C Additional Modeling Notes

C.1 Robustness of Dynamic Model

C.1.1 Symmetry of Loss Function

The loss function presented in Equation 4 of the text indicates a symmetric loss function in the distance between m_{it} and the health shock λ_{it}. This may not be desirable to the extent that individuals would prefer to over-spend, rather than under-spend, relative to a health shock; in essence, the costs of a "false positive" health visit are much smaller than the potential utility loss of a "false negative," where spending falls short of true needs.

However, careful examination of the loss function shows that whenever $c'(\cdot) \leq 1$, so that the marginal OOP cost of an additional dollar of medical spending is less than 1, the utility function in Equation (4) always leads individuals to prefer over-spending to under-spending; this is consistent with its utilization in previous studies of moral hazard. To see this formally, consider a deviation $\delta \geq 0$ from the health shock λ_{it} and notice:

$$u(\lambda_{it} + \delta) = \delta - \frac{1}{2\omega_i}\delta^2 - c_{it}(\lambda + \delta) \tag{1}$$

$$u(\lambda_{it} - \delta) = -\delta - \frac{1}{2\omega_i}\delta^2 - c_{it}(\lambda - \delta). \tag{2}$$

In this case, under-spending is strictly preferred to over-spending if and only if

$$u(\lambda_{it} - \delta) > u(\lambda_{it} + \delta) \Leftrightarrow c(\lambda_{it} + \delta) - c(\lambda_{it} - \delta) > 2\delta, \tag{3}$$

implying that the OOP costs associated with an increase in spending of 2δ must be greater than the change in total spending (of 2δ). This is impossible in all regions of a health insurance contract. Hence, the utility function preserves the desired asymmetry in which individuals tend to over-spend when their OOP costs make the first-best choice $m^*_{it} = \lambda_{it}$ infeasible (note that this is also visible in Equation 5 of the text).

C.1.2 Robustness Checks

I also show that the central model parameters (in particular, the parameters governing belief updating, $\vec{\pi}$) are robust to various modeling choices. Table 1 shows robustness to allowing $\gamma(\cdot)$ to vary more flexibly, and allowing arbitrary correlations ρ across transitory and chronic health shocks.

Table 1. Robustness of Structural Parameters to Modeling Assumptions

		Primary Model	Flexible $\gamma(\cdot)$	Correlations ρ
Panel A: Initial Beliefs				
μ_{p_0}	Prior Mean	1.433	1.671	1.418
		[0.805, 2.062]	[1.043, 2.299]	[0.537, 2.299]
$\sigma^2_{p_0}$	Prior Variance	2.389	2.170	2.369
		[2.129, 2.649]	[1.910, 2.430]	[1.735, 3.003]
ρ	$\text{Corr}(\lambda^{\text{TR}}, \lambda^{\text{CH}})$	—	—	-0.056
		—	—	[-0.062, -0.050]
Panel B: Learning from Preventive Care Investments				
μ_s	Signal Mean	0.023	0.023	0.023
σ^2_s	Signal Variance	0.982	0.914	1.282
		[0.881, 1.082]	[0.814, 1.015]	[1.180, 1.384]
γ	Average Prevention Returns	0.705	0.724	0.705
		[0.648, 0.762]	[0.699, 0.750]	[0.680, 0.730]
Panel C: Learning from Major Health Events				
π_1	Family Chronic Event	9.152	11.720	11.401
		[8.935, 9.360]	[11.511, 11.928]	[11.202, 11.600]
π_2	Own Acute Event	3.497	3.865	3.840
		[3.167, 3.829]	[3.533, 4.196]	[2.388, 5.292]
π_3	Family Acute Event	1.023	2.010	0.807
		[0.979, 1.078]	[1.960, 2.060]	[0.609, 1.005]

Notes: Table presents estimated equilibrium parameters of the model estimated via GMM on a sample of $N = 387,216$ enrollees observed in $149,938$ households between 2006 and 2013. All average parameters are expressed in terms of probabilities (%), while variances are expressed in log-odds. Signal mean μ_s is not estimated via GMM, but rather through individual-level risk predictions.

C.2 Static Model with Plan Choice

In this section, I describe an alternative structural estimation to that proposed in Section 4 of the text. The static model presented here allows households to respond to new health risk information, but also to respond to changes in spot prices of care (moral hazard) and changes in risk aversion (salience effects). Additionally, this model incorporates a stage where households choose their insurance plan. The results presented in this appendix are qualitatively similar to those presented in the main text; hence, some estimation details and results have been ommitted. Full details are available upon request.

Consider a household f comprised of individuals $i \in \mathcal{I}_f$. Individuals belong to one of two types—those without chronic illnesses and those with at least one chronic condition. I assume state-dependent preferences, so that the utility of receiving medical care differs across these types. Households and individuals are characterized by three main variables: individual beliefs about health risks (p_{ift}), household risk aversion (ψ_{ft}), and the distributions of their health shocks (as discussed in the text).

Families make two choices during each period. First, families choose their insurance coverage. Following this, both transient and chronic health shocks are realized, just as in Section 4 of the main text.[1]. Finally, individuals choose their yearly health spending. These choices are static, in the sense that both households choose plans and individuals make spending decisions on the basis of the current period's utility and type parameters only (including their beliefs about health risks). The model is static, in the sense that household decisions in period t do not affect outcomes in period $t+1$. I can therefore ignore forward-looking behavior.[2] However, individual and household type parameters—including beliefs and risk aversion—are responsive to exogenous shocks, including major health events. These parameters adjust at the end of each model period, following individual utilization choices. I model the evolution of these parameters using a Bayesian framework.

C.2.1 Utilization Choice

After choosing a health plan $j \in \mathcal{J}$ and realizing health shocks $(\vec{\lambda}_{ift}, m_{ft}^{\text{CH}})$, individuals choose spending on non-chronic medical care, m_{ift}^*. Individuals then choose m_{ift} in order to maximize their expected utility over states:

$$m_{ift}^* \equiv \text{argmax}_{m_{ift}} EU(m_{ift}; p_{ift}) = p_{ift} u_{ift,\text{C}} + (1 - p_{ift}) u_{ift,\text{H}}, \tag{4}$$

[1]Note that in this model, I restrict the demand elasticity parameter ω in my model to be homogeneous across individuals and periods, as I am incorporating additional individual-level heterogeneity in underlying health distributions.

[2]Households are, however, forward-looking within a period, as they anticipate second-stage outcomes as part of their first-stage choices. See equation 8.

where $u_{it,\mathrm{C}}$ and $u_{it,\mathrm{H}}$ represent individual utilities when diagnosed with a chronic illness and when not diagnosed, respectively. Note that Equation 4 nests the case where an individual has already been diagnosed with a chronic illness, in which case $p_{ift} = 1$. State-dependent utility functions are similar to those described in the text: individuals without chronic conditions face the typical utility function:

$$u_{ift,\mathrm{H}}(m_{ift}; \lambda_{ift}, m_{ft}^{\mathrm{CH}}, j) = (m_{ift} - \lambda_{ift}) - \frac{1}{2\omega}(m_{ift} - \lambda_{ift})^2 - c_j(m_{ift}). \tag{5}$$

Here, $c_j(m_{ift})$ represents the OOP costs associated with spending m_{ift}, conditional on the choice of plan j. Hence, individuals choose medical spending to approximately match their acute health realization λ_{ift}, accommodating the associated OOP costs of that spending.

On the other hand, individuals in the state of chronic illness face a utility function that depends on both acute and chronic health shocks, with potentially differing preference parameters. Their utility, which depends on the same model primitives as Equation 5, is given by:

$$u_{ift,\mathrm{C}} = (\alpha_1 m_{ift} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{ift}) - \frac{1}{2\omega}(\alpha_1 m_{ift} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{ift})^2 - c_j(m_{ift}). \tag{6}$$

In this state, utility is derived from both chronic and non-chronic medical spending, each of which is potentially valued at a different rate than non-chronic medical spending for healthy individuals as indicated by the parameters (α_1, α_2).

Solving the expected-utility maximization problem is the same as discuss in the text:

$$m_{ift}^* = \frac{1}{1 + p_{ift}(\alpha_1 - 1)} \left(\lambda_{ift} + \omega(1 + p_{ift}(\alpha_1 - 1) - c_j'(m_{ift}; m_{ft}^{\mathrm{CH}})) - p_{ift}\alpha_2 m_{ft}^{\mathrm{CH}}\right). \tag{7}$$

The interpretation of Equation 7 elucidates the key insights associated with this state-dependent utility framework with separate chronic care costs. In this expansion of the model, individuals choose to consume less non-chronic health care as chronic care costs increase in value, either by increases in magnitude, marginal utility, or likelihood. Equation 7 also highlights the ways that chronic care costs affect spending decisions through prices (moral hazard).

C.2.2 Plan Choice

In the first stage of the model, households choose an insurance plan to maximize their *ex-ante* expected utilities without knowing their realization of health shocks. This expected utility depends on the distributions of both health shocks as well as a household risk aversion

parameter, which depends flexibly on household demographics and is allowed to evolve over time to capture the salience effects associated with health events, as discussed in Section C.2.4. The household expected utility function for a given plan j is therefore:

$$U_{fjt} = -\sum_{i \in \mathcal{I}_f} \left[\int \int \frac{1}{\psi_{ft}(x_{ft})} \exp\{-\psi_{ft}(x_{ft}) u^*_{ift}\} \, dF_{\lambda_i} dG_{m^{\text{CH}}} \right] - c_j(m^{\text{CH}}_{ft}) - \pi_{fj} - \eta \mathbb{1}_{fj,t-1}, \quad (8)$$

where u^*_{ift} represents the optimal payoff to individual i in period t given the realization of acute and chronic health states. In addition to each individual's realized OOP costs for non-chronic medical spending, households face OOP costs for chronic care represented by $c_j(m^{\text{CH}}_{ft})$, plan premiums π_j and perceived monetary costs η for switching plans.

C.2.3 Parameter Updating

After households and individuals have made their plan and spending choices, type parameters evolve in response to health events. Of particular interest is the way that individuals update their beliefs about their unknown transition probability (p_{ift}). Additionally, households update their risk aversion parameters (ψ_{ft}) according to an adaptive framework.

I model individual learning about health risks as a Bayesian updating process in response to health events. In particular, I assume that initial beliefs depend on individual demographics, including age, sex, health risk scores, and the presence of any pre-existing conditions within the household. Prior beliefs are based on a signal x_{if0}, which is assumed to be normally distributed with mean and variance parameters $(\mu_{pi0}, \sigma^2_{pi0})$; this signal is mapped into a probability $p_{if0} \in [0, 1]$ using the standard logistic function. The center of the distribution μ_{pi0} varies with individual demographics and is potentially correlated with other household type parameters.

Major health events provide individuals with signals y_{ift} about the underlying distribution of p_{ift}, I likewise assume that these signals are normally distributed, so that the mean and variance of an individual's posterior distribution has a closed-form solution in each period. I assume that households update their beliefs *conditional* on a health event occurring. Once the individual begins evaluating their health risk beliefs (e.g., after a diagnosis has occurred within the household), they do so in a completely standard way, including updating beliefs in all following years without major health events.

C.2.4 Estimation

The unit of observation is a family f comprised of a set of individuals $\mathcal{I}_f$ in year t. Each family faces a choice of plans that varies at the firm-year-state level.[3] Households are characterized by their unobserved type variables $\{p_{ift}, \lambda_{ift}, \psi_{ft}\}_{i \in \mathcal{I}_f}$. I allow the initial parameters $(p_{if0}, \lambda_i ft, \psi_{f0})$ to be arbitrarily correlated, and link them to observable data by assuming that they are drawn from a multivariate normal distribution which depends on observed demographics:

$$\begin{bmatrix} p_{if0} \\ \mu_{\lambda if} \\ \log(\psi_{f0}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \beta_p \boldsymbol{X}^p \\ \beta_\lambda \boldsymbol{X}^\lambda \\ \beta_\psi \boldsymbol{X}^\psi \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & & \\ \sigma_{p,\lambda} & \sigma_\mu^2 & \\ \sigma_{p,\psi} & \sigma_{\lambda,\psi} & \sigma_\psi^2 \end{bmatrix} \right). \tag{9}$$

Covariates $\boldsymbol{X}$ include age, sex, health risk score, family size, and the presence of pre-existing conditions in a household. In practice, I use individuals' first year of data in $\boldsymbol{X}^p$ and $\boldsymbol{X}^\lambda$ and within-individual averages in $\boldsymbol{X}^\psi$.

Individual beliefs evolve in response to signals about their health risks as discussed in section C.2.3. I assume that these signals y_{ift} are normally distributed with variance σ_π^2 (to be estimated) and a mean given by the logit regression:

$$y_{ift} = \pi_1 \mathbb{1}\{\text{chronic}\}_{f,-i} + \pi_2 \mathbb{1}\{\text{acute}\}_{f,-i} + \pi_3 \mathbb{1}\{\text{acute}\}_{f,i} + \pi_4 x_{ift}, \tag{10}$$

where *chronic* and *acute* indicate the occurrence of chronic or acute health events within a household and x_{ift} is a variable for the number of years that have passed since the earliest major health event in the family. Hence, π_1 is the main parameter of interest, identifying the effect of a household chronic diagnosis on individual beliefs, as in the main text.

Acute health shocks at the individual level are summarized by three parameters: $(\mu_{\lambda if}, \sigma^2_{\lambda if}, \kappa_{if})$. Both $\sigma^2_{\lambda if}$ and κ_{if} are estimated as a linear projection on individual covariates.

Finally, I allow family risk aversion ψ_{ft} to evolve over time as discussed above. In particular, $\psi_{ft}(x_t)$ evolves linearly according to:

$$\psi_{ft} = \gamma_0 \psi_{f,t-1} + \gamma_1 \text{Post}_t + \gamma_2 \left\{\text{Post}_t \times m_{f0}^{\text{CH}}\right\} + \gamma_3 \left\{\text{Post}_t \times c_j(m_{f0}^{\text{CH}})\right\} + \gamma_4 \left\{\text{Post}_t \times \text{Hosp}_{f0}\right\} + \zeta_{ft}, \tag{11}$$

where m_{f0}^{CH} represents the billed spending associated with the diagnostic event, $c_j(m_{f0}^{\text{CH}})$ the OOP spending of the diagnostic event, and Hosp_{f0} indicates whether a hospitalization occurred as part of the diagnosis. I assume that $\zeta_{ft} \sim \mathcal{N}(0, \sigma_\psi^2)$.

[3] I ignore plans that have less than five percent of the overall firm-year market share in my data to avoid including executive health plans in employee choice sets.

I denote the parameters of the model by θ. These parameters include the main parameters of interest $\vec{\pi}$ and $\vec{\psi}$, including the variances σ_π^2 and σ_ψ^2. Additional parameters included in the estimation are the utility parameters $\alpha_1, \alpha_2, \omega$, and η; the five vectors of mean shifters $(\beta_p, \beta_\psi, \beta_\lambda, \beta_{\sigma_\lambda}, \beta_\kappa)$; seven variance and covariance parameters $(\sigma_p, \sigma_\mu, \sigma_\psi, \sigma_\kappa, \sigma_{p,\psi}, \sigma_{p,\mu}, \sigma_{\psi,\mu})$; and the variance of the idiosyncratic shock term σ_ε^2, which scales the choice probabilities. I assume that these idiosyncratic shocks follow the typical Type-1 Extreme Value distribution. Based on θ and the data, I am able simulate values for $p_{ift}, \mu_{\lambda if}, \sigma_{\lambda if}, \lambda_{ift}$, and ψ_{ft}.

I estimate the full static model using a maximum likelihood approach similar to Train (2009) and Revelt and Train (1998), with the appropriate extension to a discrete/continuous multi-stage choice model as discussed in Dubin and McFadden (1984). My estimation approach is similar to other models like mine, including Marone and Sabety (2021). I estimate the parameter values θ that maximize the probability density of households' observed total healthcare spending conditional on their plan choices. The estimation is done in R version 4.0.3, following the best practices laid out in Conlon and Gortmaker (2020).

My model allows for individuals to have three type-specific dimensions of unobservable heterogeneity, in addition to the typical Type 1 Extreme Value idiosyncratic shock (which can be integrated out analytically): individual health states, individual beliefs about health risks, and household risk aversion. I therefore must numerically integrate over the three dimensions $\beta_{ft} = (p_{it}, \mu_{\lambda,i}, \psi_{ft}) \in \theta$. Given a guess of θ, I use Gaussian quadrature with 27 support points (three in each dimension) to simulate underlying consumer types, yielding simulated points $\{\beta_{fts}(\theta)\}_s$ and weights W_s.

For each simulation draw s, I can then calculate the conditional density at individuals' observed total healthcare spending and the probability of households' observed plan choices.

C.2.5 Household Spending

Given data on realized choices m_{it}, I construct the distribution of healthcare spending for each individual-year implied by the model and guess of parameters θ. Based on underlying consumer types β_{fts}, I construct individual-level parameters for health states $(\mu_{\lambda,i}, \sigma_{\lambda,i}, \kappa_i)$ based on the parameters β_{fts} and the distributions outlined in Section 4.3.1 of the text.

The model predicts that given an acute-chronic health state $(\lambda_{it}, m_{ft}^{\text{CH}})$, households choose total healthcare spending m by trading off the benefit of healthcare utilization with its out-of-pocket cost, as discussed above. Given that m_{ft}^{CH} does not have individual parameters to be estimated (as these values are drawn from an empirical distribution), inverting the expression in equation 18 of the text yields the health state realization λ_{its} that would have given rise to observed spending m_{it} given m_{ft}^{CH}. Given that observed spending is truncated from below at 0, there are two possibilities for the conditional pdf:

$$f_m(m_{it}|c_{jt},\beta_{fts},\theta) = \begin{cases} \mathbf{\Phi}\left(\frac{\log(\kappa_i)-\mu_{\lambda,i}}{\sigma_{\lambda,i}}\right) & m_{it}=0 \\ \mathbf{\Phi}'\left(\frac{\log(\lambda_{its})-\mu_{\lambda,i}}{\sigma_{\lambda,i}}\right) & m_{it}>0, \end{cases} \tag{12}$$

where $\mathbf{\Phi}(\cdot)$ is the standard normal cumulative distribution function. In practice, there are iterations where the implied pdf is zero; hence, in order to rationalize the data for any parameter guess, I use a convolution of f_m with a uniform distribution over the range [-1e-75, 1e-75], as done by Marone and Sabety (2021).

C.2.6 Plan Choices

I next calculate choice probabilities for each available health insurance plan. Given θ and β_{fts}, I numerically integrate over the joint distribution of acute and chronic health care shocks using $D = 10$ support points in each dimension. The support points for the chronic health care shocks are chosen uniformly across the empirical distribution with the empirical pdf used in calculating the associated weights. For the acute health shocks, support points are calculated over the lognormal distribution as:

$$\lambda_{itsd} = \exp\left(\mu_{is} + \sigma_{is} Z_d\right) + \kappa_{is}, \tag{13}$$

where Z_d is the appropriate Gaussian quadrature vector of points (with corresponding weights W_d). The utility maximization framework discussed above (Equation 18 in the text) is then used to calculate the optimal spending levels given individual and household shocks and the underlying parameter p_{it}. Expected utility for each support point is calculated as in equation 9 of the text and summed (with weights) over all 100 points.[4] Choice probabilities for a plan j are then given by the standard logit formula

$$L_{ftjs} = \frac{\exp(U_{ftjs}/\sigma_\epsilon)}{\sum_{i\in\mathcal{J}_{ft}} \exp(U_{ftis}/\sigma_\epsilon)}. \tag{14}$$

C.2.7 Likelihood Function

Based on the choice probabilities and conditional density functions for observed spending, the likelihood function is approximated by

$$\mathcal{L}_f = \sum_{j=1}^{J} d_{fjt} \prod_{s=1}^{S} \prod_{t=1}^{T} \left[f_m(m_{it}|c_{jt},\beta_{fts},\theta) L_{ftjs}\right]^{W_s}, \tag{15}$$

[4] In practice, to speed up estimation, I ignore points with associated weights smaller than 1e-5.

where d_{fjt} is an indicator variable equal to one if household f chose plan j at time t and zero otherwise.

The log-likelihood function for family f is therefore:

$$\ell_f = \log(\mathcal{L}_f) = \sum_{j=1}^{J} d_{fjt} \sum_{s=1}^{S} W_s \sum_{t=1}^{T} \left[\log(f_m(m_{it}|c_{jt}, \beta_{fts}, \theta)) + \log(L_{ftjs})\right], \tag{16}$$

and the objective function $\mathcal{F}(\theta)$, to be maximized, is the the sum of this log-likelihood over all households:

$$\mathcal{F}(\theta) = \sum_{f=1}^{F} \ell_f. \tag{17}$$

C.2.8 Results

Table C.1 presents the estimated parameters, with additional parameters found below. I consistently find strong effects on non-diagnosed beliefs associated with household chronic diagnoses. New chronic diagnoses are associated with an average increase in an individual's belief of a major health event of 33 percentage points, an effect which is far larger than those estimated for acute events for either the individual or their family members, which are estimated to only increase risk beliefs by five and six percentage points, respectively. These increases are persistent, with little evidence that risk beliefs decrease over time (the estimated time trend coefficient is only one percentage point each year).

Table C.1 also presents parameters illustrating how the effects of new chronic illnesses alter behaviors in other meaningful ways. Major health events—both acute and chronic—are associated with strong salience effects that increase household risk aversion. On average, experiencing a major health event increases the coefficient of household risk aversion by 0.61, a 34.9% increase over the pre-diagnosis average coefficient of 1.75.[5] These effects are stronger when the household event entails either a higher amount of total billed spending or a hospitalization, suggesting that households respond differently to the intensity of an event.

[5]To put these numbers into context, I follow the results of Cohen et al. (2007) and consider the amount \$X that would make the average household in my sample indifferent between a sure payoff of \$0 and an equal-odds gamble between winning \$100 and losing \$X. Prior to a diagnosis, the average value of \$X is roughly \$85.08; after diagnosis, this value changes to \$80.85. These results are comparable with previous estimates of household risk aversion for health insurance—however, as mentioned in Einav et al. (2013), the coefficients from models incorporating both health and financial risk do not compare to those of models with pure financial risk.

		Model 1		Model 2		Model 3	
		Estimate	Std. Err.	Estimate	Std. Err.	Estimate	Std. Err.
Panel A: Dynamic Parameters							
Belief Evolution							
π_1	Family Chronic Event	0.69	(0.002)	0.17	(0.002)	0.33	(0.002)
π_2	Own Acute Event	0.07	(0.002)	0.02	(0.001)	0.05	(0.002)
π_3	Family Acute Event	0.09	(0.002)	0.03	(0.001)	0.06	(0.002)
π_4	Years since Event	-0.01	(0.000)	0.002	(0.000)	0.01	(0.000)
σ_π	Error Variance	10.29	(0.000)	0.12	(0.005)	1.52	(0.018)
Risk Aversion Evolution							
ψ_0	Persistence, Year $t-1$	–	–	–	–	0.95	(0.025)
ψ_1	Health Event (HE)	–	–	–	–	0.61	(0.015)
ψ_2	HE × Year 0 Cost	–	–	–	–	0.19	(0.020)
ψ_3	HE × Year 0 OOP	–	–	–	–	-0.88	(0.024)
ψ_4	HE × Hospitalization	–	–	–	–	1.51	(0.033)
σ_ψ	Error Variance	–	–	–	–	0.01	(0.016)
Panel B: Heterogeneity in Types							
σ^2_ε	Idiosyncratic Shock	5.92	(1.006)	6.24	(0.109)	3.56	(0.085)
σ^2_p	Initial Beliefs	16.59	(0.410)	24.43	(0.003)	14.51	(0.001)
σ^2_ψ	Initial Risk Aversion	15.22	(0.289)	5.55	(0.005)	2.57	(0.005)
σ^2_λ	Acute Shocks	–	–	0.58	(0.004)	2.03	(0.001)
$\rho_{p,\psi}$		-0.87	(0.360)	-0.43	(0.002)	-0.54	(0.002)
$\rho_{p,\lambda}$		–	–	-0.91	(0.006)	0.38	(0.002)
$\rho_{\psi,\lambda}$		–	–	0.12	(0.002)	0.09	(0.002)
Beliefs Evolve		Yes		Yes		Yes	
Acute Shock Heterogeneity				Yes		Yes	
Risk Aversion Evolves						Yes	

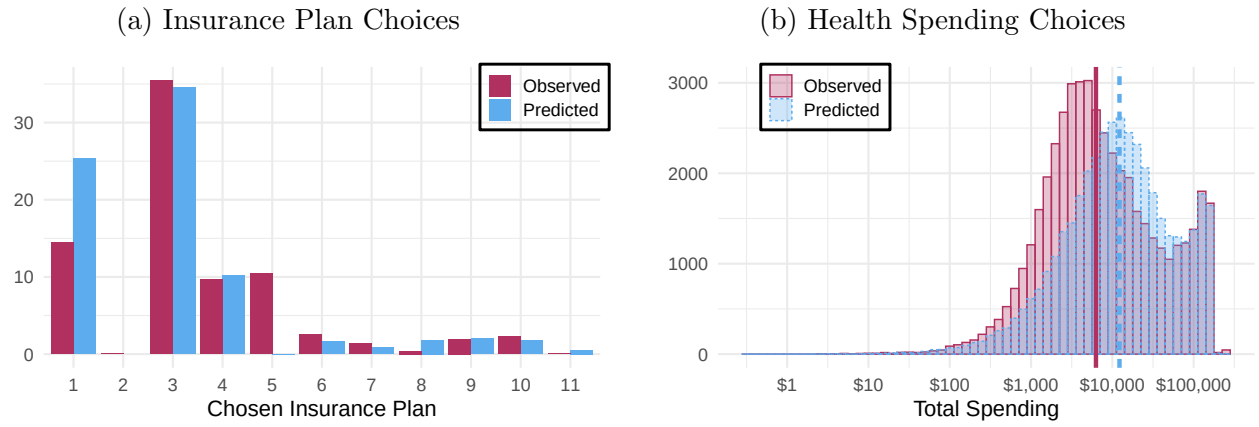
Notes: This table presents estimates for selected parameters of the structural model of health choice. Belief evolution parameters $\vec{\pi}$ are reported as marginal effects. Standard errors are derived from the analytical Hessian of the likelihood function. Column 3 presents my primary estimates used in later calculations. All models are estimated on an unbalanced panel of 179,044 households over eight years. Preference coefficients are relative to thousands of dollars.

Table C.1. Estimated Structural Parameters of Interest

C.2.9 Model Fit

I evaluate the fit of my estimated model at both the plan choice and spending stages. To evaluate plan choices, I compare plan choices for households observed in the data with those predicted by the model in Figure 1. Predicted choice probabilities are influenced by premiums, inertia, and household expectations of their acute and chronic health shocks, valued based on their level of risk aversion. At the level of household spending, I compare observed household spending distributions to those predicted by the model. As spending decisions are made after the realization of two random variables (acute and chronic health shocks), I base the model predictions off of a single draw of these underlying variables. I pool all individuals within a firm across years.

Figure 1. Predicted and Observed Insurance Plan and Health Care Spending Choices



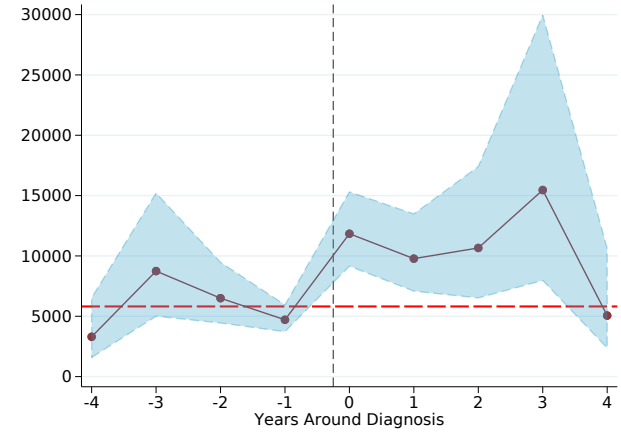
Notes: Figures show overall match between estimated model predictions and observed household choices, at both the plan choice (left) and spending (right) stages of the model. In the first panel, market shares for each insurance plan offered to employees of the single largest firm are shown. All years are pooled, so each observation is a household-year. The overall match rate is 82.2%. The second panel plots distributions of predicted and observed household health care spending, conditional on predicted/observed spending greater than zero (the observed rate of zero spending is 16.6% and the predicted rate is 13.2%). All years are pooled, so an observation is a household-year. Vertical lines represent the mean of the respective distribution.

Figure 1 presents the results. The first panel shows the observed and predicted market shares for enrollment in plans offered in the largest firm in my sample. Overall, predicted shares are closely matched. The panel on the right presents observed and estimated spending conditional on a plan choice. Here, the model predicts slightly higher levels of billed spending than are typically observed, with a difference of about $1,000 between the means of the two distributions. The model appropriately predicts the extensive margin of spending, appropriately capturing the fraction of individuals who choose zero medical spending in a given year.

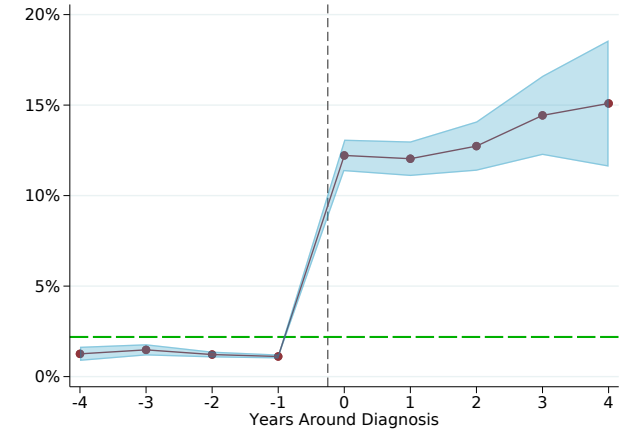
Figure 2 illustrates the model's predictions surrounding behavior following new chronic diagnoses in a household as recentered time series graphs, similar to those reported in the text. In my model, household diagnoses are associated large increases in OOP spending (about 20%, a difference which is statistically indistinguishable from the 10% reported earlier). Importantly, I predict large accompanying changes in individual health risk beliefs following a new chronic diagnosis in the family. The horizontal green line in the Panel (b) of Figure 2 depicts the pooled average risk of diagnosis within my sample, which is roughly 2.5%. Prior to health events, individuals tend to underweight their health risks by about 58%; however, following a diagnosis, individuals move to *over-weighting* their risks by over *six* times the true in-sample rates of diagnosis. Instead, these households make choices as though they perceived their risk of a chronic diagnosis to be greater than one in ten.

Figure 2. Model Predictions: Non-Diagnosed Spending and Beliefs Around a New Diagnosis

(a) Effect on Non-Diagnosed OOP Spending



(b) Effect on Health Risk Beliefs



Notes: Figures show recentered time series for model predictions of spending and beliefs for non-diagnosed household members who have experienced a diagnosis with a new chronic illness in the household. The first panel illustrates percentage changes in the inverse hyperbolic sine of OOP spending, measured in 2020 USD. The second panel illustrates estimated changes in predicted beliefs, averaged over draws from individual posterior distributions. The green horizontal line in Panel (b) illustrates the average in-sample rate of diagnosis with a new chronic condition, roughly 2.5%.