Online Appendix
Health Insurance for "Humans": Information Frictions, Plan Choices, and Consumer Welfare

Benjamin R. Handel *
Jonathan T. Kolstad †

March 6, 2015

Abstract

This online appendix provides supporting analysis for the primary manuscript "Health Insurance for "Humans": Information Frictions, Plan Choices, and Consumer Welfare" published in the *American Economic Review*. Appendix A describes the survey used in the analysis in detail. Appendix B describes the cost model setup and estimation. Appendix C describes the choice model estimation algorithm in greater detail. Appendix D discusses a specification for our primary model that structurally models beliefs about relevant choice objects using the survey question answers. Appendix E presents a range of supporting analyses. Appendix F presents some summary statistics, and a contribution model, for consumers' health savings account choices.

*Address: Department of Economics, University of California at Berkeley, 530 Evans Hall #3880, Berkeley, CA, 94720 (e-mail: handel@berkeley.edu).

†Address: Department of Health Care Management, The Wharton School, University of Pennsylvania, 3641 Locust Walk, 306 Colonial Penn Center, Philadelphia, PA, 19104 (email: jkolstad@wharton.upenn.edu).

A Online Appendix: Survey Instrument

This appendix describes the details of how our survey was administered and provides an exact description of the questions and answer options used in our analyses (described in the text in Tables 2 and 3. The survey was designed in late 2011, in collaboration with the Human Resources (HR) and Communications departments of the employer we study. The team included representatives from a variety of stakeholders within these departments. As described in Section 2, we designed separate surveys for three distinct groups of employees: (i) incumbent HDHP employees (ii) new HDHP enrollees (could have been in PPO before) and (iii) PPO enrollees (there were very few switching back from the HDHP into the PPO from 2011 to 2012). There was substantial overlap in the questions asked to the three groups, although some were irrelevant to a given group and were thus excluded (also, the wording changed to reflect the group in question). Each survey included between 20-25 questions.

The survey was released in early 2012, with electronic invitations sent to 1,500 randomly selected employees from each of the three cohorts above, totaling 4,500 employees. A small group of high-level employees (upper management) were excluded by the HR department as potential survey candidates due to their time constraints. The email was sent from a no-reply address by the employer's insurance provider, and linked to the survey, which was hosted online by this provider. All questions required the employee to choose one or more answers, and never required the employee to fill in their own answers.[1] An example screenshot of two questions from the PPO enrollee survey is given below.

7. If you had signed up for the HDHP Plan, what would your household's deductible (amount you have to pay for care before the Plan begins to pay for costs) have been this year?

- $0
- $750
- $1500
- $3000
- $3750
- $5000
- G. Not sure

8. If you had enrolled in the HDHP Plan, what is the rate of coinsurance (% of costs you pay once your deductible is reached) you would pay when visiting an in-network [redacted] provider or pharmacy?

- 0%
- 5%
- 10%
- 20%
- 30%
- Not sure

< Back Next >

All surveys were hosted and completed electronically: respondents were identified when clicking on the link to respond, so that their responses could be linked to the administrative data used in

[1]For certain questions that allowed the employee to select one or more answers, an 'Other' option was given. If the employee chose this option, they were prompted to fill in this answer. None of these questions were used in our empirical analysis.

our analysis. As described in the text, we received responses from 579 incumbent HDHP enrollees, 571 new HDHP enrollees, and 511 PPO enrollees for an average response rate of 38%. No financial incentive was given to respond (in the literature, this is quite a high response rate for this kind of survey, given the lack of financial incentive). See Table 1 and the text in Section 2 for detailed comparisons between the full population, survey recipients, and survey respondents on the basis of observable demographics and health risk. The text there discusses respondent selection into the survey, and how it seems minimal on the basis of there observable measures.

We now present the questions and answers used in our analysis, and summarized in the main text in Tables 2 and 3. We present these from the New HDHP enrollee survey and don't present the questions for all three cohorts, since they are very similar to those presented here, with slight wording / framing changes. After delineating these questions, we give a brief discussions of other questions asked but not used in this analysis explicitly. When something is in bold, the true material used was replaced to protect the identity of the firm. For many questions, the order of the answers were shuffled, here was present a specific ordering. The numbering of the questions below corresponds exactly to the numbers for each question used in the main text.

Questions on plan financial characteristics, presented in Table 2 are:

1. What is your household deductible this year in the **HDHP**?

a. $0
b. $750
c. $1,500
d. $3,000
e. $3,750
f. $5,000
g. Not sure

2. In the **HDHP**, what is the rate of coinsurance (% you pay once your deductible is reached) you would need to pay when visiting an in-network **Insurer Name Here** provider or pharmacy?

a. 0%
b. 5%
c. 10%
d. 20%
e. 30%
f. Not sure

3. What is the maximum out-of-pocket you can spend under the **HDHP**, regardless of any funds you or **the firm** may have contributed to your Health Savings Account (HSA)?

a. $0
b. $2,500
c. $5,000
d. $6,250
e. $7,500

f. I don't know

4. How much did **the firm** contribute to your Health Savings Account (HSA) this year, including the Early Adopter Incentive?

a. $0
b. $750
c. $1,500
d. $3,000
e. $3,750
f. $6,250
g. Not sure

5. Which of the following statements is true about the Health Savings Account (HSA)?

a. Funds in the Health Savings Account roll over from year to year
b. If I don't use funds in a given year, they will be lost
c. Not sure

6. Given the tax advantages of a Health Savings Account (HSA), about how much would $1,000 in an HSA be worth in pre-tax dollars in 2012?

a. $700-$999
b. $1,000
c. $1,001-$1,300
d. $1,301-$1,600
e. Greater than $1,600
f. I don't know

The following questions and answers correspond to frictions not related to plan financial characteristics (presented in Table 3 in the main text):

7. How do the medical providers you can use in the **HDHP** in-network compare to those you can use in the **PPO** plan?

a. I can access more providers in the **HDHP**
b. I can access more providers in the **PPO**
c. I can access the same providers under each plan
d. Not sure

8. With any health plan you may spend time choosing medical providers, processing bills, and administering other plan logistics. Approximately how much time do you expect to spend on these activities this year in the **HDHP** plan, assuming a "typical" health year for you and your family?

a. No time at all
b. Less than an hour
c. 1-5 hours
d. 6-10 hours
e. 10-20 hours
f. More than 20 hours

9. Which statement best represents how you feel about spending time managing your **HDHP** plan? (Select One)

a. I understand that I may need to spend time managing my health plan, and I'm not at all concerned about it
b. I accept that I may need to spend time managing my health plan, but I'm concerned with how much time I might have to spend
c. I don't like having to spend time managing my health plan at all, no matter how much time it might be

10. What do you estimate (off the top of your head) is the total cost of the medical care you and your covered dependents consumed (including both what you paid and **the firm** paid) in the last calender year of 2011, i.e. January - December 2011?

a. \$0-\$500
b. \$501-\$2,500
c. \$2,501-\$5,000
d. \$5,001-\$10,000
e. Greater than \$10,000
f. Not sure

11. How confident are you in this estimate (**reference to 10. above**)?

a. Not very confident, or not confident at all
b. Somewhat confident
c. Very confident

12. Based on the total health care needs of you and your dependent(s) in a "typical" year, do you expect to financially benefit from the **HDHP** plan this year (including the value provided by the Health Savings Account and the **firm** contribution)?

a. Yes
b. No
c. Not sure

In addition to these questions, which the analysis focuses on, we ask about 10-15 other questions covering the following topics:

–Primary reasons for enrolling in HDHP (PPO): consumers can choose several options from list of 7.

–Questions around whether you discussed health plan choice with others, and whether those discussions were informative / influential.

–Questions about consumer learning, including time spent with plan materials provided by Benefits and Communications group and effectiveness of those materials. Also, what plan aspects consumers would like to learn more about.

–Impact of cost-sharing / deductible for medical care utilization. Is utilization impacted by additional cost sharing in HDHP, and, if so, exactly how (list of options)?

Finally, we are currently in the process of running a survey in 2013 that delves more deeply into questions about consumer hassle costs in plan use, consumer medical care utilization, and the mechanisms through which consumers acquire information.

B Online Appendix: Cost Model Setup and Estimation

This appendix describes the details of the cost model, which is summarized at a high-level in section 3.[2] The output of this model, F_{kjt}, is a family-plan-time specific distribution of predicted out-of-pocket expenditures for the upcoming year. This distribution is an important input into the choice model, where it enters as a family's predictions of its out-of-pocket expenses at the time of plan choice, for each plan option.[3] We predict this distribution in a sophisticated manner that incorporates (i) past diagnostic information (ICD-9 codes) (ii) the Johns Hopkins ACG predictive medical software package (iii) a non-parametric model linking modeled health risk to total medical expenditures using observed cost data and (iv) a detailed division of medical claims and health plan characteristics to precisely map total medical expenditures to out-of-pocket expenses.The level of precision we gain from the cost model leads to more credible estimates of the choice parameters of primary interest (e.g. risk preferences and information friction impacts).

In order to most precisely predict expenses, we categorize the universe of total medical claims into four mutually exclusive and exhaustive subdivisions of claims using the claims data. These categories are (i) hospital and physician (ii) pharmacy (iii) mental health and (iv) physician office visit. We divide claims into these four specific categories so that we can accurately characterize the plan-specific mappings from total claims to out-of-pocket expenditures since each of these categories maps to out-of-pocket expenditures in a different manner. We denote this four dimensional vector of claims C_{it} and any given element of that vector $C_{d,it}$ where $d \in D$ represents one of the four categories and i denotes an individual (employee or dependent). After describing how we predict this vector of claims for a given individual, we return to the question of how we determine out-of-pocket expenditures in plan j given C_{it}.

Denote an individual's past year of medical diagnoses and payments by ξ_{it} and the demographics age and sex by ζ_{it}. We use the ACG software mapping, denoted A, to map these characteristics into a predicted mean level of health expenditures for the upcoming year, denoted θ:

$$A : \xi \times \zeta \to \theta$$

In addition to forecasting a mean level of total expenditures, the software has an application that predicts future mean pharmacy expenditures. This mapping is analogous to A and outputs a prediction λ for future pharmacy expenses.

We use the predictions θ and λ to categorize similar groups of individuals across each of four claims categories in vector in C_{it}. Then for each group of individuals in each claims category, we use the actual ex post realized claims for that group to estimate the ex ante distribution for each individual under the assumption that this distribution is identical for all individuals within the cell. Individuals are categorized into cells based on different metrics for each of the four elements of C:

$$\begin{aligned} \text{Pharmacy:} \quad & \lambda_{it} \\ \text{Hospital / Physician (Non-OV):} \quad & \theta_{it} \\ \text{Physician Office Visit:} \quad & \theta_{it} \\ \text{Mental Health:} \quad & C_{MH,i,t-1} \end{aligned}$$

For pharmacy claims, individuals are grouped into cells based on the predicted future mean phar-

[2] The model is similar to that used in Handel (2013).

[3] In the consumer choice model, this is mostly useful for estimating out-of-pocket expenditures in the HDHP, since the PPO plan has essentially zero expenditures.

macy claims measure output by the ACG software, λ_{it}. For the categories of hospital / physician (non office visit) and physician office visit claims individuals are grouped based on their mean predicted total future health expenses, θ_{it}. Finally, for mental health claims, individuals are grouped into categories based on their mental health claims from the previous year, $C_{MH,i,t-1}$ since (i) mental health claims are very persistent over time in the data and (ii) mental health claims are uncorrelated with other health expenditures in the data. For each category we group individuals into a number of cells between 8 and 12, taking into account the trade off between cell size and precision.

Denote an arbitrary cell within a given category d by z. Denote the population in a given category-cell combination (d, z) by I_{dz}. Denote the empirical distribution of ex-post claims in this category for this population $\hat{G_{I_{dz}}}(\cdot)$. Then we assume that each individual in this cell has a distribution equal to a continuous fit of $\hat{G_{I_{dz}}}(\cdot)$, which we denote G_{dz}:

$$\varpi : \hat{G_{I_{dz}}}(\cdot) \rightarrow G_{dz}$$

We model this distribution continuously in order to easily incorporate correlations across d. Otherwise, it would be appropriate to use $G_{I_{dz}}$ as the distribution for each cell.

The above process generates a distribution of claims for each d and z but does not model correlations over D. It is important to model correlation over claim categories because it is likely that someone with a bad expenditure shock in one category (e.g. hospital) will have high expenses in another area (e.g. pharmacy). We model correlation at the individual level by combining marginal distributions G_{idt} $\forall$ d with empirical data on the rank correlations between pairs (d, d').[4] Here, G_{idt} is the distribution G_{dz} where $i \in I_{dz}$ at time t. Since correlations are modeled across d we pick the metric θ to group people into cells for the basis of determining correlations (we use the same cells that we use to determine group people for hospital and physician office visit claims). Denote these cells based on θ by z_θ. Then for each cell z_θ denote the empirical rank correlation between claims of type d and type d' by $\rho_{z_\theta}(d, d')$. Then, for a given individual i we determine the joint distribution of claims across D for year t, denoted $H_{it}(\cdot)$, by combining i's marginal distributions for all d at t using $\rho_{z_\theta}(d, d')$:

$$\Psi : G_{iDt} \times \rho_{z_{\theta_{it}}}(D, D') \rightarrow H_{it}$$

Here, G_{iDt} refers to the set of marginal distributions $G_{idt} \forall d \in D$ and $\rho_{z_{\theta_{it}}}(D, D')$ is the set of all pairwise correlations $\rho_{z_{\theta_{it}}}(d, d') \forall (d, d') \in D^2$. In estimation we perform Ψ by using a Gaussian copula to combine the marginal distribution with the rank correlations, a process which we describe momentarily.

The final part of the cost model maps the joint distribution H_{it} of the vector of total claims C over the four categories into a distribution of out of pocket expenditures for each plan. For the HDHP we construct a mapping from the vector of claims C to out of pocket expenditures OOP_j:

$$\Omega_j : C \rightarrow OOP_j$$

This mapping takes a given draw of claims from H_{it} and converts it into the out of pocket expenditures an individual would have for those claims in plan j. This mapping accounts for plan-specific features such as the deductible, co-insurance, co-payments, and out of pocket maximums listed in table A-2. We test the mapping Ω_j on the actual realizations of the claims vector C to verify that our mapping comes close to reconstructing the true mapping. Our mapping is necessarily simpler

[4]It is important to use rank correlations here to properly combine these marginal distribution into a joint distribution. Linear correlation would not translate empirical correlations to this joint distribution appropriately.

and omits things like emergency room co-payments and out of network claims. We constructed our mapping with and without these omitted categories to ensure they did not lead to an incremental increase in precision. We find that our categorization of claims into the four categories in C passed through our mapping Ω_j closely approximates the true mapping from claims to out-of-pocket expenses. Further, we find that it is important to model all four categories described above: removing any of the four makes Ω_j less accurate.

Once we have a draw of OOP_{ijt} for each i (claim draw from H_{it} passed through Ω_j) we map individual out of pocket expenditures into family out of pocket expenditures. For families with less than two members this involves adding up all the within family OOP_{ijt}. For families with more than three members there are family level restrictions on deductible paid and out-of-pocket maximums that we adjust for. Define a family k as a collection of individuals i_k and the set of families as K. Then for a given family out-of-pocket expenditures are generated:

$$\Gamma_j : OOP_{i_k,jt} \rightarrow OOP_{kjt}$$

To create the final object of interest, the family-plan-time specific distribution of out of pocket expenditures $F_{kjt}(\cdot)$, we pass the total cost distributions H_{it} through Ω_j and combine families through Γ_j. $F_{kjt}(\cdot)$ is then used as an input into the choice model that represents each family's information set over future medical expenses at the time of plan choice. Figure B1 outlines the primary components of the cost model pictorially to provide a high-level overview and to ease exposition.

We note that the decision to do the cost model by grouping individuals into cells, rather then by specifying a more continuous form, has costs and benefits. The cost is that all individuals within a given cell for a given type of claims are treated identically. The benefit is that our method produces local cost estimates for each individual that are not impacted by the combination of functional form and the health risk of medically different individuals. Also, the method we use allows for flexible modeling across claims categories. Finally, we note that we map the empirical distribution of claims to a continuous representation because this is convenient for building in correlations in the next step. The continuous distributions we generate very closely fit the actual empirical distribution of claims across these four categories.

Cost Model Identification and Estimation. The cost model is identified based on the two assumptions of (i) no moral hazard / selection based on private information and (ii) that individuals within the same cells for claims d have the same ex ante distribution of total claims in that category. Once these assumptions are made, the model uses the detailed medical data, the Johns Hopkins predictive algorithm, and the plan-specific mappings for out of pocket expenditures to generate the the final output $F_{kjt}(\cdot)$. These assumptions, and corresponding robustness analyses, are discussed at more length in the main text.

Once we group individuals into cells for each of the four claims categories, there are two statistical components to estimation. First, we need to generate the continuous marginal distribution of claims for each cell z in claim category d, G_{dz}. To do this, we fit the empirical distribution of claims $G_{I_{dz}}$ to a Weibull distribution with a mass of values at 0. We use the Weibull distribution instead of the log-normal distribution, which is traditionally used to model medical expenditures, because we find that the log-normal distribution over-predicts large claims in the data while the Weibull does not. For each d and z the claims greater than zero are estimated with a maximum likelihood fit to the Weibull distribution:

$$\max_{(\alpha_{dz},\beta_{dz})} \Pi_{i\in I_{dz}} \frac{\beta_{dz}}{\alpha_{dz}}(\frac{c_{id}}{\alpha_{dz}})^{\beta_{dz}-1}e^{-(\frac{c_{id}}{\alpha_{dz}})^{\beta_{dz}}}$$

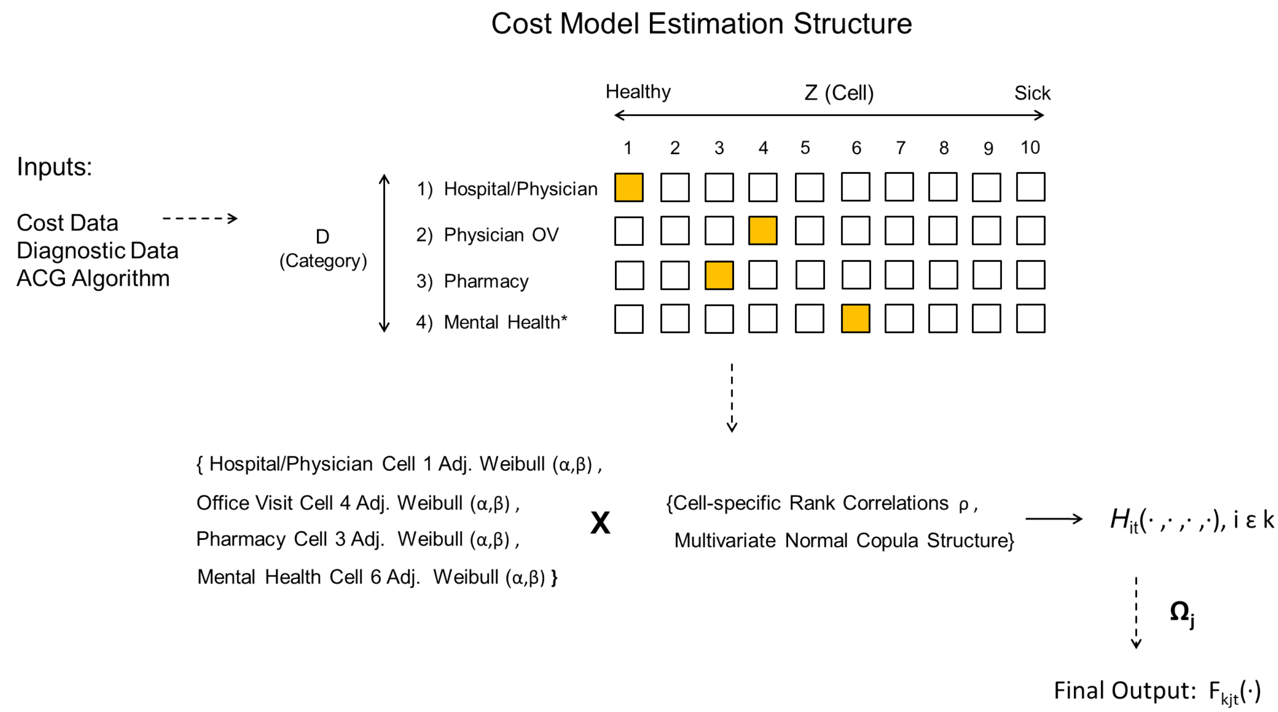

Figure B1: This figure outlines the primary steps of the cost model described in Appendix B. It moves from the initial inputs of cost data, diagnostic data, and the ACG algorithm to the final output F_{kjt} which is the family, plan, time specific distribution of out-of-pocket expenditures that enters the choice model for each family. The figure depicts an example individual in the top segment, corresponding to one cell in each category of medical expenditures. The last part of the model maps the expenditures for all individuals in one family into the final distribution F_{kjt}.

Here, $\hat{\alpha_{dz}}$ and $\hat{\beta_{dz}}$ are the shape and scale parameters that characterize the Weibull distribution. Denoting this distribution $W(\hat{\alpha_{dz}}, \hat{\beta_{dz}})$ the estimated distribution $\hat{G_{dz}}$ is formed by combining this with the estimated mass at zero claims, which is the empirical likelihood:

$$\hat{G_{dz}}(c) = \begin{cases} G_{I_{dz}}(0) & \text{if } c = 0 \\ G_{I_{dz}}(0) + \frac{W(\hat{\alpha_{dz}}, \hat{\beta_{dz}})(c)}{1 - G_{I_{dz}}(0)} & \text{if } c > 0 \end{cases}$$

Again, we use the notation $\hat{G_{iDt}}$ to represent the set of marginal distributions for i over the categories d: the distribution for each d depends on the cell z an individual i is in at t. We combine the distributions $\hat{G_{iDt}}$ for a given i and t into the joint distribution H_{it} using a Gaussian copula method for the mapping Ψ. Intuitively, this amounts to assuming a parametric form for correlation across $\hat{G_{iDt}}$ equivalent to that from a standard normal distribution with correlations equal to empirical rank correlations $\rho_{z_{\theta_{it}}}(D, D')$ described in the previous section. Let $\Phi^{i}_{1|2|3|4}$ denote the standard multivariate normal distribution with pairwise correlations $\rho_{z_{\theta_{it}}}(D, D')$ for all pairings of the four claims categories D. Then an individual's joint distribution of non-zero claims is:

$$\hat{H_{i,t}}(\cdot) = \Phi_{1|2|3|4}(\Phi_1^{-1}(\hat{G_{id_1t}}), \Phi_2^{-1}(\hat{G_{id_2t}}), \Phi_3^{-1}(\hat{G_{id_3t}}), \Phi_4^{-1}(\hat{G_{id_4t}})))$$

Above, Φ_d is the standard marginal normal distribution for each d. $\hat{H}_{i,t}$ is the joint distribution of claims across the four claims categories for each individual in each time period. After this is estimated, we determine our final object of interest $F_{kjt}(\cdot)$ by simulating K multivariate draws from $\hat{H}_{i,t}$ for each i and t, and passing these values through the plan-specific total claims to out of pocket mapping Ω_j and the individual to family out of pocket mapping Γ_j. The simulated $F_{kjt}(\cdot)$ for each k, j, and t is then used as an input into estimation of the choice model.

New Employees. For the first-stage full population model that compares new employees to existing employees to identify the extent of inertia, we need to estimate F_{kj} for new families. Unlike for existing families, we don't observe past medical diagnoses / claims for these families, we just observe these things after they join the firm and after they have made their first health plan choice with the firm. We deal with this issue with a simple process that creates an expected ex ante health status measure. We backdate health status in a Bayesian manner: if a consumer has health status x ex post we construct ex ante health status y as an empirical mixture distribution $f(y|x)$. $f(y|x)$ is estimated empirically and can be thought of as a reverse transition probability (if you are x in period 2, what is the probability you were y in period 1?). Then, for each possible ex ante y, we use the distributions of out-of-pocket expenditures F estimated from the cost model for that type. Thus, the actual distribution used for such employees is described by $\int_{x \in X} f(y|x)F(y)dy$. The actual cost model estimates $F(y)$ do not include new employees and leverages actual claims data for employees who have a past observed year of this data.

C Online Appendix: Choice Model Estimation

This appendix describes the algorithm by which we estimate the parameters of the choice model. The corresponding section in the text provided a high-level overview of this algorithm and outlined the estimation assumptions we make regarding choice model fundamentals and their links to observable data.

We estimate the choice model using a random coefficients probit simulated maximum likelihood approach similar to that summarized in Train (2009) and to that used in Handel (2013). The simulated maximum likelihood estimation approach has the minimum variance for a consistent and asymptotically normal estimator, while not being too computationally burdensome in our framework. We set up a likelihood function to predict the health choices of consumers in 2012. The maximum likelihood estimator selects the parameter values that maximize the similarity between actual choices and choices simulated with the parameters.

First, the estimator simulates Q draws for each family from the distribution of health expenditures output from the cost model, F_k for each family. The estimator also simulates D draws for each family-year from the distribution of the random coefficient γ_k, as well as from the distribution of idiosyncratic preference shocks ϵ_{kj}.

We define θ as the full set of model parameters of interest for the full / primary specification in Section 3:[5]

$$\theta \equiv (\mu_\gamma, \delta, \sigma_\gamma, \sigma_\epsilon, \eta_1, \eta_0, \beta).$$

We denote θ_{dk} as one draw derived from these parameters for each family, including the parameters that are constant across draws (e.g., for observable heterogeneity in γ or η) and those which change with each draw (unobservable heterogeneity in γ and ϵ):[6]

$$\theta_{dk} \equiv (\gamma_k, \epsilon_{kJ}, \eta_k, \beta)$$

Denote θ_{Dk} as the set of all D simulated parameter draws for family k. For each $\theta_{dk} \in \theta Dk$, the estimator uses all Q health draws to compute family-plan-specific expected utilities U_{dkj} following the choice model outlined in the text in Section 3. Given these expected utilities for each θ_{dk}, we simulate the probability of choosing plan j^* in each period using a smoothed accept-reject function with the form:

$$Pr_{dk}(j = j^*) = \frac{\left(\frac{\frac{1}{-U_{dkj^*}}(\cdot)}{\Sigma_J \frac{1}{-U_{skj}}(\cdot)}\right)^\tau}{\Sigma_{\hat{j}}\left(\frac{\frac{1}{-U_{sk\hat{j}}}(\cdot)}{\Sigma_J \frac{1}{-U_{skj}}(\cdot)}\right)^\tau}$$

This smoothed accept-reject methodology follows that outlined in Train (2009) with some slight modifications to account for the expected utility specification. In theory, conditional on θ_{dk}, we would want to pick the j that maximizes U_{kj} for each family, and then average over D to get final choice probabilities. However, doing this leads to a likelihood function with flat regions, because for small changes in the estimated parameters θ, the discrete choice made does not change. The smoothing function above mimics this process for CARA utility functions: as the smoothing parameter τ becomes large the smoothed Accept-Reject simulator becomes almost identical to the

[5]While we discuss estimation for the full model, the logic extends easily to the other specifications estimated in this paper.

[6]Here, we collapse the parameters determining γ_k and η_k into those factors to keep the notation parsimonious.

true accept-reject simulator just described, where the actual utility-maximizing option is chosen with probability one. By choosing τ to be large, an individual will always choose j^* when $\frac{1}{-U_{kj^*}} > \frac{1}{-U_{kj}} \forall j \neq j^*$. The smoothing function is modified from the logit smoothing function in Train (2009) for two reasons: (i) CARA utilities are negative, so the choice should correspond to the utility with the lowest absolute value and (ii) the logit form requires exponentiating the expected utility, which in our case is already the sum of exponential functions (from CARA). This double exponentiating leads to computational issues that our specification overcomes, without any true content change since both models approach the true accept-reject function.

Denote any choice made $\mathbf{j}$ and the set of such choices as $\mathbf{J}$. In the limit as τ grows large the probability of a given $\mathbf{j}$ will either approach 1 or 0 for a given simulated draw d and family k. For all D simulation draws we compute the choice for k with the smoothed accept-reject simulator, denoted $\mathbf{j}_{dk}$. For any set of parameter values θ_{Sk} the probability that the model predicts $\mathbf{j}$ will be chosen by k is:

$$\hat{P}_k^{\mathbf{j}}(\theta, F_{kj}, X_{kt}^A, X_{kt}^B, \mathbf{Z}') = \Sigma_{d \in D} \mathbf{1}[\mathbf{j} = \mathbf{j}_{dk}]$$

Let $\hat{P}_k^{\mathbf{j}}(\theta)$ be shorthand notation for $\hat{P}_k^{\mathbf{j}}(\theta, F_{kj}, X_{kt}^A, X_{kt}^B, \mathbf{Z}')$. Conditional on these probabilities for each k, the simulated log-likelihood value for parameters θ is:

$$SLL(\theta) = \Sigma_{k \in K} \Sigma_{\mathbf{j} \in \mathbf{J}} d_{k\mathbf{j}} \; ln \hat{P}_k^{\mathbf{j}}$$

Here $d_{k\mathbf{j}}$ is an indicator function equal to one if the actual choice made by family k was $\mathbf{j}$. Then the maximum simulated likelihood estimator (MSLE) is the value of θ in the parameter space Θ that maximizes $SLL(\theta)$. In the results presented in the text, we choose $Q = 50$, $S = 50$, and $\tau = 6$, all values large enough such that the estimated parameters vary little in response to changes.

C.1 Model Implementation and Standard Errors

We implement the estimation algorithm above with the KNITRO constrained optimization package in Matlab. One challenge in non-linear optimization is to ensure that the algorithm finds a global maximum of the likelihood function rather than a local maximum. To this end, we run each model 12 times where, for each model run, the initial parameter values that the optimizer begins its search from are randomly selected from a wide range of reasonable potential values. This allows for robustness with respect to the event that the optimizer finds a local maximum far from the global maximum for a given vector of starting values. We then take the estimates from each of these 12 runs, and select the estimates that have the highest likelihood function value, implying that they are the best estimates (equal to or closest to a global maximum). We ran informal checks to ensure that, for each model, multiple starting values converged to very similar parameters similar to those with the highest likelihood function value, to ensure that we were obtaining robust results.

We compute the standard errors, provided in Appendix E, with a block bootstrap method. This methodology is simple though computationally intensive. First, we construct 50 separate samples, each the same size as our estimation sample, composed of consumers randomly drawn, with replacement, from our actual estimation sample. We then run each model, for 8 different starting values, for each of these 50 bootstrapped samples (implying 400 total estimation runs per model). The 8 starting values are drawn randomly from wide ranges centered at the actual parameter estimates. For each model, and each of the 50 bootstrapped samples, we choose the parameter estimates that have the highest likelihood function value across the 8 runs. This is the final estimate for each bootstrapped sample. Finally, we take these 50 final estimates, across the bootstrapped samples, and calculate the 2.5th and 97.5th percentiles for each parameter and statistic (we actually use the 4th and 96th percentiles given that 50 is a discrete number). Those percentiles are then, respectively, the upper and lower bounds of the 95% confidence intervals presented in Appendix E. See e.g., Bertrand et al. (2004) for an extended discussion of block bootstrap standard errors.

Finally, it is important to note that the 95% confidence intervals presented in Appendix E should really be interpreted as outer bounds on the true 95% intervals, due to computational issues with non-linear optimization. Due to time and computational constraints, we could only run each of the 50 bootstrap sample runs 8 times, instead of 12. In addition, we could not check each of these bootstrapped runs with the same amount of informal checks as for the primary estimates. This implies that, in certain cases, it is possible that one or several of the 50 estimates for each of the bootstrapped samples are not attaining a global maximum. In this case, e.g., it is possible that 45 of the 50 final estimates are attaining global maxima, while 5 are not. As a result, it is possible that the confidence intervals reported are quite wide due to computational uncertainty, even though the 45 runs that attain the global maximum have results that are quite close together. In essence, in cases where computational issues / uncertainty lead to a final estimate for a bootstrapped sample that is not a global maximum, the confidence intervals will look wide (because of these outlier / incorrect final estimates) when most estimates are quite similar. One solution to this issue would be to run each of the models more times (say 12 or 20) for each bootstrapped sample. This would lead to fewer computational concerns, but would take 1.5 to 2.5 times as long, which is substantial since the standard errors for one model take 7-10 days to run.

As a result, the confidence intervals presented should be thought of as outer bounds on the true 95% CIs. This means that for the models where these bounds are tight, the standard error results are conclusive / compelling since the true 95% CI lies in between these already tight bounds. In cases where the CI is very wide, this means that the true 95% CI lies in that wide range, and that we cannot draw meaningful conclusions due to computational uncertainty in all likelihood. Of course, it is possible the true CI is wide, but, in cases where 46 out of 50 bootstrapped parameter estimates are tight and four are outliers (without substantial variations in the underlying samples) this suggests that computational uncertainty is at fault for the wide bounds.

D Online Appendix: Structural Information Model

The primary estimated models presented in the main text used reduced form specifications for information frictions. Specifically, as described in Section 3 in the main text, the primary models includes indicator variables derived from survey question answers that are used in the model as shifters of consumer willingness to pay for insurance. The "types" model presented aggregates these indicator variables into an information index that predicts willingness to pay.

The survey design did not explicitly target structural parameters such as exact consumer beliefs about different plan features: as described in Section 2 there is an inherent tradeoff when designing a survey between posing questions that consumers can easily understand and questions that more clearly link to direct structural objects in a model of consumer information and choice under uncertainty. There are clear advantages to asking simpler questions (e.g. consumers might not understand probability, or feel less comfortable with open response answers). There are also clear advantages to questions that elicit beliefs structurally: if consumers can answer them correctly, they may provide more precise signals to use in the context of economic analysis. After substantial debate about the more prudent form for questions together with the firm's HR department, we opted for simpler questions that we felt consumers could easily understand, at the expense of asking for more precise structural objects. The design of the survey data links directly to our primary models that include signals about consumer information in a reduced form manner.

Though the models presented in detail in the text reflect our preferred approach, we also belief that it is important and illustrative to investigate specifications that seek to structurally integrate information-based survey questions into the consumer choice problem. A more structural interpretation of survey question answers is valuable to investigate for a number of reasons. First, a structural interpretation of the information answers links the data more closely to the underlying structural objects they represent. For example, the primary specification in the text includes a "rational expectations" distribution of health risk structurally, and then includes indicator variables for whether the consumer knows the deductible or not. This implies that all consumers who do not correctly know the deductible have willingness to pay shifted by the same amount, regardless of their health risk. Our structural model allows us to more tightly link health risk expectations and information about the deductible: for example, a very healthy consumer might care less about an incorrectly perceived large deductible than a very sick consumer. Second, a fully structural specification is better able to capture the magnitude of consumer misperceptions. For example, in the primary model in the text, someone who has the deductible wrong by a little is treated the same as someone who has the deductible wrong by a lot. A fully structural specification can clearly integrate these differences more precisely (though the reduced form approach could also include a finer set of indicator variables to partially address this issue). Finally, including each information related question in a fully structural manner may lead to different (and potentially better) estimates of risk preferences by shifting the mean and variance of consumer beliefs about out of pocket expenditures in each plan design. The primary model in the text is equivalent to shifting mean consumer projected out of pocket expenditures, but does not shift the variance. For example, if a consumer beliefs the out-of-pocket maximum in the HDHP is quite high, higher than in reality, this would increase their expenditure variance as well as the mean.

In this section, both as a robustness check to our primary risk preference estimates and because it is interesting in its own right, we describe and estimate a fully structural information specification that maintains additional assumptions to structurally integrate the survey data as precise objects in the consumer decision problem. This specification is briefly discussed in the main text in Section 3, and the results are also presented in Table 4 in the main text so they can be directly compared to our primary models.

We now describe how we model and implement the structural information specification. The first three information-related survey questions we structurally integrate ask about financial characteristics of the HDHP non-linear contract design. Specifically, as discussed in the text and in Online Appendix A we use the questions that ask consumers directly about the HDHP deductible, the HDHP coinsurance rate, and the HDHP out-of-pocket maximum. Here, we link the answers consumers provide to their beliefs about out-of-pocket cost realizations in the HDHP.

The cost model presented in Online Appendix B describes the distribution of total costs (insurer + insured) H_{it} for each individual i at time t. The mapping Ω_j describes how individual total expenditures map to out-of-pocket expenditures in the HDHP. The mapping Γ_j describes how the vector of individual out-of-pocket expenditures for a set of family members map into a family out-of-pocket expenditure amount. The mapping is decomposed into individual and family components because for families of 3 or more, one cannot just map total family expenditures through the plan design into out-of-pocket costs, because the individual-level within family amounts matter as well (as described in detail in Online Appendix B). For exposition in this section we assume that there is a direct one-to-one mapping from family total expenditures H_{kt} to out-of-pocket expenditures OOP_{kjt}, and we denote this mapping Υ_j.[7]

We illustrate the model for a family with 3 or more members, though it is easy to see how it extends to individuals and employees with one dependent (we implement the model for everybody). With this simplified version of the cost model in Online Appendix B define the HDHP out-of-pocket mapping from a draw C from total family expenditure distribution H_{kt} to a family out-of-pocket payment as:

$$\Upsilon_j : C \rightarrow OOP_{k,HDHP,t}$$

The distribution of the output $OOP_{k,HDHP,t}$ is what enters the choice model described in the text as $F_{k,HDHP,t}$. For a family k with 3 or more members the mapping Υ_j is represented by:

$$OOP_{k,HDHP,t} = \begin{cases} C & \text{if} \quad 0 \leq C \leq 3750 \\ 0.1(C-3750)+3750 & \text{if} \quad 3750 \leq C \leq 28750 \\ 6250 & \text{if} \quad 28750 \leq C \end{cases}$$

As described in the text, the mapping is defined this way because for a family of 3 or more, the family deductible is \$3750, the coinsurance rate is 10%, and the family out-of-pocket maximum is \$6250.

Now, consider an employee answering the multiple choice survey questions about deductible, coinsurance, and out-of-pocket maximum. We assign that employee's answers to the entire family by taking the numerical answers given to these questions and assuming that they believe with certainty that these answers equal the actual plan characteristics.[8] For each family k define the perceived deductible, coinsurance, and out-of-pocket maximum for the HDHP as $D\hat{E}D_k$, $C\hat{I}_k$, and $M\hat{A}X_k$ respectively. Now, we define the plan design perceived by the consumer, $\widehat{\Upsilon_{HDHP}}$ as:

[7]This is for expositional purposes only. When we implement this model, we allow for the cases present in families greater than three where some family members meet their individual-level deductibles before the family meets the family-level deductible. In general, the actual ex post expenses for almost all families in the data can map total family costs into out-of-pocket costs as described here.

[8]The survey does not give a way to express numerical uncertainty, e.g. a consumer guesses the deductible is \$X, but has uncertainty about their answer. A richer model could elicit beliefs about the deductible and include that in the choice problem.

$$\widehat{OOP_{k,HDHP,t}} = \begin{cases} C & \text{if} \quad 0 \leq C \leq D\hat{E}D_k \\ C\hat{I}_k(C - D\hat{E}D_k) + D\hat{E}D_k & \text{if} \quad D\hat{E}D_k \leq C \leq D\hat{E}D_k + \frac{1}{C\hat{I}_k}(M\hat{A}X_k - D\hat{E}D_k) \\ M\hat{A}X_k & \text{if} \quad D\hat{E}D_k + \frac{1}{C\hat{I}_k}(M\hat{A}X_k - D\hat{E}D_k) \leq C \end{cases}$$

Under the assumption that consumers know their distribution of total expenditures (insured + insuree), an assumption that we relax shortly, mapping the distribution H_{kt} through $\widehat{\Upsilon_{HDHP}}$ generates the perceived distribution of out-of-pocket expenditures $\widehat{F_{k,HDHP,t}}$ for family k at time t. This is the version of the rational expectations distribution F used throughout the text that incorporates beliefs about these non-linear plan characteristics.

We note, importantly, that some people respond "not sure" to answers about these three plan financial characteristics. For these people there is no obvious perceived characteristic to include. We deal with this issue by randomly drawing an answer for these people from the distribution of answers given in the entire population, conditional on being in the same coverage tier (and thus having the same true plan financial characteristics). This is a strong assumption, that is somewhat weakened in the primary model in the text where this answer just shifts willingness to pay, rather than having a specific structural interpretation.[9]

In addition to integrating survey data on these three plan financial characteristics, we structurally integrate answers to the questions that ask consumers (i) about their perception of their own total medical expenditures (insured + insuree) and (ii) about the HDHP subsidy amount. For the question asking consumers about the HDHP plan subsidy, we take their answer as their perceived value and input it into the choice model $\widehat{HSA_k^S}$. For people answering "not sure" to that question we used a random draw from the population distribution of answers conditional on coverage tier.

For the question that asks consumers about their total family expenditures, we develop a methodology that changes consumer perceptions from the rational expectations distribution H_{kt} to a perceived distribution $\widehat{H_{kt}}$. This methodology has the following steps:

1. Group each family into a category based on their answer to question 10 in Online Appendix A asking them what their total medical expenditures for the past year were. This places consumers into five groups of total medical spending.

2. Determine the set of consumers with a specific number of covered dependents who have projected mean expenditures within each answer bucket to question 10, following the rational consumer cost model developed in Online Appendix B.

3. For an employee in a given coverage tier who answers question 10 by saying they have health expenditures in a given range (i) if they are correct, keep their distribution H_{kt} as defined in Online Appendix B (ii) if they are incorrect, draw 50 expenditure draws from other families who have *actual* projected means in that range.

These draws from step 3 serve as their perceived distribution of expenses $\widehat{H_{kt}}$ when making their plan choice, rather than their "true" distribution. This is then mapped through their true

[9] We encountered one additional issue for a small number of consumers. A few consumers answered questions in a conflicting manner whereby they stated that the out-of-pocket maximum was smaller than the deductible. For these consumers, whose answers are structurally inconsistent, we assign them to a plan where the deductible they state is the true deductible, and that deductible equals the out-of-pocket maximum. We use a similar design for consumers who answer that there is a 0% coinsurance rate, but answer that the deductible and out-of-pocket maximum are not the same.

plan design as described above. For people answering "not sure" to this question we used their true cost distribution. Since using these responses to questions about past total medical expenditures as structural measures for projected expenditures going forward requires stronger assumptions than those required for integrating the questions about plan financial characteristics, we present estimates two models, one where these structural misperceptions of total expenditures are included and one where they are not.

Finally, for the questions on time and hassle costs and information about the provider networks in each plans we use the same methodology as in the primary model and include indicator variables that reflect consumer answers to these questions.

Given this setup, the consumer choice problem is to choose the plan that maximizes their perceived utility (or perceived willingness to pay). Borrowing notation from the primary models listed in Section 3 consumer willingness to pay is described by:

$$U_{kj} = \int_0^\infty \widehat{f_{kj}(s)} u_k(W_k, x_{kj}(P_{kj}, s))ds$$

$$x_{kj} = W_k - \widehat{P_{kj}} - s + \eta(X_k^B)\mathbf{1}_{j_t=j_{t-1}} + \mathbf{Z'}_k\beta\mathbf{I}_{HDHP} + \epsilon_{kj}$$

Here, the rational expectations distribution of projected out-of-pocket expenditures F_{kj} is replaced with the perceived distribution $\widehat{F_{kj}}$ where the latter is formed as described above, leveraging the answers to four survey questions on the plan financial characteristics and total expenditures. $\widehat{P_{kj}}$ is the perceived premium difference between the two plans, given the perceived subsidy $\widehat{HSA_k^S}$. With some slight abuse of notation, Z' now includes indicator variables for only the questions asking about information on provider networks and perceptions of time and hassle costs (as in the models described in the text). We note that while the model described in this appendix is our main structural information specification, we have investigated several versions for this model and additional results are available upon request. These versions test small variations, such as including perceptions about total medical expenditures in a slightly different manner.

The estimates and 95% CIs for two structural information models (with and without the structural version of the total medical expenditures question) are presented in table D1. The model in the first column, without the structural interpretation of total expenditures, is also presented in the main text in table 4. Both structural information models estimate that consumers are slightly *less* risk averse than our primary specifications presented in the text do. Both versions here predict a gamble interpretation of approximately $X = 953$ while that in the main model presented in the main text is $X = 920.47$ (see the text in Section 4 for an extended discussion of what gamble interpretation refers to). Thus, for our welfare analysis and counterfactuals, which focus on consumer welfare loss from additional risk exposure, the structural information model would lead to an even larger difference in welfare predictions relative to the baseline models that do not include measures of information frictions and hassle costs. Table D1 also presents the 95% CI for the structural information model without the structural interpretation of total expenditures: the 95% CI for the mean risk aversion gamble interpretation is $[922.21, 957.50]$. This means that even for the highest risk aversion in this confidence interval, there would be substantial welfare implications for additional risk exposure relative to the baseline model. It is worth mentioning that both the risk aversion results, and almost all other coefficients, are very near each other for both of the structural information models estimated.

The table also presents estimates for the variables included for provider network information and perceived time and hassle costs: these estimates remain large and similar in spirit to those presented in our primary specification in the main text (both remain significantly different than 0 at the 95% level). The estimated coefficients are larger in magnitude for both (i) the implied willingness to pay for each incremental hour of perceived time and hassle costs and (ii) the willingness to pay for the PPO for people who believe it allows access to more providers. These larger coefficients likely reflect the fact that, in the structural information model, the value actually at stake in the consumer insurance decision may not be the same as the perceived value at stake which can be larger. Since, as shown in section 2, both perceived time and hassle costs and the belief that the PPO allows access to more providers are correlated with PPO choice, the increased coefficients in the structural information model suggest that those individuals leave more *perceived* money on the table to join the PPO than the *actual* money they leave on the table. Aggregating over the survey measures that remain in the structural information model reveals an overall mean effect of -$2,907 of limited information about provider networks and perceived relative HDHP time and hassle costs on willingness to pay for the HDHP relative to the PPO (for the model with structural total expenditure measures included). See Section 4 in the main text for an extended discussion of how to interpret these coefficients and how they compare to those in the primary models presented in the main text.

Structural Information Models **Estimates & 95% CIs**	(4) Main Version	95% CIs for (1)	(18) Version w/ TME included
Average μ_γ	$4.94 \cdot 10^{-5}$	$[4.44 \cdot 10^{-5}, 8.43 \cdot 10^{-5}]$	$4.94 \cdot 10^{-5}$
Std. Dev. μ_γ	$6.16 \cdot 10^{-6}$	$[4.70 \cdot 10^{-6}, 1.01 \cdot 10^{-5}]$	$6.14 \cdot 10^{-6}$
Gamble Interpretation of Average μ_γ	952.89	[922.21,957.50]	953.09
σ_γ	$5.80 \cdot 10^{-6}$	$[1.31 \cdot 10^{-11}, 4.11 \cdot 10^{-5}]$	$6.51 \cdot 10^{-6}$
σ_ϵ, HDHP	0.13	[0.01,431.27]	0.09
Time cost hrs. X prefs:			
Time cost hrs.	-12.31	[-114.20,131.17]	11.53
... X Accept, concerned	-195.11	[-360.20,-113.27]	-219.61
... X Dislike	-220.38	[-384.01,-134.42]	-242.68
Provider networks:			
HDHP network bigger	-1424.55	[-3015.46,492.26]	-1359.46
PPO network bigger	-4513.65	[-6337.77,-2153.47]	-4620.94
Not sure	-590.95	[-1300.25,658.39]	-567.00
TME guess:			
Overestimate	1178.21	[703.39,2107.89]	-
Underestimate	-1175.57	[-1535.98,78.91]	-
Not sure	-1552.80	[-3312.02,-30.72]	-
Average Survey Effect	-3022.12	[-3732.14,-1180.59]	-2907.06
σ Survey Effect	2617.07	[1988.31,3292.41]	2642.65

Table D1: This table presents the results from the structural information models described in depth in Online Appendix D. Column (4), repeated from Table 4 presents the version that does not treat the survey question total medical expenditure as a structural object, and the next column presents the 95% confidence intervals for those estimates. Column (18) presents the version of the model that does have a structural treatment of total medical expenditure perceptions. The results on risk preferences are similar to those in the main specifications in the text, though indicative of slightly *less* risk aversion that those specifications. See the Online Appendix text for a further discussion.

E Online Appendix: Additional Analysis

This appendix presents results from additional analyses referred to in the main text. It includes (i) some additional descriptive analysis (ii) several robustness checks for the primary model specifications and (iii) standard errors for all model estimates presented in the main text.

Table E1 presents the specific financial and non-financial characteristics for the PPO and HDHP plan options offered at the firm. This table is described in the main text in Section 2 and the relationship between the financial aspects of the two insurance contracts are shown graphically in Figure 1 in that section.

Table E2 presents raw correlations between pairs of binary friction variables derived from the survey. Each entry represents the correlation between the variables listed for the relevant row / column pair. For example, the correlation between correctly knowing one's deductible and correctly knowing one's coinsurance rate is 0.35. As discussed in Section 2, the high correlations between several of the friction measures for information about plan financial characteristics suggests that a types specification, which we investigate in Section 3 might be interesting.

Table E3 presents the raw correlations between all other frictions measures, including, e.g., perceived time and hassle costs and provider network knowledge. In this table, we include an aggregated measure for knowledge of plan financial characteristics, reflecting the substantial correlations in those measures shown in Table E2 (this is also done here to make the exposition clearer and more parsimonious). The correlations between these measures are lower, suggesting real heterogeneity in the population across these dimensions. Multi-dimensional heterogeneity in frictions for consumers is suggestive of nuanced and informative answers. The limited pairwise correlations between the aggregated measure of plan financial characteristic knowledge, knowledge about provider networks, expected time and hassle costs, and knowledge of own past medical expenditures suggests that if confirmation bias were present, it would have to manifest on different dimensions for different consumers, which we believe is less likely than the case where it is present on similar dimensions across consumers. There is also additional suggestive evidence of limited confirmation bias, seen in other tables. First, for many information-related questions, such as those on plan financial characteristics, PPO enrollees are more likely to answer "not sure" relative to HDHP enrollees: both groups are similarly likely to answer these questions incorrectly. "Not sure" suggests a lack of knowledge, but does not suggest validation of the PPO choice. Furthermore, for many more factually based questions (e.g., what was the deductible) it is not obvious that one answer is more preferential to a specific plan. Second, there is meaningful variation across questions in the proportion of consumers choosing answers that are favorable to the plan they chose. For example, 71% of PPO enrollees know that you can roll over HSA funds, an answer that is favorable to the HDHP plan, while only 6% believe the HDHP plan provides access to more doctors.

Table E4 presents descriptive statistics for all new employees in 2011, and compares that population to the permanent set of existing employees studied in our full population analyses. The comparison between these two groups is especially relevant to identification of the inertia parameter η in models where it is relevant / included (such full population baseline choice model with inertia, the survey respondent analog to that model, and the sequence of models with friction measures that include η. The table shows that new employees are relatively likely to be younger, lower income, and single. However, they do cover the range of demographics on each of these dimensions in large enough quantities to identify inertia conditional on observable heterogeneity, which mitigates any concerns of selection on these characteristics into the new employee sample for the purposes of estimating inertia. Also, for new employees we include projected health risk distributions that backdate their future (ex-post) claims in a Bayesian manner: i.e. if you have health status x ex post we construct ex ante health status y as an empirical mixture distribution $f(y|x)$. Then, for each possible ex ante y, we use the distributions of out-of-pocket expenditures F estimated from the cost model for that type. Thus, the actual distribution used for such employees is described by $\int_{x \in X} f(y|x)F(y)dy$ (see Appendix B for more details).

Table E5 presents the full results for the first-stage model that estimates risk preferences, health risk, and inertia for the full permanent population references in column 1 of Table 1. The main estimates from this model are discussed in the text in Section 4. Figure E1 presents a histogram of the inertia parameter η in the population, where it varies from person to person as a function of observable heterogeneity. The impact of inertia is larger for families than for single employees, reflecting the fact that the former have more money at stake in the health insurance decision. These inertia estimates are used as inputs into the primary models with frictions that we estimate, as described in Section 3.

Table E6 shows the results for our baseline model with inertia, estimated only on the most informed patients in our primary sample. To do this, we restrict the sample to those for who our information type index $q \geq 4$ (8 is the maximum value). This restricted sample corresponds to the approximately 30% of consumers who are most informed (see Figure 4 in Section 3 in the main text for a histogram of types, and the corresponding discussion in the text in that section for a discussion of how the types are constructed). We estimate this model to assess the identification assumption that the choices of fully informed consumers identify the distribution of risk preferences in our full model. The resulting distribution of risk preference parameters are similar between the full model and baseline model with informed consumers, while both are quite different from the baseline model with inertia estimated for all consumers. This helps to validate the assumption that risk preferences in the full model are appropriately identified based on the choices that fully informed consumers are making (where frictions don't play a role in choices). This table compares the relevant risk preference coefficients in the baseline model with inertia, for informed consumers, to those from the same model estimated for all consumers ((2) in Table 4) and those from our full models with frictions ((3) in Table 4).

Table E7 presents results from the incremental friction models where we add friction measures one at a time. Compared to average consumer "gamble interpretation" of $X = 812.61$ for the baseline model with inertia, the mean "gamble interpretation" in the incremental models are $X = 895.35$ for the model with plan financial characteristic frictions, 852.14 with total medical expenditure frictions, 890.42 with provider network / medical access frictions, and 891.16 with time and hassle cost measures included. Except for the model that incorporates total medical expenditure frictions, all incremental models have gamble interpretations for the average consumer that lie outside the 95% confidence interval for that estimated in the baseline model with inertia. Moreover, likelihood ratio tests of the incremental models relative to the baseline model with inertia reject the null hypothesis that each of the incremental models is equivalent to that model. These models, also discussed in the main text, illustrate how survey data on even one or two questions can be valuable additions to typical administrative datasets.

Table E8 studies the role of inertia in the context of information frictions. The first column restates the results from the baseline model without inertia or information frictions. The second column restates the results from the baseline model with inertia, identified by the choices made by new employees vs. existing employees. The third column presents results from the full model *without* inertia included from the first-stage estimates, while column four repeats the results from the full model with inertia. We include a discussion of this table, and the implications of the results, in the text in Section 4. The main takeaways are that (i) adding inertia to the baseline model substantially changes risk preference estimates and (ii) when imputed inertia is removed from the full model, the the choice friction estimates become much stronger and replace much of the magnitude of inertia (indicating that inertia is closely related to information frictions). This suggests the our friction measures are good proxies for inertia in our environment. The impact on specific frictions is quite interesting: excluding the first-stage inertia estimates substantially increases the impact of both plan financial knowledge measures and total medical expenditure knowledge measures, while moderately impacting other estimates. This suggests that these two frictions are the most tightly linked to inertia. Finally we note that with or without inertia, the full model has similar risk preference estimates that differ from those in the baseline models.

Figure E2 presents a histogram for an alternative one-dimensional type index to that discussed in the main text (the analogous figure for the primary type index is Figure 4 in Section 3). The alternative index gives consumer more credit if they get "hard" questions correct: specifically, it gives a consumer X points for a correct answer to an information-based question if a (1-X) proportion of the total respondents get that question correct. Thus, for getting a question that no one else gets right correct one gets 1 point, while if everyone else gets the question correct you get 0 points. The two indices are similarly skewed towards uninformed consumers while both have some meaningful mass of informed consumer. Table 5 in the main text includes estimates from a model that includes this alternative type index, along with measures of time and hassle costs. As with the primary type model estimates, there is a monotone relationship between level of information as represented by the index score and consumer valuation of the HDHP plan.

Tables E9 and E10 present two sets of "placebo" models designed as robustness checks to verify that our primary conclusions about the impact of including friction measures on risk preferences are not artifacts of the model setup. I.e., we use placebo models to verify that adding variables that *should be meaningless* don't impact risk preference estimates in any systematic way. The results in these two tables support our framework: adding meaningless placebo variables has little to no impact on risk preference estimates. We use three placebo measures: the first is a random number associated with an employee's actual building. This enters as an actual number that can be related to plan valuation directly: if these numbers are not related to health plan choices and valuations, this variable should not impact risk preference estimates. The additional two placebo measures are (i) a high-level measure of the division of the firm the employee works in (5 such divisions for over 50,000 employees) and (ii) a completely random number. Table E9 presents results when each of these placebo variables is added to the baseline model, without friction measures, while Table E10 presents the results when the placebo variables are added to the full model with all friction measures. Relative to the baseline models, the placebo variables have small coefficients, don't markedly impact risk preference estimates, and actually have negative likelihood ratio test statistics values relative to the baseline model suggesting that these variables add no explanatory power (this reflects estimation uncertainty, in theory, this number should only be positive). The same general conclusions hold for the placebo models relative to the full model, though including these extra variables introduces some estimation uncertainty / difficulties that lead to noisy results.

The remainder of the tables in this appendix present bootstrapped standard errors for all models estimated and discussed in the main text. See Appendix C for a detailed discussion of how standard errors are computed. Here, we present 95% confidence intervals using the block bootstrapped method discussed in that Appendix. We note, as discussed there, that these confidence intervals should be interpreted as bounds on the actual 95% confidence intervals due to estimation uncertainty. For our primary estimates, we ran the estimation routine many times and found the best likelihood function values and also verified that other nearby likelihood results provided essentially identical estimates. For the standard errors, due to computational constraints, we were not able to run as many estimation runs per sub-sample, leading to additional computational uncertainty. In certain cases, this issue leads to outlier estimation runs (due to finding local maxima rather than global) so it is natural to interpret our intervals as outer bounds on the true CIs in such cases. For many of the specifications, the 95% CI is still quite tight, supporting our main results and allowing

Health Plan Characteristics **Family Tier**	PPO	HDHP
Premium	0	0
Health Savings Account (HSA)	No	Yes
HSA Subsidy	-	$3,750*
Max. HSA Contribution	-	$6,250**
Deductible	0	$3,750*
Coinsurance (IN)	0%	10%
Coinsurance (OUT)	20%	30%
Out-of-Pocket Max.	0**	$6,250*
Provider Network	Same as HDHP	Same as PPO

* Single employees (couples) have value equal to .4 (.8) of family tier
**Single employees have a maximum of of$3,100 is max. contribution
while those over 55 can contribute an extra $1,000
***For out-of-network spending, PPO has a deductible of $100 per person (up to $300)
and an out-of-pocket max. of $400 per person (up to $1200)

Table E1: This table presents key characteristics of the two primary plans offered at the firm we study. The PPO option has more comprehensive risk coverage while the HDHP option gives a lump sum payment to employees up front but has a lower degree of risk protection. The numbers in the main table are presented for the family tier (the majority of employees) though we also note the levels for single employees and couples below the main table.

meaningful conclusions to be drawn.

Table E11 presents 95% CIs for the set of baseline models, while Tables E14, E15, and E12 presents 95% CIs for all incremental models (with one friction added) and the full model. Finally, Table E13 presents 95% CIs for the two types specifications, and Table E16 presents 95% CIs for the counterfactual simulations run in Section 5. The standard errors and their implications are discussed in the relevant locations in the main text.

Variable	Deductible: Correct	Deductible: Incorrect	Deductible: Not sure	Subsidy: Correct	Subsidy: Incorrect	Subsidy: Not sure	Coinsurance: Correct	Coinsurance: Incorrect	Coinsurance: Not sure
Deductible:									
Correct	–	–	–	0.446	-0.065	-0.364	0.350	0.156	-0.418
Incorrect	–	–	–	-0.091	0.262	-0.124	0.034	0.256	-0.253
Not sure	–	–	–	-0.328	-0.162	0.435	-0.347	-0.357	0.593
Subsidy:									
Correct	0.446	-0.091	-0.328	–	–	–	0.329	0.073	-0.328
Incorrect	-0.065	0.262	-0.162	–	–	–	0.033	0.127	-0.138
Not sure	-0.364	-0.124	0.435	–	–	–	-0.333	-0.169	0.416
Coinsurance:									
Correct	0.350	0.034	-0.347	0.329	0.033	-0.333	–	–	–
Incorrect	0.156	0.256	-0.357	0.073	0.127	-0.169	–	–	–
Not sure	-0.418	-0.253	0.593	-0.328	-0.138	0.416	–	–	–
Out-of-pocket maximum:									
Correct	0.301	0.050	-0.315	0.324	0.051	-0.343	0.324	0.061	-0.313
Incorrect	0.066	0.305	-0.317	0.104	0.129	-0.200	0.121	0.198	-0.272
Not sure	-0.299	-0.295	0.520	-0.349	-0.150	0.445	-0.364	-0.214	0.480
HSA roll over:									
Correct	0.268	0.009	-0.251	0.280	0.058	-0.307	0.169	0.179	-0.293
Incorrect	-0.115	0.080	0.037	-0.127	0.050	0.079	-0.040	-0.051	0.076
Not sure	-0.229	-0.076	0.272	-0.233	-0.110	0.306	-0.171	-0.174	0.290

Variable	OOP maximum: Correct	OOP maximum: Incorrect	OOP maximum: Not sure	HSA roll over: Correct	HSA roll over: Incorrect	HSA roll over: Not sure
Deductible:						
Correct	0.301	0.066	-0.299	0.268	-0.115	-0.229
Incorrect	0.050	0.305	-0.295	0.009	0.080	-0.076
Not sure	-0.315	-0.317	0.520	-0.251	0.037	0.272
Subsidy:						
Correct	0.324	0.104	-0.349	0.280	-0.127	-0.233
Incorrect	0.051	0.129	-0.150	0.058	0.050	-0.110
Not sure	-0.343	-0.200	0.445	-0.307	0.079	0.306
Coinsurance:						
Correct	0.324	0.121	-0.364	0.169	-0.040	-0.171
Incorrect	0.061	0.198	-0.214	0.179	-0.051	-0.174
Not sure	-0.313	-0.272	0.480	-0.293	0.076	0.290
Out-of-pocket maximum:						
Correct	–	–	–	0.197	-0.078	-0.173
Incorrect	–	–	–	0.089	0.058	-0.155
Not sure	–	–	–	-0.233	0.014	0.269
HSA roll over:						
Correct	0.197	0.089	-0.233	–	–	–
Incorrect	-0.078	0.058	0.014	–	–	–
Not sure	-0.173	-0.155	0.269	–	–	–

Table E2: Correlation matrix for responses to information questions on plan financial characteristics. Question responses presented in Table 2.

.Variable	Ben. know.: Any incorr.	BK: Any 'not sure'	Time cost hrs.	TC hrs. X Accept	TC X Dislike	Prov. net.: Same	PN: HDHP bigger	PN: PPO bigger	PN: Not sure
Benefits knowledge:									
Any incorrect	–	–	0.104	0.070	0.052	0.148	0.101	0.075	-0.243
Any 'not sure'	–	–	-0.131	-0.029	-0.053	-0.177	-0.090	0.002	0.215
Time cost hrs. X prefs:									
Time cost hrs.	0.104	-0.131	–	–	–	0.027	0.072	0.156	-0.164
Time cost hrs. X Accept	0.070	-0.029	–	–	–	0.050	0.070	0.038	-0.107
Time cost hrs. X Dislike	0.052	-0.053	–	–	–	-0.022	0.032	0.132	-0.080
Provider networks:									
Same	0.148	-0.177	0.027	0.050	-0.022	–	–	–	–
HDHP network bigger	0.101	-0.090	0.072	0.070	0.032	–	–	–	–
PPO network bigger	0.075	0.002	0.156	0.038	0.132	–	–	–	–
Not sure	-0.243	0.215	-0.164	-0.107	-0.080	–	–	–	–
TME guess:									
Correct	-0.016	-0.152	0.109	0.049	0.036	0.031	0.042	0.008	-0.056
Overestimate	0.032	0.041	0.021	-0.019	0.026	0.061	-0.038	-0.037	-0.017
Underestimate	0.080	0.027	-0.028	0.038	-0.035	0.021	0.027	-0.012	-0.026
Not sure	-0.131	0.148	-0.166	-0.104	-0.048	-0.167	-0.049	0.057	0.149
Tax benefits:									
Understands	0.095	-0.090	0.035	-0.048	0.072	0.135	-0.042	0.001	-0.112
Misunderstands	-0.196	0.259	-0.092	-0.046	-0.029	-0.265	-0.026	0.082	0.217
Not sure	0.130	-0.197	0.068	0.079	-0.022	0.171	0.055	-0.083	-0.140

Variable	TME: Correct	TME: Overestimate	TME: Underestimate	TME guess: Not sure	Tax ben.: Understand	TB: Misunderstand	TB: Not sure
Benefits knowledge:							
Any incorrect	-0.016	0.032	0.080	-0.131	0.095	-0.196	0.130
Any 'not sure'	-0.152	0.041	0.027	0.148	-0.090	0.259	-0.197
Time cost hrs. X prefs:							
Time cost hrs.	0.109	0.021	-0.028	-0.166	0.035	-0.092	0.068
Time cost hrs. X Accept	0.049	-0.019	0.038	-0.104	-0.048	-0.046	0.079
Time cost hrs. X Dislike	0.036	0.026	-0.035	-0.048	0.072	-0.029	-0.022
Provider networks:							
Same	0.031	0.061	0.021	-0.167	0.135	-0.265	0.171
HDHP network bigger	0.042	-0.038	0.027	-0.049	-0.042	-0.026	0.055
PPO network bigger	0.008	-0.037	-0.012	0.057	0.001	0.082	-0.083
Not sure	-0.056	-0.017	-0.026	0.149	-0.112	0.217	-0.140
TME guess:							
Correct	–	–	–	–	0.023	-0.067	0.051
Overestimate	–	–	–	–	0.009	-0.053	0.047
Underestimate	–	–	–	–	0.007	-0.019	0.015
Not sure	–	–	–	–	-0.060	0.212	-0.171
Tax benefits:							
Understands	0.023	0.009	0.007	-0.060	–	–	–
Misunderstands	-0.067	-0.053	-0.019	0.212	–	–	–
Not sure	0.051	0.047	0.015	-0.171	–	–	–

Table E3: Correlation matrix for responses to plan financial frictions (an aggregated measure) and all other friction measures. Answers to these questions are presented in text in Tables 2 and 3.

New vs. Existing Employees	Existing Employees	New employees
No. Employees	41,361	2339
2011 PPO%	88.8	85.7
Gender (% Male)	76.4	77.5
Age		
18-29	8.6%	36.7%
30-39	41.1%	36.3%
40-49	38.1%	20.4%
50-59	10.9%	6.2%
≥60	1.3%	0.5%
Income		
Tier 1 (< $75K)	2.7%	7.1%
Tier 2 ($75K-$100K)	10.1%	28.1%
Tier 3 ($100K-$125K)	35.3%	36.3%
Tier 4 ($125K-$150K)	30.5%	20.8%
Tier 5 ($150K-$175K)	12.0%	5.5%
Tier 6 ($175K-$200K)	4.7%	1.3%
Tier 7 ($200K-$225K)	2.0%	0.4%
Tier 8 ($225K-$250K)	0.7%	0.3%
Tier 9 (> $250K)	0.8%	0.2%
Family Size		
1	23.0%	44.0%
2	19.0%	17.8%
3+	58.0%	38.2%

Table E4: This table compares employees who are new to the firm in 2011 to those present in 2011 who joined the firm prior to 2011. The distinction between new employees and existing employees is central to the identification of inertia in the models described in Section 3.

Full Sample	
Inertia Estimates	
μ_γ - Intercept	$2.01 \cdot 10^{-3}$
μ_γ - Slope, Age	$3.92 \cdot 10^{-7}$
μ_γ - Slope, Female	$5.75 \cdot 10^{-5}$
μ_γ - Slope, Income	$9.83 \cdot 10^{-7}$
Average μ_γ	$2.05 \cdot 10^{-3}$
Gamble Interpretation of Average μ_γ	305.99
σ_γ	$1.70 \cdot 10^{-3}$
σ_ϵ, HDHP	440.29
Inertia - Intercept	828.16
Inertia - Slope, Age	22.10
Inertia - Slope, Female	-36.69
Inertia - Slope, Income	-32.80
Inertia - Slope, Family size = 2	738.69
Inertia - Slope, Family size > 2	1141.11
Average Inertia	2,396
σ Inertia*	502

*The standard deviation reported of inertia in the population is based on observable heterogeneity.

Table E5: This table presents the results the of full population model used to estimate inertia. The identification for inertia in this model comes from comparing the choices made by new employees (with no default option) to those made by existing employees who do have a default option. These inertia estimates are used as inputs into the primary models with frictions, so that the friction impacts are in addition to that linked to inertia. We also estimate a model with friction measures but no inertia in Table E8 which illustrates that, when inertia is not netted out, the friction estimates increase in magnitude, indicating a tight link to inertia, though the change in risk preference estimates are robust to this modeling choice.

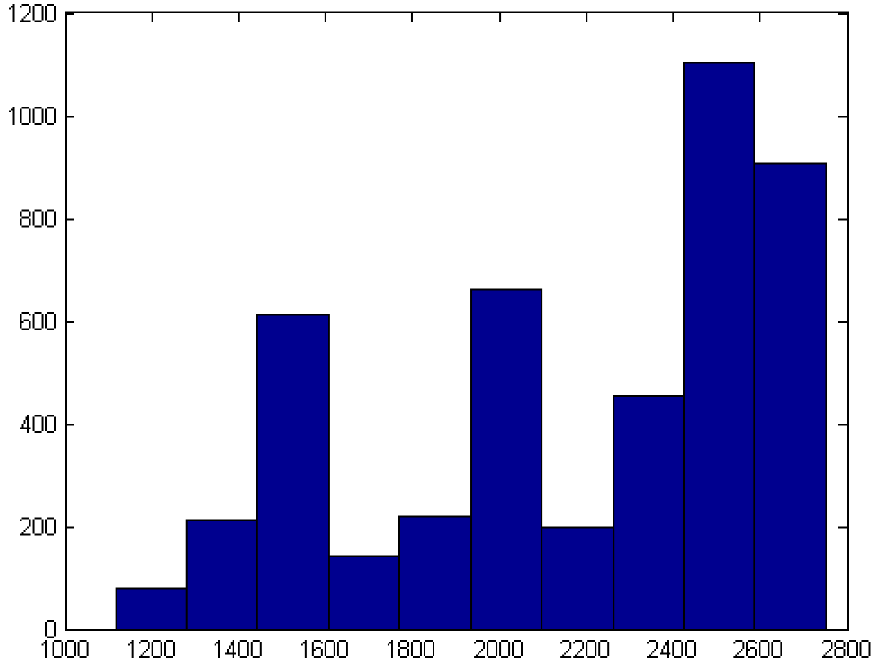

Figure E1: Histogram of Inertia Estimates from full population model, for the full population sample used in that model. Differentiation is based on observable heterogeneity, as seen in Table E5.

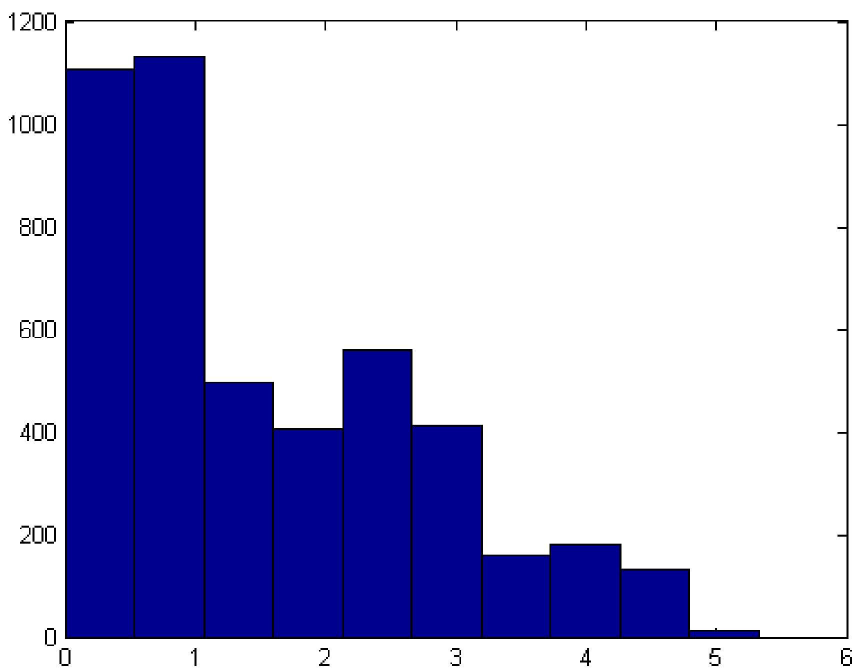

Figure E2: Histogram of weighted information type index q' for the sample of survey respondents.

Baseline Model w/ Informed Consumers			
Model	(2I) Base + Inertia (Informed)	(3) Full Model (All)	(2) Base + Inertia (All)
Average μ_γ	$8.7 \cdot 10^{-5}$	$8.6 \cdot 10^{-5}$	$2.3 \cdot 10^{-4}$
Std. Dev. μ_γ	$2.3 \cdot 10^{-5}$	$1.4 \cdot 10^{-5}$	$3.6 \cdot 10^{-5}$
Gamble Interpretation of Average μ_γ	919.15	920.47	812.61
σ_γ	$3.1 \cdot 10^{-5}$	$2.2 \cdot 10^{-9}$	$1.6 \cdot 10^{-4}$
Total Std. Dev. γ	$3.9 \cdot 10^{-5}$	$1.4 \cdot 10^{-5}$	$1.7 \cdot 10^{-4}$

Table E6: This table presents the results the of our baseline model with inertia, estimated only on the most informed patients in our primary sample. To do this, we restrict the sample to those for who our information type index $q \geq 4$ (8 is the maximum value). This restricted sample corresponds to the approximately 30% of consumers who are most informed (see Figure 4 in Section 3 in the main text for a histogram of types, and the corresponding discussion in the text in that section for a discussion of how the types are constructed). We estimate this model to assess the identification assumption that the choices of fully informed consumers identify the distribution of risk preferences in our full model. This table compares the relevant risk preference coefficients in the baseline model with inertia, for informed consumers, to those from the same model estimated for all consumers ((2) in Table 4) and those from our full models with frictions ((3) in Table 4).

Incremental Model Estimates				
Model	(7) Benefits Knowledge	(8) Time/Hassle Costs	(9) Provider Networks	(10) TME Info
Average μ_γ	$1.2 \cdot 10^{-4\dagger}$	$1.2 \cdot 10^{-4\dagger}$	$1.2 \cdot 10^{-4\dagger}$	$1.7 \cdot 10^{-4}$
Std. Dev. μ_γ	$1.8 \cdot 10^{-5\dagger}$	$1.7 \cdot 10^{-5\dagger}$	$1.9 \cdot 10^{-5\dagger}$	$2.5 \cdot 10^{-5\dagger}$
Gamble Interp. of Average μ_γ	$895.35^\dagger$	$891.16^\dagger$	$890.42^\dagger$	852.14
σ_γ	$3.9 \cdot 10^{-5\dagger}$	$4.8 \cdot 10^{-5\dagger}$	$7.4 \cdot 10^{-5}$	$1.0 \cdot 10^{-4}$
σ_ϵ, HDHP	0.40	30.77	0.52	0.11
Benefits knowledge:				
Any incorrect	-340.79	-	-	-
Any 'not sure'	-777.49^{**}	-	-	-
Time cost hrs. X prefs:				
Time cost hrs.	-	1.65	-	-
... X Accept, concerned	-	-108.70**	-	-
... X Dislike	-	-137.08^{**}	-	-
Provider networks:				
HDHP network bigger	-	-	-1015.50^{**}	-
PPO network bigger	-	-	-2485.84^{**}	-
Not sure	-	-	-547.85	-
TME guess:				
Overestimate	-	-	-	-579.50
Underestimate	-	-	-	-674.43^{**}
Not sure	-	-	-	-759.77
Average Survey Effect	-794.25^{**}	-1045.43^{**}	-604.21	-391.33
σ Survey Effect	328.40	1022.85	757.29	319.83
Likelihood Ratio Test Stat vs. (2)	64.30	228.63	96.40	44.32

Standard errors for all parameters presented in Online Appendix E.
$\dagger$ Point estimate outside of 95% CI for same parameter in model (2).
** 95% CI for parameter does not include 0.

Table E7: This table presents the estimates for the incremental models that add either a specific information friction or hassle costs to the inertial baseline model, as described in Section 3. These results are discussed in the context of our primary specifications in Section 4.

Information Frictions and Inertia **Model w/o Explicit Inertia**				
Model	(1) Baseline	(2) Baseline Inertia	(11) Full Model No Inertia	(3) Full Model Inertia
Average μ_γ	$1.60 \cdot 10^{-3}$	$2.30 \cdot 10^{-4}$	$9.36 \cdot 10^{-5}$	$8.64 \cdot 10^{-5}$
Std. Dev. μ_γ	$3.09 \cdot 10^{-4}$	$3.64 \cdot 10^{-5}$	$1.31 \cdot 10^{-5}$	$1.39 \cdot 10^{-5}$
Gamble Interpretation	366.74	812.61	914.40	920.47
σ_γ	$1.79 \cdot 10^{-3}$	$1.57 \cdot 10^{-4}$	$3.59 \cdot 10^{-9}$	$2.19 \cdot 10^{-9}$
σ_ϵ, HDHP	149.23	5.01	0.05	17.70
Benefits knowledge:				
Any incorrect	-	-	-457.59	98.04
Any 'not sure'	-	-	-1231.69	-467.48
Time cost hrs. X prefs:				
Time cost hrs.	-	-	-58.78	-9.72
... X Accept, concerned	-	-	-99.51	-118.15
... X Dislike	-	-	-95.67	-128.98
Provider networks:				
HDHP network bigger	-	-	-777.75	-594.38
PPO network bigger	-	-	-2559.40	-2362.85
Not sure	-	-	-518.66	-201.81
TME guess:				
Overestimate	-	-	-33.18	62.98
Underestimate	-	-	-563.80	-108.30
Not sure	-	-	-1067.37	-688.91
Average Survey Effect	-	-	-3356.28	-1787.40
SD Survey Effect	-	-	1707.11	1303.64
Likelihood Ratio Test Stat vs. (1)	-	1172.75	840.39	1552.29

Standard errors for all parameters presented in Online Appendix E.

Table E8: This table studies the role of inertia in the context of information frictions. The first column presents the results from the baseline model without inertia or information frictions. The second column restates the results from the baseline model with inertia, identified by the choices made by new employees vs. existing employees. The third column presents results from the full disaggregated model *without* inertia imputed from the new employee choices, while column four repeats the results from the full model with inertia. The main takeaways are that (i) adding inertia to the baseline model substantially changes risk preference estimates and (ii) when imputed inertia is removed from the full model, the the choice friction estimates become much stronger and replace much of the magnitude of inertia (indicating that inertia is closely related to information frictions). Finally we note that with or without inertia, the full model has similar risk preference estimates that differ from those in the baseline models.

Placebo Tests: Uninformative Variables Relative to Baseline				
	(1)	(12)	(13)	(14)
Model	Baseline	Placebo 1 Job Division	Placebo 2 Building ID	Placebo 3 Random Number
Average μ_γ	$1.60 \cdot 10^{-3}$	$1.87 \cdot 10^{-3}$	$1.62 \cdot 10^{-3}$	$1.55 \cdot 10^{-3}$
Std. Dev. μ_γ	$3.09 \cdot 10^{-4}$	$4.25 \cdot 10^{-4}$	$3.53 \cdot 10^{-4}$	$3.40 \cdot 10^{-4}$
Gamble Interpretation	366.74	327.72	363.97	374.56
σ_γ	$1.79 \cdot 10^{-3}$	$2.20 \cdot 10^{-3}$	$1.88 \cdot 10^{-3}$	$1.80 \cdot 10^{-3}$
σ_ϵ, HDHP	149.23	32.61	236.62	152.42
Placebo 1: Job Division*				
Group 1	-	-147.81	-	-
Group 2	-	167.36	-	-
Group 3	-	-132.37	-	-
Group 4	-	5.59	-	-
Group 5	-	-8.24	-	-
Placebo 2: Building ID*				
Group 1	-	-	-178.86	-
Group 2	-	-	-264.49	-
Group 3	-	-	-233.57	-
Placebo 3: Random Number*				
Group 1	-	-	-	-182.49
Group 2	-	-	-	-211.86
Group 3	-	-	-	-93.60
Average Survey Effect	-	-0.61	-142.87	-121.99
SD Survey Effect	-	75.45	109.42	TBD
LR Test Statistic (x) vs. (1)	-	-17.82	-15.54	-21.83

*One category is omitted for each set of placebo variables.

Table E9: This table investigates several 'placebo' models that add what should be meaningless variables to the baseline model. Column 1 repeats the baseline model results, and columns 2-4 describe the placebo model and results, which are discussed in more detail in the text of this appendix. The bottom part of the table investigates hypothesis tests to illustrate that these models are rejected against the models with survey effects. Crucially, the risk preference estimates are unchanged with the addition of placebo variables.

Placebo Tests: Uninformative Variables Relative to Full Model				
	(3)	(15)	(16)	(17)
Model	Full Model	Placebo 1 Job Division	Placebo 2 Building ID	Placebo 3 Random Number
Average μ_γ	$8.64 \cdot 10^{-5}$	$3.45 * 10^{-4}$	$1.48 * 10^{-4}$	$1.44 * 10^{-4}$
Std. Dev. μ_γ	$1.39 \cdot 10^{-5}$	TBD	TBD	TBD
Gamble Interpretation	920.47	741.93	871.13	872.05
σ_γ	$2.19 \cdot 10^{-9}$	$2.72 * 10^{-4}$	$6.49 * 10^{-5}$	$5.86 * 10^{-5}$
σ_ϵ, HDHP	17.70	583.89	551.21	620.39
Placebo 1: Job Division*				
Group 1	-	549.52	-	-
Group 2	-	878.39	-	-
Group 3	-	-117.31	-	-
Group 4	-	33.92	-	-
Group 5	-	-85.89	-	-
Placebo 2: Building ID*				
Group 1	-	-	-181.26	-
Group 2	-	-	-252.03	-
Group 3	-	-	464.19	-
Placebo 3: Random Number*				
Group 1	-	-	-	732.81
Group 2	-	-	-	618.73
Group 3	-	-	-	552.14
Average Survey Effect	-1787.40	-1789.73	-2406.10	-2141.23
SD Survey Effect	1303.64	TBD	TBD	TBD
LR Test of vs. (7)	-	-838.55	-756.99	-787.57

*One category is omitted for each set of placebo variables.

Table E10: This table investigates several 'placebo' models that add what should be meaningless variables to the full model with inertia. Column 1 repeats risk preference results from the full model, and columns 2-4 describe the placebo model and results for risk preferences and placebo effects, which are discussed in more detail in the text of this appendix. All friction coefficients are omitted here for brevity, but are available upon request. The bottom part of the table investigates hypothesis tests vs. the full model w/o placebos. The highly negative LR test statistics suggest a lot of estimation uncertainty for these placebo models relative to the full model: in theory these statistics should always be positive though with uncertainty because of complex non-linear optimization this need not be the case in practice.

95% Confidence Intervals **Baseline Models** **No Information Frictions**	(1) Baseline	(2) Baseline + Inertia
μ_γ - Intercept	$[3.09 \cdot 10^{-4}, 1.14 \cdot 10^{-2}]$	$[2.73 \cdot 10^{-4}, 6.11 \cdot 10^{-4}]$
μ_γ - Slope, Age	$[-1.38 \cdot 10^{-4}, -2.57 \cdot 10^{-6}]$	$[-7.30 \cdot 10^{-6}, -2.71 \cdot 10^{-6}]$
μ_γ - Slope, Female	$[-1.82 \cdot 10^{-3}, 1.31 \cdot 10^{-3}]$	$[-4.75 \cdot 10^{-5}, 5.06 \cdot 10^{-5}]$
μ_γ - Slope, Income	$[-1.20 \cdot 10^{-4}, 3.87 \cdot 10^{-4}]$	$[-1.31 \cdot 10^{-5}, 1.51 \cdot 10^{-5}]$
Average μ_γ	$[1.69 \cdot 10^{-4}, 7.14 \cdot 10^{-3}]$	$[1.56 \cdot 10^{-4}, 3.64 \cdot 10^{-4}]$
Std. Dev. μ_γ	$[2.64 \cdot 10^{-5}, 1.25 \cdot 10^{-3}]$	$[2.35 \cdot 10^{-7}, 6.38 \cdot 10^{-7}]$
Gamble Interpretation of Average μ_γ	[97.05,855.12]	[733.63,864.68]
σ_γ	$[8.45 \cdot 10^{-5}, 1.04 \cdot 10^{-2}]$	$[6.51 \cdot 10^{-5}, 3.12 \cdot 10^{-4}]$
σ_ϵ,	[0.00,1513.89]	[0.00,545.13]

Table E11: This table presents the 95% confidence intervals for the baseline models presented in Table 4 in the main text. The implications of these standard errors are discussed further in Section 4.

95% Confidence Intervals **Full Model** **Disaggregated** Model	(3) Full Model
Average μ_γ	[$8.19 \cdot 10^{-5}$,$2.23 \cdot 10^{-4}$]
Std. Dev. μ_γ	[$9.41 \cdot 10^{-6}$,$4.41 \cdot 10^{-5}$]
Gamble Interpretation of Average μ_γ	[822.51,924.23]
σ_γ	[$5.98 \cdot 10^{-6}$,$1.55 \cdot 10^{-4}$]
σ_ϵ, HDHP	[1.58,666.04]
Benefits knowledge:	
Any incorrect	[-614.70,377.52]
Any 'not sure'	[-1670.66,127.94]
Time cost hrs. X prefs:	
Time cost hrs.	[-90.07,118.86]
... X Accept, concerned	[-282.81,-55.79]
... X Dislike	[-293.99,-70.02]
Provider networks:	
HDHP network bigger	[-1842.45,562.52]
PPO network bigger	[-3957.68,-1286.62]
Not sure	[-937.44,303.21]
TME guess:	
Overestimate	[-810.72,704.28]
Underestimate	[-1154.63,837.19]
Not sure	[-1987.28,320.99]
Average Survey Effect	[-2148.63,-906.96]
σ Survey Effect	[1264.29,2329.12]

Table E12: This table presents the 95% confidence intervals for the full model presented in Table 4 in the text. Implications of these SEs are discussed further in Section 4.

95% Confidence Intervals Aggregated Information Types & Hassle Costs		
Model	(5) Types Unweighted	(6) Types Weighted
Average μ_γ	$[6.78 * 10^{-5}, 4.26 * 10^{-3}]$	$[6.65 * 10^{-5}, 9.02 * 10^{-5}]$
Std. Dev. μ_γ	$[7.58 * 10^{-6}, 3.74 * 10^{-4}]$	$[3.28 * 10^{-5}, 2.47 * 10^{-3}]$
Gamble Interpretation	[161.7, 936.6],	[917.25, 937.7]
σ_γ	$[3.46 * 10^{-8}, 1.1 * 10^{-2}]$	$[0, 1.72 * 10^{-5}]$
σ_ϵ, HDHP	[0.1,5146]	[0,529.56]
Unweighted Information Index*		
Lowest Quartile	[-8799,-4642]	-
Second Quartile	[-4578,-2613]	-
Third Quartile	[-1879,-625]	-
Weighted Information Index*		
Lowest Quartile	-	[-5334,-3027]
Second Quartile	-	[-3291,-1538]
Third Quartile	-	[-600,410]
Time cost hrs. X prefs:		
Time cost hrs.	[-594,155]	[-123,95]
... X Accept, concerned	[-347,-12]	[-225,-9]
... X Dislike	[-756,-55]	[-245,-27]
Average Survey Effect	[-11,705,-2166]	[-3501,-1980]
SD Survey Effect	[1948,9377]	[1482,2496]

*The omitted category is the fourth quartile, i.e. the most informed consumers.

Table E13: This table presents the 95% confidence intervals for the one-dimensional information types models presented in Table 5. The two models correspond to two different ways to construct the type index, as discussed in the main text. Implications of these SEs are discussed further in Section 4.

95% CIs		
Incremental Models		
Frictions		
	(7)	(8)
Model	Plan Design Knowledge	Time/Hassle Costs
Average μ_γ	$[8.9 \cdot 10^{-5}, 2.2 \cdot 10^{-4}]$	$[1.0 \cdot 10^{-4}, 1.8 \cdot 10^{-4}]$
Std. Dev. μ_γ	$[1.2 \cdot 10^{-5}, 4.1 \cdot 10^{-5}]$	$[1.3 \cdot 10^{-5}, 2.9 \cdot 10^{-5}]$
Gamble Int. of Average μ_γ	[821.25,918.65]	[846.37,907.10]
σ_γ	$[0, 1.56 \cdot 10^{-4}]$	$[2.90 \cdot 10^{-5}, 1.07 \cdot 10^{-4}]$
σ_ϵ, HDHP	[0.00,63.73]	[0.00,319.45]
Benefits knowledge:		
Any incorrect	[-718.23,163.86]	–
Any ‘not sure’	[-1191.18,-186.59]	–
Time costs hrs. X Prefs:		
Time cost hrs.	–	[-55.61,108.39]
... X Concerned	–	[-206.98,-35.78]
... X Dislike	–	[-246.48,-63.58]
Average Survey Effect	[-1278.65,-283.61]	[-1367.36,-601.63]
σ Survey Effect	[150.43,521.47]	[710.47,1368.95]

Table E14: This table presents the 95% confidence intervals for the incremental friction model estimates for hassle costs or knowledge of plan financial characteristics, presented in Table E7 in this Online Appendix. Their implications are discussed further in Section 4.

95% CIs		
Incremental Models		
Frictions		
	(9)	(10)
Model	Provider Networks	TME Info
Average μ_γ	$[1.8 * 10^{-4}, 6.9 * 10^{-3}]$	$[1.8 * 10^{-4}, 8.0 * 10^{-3}]$
Std. Dev. μ_γ	$[2.8 * 10^{-5}, 2.8 * 10^{-3}]$	$[3.3 * 10^{-5}, 2.5 * 10^{-3}]$
Gamble Int. of Average μ_γ	[100.97,849.64]	[88.13,845.48]
σ_γ	$[9.16 * 10^{-5}, 1.12 * 10^{-2}]$	$[1.14 * 10^{-4}, 1.02 * 10^{-2}]$
σ_ϵ, HDHP	[0,2953]	$[23.8, 4226.6]$
Provider networks:		
HDHP network bigger	[-1188.24,-82.59]	–
PPO network bigger	[-2865.64,-969.25]	–
Not sure	[-1232.05,887.46]	–
TME guess:		
Overestimate	–	[-978,579]
Underestimate	–	[-1082,-63]
Not sure	–	[-926,33]
Average Survey Effect	[-928.99,234.27]	[-566.28,39.48]
σ Survey Effect	[450.05,1052.35]	[144.40,533.26]

Table E15: This table presents the 95% confidence intervals for the incremental friction model estimates for provider network knowledge or total medical expenditure knowledge, presented in Table E7 in this Online Appendix. Their implications are discussed further in Section 4.

95% CIs
Forced HDHP Enrollment
Welfare Analysis

Model	Mean Welfare Impact Point Estimate	Mean Welfare Impact 95% CI
Baseline model, no inertia	-1237.61	[-1400.70, -807.29]
Baseline model	-874.46	[-970.04, -807.36]
Full model	-788.94	[-923.46 , -695.02]
Risk neutral	-726.09	NA

Table E16: The table presents the 95% CIs for the mean consumer welfare impact of the menu design counterfactual considered in Section 5. See Table 6 for the primary results / discussion in the text.

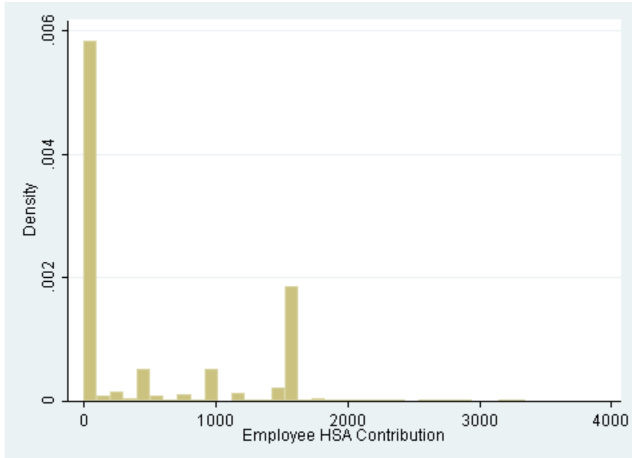

Figure F1: Histogram of HSA contributions by single HDHP enrollees in 2011.

F Online Appendix: Model for Incremental HSA Contributions

This appendix describes how we model incremental employee contributions to their health savings accounts (HSA). The primary model, described in Section 3, incorporates these estimated incremental HSA contributions as inputs into the fixed value / premium that consumers get with the HDHP plan (consumers who enroll in the PPO cannot enroll in or derive value from an HSA).

The primary reason for why we model HSA contributions, rather than use the exact values contributed by each employee (which we observe), is that we need to model the counterfactual contributions PPO enrollees would make if they enrolled in the HDHP. To this end, we train a model of contribution choice based on 2011 HDHP enrollees' actual contributions, and use this model to predict what PPO enrollees might have contributed, were they to enroll in the HDHP. We do the same for actual HDHP enrollees to maintain consistency.

Figures F1, F2, and F3 present the distributions of actual HDP enrollee contributions in 2011, for single employees, employees with one dependent, and employees with more than one dependent respectively. The figures reveal that the distribution of contributions is quite bimodal: either employees choose to forgo contributions altogether, or contribute near the maximum, with very few in between. We note that, for 2013, when all employees were forced to enroll in the HDHP, approximately 60% of employees make positive incremental HSA contributions, a similar proportion to what we see for HDHP enrollees in 2011.

Given their bimodal nature, we model HSA contributions as a two-stage choice. In the first stage, the employee decides whether or not to contribute. Then, if they do decide to contribute, they choose a non-zero amount, which in our model depends on their observable demographics. To estimate the parameters of this model, we first run a probit regression on the decision of 2011 HDHP enrollees to contribute a non-zero amount to their HSA, based on age, gender, income, and family size. We also include a dummy for whether or not their age is above 55, as employees older than that were allowed to contribute an extra "catch-up" $1000 above the normal contribution maximum. We then take those who actually did contribute a nonzero amount, and run a linear regression on their contribution, based on these same demographics. Since employees in different tiers have different maximum contributions, and different incentives to contribute, we run three separate regressions for each coverage tier (single, with spouse, family). The estimates from this model are presented in Table F1.

Based on these estimates, to simulate contributions when estimating the choice model, we

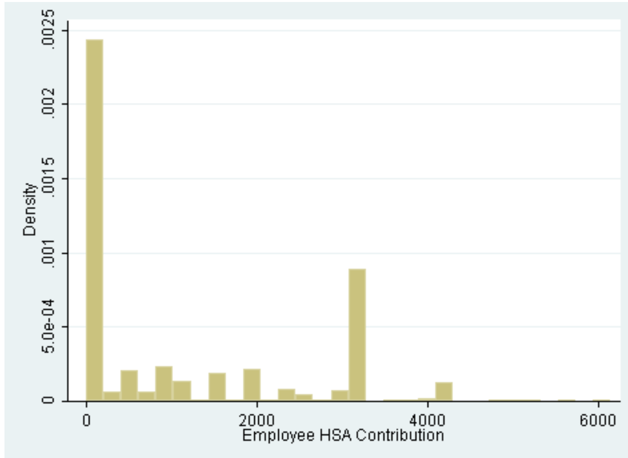

Figure F2: Histogram of HSA contributions by employees with one dependent who enroll in the HDHP in 2011.

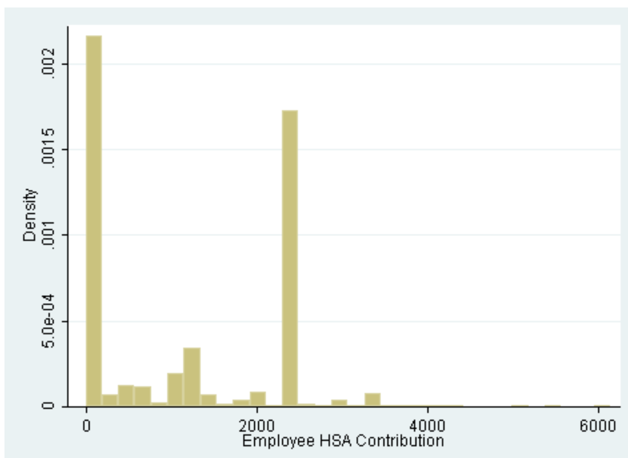

Figure F3: Histogram of HSA contributions by employees with more than one dependent who enroll in the HDHP in 2011.

generate a family-specific probability of an HSA contribution based on the first stage. We then draw a Bernoulli random variable with this probability for each family, which determines whether or not they contribute. For those who contribute, their contribution is given by the coefficients coming from second stage associated with their family tier. This output is HSA_k^C is Section 3. Then, the tax benefits from these contributions are obtained by multiplying this contribution by the marginal tax rate τ_k facing the employee, which depends on their observed income level.

Model	First Stage	Second Stage	Second Stage	Second Stage
		Tier 1	Tier 2	Tier 3
	Probit	OLS	OLS	OLS
Dep. Variable:	HSA > 0	HSA cont.	HSA cont.	HSA cont.
Age	-0.036	11.390	21.823	7.593
Age ≥ 55	0.212	439.081	544.984	852.849
Female	0.117	47.305	12.694	19.571
Income	-0.108	68.604	169.556	104.814
Family Tier 2	-0.094	-	-	-
Family Tier 3	-0.062	-	-	-
Intercept	1.279	612.064	849.499	1289.726

Table F1: This table presents the coefficients from the model predicting incremental consumer HSA contributions.

References

Bertrand, M., E. Duflo, and S. Mullainathan, "How Much Should We Trust Differences-In-Differences Estimates?," *Quarterly Journal of Economics*, 2004, *119* (1), 249–275.

Handel, B., "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *American Economic Review*, 2013, *103* (7), 2643–2682.

Train, K., *Discrete Choice Methods with Simulation, 2nd ed.*, Cambridge University Press, 2009.