance-covariance matrix				
	μ_λ_bar	$ln(\omega)$	$ln(\psi)$	
μ_λ_bar	0.66 (0.046)	-0.28 (0.038)	-0.49 (0.083)	
$ln(\omega)$	--	1.76 (0.115)	1.67 (0.208)	
$ln(\psi)$	--	--	3.47 (0.805)	
Additional parameters				
σ_μ	0.11 (0.052)			
σ_κ	274 (13.0)			
γ_1	0.204 (0.06)			
γ_2	7.17 (1.70)			
Unconditional statistics				
	$E(\lambda)$	ω	ψ	
Average	4,570	0.313	0.796	
Std. Deviation	12,500	0.765	8.350	

The table reports results from the estimation of the multiplicative model described in Appendix E. The table parallels Table 7(a) of the paper, and the bottom panel parallels the top panel of Table 7(b) of the paper.

Appendix Table A7: Estimates of moral hazard heterogeneity from a multiplicative model

	Mean	Std. Dev.	10th	25th	50th	75th	90th
Spending difference as we move from no to high deductible plan	112	289	0	0	4	88	315
Spending difference as we move from full to no insurance	961	4,789	0	25	138	564	1,802

The table reports results from the estimation of the multiplicative model described in Appendix E. The table parallels Table 9 of the paper.

Appendix Table A8: Robustness

	Average moral hazard effect	Moral hazard: coefficient of variation	90-10 difference in probability choose high deductible plan: Quantiles of risk	90-10 difference in probability choose high deductible plan: Quantiles of moral hazard	Average moral hazard effect for "selected" group	Welfare effect of "Perfect Screening"	Welfare effect of "Imperfect Screening"
	(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 described in Appendix F. 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 9, 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% 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% 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 10, 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 (8) in the main text) to depend on the covariates. Row 6 allows the joint distribution of the latent variables (in equation (8) in the main text) to be more flexible by allowing it to follow a mixture of two normal distributions.