An Ounce of Prevention or a Pound of Cure? The Value of Health Risk Information*

Alex Hoagland†

July 13, 2021

Preliminary and Incomplete.
Please do not cite or circulate.

Abstract

Health care consumers frequently fall short of choices that are objectively utility-maximizing, a fact typically attributed to insufficient information. However, even as access to reliable health information is increasing, individuals may continue to misinterpret that information. I assess how health risk information disclosed through health shocks may lead households to transition from under-weighting to over-weighting their risks. This tradeoff between inattention and salience in individual risk assessment has not been well-studied, and generates new distortions in choices and perpetuates the use of low-value care. I use household major medical events as risk signals that generate informational spillovers within the family. I show that these events generate strong spending responses of about 10% annually, and that they are consistent with households reevaluating their health risks. However, these responses conflate increased investment in both high- and low-return services. To evaluate welfare effects, I write and estimate a structural model of health choices in which individuals learn about their health risks over time. The model suggests that consumers over-respond to health information, resulting in welfare losses averaging $2,788 from the new information. Placing bounds on how consumers update their beliefs in response to diagnoses improves welfare for 80% of households and suggests the revelation of health risk information can be optimally targeted on household demographics to improve social welfare gains.

Keywords: Health care choice, learning, behavioral health economics, discrete choice models, chronic illness
JEL codes: I12, I13, D83, D91, D12

*I am grateful to my advisors, Randall Ellis, Tal Gross, Marc Rysman, and Jihye Jeon, as well as Kristen Blair, Corinne Andriola, Jason Abaluck, Marshall Drake, Keith Ericson, James Feigenbaum, Sarah Gordon, Timothy Layton, Patrick Power, and Paul Shafer for very helpful comments and support. I am also grateful to participants in sessions at the ASHEcon and iHEA 2021 conferences, as well as those in numerous Boston University seminars for useful feedback.

†Boston University, Department of Economics, 270 Bay State Road, Boston, MA, 02215. Email: alcobe@bu.edu. Website: alex-hoagland.github.io.

1 Introduction

Consumers face many difficulties in optimizing their health care behaviors, whether choosing a health insurance plan (Abaluck and Gruber, 2011, 2016a; Handel, 2013) or selecting the appropriate use of medical care (Baicker et al., 2015; Iizuka et al., 2021). Regulators and policy experts typically point to a lack of consumer information—of health risks, costs of care, or the health care system more generally—as a fundamental contributor to these sub-optimal decisions (Kenkel, 1991; Sloan et al., 2003). Even as access to high-frequency health information is on the rise in developed countries (Handel and Kolstad, 2017; Krummel, 2019), questions remain about the extent to which increasing consumer *access* to health information, rather than improving the *interpretation* of health information, improves behaviors and, ultimately, welfare.

While health information can take on many forms, one of its most salient dimensions is knowledge about one's own health risk, including both current and expected future health care needs.[1] Signals associated with this type-specific information are proliferating rapidly in developed countries, both inside brick-and-mortar health facilities through increased access to genetic screening and preventive care,[2] and beyond the doctor's office, including through the expansion of workplace wellness programs (Jones et al., 2019; Song and Baicker, 2021, 2019) and the availability of health data devices (Handel and Kolstad, 2017). However, there is a lack of clear connection both between new risk information and health *behaviors* (Kim et al., 2019; Dupas, 2011) and between increased utilization and improved health *outcomes* (Alalouf et al., 2019; Iizuka et al., 2021). Without a clearer picture of what risk information consumers choose to respond to and how they internalize it, simply increasing the frequency of health signals may reduce consumer welfare either by proliferating the utilization of low-value care or targetting the wrong population altogether. Addressing the ambiguity of these welfare effects is important when addressing policy questions such as genetic testing (Posey and Thistle, 2021), health information campaigns, (Cairncross et al., 2005), and insurance design more generally (Pauly and Blavin, 2008).

[1] Previous work has identified and explored the effects of multiple other dimensions of consumer health information. These include information about providers such as hospital or surgeon "report cards" (Dranove and Sfekas, 2008; Kolstad, 2013), how to receive medical care (Margolis et al., 2014), or learning about the structures and organizations making up the health care system (Sorensen, 2006). In addition, although this paper focuses on health information available to consumers, similar discussions have taken place in expanding physicians' information, for example through the use of electronic health records (DesRoches et al., 2010).

[2] Many developed countries have recently prioritized access to preventive care. In the United States, the Affordable Care Act (ACA) requires nearly all types of insurance coverage to provide many preventive services at no cost (Strokoff and Grossman, 2010; Shafer et al., 2021); additionally, Medicare has covered annual wellness visits without cost-sharing since 2011 (Chung et al., 2015). Other nations have implemented similar nationwide screening programs, including the United Kingdom (Dalton and Soljak, 2012), South Korea (Lee and Lee, 2010), Japan (Kohro et al., 2008), and Austria (Hackl et al., 2015).

In this paper, I examine how privately-insured consumers who receive these health risk signals through witnessing a major health event—such as a new chronic diagnosis—within their household modify their health care choices. I show that affected individuals alter their behaviors in ways consistent with learning about risk rather than other potential mechanisms, suggesting the presence of informational spillovers arising from these events. When an individual is diagnosed with a chronic condition, non-diagnosed household members significantly and persistently increase their own health care utilization. The magnitude of this increase is constant over various forms of insurance plan designs, including plans without deductibles, suggesting that financial or moral hazard concerns are not driving changes.[3] These household members further increase their use of preventive care, particularly services that are specific to the particular illness affecting the diagnosed household member. Additionally, I assess the possibility that salience effects may be driving these changes (Dalton et al., 2020; Fadlon and Nielsen, 2019) by comparing my results to responses to acute major health events—such as hospitalizations—for which no health risk information is transmitted. I find that chronic events induce stronger and more persistent behavior changes, suggesting that individuals are responding to more than only the trauma associated with a health event. Finally, to rule out other forms of learning, including systematic learning about health care organizations or building physician relationships (Sabety, 2020), I show that chronic diagnoses within a household engender increased adherence to existing preventive prescriptions, such as statins and other cardiovascular medications.

These observed changes in individual behavior are reconcilable with a model of health capital where agents have imperfect knowledge about their health risk, and alter their behaviors as signals improve their knowledge (Bundorf et al., 2021a; Grossman, 1972). However, I demonstrate that these changes in behavior may not be inherently welfare improving. In particular, I observe that affected household members increase their use of "low-value" health services, services that are generally agreed to be cost ineffective due either to their reach (e.g., benefiting few patients) or their average returns (e.g., low levels of benefits relative to costs) (Colla et al., 2015). Households responding to chronic diagnoses are most likely to increase their utilization of low-value screenings and imaging services, such as cardiac screenings before low-risk surgeries or imaging services for headaches or lower back pain. In addition, while households noticeably change their medical spending decisions following a chronic event, they do not meaningfully alter their insurance plan choices. Both of these cast doubt on the extent to which new health information ultimately improves choice quality.

[3]As is common in the health economics literature, I use the phrase "moral hazard" to denote induced-demand effects arising from changes in the price an individual faces for care, effects which have been explored in-depth through previous work (Guo and Zhang, 2019; Kowalski, 2016). For a more in-depth discussion of this abuse of notation, see Einav et al. (2013).

These findings motivate a structural approach to model the evolution of household decisions following health events and quantify the associated welfare effects of receiving health information. I write and estimate a structural model in which households make health insurance coverage and spending decisions while learning about their health risks (Cardon and Hendel, 2001; Einav et al., 2013; Marone and Sabety, 2020). In addition to providing consumers with health risk information, major health events affect consumer choices by altering the conditional cost of non-chronic care and household risk aversion. My model therefore allows me to separate the welfare effects of receiving health information separate from these potentially confounding effects.

Households typically possess multiple dimensions of private information (Finkelstein et al., 2009), and major health events may evoke simultaneous changes across these dimensions that confound my ability to attribute changes in observed behavior to changes in risk beliefs. Hence, a key identification challenge in my model is inferring changes in individual beliefs about their health risks separate from other type parameters. To address this concern, I first exploit variation in the availability and generosity of plans offered to households to separately identify changes in household risk aversion at the time of plan choice. Here, the intuition is that individual beliefs about risk determine optimal medical spending and coverage levels, while household risk aversion also determines the gradient of preferred coverage as the price or generosity of plans varies (Ericson et al., 2020). I complement this approach with additional information about the circumstances of a diagnosis (e.g., whether a hospitalization occurred) to further model risk preferences and risk beliefs separately. Second, I use a large set of major medical events to leverage variation in the relative riskiness and costliness of a diagnosis, which allows me to isolate the causal effect of new *risk* information on beliefs, rather than more general systematic learning. Finally, the level of needed follow-up care associated with a chronic illness separately identifies the price effects from a new diagnosis within the household.

My analysis suggests that for many households, information from health events may not be welfare-improving. Observed consumer choices are rationalized by large changes to individual beliefs following health risk information; however, these changes do not result in utility gains when compared to scenarios in which the information is not revealed. Households would be willing to pay an average of about $2,788 per household per year in order to *avoid* the health risk information communicated by a major medical event. The central insight here is that there is a tension between the seriousness of a major health event and the extent to which individuals update their beliefs about their risk: new diagnoses in a household spur large changes in an individual's assessment of their health risks, resulting in average posterior beliefs that are well above the average in-sample risk of diagnosis.

Counterfactual simulations suggest that bounding these changes in risk beliefs substantially increases consumer welfare. 80% of the households in my sample would find health information welfare-improving were their responses to such information mitigated. Finally, I explore the substantial heterogeneity associated with the returns to new information to show that the social gains from revealing this information can be improved by targeting its revelation based on household demographics, such as *ex-ante* risk. This has important implications for important health policy proposals, such as genetic testing and other important screenings, such as for Covid-19 (Oster et al., 2013).

My analysis contributes to a growing literature estimating learning and preferences in structural models of health behavior (Barseghyan et al., 2018; Bundorf et al., 2021a). My model makes use of previous identification results to simultaneously estimate weighted probabilities and standard risk aversion parameters in a nonlinear framework (Ericson et al., 2020).[4] Additionally, I incorporate findings from the behavioral economics literature that highlight the role of mis-weighted probabilities and overconfidence in models of consumer choices (Moore and Healy, 2008; Tiefenbeck et al., 2016; Friehe and Pannenberg, 2019), including their role in rationalizing choices that would otherwise require unreasonably high levels of risk aversion (Paserman, 2008; Spinnewijn, 2015). My model also provides insight into the development of subjective beliefs; in particular, my model illustrates that individual responses to health information may help to explain why consumers are better at predicting their relative risk rather than their absolute risk (Bundorf et al., 2021b) and may be biased when assessing their own health risks (Arni et al., 2021). I extend previous models by separately identifying the value of the rich information communicated by a health event and highlighting the particular behaviors—such as information misinterpretation—that dampen potential gains. Finally, I highlight the role that differences in modeling choices play in the generation of economically-meaningful parameter estimates.

I also contribute to a literature assessing the role of health risk information in improving decisions and welfare. A rich literature has highlighted how individuals respond to information about their own health risks, including their own diagnosis.[5] I join a newer discussion on the spillover effects of information, including for family members responding to health events (Fadlon and Nielsen, 2019; Song, 2021; Bouckaert et al., 2020) and community members responding to community-level infectious disease outbreaks (Agüero and Beleche, 2017). Major health events have also been used to infer the demand effects for health care within

[4]See Barseghyan et al. (2013) and their later review paper Barseghyan et al. (2018) for a more thorough discussion of the literature estimating models of probability weighting in other settings in economics.

[5]For an in-depth review of this literature, see Alalouf et al. (2019). Some previous work has demonstrated that certain diagnoses can have dramatic impacts (Almond et al., 2010; Oster et al., 2013); however, examinations of other diagnoses revealed a lack of noticeable responses (Kim et al., 2019; Dupas, 2011).

households following an effective change in spot prices (Eichner, 1997; Kowalski, 2016). I incorporate the findings of this literature into the first dynamic structural model addressing this question, which accommodates both learning and salience effects. In addition, I expand the scope of analysis to a wide array of health events that affect many families across the United States every year.

Finally, my work is relevant to the well-established literature exploring suboptimal health decisions made by most consumers (Abaluck and Gruber, 2011, 2016a; Abaluck and Compiani, 2020; Ketcham et al., 2012; Handel, 2013; Handel and Kolstad, 2015). This literature includes an ongoing discussion about the extent to which improving health information generally may improve decision-making (Abaluck and Gruber, 2016b; Gruber et al., 2020; Cutler and Zeckhauser, 2004). My analysis reveals that some health signals—such as major health events—do little to align household choices with the value of medical care, and may instead lead to an increase in the over-utilization of services that provide little or no benefit to households. Hence, simply improving access to health information may shift consumers only from one type of poor decision-making to another, while increasing total health spending. Additionally, my paper underscores the role of behavioral economics in structural models assessing the quality of consumer choices. I show that including factors such as belief discounting may help to explain why overcoming information frictions is not simply a matter of increased access to health information.[6]

I present the empirical setting and my data in Section 2. Following a discussion of major health events, I provide preliminary evidence documenting the responses these events generate and evaluating the potential mechanisms driving them in Section 3. Then, to evaluate the welfare effects associated with these responses, I present the details of my model in Section 4 and estimate it in Section 5. The model output informs several counterfactual analyses assessing the role of consumer responsiveness to information, which I present in Section 6. Finally, I discuss the relevance of my findings and directions for future work in Section 7.

2 Empirical Setting & Data

My primary data on household plan choice, health utilization, and major medical events come from the IBM/Truven Marketscan *Commercial Claims and Encounters* Data®. These data contain detailed inpatient, outpatient, and pharmaceutical claims for a sample of households who are actively enrolled in employer-sponsored insurance (ESI) through large U.S.

[6]In doing so, this study sheds additional light on the value of preventive health care for individuals who are at-risk of developing a chronic condition (Mehrotra and Prochazka, 2015; Goroll, 2015; Rubin, 2019).

firms who have contracted with participating payers. Each observation includes diagnostic, procedural, and payment information, as well as household, firm, and insurance plan identifiers. I obtained data from 2006 to 2018, with the exception of plan identifiers which are only available until 2013. Throughout, spending data has been normalized to 2020 USD using the CPI-U series. This study was approved as exempt by the Boston University institutional review board.

My final sample includes households with two or more members observed for two or more years, each of which is employed by (and insured with) one of eight large firms. I required that each household have full eligibility and continuous enrollment across their window of observation. My final sample consists of 353,403 households and 5,439,482 individual-year observations.

Table 1 presents summary statistics for the full sample as well as the subset of the sample with insurance plan identifiers. In general, the two groups have similar demographics, spending trends, and health states. A notable exception is that households in the plan-identified sample incur lower OOP costs than the full sample, suggesting that they possess more generous insurance coverage on average. However, this is likely due to time trends rather than substantive differences. Medical spending, as expected, is highly skewed, with average annual household spending in the range of $2,500 compared to a median of about $400. Observed switches in plan choices is low, as documented in the literature on plan choice (Handel, 2013).

2.1 Plan characteristics.

Heterogeneity in each household's choice of plans provides a plausibly exogenous source of variation in how major medical events and chronic health costs impact household spending decisions. I exploit the claims data to estimate the characteristics of each plan in my households' choice sets, which will be important inputs in my theoretical model.

I define a household's plan choice set at the firm-state-year level, and limit attention to plans covering at least five percent of all covered lives within a firm-year to rule out executive plans. In reality, health plans are defined by a complicated set of cost-sharing measures, including copayment and coinsurance rates that vary widely across provider specializations, networks, and procedures. For tractability, my structural model takes in a simplified version of these measures: a family deductible, a simplified non-specialist coinsurance rate, and a family OOP maximum. I construct measures for each plan's individual and family deductibles based on the empirical distribution of payments in the claims data (Zhang et al., 2018). I then estimate the other two cost-sharing parameters as those that minimize the

Table 1. Household Summary Statistics

	Full Sample	Plan-Identified Sample
Family size	3.0 (0.00)	3.0 (0.00)
Employee age	45.0 (0.01)	44.4 (0.01)
Enrollee age	30.9 (0.01)	30.4 (0.01)
% female employees	41.6 (0.00)	42.4 (0.00)
% female enrollees	50.2 (0.00)	50.3 (0.00)
Total medical spending	\$2,504.41 [\$679.75] (4.51)	\$2,454.88 [\$624.16] (7.12)
OOP medical spending	\$443.07 [\$109.66] (0.53)	\$337.98 [\$80.33] (0.89)
% individuals w/ zero spending	15.4 (0.00)	16.6 (0.00)
% individuals w/ zero OOP	21.0 (0.00)	22.2 (0.00)
% switching plans	–	5.3 (0.00)
% experiencing chronic diagnosis	6.3 (0.00)	5.2 (0.00)
% experiencing acute event	1.0 (0.00)	0.6 (0.00)
Diagnosis OOP, chronic illnesses	\$1,082.05 [\$464.69] (11.59)	\$854.62 [\$329.90] (17.72)
Diagnosis OOP, acute events	\$2,494.42 [\$1,419.91] (68.05)	\$2,107.09 [\$964.62] (122.50)
Years	2006–2018	2006–2013
N_{families}	353,403	179,044
$N_{\text{individuals}}$	1,087,353	555,733

Notes: Values based on Marketscan claims data, 2006–2018. Enrollees are employees plus their covered dependents. Spending values are reported in 2020 USD. Standard errors are reported in parentheses and sample medians (when reported) are in brackets.

sum of squared residuals between predicted and observed OOP cost for households within each plan year (Marone and Sabety, 2020). Appendix A.1 describes this methodology in more detail and evaluates the quality of these inferences. Finally, I estimate each plan-year's family premium as the average cost of all households enrolled in the plan over a year, and assume that employee premium contributions are consistent with the national averages for household coverage (Foundation, 2020).

There is substantial variation across firms, regions, and years in the generosity of coverage offered to employees, which I describe in Table 2. The average firm offers enrollees between 2 and 4 plans in a given year, with a wide degree of variation in the average family deductible. This variation is constituted by heterogeneity in the frequency with which firms offer zero-deductible health plans as well as in the size of nonzero deductibles. Similar variations exist in other plan characteristics, including copayment rates and OOP maxima.

	Firm							
	A	B	C	D	E	F	G	H
# of plans offered	3.5	2.5	3.0	2.0	2.0	2.6	2.8	3.0
Family premium ($000s)	12.7	9.8	9.7	10.2	9.3	8.9	9.1	11.5
Family deductible ($000s)	0.4	0.4	2.1	1.0	1.0	0.7	0.9	0.5
% of 0-deductible plans	64.3	46.7	0.0	0.0	0.0	22.2	31.8	38.9
Family OOP max. ($000s)	3.5	4.6	5.1	5.9	4.3	4.1	5.2	3.9
HHI of all plans	0.4	0.6	0.4	0.6	0.9	0.6	0.7	0.4

Notes: Averages are pooled across all plans and years in a given firm.

Table 2. Average Plan Characteristics, 2006–2013

2.2 Major medical events and costs.

I model the ways households respond to information about their health risk communicated through major health events within the family. I identify these events based on observed diagnostic codes in the claims data, using a subset of the Department of Health and Human Services' Hierarchical Condition Categories (HCCs). These HCCs, which are typically used in risk adjustment models, identify a basic set of chronic illnesses that may alter overall health utilization and spending. I limit my classification of health events to non-pregnancy HCCs that occur with relatively consistent frequencies, as discussed in Appendix A.2.

To ensure that I identify new diagnoses, I require that relevant diagnosis codes appear during or after an individual's second observed year. Additionally, I drop households for which the diagnosed individual is not present for at least a full year after their medical event to exclude individuals who might have passed away during or shortly after their event.

An important feature of my analysis is the separate treatment of health costs for major medical events, including the costs associated with maintaining the health of someone with a chronic condition. To measure these costs, I collaborated with a physician to identify a set of disease-specific procedures and prescriptions associated with each health condition in my sample. I then identify household spending on these health events based on the claims for these procedures and prescriptions, both in the year of diagnosis and following years. Appendix A.3 lists the relevant codes used for each diagnosis.

3 Spillover Effects of Household Health Events

This section presents my main reduced-form empirical results. I first show that households increase their overall medical utilization by about 10% annually, as well as increasing their investment in billed spending on preventive care. I demonstrate that these responses are largest among those with the least financial incentives created by the major health event, suggesting that moral hazard concerns are not a main driver of my results. Finally, I illustrate that the observed responses are consistent with a reevaluation of one's own risk. I do this by showing that households are more likely to invest in preventive care that is specific to the illness their family member experienced; that affected individuals increase their adherence to preventive medications they were already taking prior to the health event; and that household members increase their utilization of "pseudo-preventive" low-value services, such as extraneous screenings and imaging services.

3.1 Induced Spending Changes

To estimate the causal impact of health shocks on health choices, I first estimate two-way fixed effects "event study" regressions of the following form:

$$\sinh^{-1}(y)_{ft} = \alpha_f + \tau_t + \sum_{k=-T}^{T} \gamma_k \mathbb{1}\left\{t - E_{ft} = k\right\} + \epsilon_{ft}. \tag{1}$$

The variable y_{ft} represents a spending outcome for a household f in year t, such as the annual amount a family paid out of pocket for medical care. I adjust for highly-skewed distributions of spending variables by using the inverse hyperbolic sine transformation.[7] An

[7] I use the inverse hyperbolic sine transformation to accommodate the high-frequency of zero-spending individuals and households in my data (Harris and Stöcker, 1998). Bellemare and Wichman (2020) show that for a model with continuous variables x and y and specification $\sinh^{-1}(y) = \beta x + \varepsilon$, the elasticity of y with respect to x is $(\beta x/y)\sqrt{y^2+1} \approx \beta x$ whenever $y \geq 2$. Bellemare and Wichman (2020) also discuss the ways using this measure may refine estimates using the more common $\log(y+1)$ transformation. I show in Appendix B.5 that my results are not substantively altered when using the logarithm transformation.

added advantage of this transformation is that the resulting regression coefficients can be interpreted as approximate percentage changes in the outcome variable, relative to the year prior to the shock. I include household and year fixed effects, as well as dummy variables indicating when an observation occurred relative to E_{ft}, a household's event year.[8] The coefficients on these indicator variables, $\{\gamma_k\}$, are the coefficients of interest. I also adjust for potentially correlated responses within a household by clustering standard errors at the household level.

This approach allows me to identify the potentially time-varying effects of health shocks—which might have decaying influence on household choices over time—while simultaneously controlling for any unobserved household- or year-specific deviations in behavior. However, recent work has highlighted that two-way fixed-effects estimators can be difficult to interpret without strong modeling assumptions (Imai and Kim, 2020). In particular, the presence of heterogeneous treatment effects may obscure interpretation of both pre-trends and estimated time-varying treatment effects (Sun and Abraham, 2020; de Chaisemartin and D'Haultfoeuille, 2019; Goodman-Bacon, 2018). In addition to the traditional event study results shown here, I present a number of additional specifications in B.5, including simple recentered time series graphs, standard difference-in-differences regression coefficients, and the robust alternative estimator proposed by de Chaisemartin and D'Haultfoeuille (2019). Taken together, these results provide a consistent picture of a strong household response to health shocks.

Figure 1 presents the time-varying causal effect of a health shock on household OOP spending for all non-diagnosed individuals. The first panel illustrates that across all plan types, non-diagnosed household members increase their annual OOP spending by about 10%, or about $50 annually, starting in the year of the health event and persisting across multiple years. The second panel limits the scope of the analysis to households enrolled in plans with no deductible at the time of the health event; among this sample, I find a similarly-sized response, with an average effect size of about 20% or $40 annually. Additional results in Appendix B.5 corroborate this finding in other outcome variables, including total reported billed spending or visit frequencies.[9]

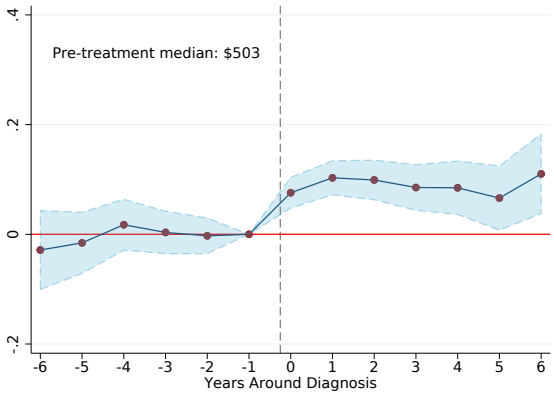
A natural response to observing the phenomenon illustrated in Figure 1, panel (a) is to conclude that the spending increase is driven by moral hazard responses among the non-

[8]Important for this approach, I utilize a large control group in my sample, allowing me to separately identify the time-varying treatment effects from yearly fixed effects (Borusyak and Jaravel, 2016).

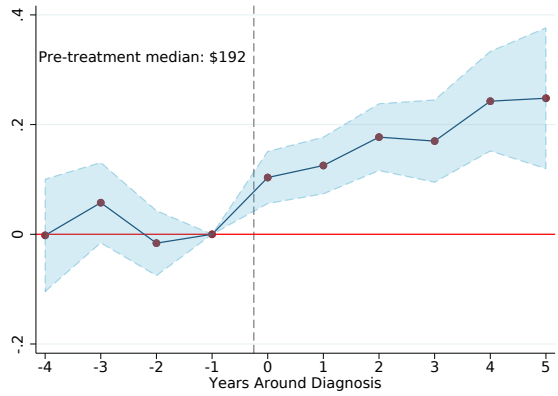
[9]Note that in my data set, the billed spending variable represents the sum of what individuals pay out of pocket for care and what the insurance plan ultimately pays to the provider; it does not reflect any price negotiations or other discounts that were provided at the time of service, and therefore reflects the final billed price of care, not the listed price.

Figure 1. Chronic Diagnoses Increase Non-Diagnosed Out-of-Pocket Spending

(a) Full sample



(b) Households facing zero deductible



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new diagnosis on medical spending for different populations. For each panel, the dependent variable is the inverse hyperbolic sine of total OOP spending for all non-diagnosed individuals in a household. The first panel includes the full sample, while the second panel limits the sample to those enrolled in plans with zero deductible at the time of the major health event. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

diagnosed individuals. Here, the intuition is that a chronic diagnosis, such as diabetes mellitus (diabetes), implies consistent and predictable costs on a household, such as through insulin prescriptions and endocrinologist visits. These additional costs, which are largely fixed for the individual, shift the cost-sharing characteristics of a health plan for the rest of the household, effectively lowering their spot price of future (non-chronic) health care. Two features of my results suggest that these induced demand responses are unlikely to be the principal driver of my results. First, the costs of a chronic diagnosis are typically larger in the year of diagnosis than in future years, especially when a hospitalization is required to diagnosis the illness or there are acute complications that must be dealt with. This would suggest that if other household members were responding to changes in care prices alone, their responses would be much larger in the period close to the diagnostic event, and more muted in following years. Panel (a) of Figure 1 does not show this to be the case. Second, panel (b) of Figure 1 illustrates that non-diagnosed individuals respond to health shocks even when those shocks do little to change their spot price of medical care. Were moral hazard responses the principal mechanism of response, households in these plans would have much weaker incentives to adjust their choices.[10] Appendix B.5 corroborates this finding, showing that families who are closer to meeting their deductibles prior to a health event are not more likely to increase their spending than those for whom chronic care costs may not

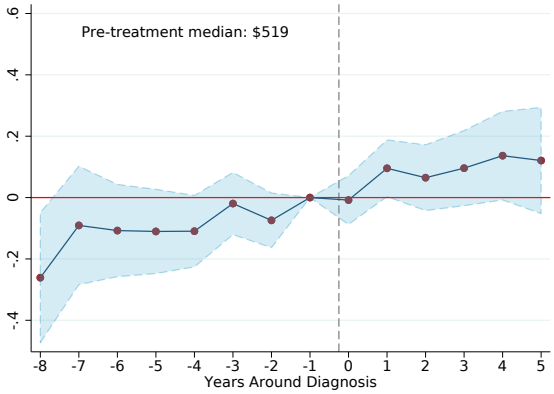
[10] A corresponding result for the subset plans with nonzero deductibles is included in Appendix B.5.

meaningfully change family cost-sharing rates.

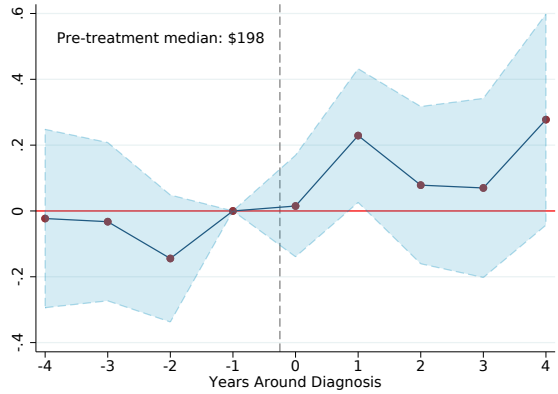
It may also be the case that the intensity of major health events realigns household preferences to prioritize medical care. That is, rather than altering individual *beliefs* about health risks, it is possible that health events alter an household's risk *preferences* by affecting their marginal utility of medical care relative to other forms of consumption. While these salience effects almost certainly play some rule in future household decisions, I find that households alter their spending patterns only in response to certain types of major health events. In particular, Figure 2 presents the results of similar regressions estimating the effect of acute hospitalizations on spending patterns, rather than the effect of new chronic diagnoses. Here, household responses to events are estimated to be both smaller and noisier; however, these hospitalizations make health care at least as salient—if not more so—than chronic diagnoses. This suggests that changes in risk preferences arising from a "health scare" are insufficient to entirely explain changes in behavior.

Figure 2. Acute Health Events Have no Effect on Non-Diagnosed OOP Spending

(a) Full sample



(b) Households facing zero deductible



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of an acute hospitalization within the household on medical spending for different populations. For each panel, the dependent variable is the inverse hyperbolic sine of total OOP spending for all non-diagnosed individuals in a household. The first panel includes the full sample, while the second panel limits the sample to those enrolled in plans with zero deductible at the time of the major health event. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

3.2 Responding to Risk Information

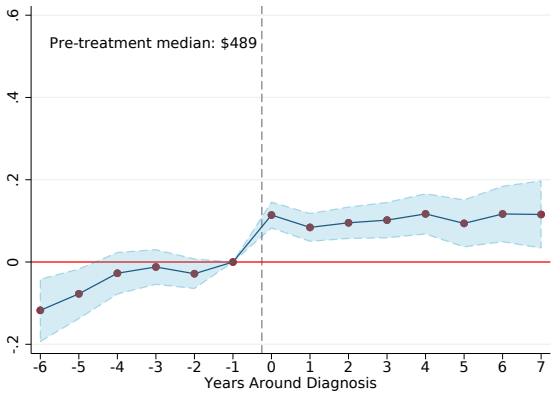
These results suggest a profound and persistent change in how non-diagnosed household members engage with the health care system. In this section, I argue that these responses are consistent with household members updating their beliefs about their own health risks

following the receipt of health information from a major event. To do this, I assess how household investments in preventive care evolve following major health events.

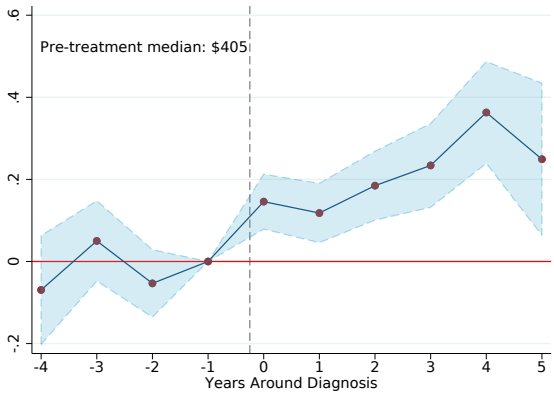
I focus first on the general use of preventive care by measuring the impact of individual diagnoses on the use of wellness visits by other household members. Wellness visits are non-problem-based visits with a family or primary care physician that are generally recommended about once a year; these visits include routine screenings for important chronic conditions including cancers and mental health conditions. These visits constitute an important jumping-off point for the use of other preventive services (Jiang et al., 2018) and are therefore generally considered to be an important form of high-value care (Tong et al., 2021).

Figure 3. Chronic Diagnoses Increase Billed Spending for Wellness Visits

(a) Full sample



(b) Households facing zero deductible



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new chronic diagnosis within the household on medical spending on wellness visits for different populations. For each panel, the dependent variable is the inverse hyperbolic sine of total spending (insurer spending + OOP spending) on wellness visits for all non-diagnosed individuals in a household. The first panel includes the full sample, while the second panel limits the sample to those enrolled in plans with zero deductible at the time of the major health event. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

Figure 3 presents regression estimates from an identical specification to the above figures, this time using overall spending on wellness visits as the outcome variable.[11] New household chronic diagnoses evoke similarly strong responses in the use of wellness visits as for overall

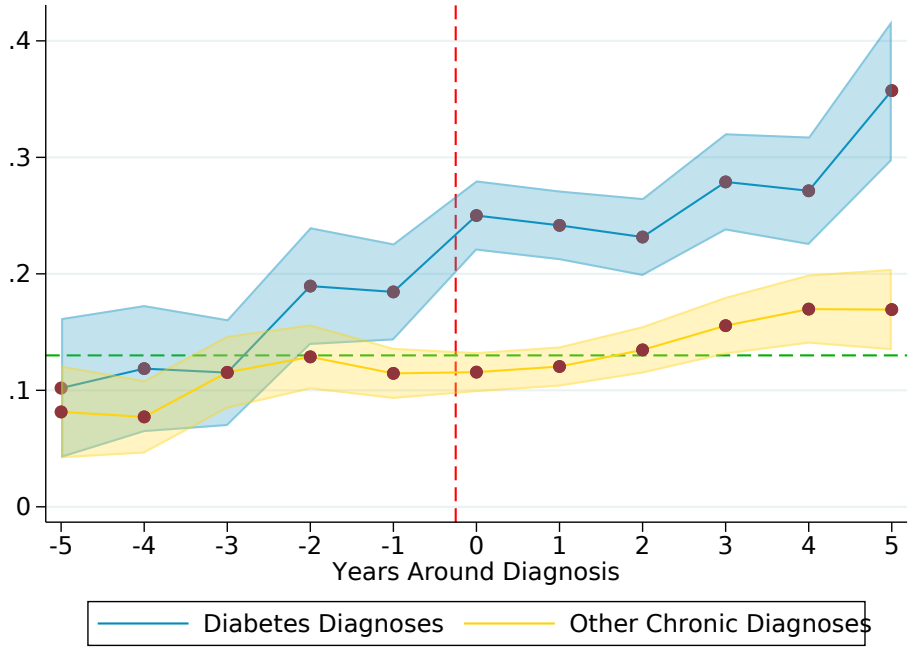
[11]Even before the ACA's cost-sharing exclusion took effect in 2010 (or 2012 for certain women's health services), OOP costs for preventive care were steadily declining for those with ESI (Hong et al., 2017). Once the ACA took effect, the majority of wellness visits for those with ESI should be free to consumers (Shafer et al., 2021), a feature I observe in the data. Year fixed effects in the regression specification should absorb these time trends for both pre- and post-ACA trends. However, to avoid confounding changes in cost-sharing rates with responses to health events, I use billed spending rather than OOP spending as my outcome variable of interest.

spending, inducing a 10% increase in overall spending on wellness visits for the full sample and even larger effects for households without deductibles.

Such observed responses could be driven by factors beyond changes in a household's assessment of their health risks, including improved physician relationships (Sabety, 2020), salience effects, or general exposure to the health care system. To more explicitly explore the link between major health events and risk beliefs, I estimate the causal effects of health shocks on preventive services that are specific to an affected household's diagnosis. Here, the intuition I rely on is that household exposure to risk information is more targeted than other forms of health information; hence, the extent to which I observe households selecting into preventive services that are disease-specific rather than general provides evidence of responses specifically to new risk information.

Figure 4. Diabetes Diagnoses are Associated with Higher Rates of Diabetes Screenings

(a) Diabetes Screenings (Rate/1,000 Children)



Notes: Figure shows average utilization rates of diabetes screenings for non-diagnosed household members 18 years of age and younger, measured in rates per 1,000 children. Point estimates and 95% confidence intervals are presented. The top (blue) line indicates average rates for households who experience a diabetes diagnosis, and the bottom (gold) line indicates rates for those affected by other chronic diagnoses. The horizontal green dashed line indicates the average utilization rate for all other households in the sample (who do not experience a diagnosis).

For example, I find that being exposed to a diabetes diagnosis in one's household drastically increases their chances of seeking out screening for diabetes. Figure 4 plots re-centered

time series that depict the associations between household diagnoses and the takeup of diabetes screenings for children within a household.[12] The figure plots average utilization rates of diabetes screenings among household members under 18 years old for two groups: those who are exposed to a diabetes diagnosis in their home and those who are exposed to a different new chronic diagnosis. The horizontal dashed line indicates the average utilization of these screenings in unaffected households. Children whose family members are diagnosed with conditions other than diabetes do not appear to significantly alter their screening behaviors; on the other hand, siblings of those diagnosed with diabetes increase screenings dramatically following the diagnosis and are over twice as likely to be screened for diabetes as other children. Rather than utilizing more preventive care as a general response to a major health event, households appear to seek investments that are specific to the health risk information such events afforded them.[13]

This analysis shows associations only; to identify the causal effect of specific diagnoses on the utilization of disease-specific preventive services, I use a triple differences approach. This approach disentangles two competing effects: effects arising from experiencing any chronic illness occurring (e.g., salience effects) and the disease-specific informational effect. I estimate the effect of a new chronic diagnosis on a household f's decision to screen for a specific diagnosis d during time t, as summarized in Equation 2:

$$\begin{aligned} Pr(\text{Screening})_{fdt} =& \beta_{\text{DD}}(\text{post}_t \times \text{chronic}_f) + \beta_{\text{DDD}}(\text{post}_t \times \text{chronic}_f \times \mathbb{1}\{\text{chronic}_f = d\}) \\ &+ \alpha_f + \tau_t + \varepsilon_{fdt}, \end{aligned} \tag{2}$$

where $chronic_f$ is a dummy variable indicating whether any chronic diagnosis occurred within the household and $post_t$ indicates all periods including and after the diagnosis. Hence, β_{DD} identifies the effect of any chronic diagnosis on screening, while the triple interaction β_{DDD} identifies the change in that effect for the specific diagnosis of interest, relative to other diagnoses.[14] For example, one such regression estimates the probability of a household

[12] I focus on the diagnoses and screenings of children to highlight the transmission of genetic risk information associated with Type 1 Diabetes Mellitus diagnoses. Similar results hold for the screenings of spouses following a Type 2 Diabetes diagnosis, which transmits health risk information more about lifestyle choices than genetic risks.

[13] Note that the figure suggests significantly different screening behaviors in the year prior to a family member's diagnosis. Although these differences are smaller, such pre-event differences could be the result of households responding to early symptoms of individuals who have not yet received a diagnosis. For example, households may choose to screen all children in a home when one is displaying diabetes-like symptoms, thereby anticipating some form of the health risk information even before the event itself (the diagnosis) takes place.

[14] The sum of the coefficients $\beta_{DD} + \beta_{DDD}$ identifies the diagnosis-specific effect of receiving a diagnosis, relative to all non-diagnosed households in my sample. Notice that, in Equation 2, all requisite interaction terms for the triple differences are either subsumed in the fixed-effects or colinear with the included variables given the unique structure of my treatment variables.

screening its members for diabetes after a diabetes diagnosis occurs within the household ($\beta_{\text{DD}} + \beta_{\text{DDD}}$) as opposed to after any other chronic diagnosis (β_{DD}) .

The triple difference approach is advantageous because it allows me to compare the causal effect of diagnoses on the use of preventive care across multiple control groups. When the outcome variable of interest is a screening for a specific service (e.g., diabetes), this approach estimates the effect of a corresponding diagnosis relative to all other diagnoses, for which the screening reveals no information. In this context, the identifying assumption for the triple differences approach is the same as the identifying assumption for the simpler difference-in-differences regressions: that spending differences between diagnosed and undiagnosed households would have evolved similarly over time in the absence of treatment.[15]

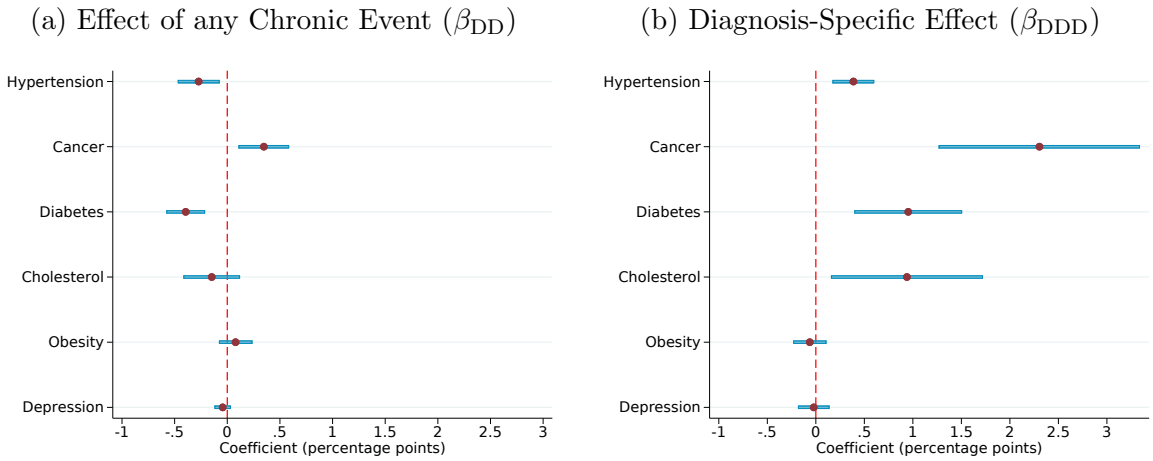
I estimate several versions of this regression for various diagnosis-screening pairs. I select diagnoses and screenings which are commonly utilized and for which there are clear diagnostic codes available.[16] I examine the impact of new diabetes, and cancer diagnoses on their respective screenings, as well as the effect of diabetes diagnoses on cholesterol screenings. I also assess the impact of any new chronic diagnosis in a household on the rate of new hypertension diagnoses, relative to all major health events. Given that there is no procedure code for hypertension screenings, this approach proxies the effect of the risk information associated with chronic diagnoses on new general wellness screenings, relative to the other forms of health information accompanying acute events.

Finally, I utilize this approach to further estimate the effect of chronic diagnoses on "placebo" screenings, including services that do not necessarily require medical expertise to identify risk. These include the effect of new diabetes diagnoses on obesity screenings and the effect of new mental health disorder diagnoses on screenings for depression. In the case of the first, medical expertise is rarely needed to identify that one is at risk for obesity (although associated counseling services are still vital for management and treatment). The case of depression screenings has a similar intuition; in addition, household members who may feel they are similarly at risk for mental health disorders may simply bypass this preventive screening and seek medical services directly from a mental health professional. Hence, in each of these cases, I would expect to observe little, if any, impact of a major diagnosis on screening behavior.

[15] When adding the triple interaction, the identifying assumption is modified only to include the assumption that spending differences between households diagnosed with one condition and households diagnosed with another would have evolved similarly in the absence of treatment, a statement which is subsumed in the initial identifying assumption. Appendix B.5 includes standard difference-in-differences regression results corroborating the findings reported here.

[16] Coding practices reduce my ability to test this finding for each individual diagnosis. For example, given that there are no well-used diagnostic or procedure codes for asthma screenings, I am unable to assess the impact of a new household asthma diagnosis on screenings.

Figure 5. Chronic Diagnoses Increase Take-up of Disease-Specific Preventive Care



Notes: The figure shows estimated coefficients and 95% confidence intervals for the triple differences specification (equation 2). Each row identifies a separate diagnosis/screening pair, listed at the far left. Outcome variables are measured as a binary variable indicating whether that screening or diagnosis occurred at least once within a year among the non-diagnosed household members. Coefficients in the first panel show the difference-in-difference coefficient representing the effect of the diagnosis compared to those who did not experience a major event. Coefficients in the second panel show the triple-difference coefficients, or the effect of the specific diagnosis relative to households exposed to other chronic diagnoses. Results for cholesterol and obesity screenings are shown for a new diabetes diagnosis. Standard errors are clustered at the household level.

Figure 5 presents the estimation results from these six regressions in two panels. The first visualizes the traditional difference-in-differences coefficients, or the general effect of any chronic diagnosis on screenings; the second panel identifies the diagnosis-specific effect from the triple differences coefficient. The first panel suggests that general effects of diagnoses on screenings are minimal. Exposed household members may be more likely to seek out cancer screenings, perhaps in response to the salience effect of a chronic diagnosis occurring; however at the same time, I find that chronic diagnoses are associated with fewer hypertension and diabetes screenings overall. These responses, in conjunction with the disease-specific effects presented in panel (b), suggest that households may trade off commonly-used preventive services (e.g., diabetes screenings) for more specific ones (e.g., cancer screenings) as they learn new information about their specific health risks.

However, I consistently find that household responses to major health events align with the particular type of event they experienced. In particular, households are most likely to engage in hypertension screenings, diabetes and cholesterol screenings, and cancer screenings in response to the related health event. However, I do not observe that even targeted diagnoses such as diabetes or mental health diagnoses impact less-specialized screenings such as obesity or depression screenings.

I also find that in addition to selecting screenings based on the health risk information

Screening	Diagnosis	Pre-Diagnosis Mean	General Effect (β_{DD})	Diagnosis Effect (β_{DDD})
Panel A: Main Effects				
Hypertension[1]	Any Chronic[2]	2.01	-0.27**	0.39***
		(0.007)	(0.102)	(0.110)
Cancer	Cancer	20.72	-0.01	2.74***
		(0.021)	(0.113)	(0.509)
Diabetes	Diabetes	6.21	-0.46***	1.31***
		(0.012)	(0.086)	(0.279)
Cholesterol	Diabetes	17.01	-0.22	1.23***
		(0.019)	(0.126)	(0.389)
Panel B: Placebo Regressions				
Obesity[1]	Diabetes	1.04	0.02	0.10
		(0.005)	(0.035)	(0.110)
Depression	Depression	0.36	-0.01	-0.08
		(0.003)	(0.037)	(0.077)

Notes: Table presents results from six triple-difference regressions highlighting the role of household investments in disease-specific preventive care following adverse health events. Each regression uses as its outcome variable a binary indicator for the screening listed in the first column, and a binary indicator for the event in the second column as its treatment variable (see Equation 2 for the full specification). Regression coefficients for the typical difference-in-difference effect (β_{DD}) indicate the effect of any chronic health event on screenings; the triple differences coefficients (β_{DDD}) indicate the effect of the specific diagnosis on screening choices. Robust standard errors clustered at the household level shown in parentheses. [1] Due to unavailability/low-use of CPT-4 procedure codes for screenings, these outcomes are measured as new ICD-9-CM/ICD-10-CM diagnosis codes. [2] Here, the reference group is all acute major health events. $^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$

Table 3. Chronic Diagnoses Increase Take-Up of Disease-Specific Preventive Care

they receive, households are selective in which members they choose to screen, choosing based on the relative importance of genetic and lifestyle factors associated with a specific diagnosis. Appendix B.5 utilizes variation in intra-familial relationships and genetic risks to show that when households are affected by a chronic illness with a strong genetic component, household members such as children and siblings of the affected individual are more likely to be screened for the same illness. On the other hand, diagnoses such as Type 2 Diabetes Mellitus, which has a stronger lifestyle component than a genetic one, are associated with more frequent screenings for spouses. Taken together, the observed ways in which major health events affect the use of preventive care are all consistent with a model where households interpret new diagnoses as signals of their own health risk, altering their behaviors accordingly.

3.3 Separating Risk Information from Other Informational Effects

The estimates above suggest that household responses to major health events correspond to a reassessment of their health risks. However, new diagnoses may also alter spending patterns by providing families with more general health information, such as information about the value of medical care, the process of obtaining covered care through an insurer, or how to establish strong provider relationships. Generally, learning about health risk and this more systematic learning imply similar responses among affected individuals, making their effects difficult to disentangle. In this section, I focus on a particular case where new diagnoses provide risk information without more systematic information: non-diagnosed household members who were taking medications to prevent cardiovascular disease *prior* to the diagnosis within their family. I show that new diagnoses alter adherence to these prescriptions, a strong indicator that health events affect choices specifically by providing information about health risks.

Cardiovascular preventive drugs, including statins and other cholesterol-lowering drugs, are an extremely commonly used class of medications, and are known to be effective in preventing future health problems when used with a high level of adherence over many years (O'Connor, 2006).[17] In this analysis, I limit my sample to those who have filled a prescription for these medications at least once per year during their first two years in the sample. I then measure the effects of chronic diagnoses on utilization and adherence among refills of these prescriptions.

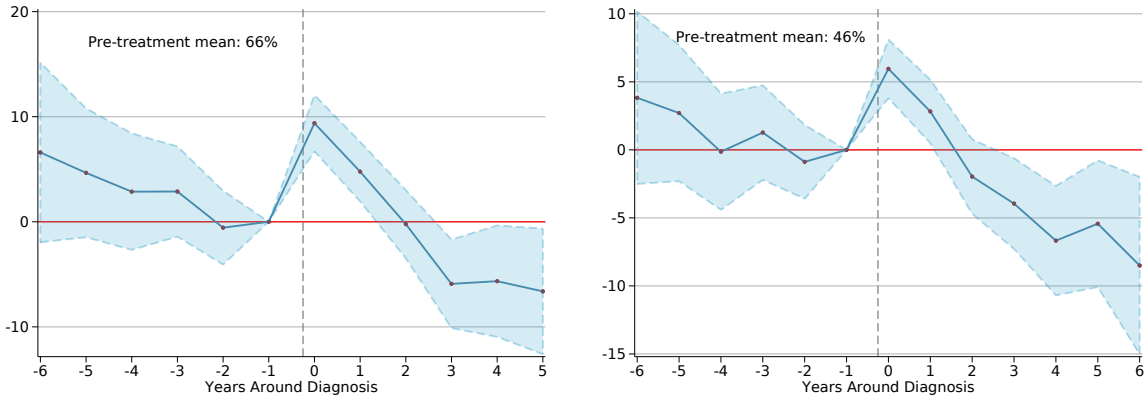
This setting provides a unique environment in which to disentangle the effects of general learning about health systems and learning about one's own health risk. Individuals on existing prescriptions already have sufficient knowledge about the health care system to

[17] Appendix Table 8 contains a detailed list of the therapeutic classes used in my sample.

receive this care from their provider and insurer. Hence, while major health events provide them with information about the potential value of adherence to their medication (along with the potential consequences for not doing so), these events are unlikely to provide new knowledge about how to obtain this medication.

Figure 6. Chronic Diagnoses Spur Re-Adherence of Existing Preventive Medications

(a) Pr(Refill Old Prescription) (b) Proportion of Days Covered



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new diagnosis on adherence to preventive medications whose prescriptions were first written prior to the major health event. The sample is limited to all non-diagnosed individuals who filled preventive cardiovascular medications at least once per year during their first two years in the sample. In the first panel, the dependent variables is a binary indicator for whether the prescription was refilled at all. The second panel uses the proportion of days covered by any preventive cardiovascular medication as the outcome variable (Choudhry et al., 2009). Standard errors are clustered at the household level.

I estimate the effect of a household medical event on both the likelihood of any use of the medication and general adherence, measured using the proportion of days covered in a year (Choudhry et al., 2009). This measure is standard in the literature on drug adherence, and corresponds to the fraction of the year after a patient's first prescription fill for which the patient has a supply of the medication. One concern in identifying the causal effect of new diagnoses on adherence is that prescription adherence may decay over time in response to barriers such as financial concerns or apathy (Slejko et al., 2014). Importantly, this decay may occur at different rates for different individuals both within and across households, meaning that such downward trends would not be accounted for using only household and year fixed effects. I therefore add controls for the number of years an individual has been in the sample to the event study specification used previously (equation 1).

Figure 6 presents the estimated dynamic treatment effect of a new chronic diagnosis on adherence to existing preventive prescriptions. As expected, individuals become less adherent to prescriptions generally. However, major medical events spur a resurgence in both the likelihood that individuals will fill their prescriptions at all and the proportion of

days covered: affected individuals are around ten percentage points more likely to refill their prescription in the year of a major medical event than in the year before, translating to an additional eight percentage point increase in the average proportion of the year for which they are covered by the prescription. Interestingly, the effects of diagnoses on adherence are not as persistent as general effects on new utilization and spending. This may simply be a feature of individuals' declining adherence over time, with health events shifting the level of overall adherence if not the trend. Alternatively, this may suggest the presence of salience effects specific to the use of preventive medications.

Overall, however, the fact that new diagnoses change individual adherence to prescriptions even among a population which has access to and knowledge of specific preventive care illustrates that individuals are learning about more than just how to obtain care. Were the effects of a new diagnosis limited only to systematic health learning, I would have expected to see muted effects among this sub-population. The estimated causal "re-adherence" to prescriptions is more consistent with individuals reevaluating the value of their medication, given new information about their own health risks.

3.4 Evaluating the Quality of Spending Responses

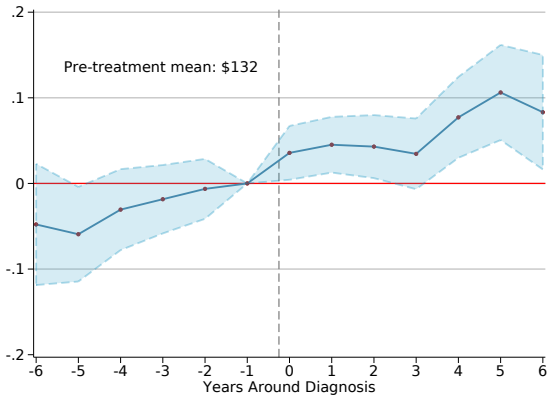
Major health events generate strong spillover effects within a household on both overall utilization patterns and preventive care investments. It is natural, therefore, to ask how these responses are distributed within a larger framework of health spending. Do major health events contribute to more informed decisions about the type of care consumers choose to utilize? Or does the salience associated with health trauma lead to further over-utilization of low-return services? I address these questions by examining household use of services typically deemed as "low-value" by medical professionals and health officials (Chua et al., 2016; Colla et al., 2015).[18] Low-value services include both services whose cost typically outweighs any benefits to an average patient (e.g., unnecessary surgeries) as well as services which are chronically over utilized in ways that dramatically lower their return (e.g., imaging services). Avoiding the use of these services can result in an overall higher quality of health care through both cost reductions and the avoidance of unnecessary risks.

Figure 7 presents estimates for the effect of new chronic diagnoses on the overall utilization of low-value services, including both total spending and overall utilization rates. Major health events are associated with a small increase in overall low-value spending of about 5 percent.

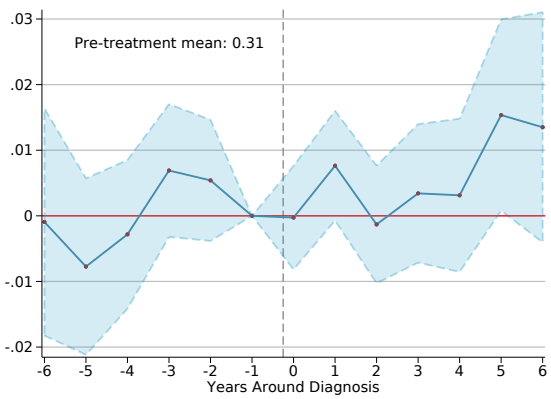
[18] These health services are based on recommendations made with the Choosing Wisely initiative, directed by the American Board of Internal Medicine (ABIM) Foundation and other physician specialty organizations (Bhatia et al., 2015; Wolfson et al., 2014). Appendix A.5 contains more detail about the specific services included in each measure.

Figure 7. Chronic Diagnoses Increase Utilization of Low-Value Care

(a) Billed Low-Value Spending



(b) Low-Value Utilization (Rate/Person)



Notes: This figure shows estimated coefficients and 95% confidence intervals for the effect of major health events on the use of low-value services (see Appendix A.5 for definitions). In the first panel, the outcome is the inverse hyperbolic sine of billed spending. In the second panel, the outcome is the number of low-value services used per household member. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

In contrast, the average rate of service use among non-diagnosed household members does not change meaningfully following a diagnosis.

However, these results mask significant heterogeneity across different types of low-value services. These services, which differ not only in their cost but also in their perceived value to each household, may be differentially affected by the ways new chronic diagnoses affect household decision making about health utilization. For example, households which receive health risk information from these diagnoses may reevaluate the value of low-value services that prioritize screenings, such as imaging services and preoperative visits. On the other hand, households that respond more to the price effects induced by a major health event may be more likely to seek out high-cost, low-return services such as elective surgeries. To explore these differences in-depth, I separate my sample of low-value services into five categories: pediatric services, including imaging services and the early use of medications such as antibiotics; adult prescription drugs, such as the use of opiates to treat migraines; unnecessary imaging services for adults, including for lower-back pain; extraneous screening services for adults, including cardiac testing before low-risk surgeries; and adult surgical procedures, such as arthroscopy for knee pain.

Table 4 presents results for both standard difference-in-difference and event study regressions for each of the five categories. New chronic diagnoses shift households spending and utilization into specific low-value service categories including screening services, pediatric care, and imaging services. The effect sizes range from an increase as large as ten percent for

Service Category	All Pediatric		Adult Drugs		Adult Imaging		Adult Screening		Adult Surgery	
Outcome Variable	Spending	Rate	Spending	Rate	Spending	Rate	Spending	Rate	Spending	Rate
DiD										
$\text{Post}_t \times$ Diagnosis$_f$	0.05***	0.02***	-0.00	-0.00	0.03***	0.01***	0.10***	0.03***	-0.10***	-0.04***
	(0.017)	(0.003)	(0.000)	(0.000)	(0.013)	(0.002)	(0.014)	(0.005)	(0.012)	(0.002)
Adjusted R^2	0.192	0.228	0.143	0.259	0.123	0.141	0.163	0.151	0.230	0.255
Event Study										
$t-4$	-0.04**	-0.02*	0.01	0.00*	0.01	-0.00	-0.10***	-0.05***	0.09***	0.03***
	(0.014)	(0.008)	(0.003)	(0.002)	(0.016)	(0.005)	(0.021)	(0.011)	(0.012)	(0.004)
$t-3$	-0.02	-0.01	0.00	0.00	-0.01	-0.01	-0.03	-0.09	0.04***	0.02***
	(0.012)	(0.007)	(0.002)	(0.001)	(0.013)	(0.004)	(0.019)	(0.010)	(0.010)	(0.003)
$t-2$	-0.01	-0.01*	0.00	0.00	0.01	0.00	-0.02	0.00	0.01	0.01**
	(0.010)	(0.005)	(0.002)	(0.001)	(0.016)	(0.004)	(0.016)	(0.010)	(0.009)	(0.002)
$t-1$	–	–	–	–	–	–	–	–	–	–
t	0.02**	0.008	0.00	0.00	0.01	0.01	0.03*	0.008	-0.03***	-0.01***
	(0.009)	(0.004)	(0.002)	(0.001)	(0.010)	(0.003)	(0.015)	(0.008)	(0.008)	(0.002)
$t+1$	0.03***	0.01***	0.00	0.00	0.03***	0.01***	0.07***	0.04***	-0.07***	-0.02***
	(0.009)	(0.005)	(0.002)	(0.001)	(0.011)	(0.003)	(0.015)	(0.008)	(0.009)	(0.003)
$t+2$	0.04***	0.02***	-0.00	0.00	0.02*	0.01**	0.06***	0.02	-0.08***	-0.03***
	(0.010)	(0.005)	(0.002)	(0.00)	(0.012)	(0.003)	(0.016)	(0.009)	(0.011)	(0.003)
$t+3$	0.05***	0.02***	-0.00	-0.00	0.03**	0.02***	0.07***	0.03**	-0.11***	-0.05***
	(0.011)	(0.006)	(0.002)	(0.001)	(0.013)	(0.004)	(0.018)	(0.011)	(0.013)	(0.005)
$t+4$	0.04***	0.02***	0.00	0.00	0.06***	0.02***	0.10***	0.03*	-0.10***	-0.05***
	(0.013)	(0.007)	(0.003)	(0.002)	(0.016)	(0.005)	(0.021)	(0.012)	(0.016)	(0.005)
Adjusted R^2	0.192	0.228	0.143	0.259	0.123	0.141	0.163	0.151	0.230	0.255
N	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161	1,538,161

Notes: Table shows estimated difference-in-difference and event study regression coefficients for the effect of a new chronic diagnosis. Two outcome variables are reported for each category: the inverse hyperbolic sine of billed spending and the number of low-value services used per household member. See Appendix A.5 for service definitions. Spending is measured in 2020 USD. Standard errors clustered at the household level are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4. Estimated Effects of Chronic Illness on Low-Value Care Utilization, by Category

low-value screenings to three percent for imaging services. Additionally, the results provide suggestive evidence that major health events provide a deterrent from low-value elective surgeries; notice, however, that the strong presence of pre-trends obfuscates the true causal effect of the diagnosis. I find no effect on the misuse of prescription drugs among adults.

These results suggest that households seek out care that they see as useful in preventing or identifying future illness, even if those services are generally accepted as being low return by health professionals. Although I observe households utilizing more of these services—such as preoperative screenings or imaging services—it is unclear whether these are decisions made at the household level or by a physician who knows the family history and hence deems these services as appropriate. What's more, there is a growing overlap of utilization of high- and low-value services, as low-value screenings become more prevalent during annual wellness visits (Ganguli et al., 2020). Regardless, the results suggest that new diagnoses prompt households to re-assess their own health risks. This provides new suggestive evidence that the utilization of low-value care may be tied more to risk beliefs rather than ignorance about the actual returns of a service, providing similar evidence to recent work (Finkelstein et al., 2021).

In addition to the utilization of low-value care, I explore other ways health events alter the quality of consumers' health care decisions, including their plan choices, in Appendix B.5. In general, I do not find that major health events prompt households to switch their health insurance plans. While new diagnoses in a household are associated with marked differences in observed spending behavior, it is still unclear whether these choices are *ex-post* more optimal for affected households.

4 Empirical Model of Belief Formation

Adverse health events—such as the onset of chronic illness—disrupt an already rocky landscape for families attempting to navigate their health care needs. These events transform a family's view of health insurance from an optional means of protecting themselves from future cost shocks to a necessary method of cost containment for a permanently increased level of needed care. Hence, families that undergo these major events experience a change in "type", from a healthy collective of individuals to a group supporting at least one relatively more expensive individual. In addition, non-diagnosed household members learn about their own risks and types from observing a family member's diagnosis.

I estimate the impact of health risk information on consumer choices as well as its implied welfare effects in a structural model of health utilization. I build on a canonical two-stage model of health spending (Cardon and Hendel, 2001; Einav et al., 2013; Marone and Sabety,

2020). In the first step, households choose an insurance plan to maximize their *ex-ante* expected utility, based on their available information about the distributions of future shocks. In the second, individuals within the household choose their spending and utilization based on the realized health shocks and the chosen plan's features.

I extend the existing model in two important ways. First, I allow consumers' types to be adaptive in response to both new information and new experiences. In my model, individuals learn about their probability of adverse health events; their degree of risk aversion may also change over time as they experience health events in the home (to accommodate any salience effects). Second, I explicitly model the differences between acute and chronic health shocks. Chronic health shocks impose recurring costs on a family, potentially altering non-chronic care prices and inducing moral hazard for other household members.

4.1 Model Primitives

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 parameters: individual beliefs about health risks (p_{it}), household risk aversion (ψ_{ft}),[19] and the distributions of their health shocks. New health events—including both new chronic diagnoses and acute hospitalizations—cause both non-diagnosed and diagnosed individuals to update their beliefs about their health risks, as well as potentially altering household risk aversion.

In each period, two types of shocks are realized. Following the typical convention, each individual has an acute health realization λ drawn from an individual-specific distribution $F_\lambda(\cdot)$. Acute health realizations model the uncertain aspect of demand for healthcare, with individuals with higher λ_i being sicker and hence demanding greater healthcare consumption.[20] Second, households in each period receive a chronic health realization, m_{ft}^{CH}. For households without a chronic illness in the family, this amounts to the expected cost of a new diagnosis. For households living with chronic conditions, these shocks are the health costs associated with maintaining health for those affected by the conditions.

Moral hazard effects in this literature are typically allowed to be individual specific by the inclusion of a parameter ω_i that determines the rate at which cost-sharing deters an

[19] Household risk aversion plays a role in how families select coverage levels, but play no direct role in utilization decisions.

[20] Rather than simply having families draw their health expenditure m_i following a plan choice (Handel, 2013; Layton, 2017), I explicitly model these health shocks in order to separately identify how spending choices are reflective of beliefs about major health events, as well as to estimate the effects financial distortions caused by health events contribute to moral hazard in spending.

individual's optimal levels of spending. However, in order to separately identify the moral hazard effects induced by changes in care prices arising from major health events, I restrict ω to be homogeneous across individuals and periods.

4.2 Model Stages

In each period, families choose their insurance coverage to maximize the sum of their expected utilities based on their type parameters (including their beliefs about health risks); after this choice, both acute and chronic health shocks are realized, and individuals choose their medical utilization, trading off health status and wealth; finally, utilities and changes to type parameters are realized. The model is best conceptualized (as usual) in reverse order.

4.2.1 Transition Probabilities

In the final stage of the model, individuals update their beliefs about the unknown transition probability, p_{it}. This transition probability includes both individual beliefs about health risks as well as more general responses to health information, separate from moral hazard and salience effects. I discuss this more in section 4.3.2.

I model this learning 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. Individuals' prior beliefs about their true risk probability are assumed to be normally distributed with mean and variance parameters $(\mu_{p,i,0}, \sigma^2_{p,i,0})$. The center of the distribution $\mu_{p,i,0}$ varies with individual demographics and is potentially correlated with other household type parameters (see section 4.3.1 for details on parameterization).

Major health events provide individuals with signals y_{it} about the underlying distribution of p, I likewise assume that these signals are normally, so that the mean and variance of an individual's posterior distribution has a closed-form solution in each period. Specifically, if $y_{it} \sim \mathcal{N}(\tilde{\mu}_{it}, \tilde{\sigma}^2_{it})$, the evolution of the mean and variance parameters can be written as:

$$\sigma^2_{p,i,t+1} = \frac{\tilde{\sigma}^2_{it}\sigma^2_{p,i,0}}{\tilde{\sigma}^2_{it} + s_{it}\sigma^2_{p,i,0}} \tag{3}$$

$$\mu_{p,i,t+1} = \frac{\tilde{\sigma}^2_{it}\mu_{p,i,t} + \sigma^2_{p,i,t}\tilde{\mu}_{it}}{\tilde{\sigma}^2_{it} + \sigma^2_{p,i,t}}, \tag{4}$$

where the variable s_{it} indicates how many health signals an individual has received by the end of period t.

An important potential difficulty when using a Bayesian framework with rare events is

the choice of updating frequency. Given the relative rarity with which these shocks occur, continuous updating of probabilities in a typical Bayesian setting would result in posterior beliefs that are tightly centered around the initial mean and vary little with new information. In such a regime, individuals would have to perceive health shocks as having an extremely high mean (e.g., $\tilde{mu}_{it}$ much greater than 1) in order for health shocks to meaningfully change health beliefs.

In my preferred specification, I address this inconsistency by reducing the number of uninformative signals individuals process, assuming instead that belief updating begins only with the occurrence of a health event, with updating taking place every year after that. Such a framework is consistent with a notion of individuals who form beliefs about their health risk once, and then only revisit those beliefs once they have been called into question. Once the individual begins evaluating her health risk beliefs, however, she does so in a completely standard way (including updating her beliefs in the following years, even in the absence of high-information health signals).

Such an approach is an intuitively appealing way to deal with the issue of Bayesian updating when signals are frequently uninformative. However, my results are relatively consistent when using an adaptive learning framework (where individual beliefs change linearly in each period with some dependence $\rho < 1$ on the previous period's beliefs).[21] Additional modeling possibilities include the use of quasi-Bayesian modeling where individuals disregard less salient signals, but still update beliefs in each period (Rabin, 2013).

4.2.2 Utilization Choice

After choosing a health plan $j \in \mathcal{J}$ and realizing acute and chronic health shocks $(\lambda_{it}, m_{ft}^{\text{CH}})$, individuals choose their optimal level of spending on non-chronic medical care, m_{it}^*. As is typical for these models, individuals trade off health production and wealth. In my extension of the model, individuals face residual uncertainty as to the likelihood of their own major medical events.[22] I therefore assume that they choose m_{it} in order to maximize their expected utility over states:

$$m_{it}^* \equiv \text{argmax}_{m_{it}} EU(m_{it}; p_{it}) = p_{it} u_{it,\text{C}} + (1 - p_{it}) u_{it,\text{H}}, \tag{5}$$

[21] For a more in-depth review of the relative strengths and weaknesses of Bayesian or adaptive learning in structural modeling, see Aguirregabiria and Jeon (2020).

[22] Although the value of chronic care costs are assumed to be made known to a household before they choose their non-chronic spending, the model abstracts away from the specific timing of individual costs within a year. Hence, even within a period, individuals have not learned whether they are facing a chronic illness, and hence maximize an expected utility across both states of the world. It is not until the end of the period that individuals know their true state and update their beliefs p_{it}.

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 5 nests the case where an individual has already been diagnosed with a chronic illness, in which case their transition probability $p_{it} = 1$. I assume that each individual's utility function is separable in health and wealth for both chronic and healthy individuals:

$$u_{it,\mathrm{H}}(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}) = h_1(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}) + y(m_i; m_{ft}^{\mathrm{CH}}) + \varepsilon_1(m_{ft}^{\mathrm{CH}}, \lambda_{it}) \tag{6}$$

$$u_{it,\mathrm{C}}(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}) = h_2(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}) + g(m_{ft}^{\mathrm{CH}}; \lambda_{it}) + y(m_{it}; m_{ft}^{\mathrm{CH}}) + \varepsilon_2(m_{ft}^{\mathrm{CH}}, \lambda_i). \tag{7}$$

The returns to medical spending $h_1(\cdot), h_2(\cdot)$, and $g(\cdot)$ are assumed to be concave, so that within-year health fluctuations λ_{it} alter the optimal level of utilization m_{it}^*. Remaining annual income is denoted by $y(m_{it}; m_{ft}^{\mathrm{CH}})$. $\varepsilon_1(\cdot)$ and $\varepsilon_2(\cdot)$ are preference shocks to capture unobserved changes in preferences due to major medical events.

I parameterize these utility functions as quadratic loss functions in the difference between medical spending and acute health status, in keeping with past work (Marone and Sabety, 2020; Einav et al., 2013), but allow for a potentially state-dependent utility function in which health status potentially alters the marginal utility of medical spending. Finkelstein et al. (2009) and later work (Finkelstein et al., 2013) discuss and provide evidence for state-dependence in the utility of *non-medical* consumption; this model introduces suggestive evidence for the state-dependence of non-chronic medical consumption as well.

Individuals without chronic conditions face the typical utility function:

$$u_{it,\mathrm{H}}(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}, j) = (m_{it} - \lambda_{it}) - \frac{1}{2\omega}(m_{it} - \lambda_{it})^2 - c_j(m_{it}). \tag{8}$$

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

$$u_{it,\mathrm{C}}(m_{it}; \lambda_{it}, m_{ft}^{\mathrm{CH}}, j) = (\alpha_1 m_{it} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{it}) - \frac{1}{2\omega}(\alpha_1 m_{it} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{it})^2 - c_j(m_{it}). \tag{9}$$

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 straightforward (Appendix C.2),

with multiple possible solutions based on the cost structure of the family's insurance plan j. In general, small or negative values of λ_{it} will result in an individual choosing $m^*_{it} = 0$; otherwise, optimal spending follows the condition:

$$m^*_{it} = \frac{1}{1 + p_{it}(\alpha_1 - 1)} \left(\lambda_{it} + \omega(1 + p_{it}(\alpha_1 - 1) - c'_j(m_{it}; m^{\text{CH}}_{ft})) - p_{it}\alpha_2 m^{\text{CH}}_{ft} \right). \tag{10}$$

The interpretation of Equation 10 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. As discussed in Bleichrodt and Eeckhoudt (2006), the extent to which these probabilities are mis-measured may artificially alter optimal spending decisions based on both the level of actual risks and the extent of the measurement error. Under the assumptions that households begin with p_{i0} close to zero, major health events could be associated with large (relative) increases in p_{it}, potentially explaining the dramatic and persistent shifts observed in Section 3.[23]

Equation 10 also highlights the ways that chronic care costs affect spending decisions through prices. The OOP cost function $c_j(m_{it}; m^{\text{CH}}_{ft})$ is assumed to account for the price of chronic care first in the timing of health spending, before any other non-chronic spending. This anticipation of chronic care costs shifts the boundaries between optimal spending solutions by depressing the rate at which discretionary medical spending translates into OOP costs. This is the method by which moral hazard effects arise from major health events.

4.2.3 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 individual health shocks λ_{it} or major health costs m^{CH}_{ft}. The household expected utility function for a given plan j is integrated

[23] Generally, small risk probabilities are more likely to be over-weighted than under-weighted (Kahneman and Tversky, 1972); this, in conjunction with the size of baseline risk probabilities, suggests that households which experience major health events experience a shift in p_{it} from a baseline number close to zero.

over the distribution of both health shocks and then summed across individuals[24]:

$$U_{fjt}\left(F_{\lambda,i}, G_{m^{\text{CH}}}, \psi_{ft}\right) = \sum_{i\in\mathcal{I}_f}\left[-p_{it}\int\int w(u^*_{it,\text{C}})dF_{\lambda,i}dG_{m^{\text{CH}},i} - (1-p_{it})\int\int w(u^*_{i,\text{H}})dF_{\lambda,i}dG_{m^{\text{CH}},i}\right] - \left[y_{ft} - c_j(m^{\text{CH}}_{ft}) - \pi_j - \eta\mathbb{1}_{j,t-1}\right]. \tag{11}$$

In addition to each individual's realized OOP costs for non-chronic medical spending, households face OOP costs for chronic care as well as plan premiums π_j, and a (perceived) cost η for switching plans ($\mathbb{1}_{j,t-1}$ is an indicator for whether the family chose plan j in year $t-1$).[25]

I assume that the von Neumann Morgenstern (vNM) utility index for this decision possess a constant coefficient of absolute risk aversion, a common choice for these models as it implies no wealth effects. Specifically, I parametrize the vNM utility index for each family as

$$w(u^*_{it}) = -\frac{1}{\psi_{ft}(x_t)}e^{-\psi_{ft}(x_t)u^*_{it}} \tag{12}$$

for a given level of *ex-post* utility u^*. Risk aversion is allowed to be evolve over time to capture the salience effects associated with health events, as discussed in Section 4.3.1.

4.3 Estimation

4.3.1 Parametrization

The unit of observation is a family f comprised of a set of individuals $\mathcal{I}_f$ in year t. Each family is characterized by their unobserved type variables $\{p_{it}, \lambda_i, \psi_{ft}\}_{i\in\mathcal{I}_f}$. Each family faces a choice of plans that varies at the firm-year-state level.[26]

[24] One concern with a utilitarian index here is that households may have little incentive to diversify their medical spending across household members. However, the choice of the utility function used in the second (spending) stage of the model makes it optimal for families to allocate care according to each individual's realization of λ_{it}; hence, this modeling choice does not give rise to families allocating all of their care to a single individual, for example. An alternative approach might be to use a CES function for utilities; however, this introduces more nuisance parameters into the estimation framework.

[25] I do not observe premiums or contributions in my data and therefore follow the methodology of Layton (2017). In particular, I assume that premiums are equal to the average cost among the employees with dependents enrolled in the plan during the prior year plus a fixed overhead cost, and then assume that employee contributions are 25% of that value (Foundation, 2020). Note that as Layton discusses, identification of the structural parameters in this model do not depend on accurate estimation of premiums, but rather require that the premium differential across firms is correct.

[26] 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.

Initial type parameters, including those governing the prior distribution of health risk beliefs as well as parameters for health shocks and risk aversion, are assumed to be arbitrarily correlated. I link these parameters $(p_{i,0}, \lambda_i, \psi_{f,0})$ to observable data by assuming they are drawn from a multivariate normal distribution that depends on observed demographics, including age, sex, health risk score, family size, and the presence of pre-existing conditions in a household.

$$\begin{bmatrix} p_{i,0} \\ \mu_{\lambda,i} \\ \log(\psi_{f,0}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \beta_p \boldsymbol{X}^p \\ \beta_\lambda \boldsymbol{X}^\lambda \\ \beta_\psi \boldsymbol{X}^\psi \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & & \\ \sigma_{p,\lambda} & \sigma_\mu^2 & \\ \sigma_{p,\psi} & \sigma_{\lambda,\psi} & \sigma_\psi^2 \end{bmatrix} \right). \tag{13}$$

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 4.2.1. I assume that these signals y_{it} are normally distributed with variance σ_ν^2 (to be estimated) and a mean given by the logit regression:

$$y_{it} = \pi_1 \mathbb{1}\{\text{chronic}\}_{f,-i} + \pi_2 \mathbb{1}\{\text{acute}\}_{f,-i} + \pi_3 \mathbb{1}\{\text{acute}\}_{f,i} + \pi_4 x_{it}, \tag{14}$$

where *chronic* and *acute* indicate the occurrence of chronic or acute health events within a household and x_{it} is a variable for the number of years that have passed since the earliest major health event in the family.[27] Hence, π_1 is the effect of a household chronic diagnosis on individual beliefs. On the other hand the variance of the signals, σ_ν^2, reveals the magnitude of unobserved information affecting individual health risk probabilities.

I parameterize the distribution of acute health shocks, I assume that $F_\lambda(\cdot)$ is a shifted lognormal distribution. Assuming a lognormal distribution for λ is natural, as the distribution of annual health expenditures is highly skewed. The choice of shifting the distribution accommodates the approximately 15% of individuals in my sample who choose zero medical spending in a given year. I therefore assume that an individual's (correct) beliefs about their transient health shocks are given by

$$\ln(\lambda_{it} - \kappa_i) \sim \mathcal{N}(\mu_{\lambda,i}, \sigma_{\lambda,i}^2). \tag{15}$$

When κ_i sufficiently large (and negative), small and negative values of λ may lead to zero spending being the utility-maximizing solution for an individual.[28]

[27] I assume that chronic diagnoses are absorbing states so that $p_{i,t+\tau} = 1$ for all individuals who have a chronic health event in year t (and $\tau \geq 0$).Hence, the case of $\mathbb{1}\{\text{chronic}\}_{f,i}$ is not included in the regression.

[28] Previous work has allowed the distributions of these shocks to evolve over time (Marone and Sabety,

Acute health shocks at the individual level are therefore summarized by three parameters: $(\mu_{\lambda,i}, \sigma^2_{\lambda,i}, \kappa_i)$. The parameter $\sigma^2_{\lambda,i}$ reflects the precision in an individual's beliefs about their transient health state. Both $\sigma^2_{\lambda,i}$ and κ_i are estimated as a linear projection on individual covariates. Details on the specific covariates used in each linear regression are provided in Appendix C.2.

In contrast to acute health shocks, which are parameterized based on individual demographics, I use the empirical distributions for chronic care costs in my data in household expected utility. I assume that individuals have rational expectations over the distributions of their chronic health care costs. This is a simplifying assumption employed for tractability, as my model already allows for the identification of heterogeneous parameters governing individual expectations about acute health shocks, as is typical in this literature (Einav et al., 2013). However, although there is evidence that consumers do not know the price of health care before they consume it (Lieber, 2017), this is less of a concern with chronic care costs, which are typically stable over time and more easily internalized by household members, particularly in later years. Variation in distributions of these costs across illnesses is useful for the separate identification of any induced demand responses arising from changes in predictable health care spending.

I use separate empirical distributions for diagnosis years and subsequent years to accommodate the drastic differences between diagnostic costs and maintenance costs for a given condition. I assume these distributions are stable over time. Households therefore face different distributions as they experience one or more major medical events over time.

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:

$$\psi_{ft}(x_t) = \psi_{f0} + \gamma_1 \left\{ \text{Post}_t \times m^{\text{CH}}_{f0} \right\} + \gamma_2 \left\{ \text{Post}_t \times c_{j0}(m^{\text{CH}}_{f0}) \right\} + \gamma_3 \left\{ \text{Post}_t \times \text{Hosp}_{f0} \right\} + \zeta_{ft}, \quad (16)$$

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

I denote the parameters of the model by θ. These parameters include the main parameters of interest, including the utility parameters $\alpha_1, \alpha_2, \omega$, and η and the regression coefficients in Equations 14 and 16, including the variances σ^2_ν and σ^2_ζ. Additional parameters included in θ are five vectors of mean shifters, $(\beta_p, \beta_\psi, \beta_\lambda, \beta_{\sigma_\lambda}, \beta_\kappa)$; seven variance and covariance

2020). In my model, which separates acute and chronic health shocks, such variation would amount to shifts in the need for non-chronic health spending, such as variation in an individual's anticipated office-visit spending from year to year. In addition to being of second-order concern to my setting, such variation seems indistinguishable from the random variation in the draws of λ_{it} already present.

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 $p_{it}, \mu_{\lambda,i}, \sigma_{\lambda,i}, \lambda_{it}$, and ψ_{ft}.

I estimate the model via maximum likelihood, following the approach laid out by Revelt and Train (1998) and Train (2009), with the appropriate adaptation for modeling a discrete choice followed by a continuous one (Dubin and McFadden, 1984). For a given household, likelihood functions are constructed as the density of their observed health spending conditional on their observed plan choices. The estimation is done in R version 4.0.3, following some of the best practices laid out in Conlon and Gortmaker (2020). I provide additional estimation details in Appendix D.4.

4.3.2 Identification & Interpretation

My model utilizes multiple sources of variation to separate multiple effects arising from major medical events. Health events may reorient households towards their health expenditures by increasing the salience of health consumption, provide experiential learning about how to obtain high-quality health care, or alter preferences for medical care in other ways, in addition to any changes in individual risk beliefs. The critical challenge is that changes in risk preferences, salience, or systematic health learning may also increase the willingness to purchase insurance and utilize medical care.

In order to separate risk aversion and preferences from beliefs about risk, I use variation in insurance plan characteristics and choice sets faced by different households in my data set. These choice sets vary at the firm-state-year level, and typically include plans with a ride range of cost-sharing parameters, including plans with zero deductible, high-deductible plans, and more flexible options. Under the assumption that risk aversion drives plan choice and not medical spending, and that households with high risk aversion seek to reduce the incidence of high OOP expenditures (Erikstrup et al., 2020), highly risk-averse households will gravitate towards the plans in their choice sets that most limit high expenses (e.g., low-deductible plans). I complement this with variation in major medical events, including their characteristics, to incorporate the role of salience associated with health trauma in changing household risk aversion. Modeling household risk aversion as responsive to hospitalizations or other high-expenditure events allows me to further separate changes in household risk preferences from their implied beliefs about major health events recurring.

I further separate moral hazard effects from other informational effects. This separate identification relies on variation in the costliness across different diagnoses as well as differences in diagnostic and maintenance costs. Each of these differences contributes to different

changes in care prices, allowing for the identification of price responses.

The principal estimated structural parameters of interest in my model are those governing the evolution of the transition probabilities p_{it}. Changes in these parameters that arise from new chronic diagnoses encompass both a reevaluation of individual health risk beliefs and other informational effects unaccounted for in the model, which may load onto this parameter. These effects include learning about the health care system more generally or forging better relationships with health care providers. Although section 3 suggests that these factors are not the principal mechanisms for responses, they may influence how p_{it} responds to new diagnoses. I therefore interpret changes in p_{it} as resulting from an aggregate informational effect, rather than from moral hazard or salience effects. Appendix C.2 discusses an interpretation of p_{it} as a preference weighting across states rather than explicitly health beliefs.

5 Structural Results

Table 5 presents the estimated parameters resulting from maximum likelihood estimation. Column 3 shows the preferred specification described in Section 4, while columns 1 and 2 present simplifications of the model that are useful both in building intuition and validating the estimated parameters. Additional parameters not relevant to the welfare effects of health information—including incidental parameters such as switching costs and individual mean-shifting regression coefficients—can be found in Table 14 in Appendix D.4.

I consistently find strong effects associated with chronic diagnoses in a household on non-diagnosed beliefs. Major health events are associated with an average increase in an individual's mean probability of a major health event of 31.8 percentage points. The bulk of this effect is due to family chronic events; in particular, the marginal effect of a household chronic event is 0.73, so that the probability that an individual's updated health risk probability is 1 increases by 73 percentage points. These increases are persistent, decreasing by only 0.4 percentage points each year at the margin. Other health events, including acute health events for both the individual and their family, result in much smaller changes to beliefs. The estimated variance for the unobserved dimension of belief changes is only 1.52, suggesting that observed events are dictating most of the changes in risk assessments.

Table 5 presents additional 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. 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.

Parameter	Description	Estimated Values		
Panel A: Dynamic Parameters				
Belief Evolution				
π_1	Impact of Family Chronic Event			0.73
π_2	Impact of Own Acute Event			-0.02
π_3	Impact of Family Acute Event			-0.04
π_4	Years since Chronic Event			-0.004
σ_π	Error Variance, Beliefs			1.52
Risk Aversion Evolution				
ψ_0	Risk Aversion Persistence	–	–	0.95
ψ_1	Major Health Event	–	–	0.61
ψ_2	Major Health Event × Billed Diagnostic Cost	–	–	0.19
ψ_3	Major Health Event × OOP Diagnostic Cost	–	–	-0.88
ψ_4	Major Health Event × Hospitalization	–	–	1.51
σ_ψ	Error Variance, Risk Aversion	–	–	0.01
Panel B: Heterogeneity in Types				
σ^2_ε	Idiosyncratic shock variance			3.56
σ^2_p	Initial beliefs variance			14.51
σ^2_λ	Acute shocks variance	–		2.03
σ^2_ψ	Initial risk aversion variance			2.57
$\rho_{p,\lambda}$	Correlation, initial beliefs and acute shocks			0.38
$\rho_{p,\psi}$	Correlation, initial beliefs and initial risk aversion			-0.54
$\rho_{\psi,\lambda}$	Correlation, acute shocks and initial risk aversion			0.09
Beliefs Evolve		X	X	X
Heterogeneity in Acute Health Shocks			X	X
Risk Aversion Parameters Evolve				X

Notes: This table presents estimates for selected parameters of the structural model of health choice; Table 14 presents estimates for the remaining parameters. 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 5. Estimated Main Structural Parameters of Interest

Panel B reports additional information regarding the distribution of household types and the value of incorporating the full richness of the model in rationalizing observed plan choices and spending. In particular, I estimate a high degree of variance in individual health risk beliefs (prior to any health event). These beliefs are strongly (positively) correlated with acute health status and (negatively) correlated with household risk aversion. These facts suggest that variation in individuals' estimated beliefs reflects variation in individual health status, as expected. Finally, in the full version of the model, the variance of the idiosyncratic error term is small, suggesting that most of the observed variation in consumer behavior can be explained by heterogeneity in individual types, responses to major health events, or both.

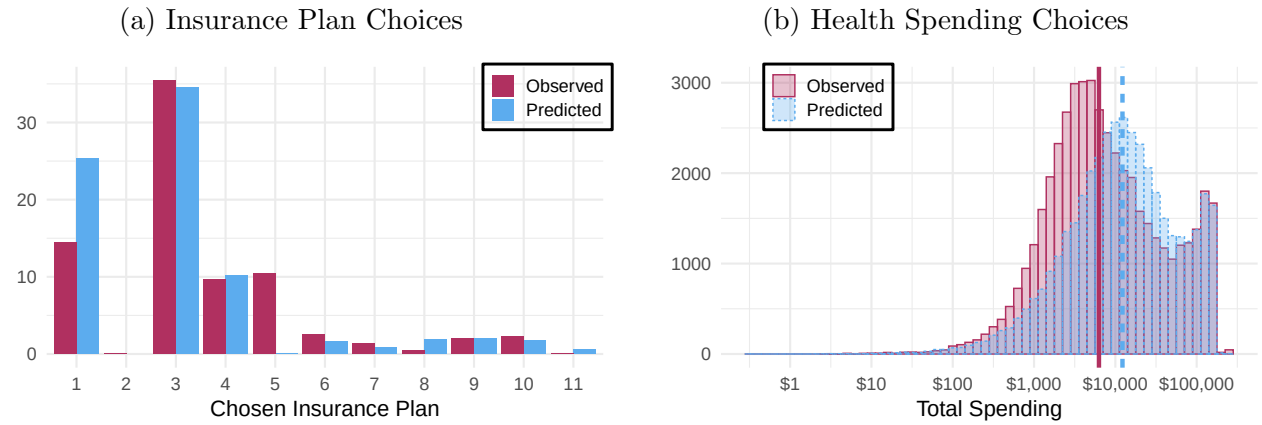
Columns 1 and 2 of Table 5 illustrate simplifications of the model that help validate the estimated parameters and build intuition. In column 1, I estimate a version of the model with no heterogeneity in acute health shocks or changes in household risk aversion. That is, $\mu_{\lambda,i}, \sigma_{\lambda,i}$, and $\kappa_{\lambda,i}$ are not allowed to vary based on individual covariates, and ψ_{ft} is fixed over time. Column 2 adds variation in acute health status to the model while continuing to hold household risk aversion constant over time. **Discussion here.**

5.1 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. 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. Since 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 8 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. Overall, however, individual deviations between observed and predicted spending are small. Additionally, 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 8. 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 (see Appendix D.4 for other firms). 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.

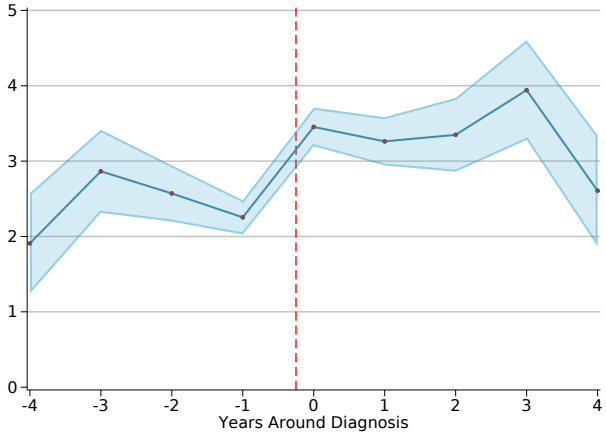
5.2 Spending Response to Major Health Events

Figure 9 illustrates the model's predictions surrounding behavior following new chronic diagnoses in a household as recentered time series graphs. Similar to the results in Section 3, I examine how these diagnoses alter the spending patterns of other household members in the panel (a). I also present estimates for how diagnoses affect estimates for individuals' underlying transition probabilities p_{it} in panel (b). I find that major health events are associated with about a 20% increase in OOP spending—this is slightly higher than the 10% estimated changes reported in Figure 1, although the differences are not statistically significant.

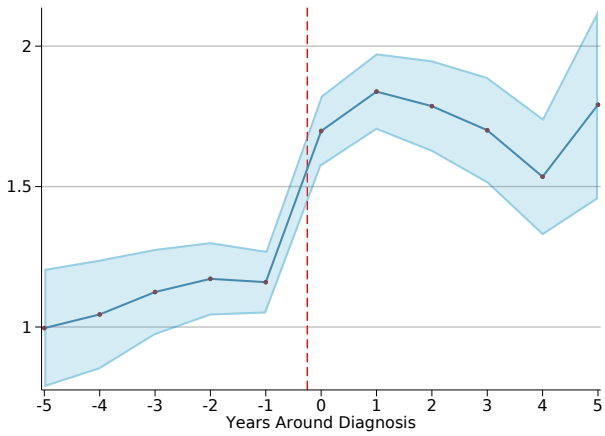
Importantly, I predict large accompanying changes in individual health risk beliefs following a new chronic diagnosis in the family. The horizontal dashed line in the figure depicts the pooled average risk of diagnosis within my sample. Prior to health events, individuals tend to underweight their health risks by about **XX**%; however, following a diagnosis, individuals move to *over-weighting* their risks by **XX**%. This provides suggestive evidence that individuals in affected households may over-respond to these events. I explore the welfare implications of these facts in the following section.

Figure 9. Predicted Effects of New Diagnosis on Non-Diagnosed OOP Spending and Risk Beliefs

(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.

6 Welfare & Counterfactual Simulations

Based on the estimated model parameters, I am able to construct a measure of each household's willingness to pay for information associated with their own health risks. I use this measure to provide a benchmark for the value associated with this information, with particular focus on whether major health events meaningfully alter individual expected utility and social surplus.

6.1 Welfare Effects of Information

Households who receive health information alter their plan choice and medical spending decisions, thereby altering their *ex-ante* expected payoffs from care. I estimate the value of this information by comparing household certainty equivalents over the as yet unrealized lottery of health states for two regimes: one in which the health information is revealed as observed and another benchmark regime in which health costs change no such event occurs. Specifically, in the benchmark state of the world, households experience changes to their health states without corresponding changes to their beliefs, risk aversion, or the conditional costs of non-chronic care. I perform this analysis only on non-diagnosed family members within an affected household to examine only changes in welfare associated with

information, rather than true changes in underlying health status.

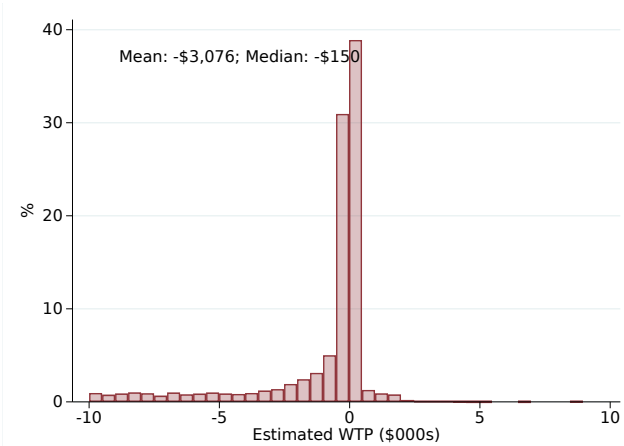
A household's willingness to pay for health information is equal to the difference in certainty equivalents across these two regimes. Certainty equivalents are given by

$$CE_{fjt} = -\psi_{ft}^{-1}\log(-U_{fjt}), \tag{17}$$

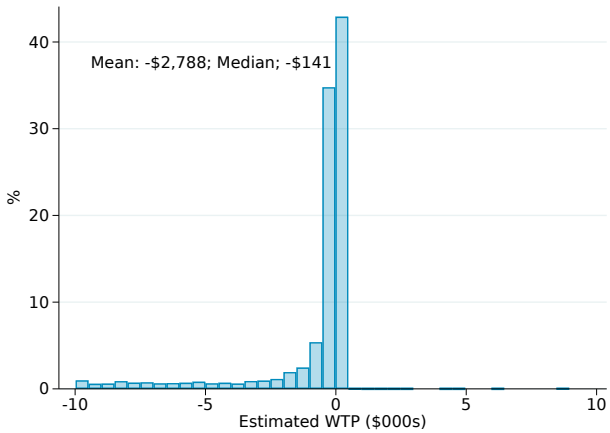
where U_{fjt} is the *ex-ante* expected utility family f expects when enrolling in plan j at time t, as defined in equation 11. I assume that conditional on the estimated parameters, households are fully rational and enrol in the plan that gives the highest expected utility at the time of choice.[29] Throughout, I report differences between CE_{fjt} across the benchmark state of the world and regimes where information is partially or fully revealed; hence, reported values are "marginal" willingness to pay measures.

Figure 10. Variation in Welfare Effects Associated with Health Events and Health Information

(a) Full Welfare Effect



(b) Health Information Effect Only



Notes: Figures show estimated changes in household willingness to pay associated with major health events. The panel on the left shows differences in household certainty equivalents in the case of a full response to a new diagnosis, including adjustments to risk aversion and moral hazard effects; the panel on the right shows only differences arising from adjustments to household risk assessments. Welfare effects are calculated in the year of the diagnosis relative to a benchmark in which no information is transmitted.

Figure 10 presents the main welfare results in the year of the new chronic diagnosis.[30] Household members who are exposed to a new chronic diagnosis experience a welfare penalty

[29]The model allows for rich heterogeneity in the prediction of health states as well as rationalizations for common choice mistakes, including switching costs. Hence, such an assumption is reasonable. Similarly, I assume that the idiosyncratic shock parameter is not relevant for the context of estimating welfare gains from health information.

[30]These welfare effects are relatively stable in the first few years following the diagnosis. For details on this, see Appendix D.4.

that averages $3,076 per household per year. However, there is substantial heterogeneity in these effects, including 28% of treated families who have a higher resulting expected utility following the realization of health information.

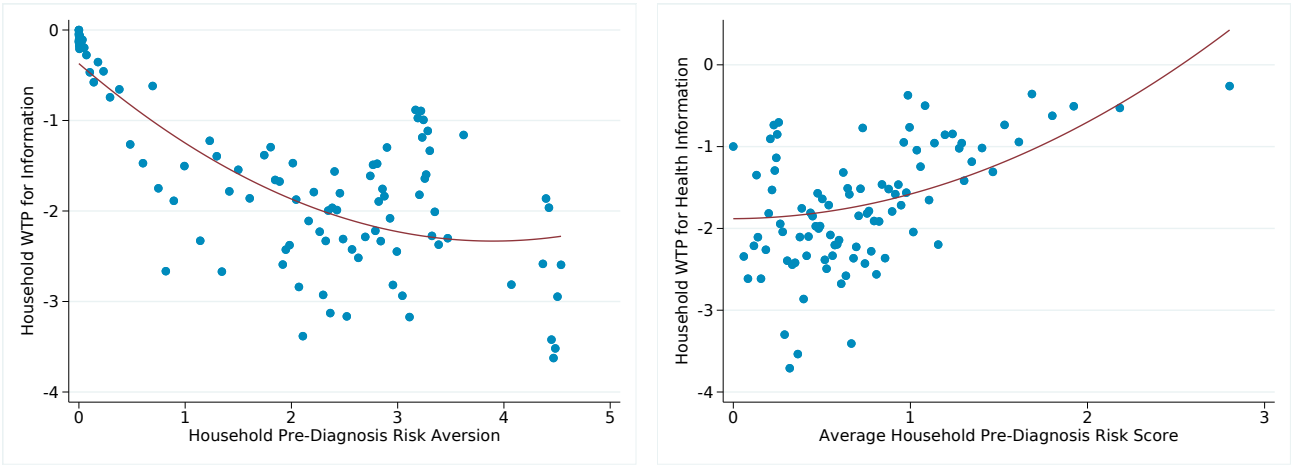
The right panel of Figure 10 shows the distribution of welfare effects associated solely with receiving new health information. A novel feature of my structural model is the ability to separate changes to household welfare that arise from dimensions of a health event other than the realization of health information. I recalculate welfare changes associated with *only* changes to household risk beliefs by holding constant changes to both household risk aversion and any moral hazard effects that arise from changes to spot prices. My analysis reveals that these dimensions contribute little to overall changes in household welfare, with 90% of welfare changes being explicitly attributable to changes in household risk assessments. The average household experiences a welfare penalty of $2,788 associated with changes to how they evaluate their risk of developing a chronic condition.

Although at first glance associating new information with a welfare penalty seems counter-intuitive, my results are consistent with a story of household over-responsiveness to information. The observed choice data which informed the estimated model parameters suggests that new chronic diagnosis spur large swings in household members' assessments of their health risks; however, these welfare calculations make clear that in many cases, households would be better off if they had acted as though they had not received the information. This is precisely because of the magnitude of the shifts in household beliefs, as I illustrate in the following section.

Importantly, the returns to health information vary with key household characteristics, including household risk levels and estimated risk aversion. Figure 11 presents these results. Households who are less averse to negative outcomes prior to the diagnosis experience lower welfare penalties, on average, than those with higher risk aversion. Differences in this parameter are intuitively meaningful: households with greater risk aversion experience greater "translation" of new health information into changes in insurance plan choices and, subsequently, health spending. Hence, households with lower levels of risk aversion tend to respond less to new information, presumably contributing to the lower estimated welfare penalties associated with the event. Similarly, households with high expected health risks prior to a new diagnosis experience lower welfare penalties. This, too, is related to overall muted responses to health information. However, this low level of responsiveness is attributable not to low variation in expected utility but an already high level of expected spending, meaning new health events change outcomes (in percentage terms) less.

Figure 11. Value of Health Information Varies with Household Risk Level and Risk Aversion

(a) Household Risk Aversion, $\overline{\psi}$ (b) Household Risk Scores



Notes: Figures show binscatters depicting the association between the estimated welfare effects of receiving health risk information and pre-diagnosis household health characteristics. These include (a) average household risk aversion and (b) average household risk scores (calculated using the Johns Hopkins ACG System). Welfare effects are calculated in the year of the diagnosis relative to a benchmark in which no information is transmitted; see Figure 10 for details. Binscatters are constructed using 100 bins and a quadratic fit line.

6.2 Evaluating Household Over-Responsiveness to Information

The results above imply that while households respond meaningfully to new health information, they may not be doing so in ways that are welfare improving. Given these estimated welfare penalties, in this section I assess the extent to which consumers' over-responsiveness to health information dampens potential welfare gains. The model predicts large swings in consumer risk beliefs when exposed to chronic diagnoses in a household. I therefore first assess the extent to which limiting the magnitude of these changes affects estimated welfare differences. I then turn to practical policy questions surrounding when information revelation is optimal, and whether targeted revelation can improve social outcomes.

6.2.1 Bounding Belief Updating

I first consider how limiting household responsiveness to adverse health events alters estimated welfare gains or penalties from new health risk information. Here, I present estimated effects from imposing arbitrary bounds on an individual's beliefs about their own health risks, illustrating that if consumers' responses more closely matched their true expected risk (conditional on the household member's diagnosis), health information would be associated with welfare gains rather than losses. I present alternative simulations expressing the same sentiment in Appendix D.4; these include imposing bounds on the model parameter π_1 controlling

the impact of chronic events on individual beliefs.

	Willingness to Pay for Health Information		
Upper bound, $\overline{p}$ (%)	Mean	Median	% with Welfare Gains
–	−$2,788	−$141	00.2
10	$2,439	$207	78.9
5	$2,472	$223	80.3
2.5	$2,484	$230	81.5
1	$2,492	$234	83.1
0.5	$2,495	$234	84.4
$\hat{p}_{it}$	$2,385	$219	83.9

Notes: Welfare effects are calculated for the year of the diagnosis, relative to a benchmark state where no information is transmitted. Moral hazard and risk aversion do not change across states.

Table 6. Bounding Consumer Responsiveness Improves Value of Health Information

Table 6 presents the results. In general, even imposing a bound on consumers' beliefs such that no individual rates their *ex-post* risk of developing a chronic condition higher than 10% increases the percentage of households who gain from health information from 0.2% to 78.9%. Once this bound is imposed, households would be willing to pay an average (median) of $2,439 ($207) to have this health information revealed to them, rather than experiencing a welfare penalty in the unrestricted context. Further restrictions on this bound improve welfare estimates and the percentage of households with positive welfare changes, but only slightly. This suggests that welfare losses are arising principally from individuals responding to diagnoses with new risk beliefs well above 10%. To the extent that these chronic conditions remain rare and genetic information does not inform risk probabilities above 10%, welfare gains from health information are hence being eliminated as individuals overweight their *ex-post* probabilities of illness.

One concern with this approach is that individuals within households may have private information compelling them to have these large changes in their risk beliefs. Such private information, including about underlying health status, could rationalize observed welfare penalties. The extent to which private information could inform patients of health risks of developing a chronic condition higher than odds of one in ten seem unlikely. However, to address this concern, I estimate individual-specific health risks $\hat{p}$ based on demographics including age, sex, and relationship with diagnosed household members. Although these predicted health risk probabilities do not capture the full range of private information, they address individual differences in potential responsiveness to new information.

I estimate predicted health risk probabilities on a validation sample constructed from all

Marketscan households not in my main sample who experience at least one chronic event during their observed period. Additional details about this estimation and summary statistics for the resulting probabilities are provided in Appendix D.4. The predicted probabilities are small, averaging about 2%.

The bottom row of Table 6 uses these predicted probabilities as individual-specific upper bounds on changes to individual beliefs. Hence, this simulation predicts the welfare effects associated with hypothetical health events that perfectly inform consumers of their risk in the absence of private information. Even accommodating individual-specific deviations in information bounds, I continue to find that restraining the extent to which individuals respond to health information corresponds to higher overall valuations of health information. The estimated effect sizes are similar to arbitrary uniform bounds, with an average welfare gain of $2,385 per household-year and 83.9% of households experiencing gains from health information.

6.2.2 Optimal Information Revelation

In addition to concerns about individual over-responsiveness to health information, policy guiding the revelation of health information must also balance the potentially heterogeneous returns from such revelation. In the face of such variation, full information revelation may not be socially optimal. This includes cases where a full screening regime is not financially feasible, where the information itself may result in consumers declining actuarily fair insurance (Posey and Thistle, 2021), or where there is a disconnect between privately and socially optimal information revelation (Oster et al., 2013). In these cases, the ability to target policies that reveal health risk information may improve the social returns as well as the fraction of households who benefit from these programs.

I estimated strong heterogeneous returns to health information as presented in Figure 11. Based on these results, I consider the effects of targeting information revelation based on observable characteristics, such as individual risk measured by risk scores.[31] I consider a scenario in which individuals can receive a one-time update to information about their health risks, modeled as changes to their probability of adverse health events p_{it}. When individuals receive this information, this probability is adjusted to be equivalent to their predicted risk probability $\hat{p}_{it}$, as discussed in the preceding subsection. I assume that following this information, individual beliefs are constant at their predicted risk level, with no residual uncertainty or updating across periods.[32] I further assume that this information does not

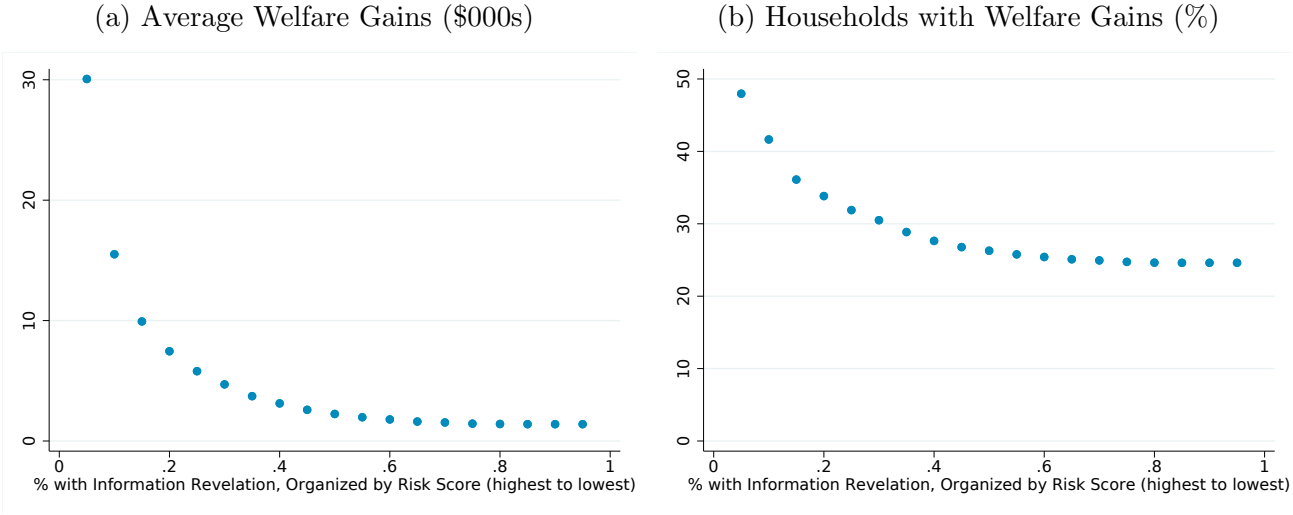
[31] My model predicts variation in the value of health information based on less-easily observable characteristics, such as household risk aversion. Although perhaps more costly to observe, information targeting along this dimension may also prove beneficial. I present these results in Appendix D.4.

[32] In reality, individuals are more likely to behave as though this information revelation were a single

entail any salience or moral hazard effects; throughout, I present results only for the welfare effects associated with new health risk information.

In this section, I present results estimating the individual and social value of such a policy for a sample of 10,000 households that do not experience major health events in my sample. Although they do not receive information through natural exposure to major health events, individuals in this group may have erroneous beliefs about their health risks and may benefit from new health information. Further, estimating the effects of this policy on this group provides a validating sample for the results presented earlier documenting the value of information transmitted in a more quasi-random setting.

Figure 12. Targeted Revelation of Information by Risk Improves Welfare Gains



Notes: Figures illustrate estimated welfare gains associated with revelation of health information, based on individual risk scores. Each point represents a different scenario, in which a different fraction of individuals are given information about their predicted health risks, $\hat{p}$, as described in the text. Information revelation is targeted by average individual risk score; hence for each point, the individuals with the highest x% of risk scores are given health information. Returns to health information are presented as (a) average expected welfare changes, measured as willingness to pay in thousands of dollars, and (b) the percentage of households with positive welfare gains.

Figure 12 presents the results. Across all the households in my sample, the average household would be willing to pay approximately $1,500 per year for such information (the right-most point in Panel (a)); however, only about a quarter of households would be willing to pay anything for this information, as others would make worse choices with updated beliefs (the right-most point in Panel (b)). Both panels of the figure present alternative

non-definitive signal, although perhaps one to which they ascribe a large weight relative to the others. This would imply that individuals do not update their beliefs fully to the revealed target $\hat{p}_{it}$ in the year of the signal nor hold their beliefs constant in the years after the signal. However, for the purposes of this exercise, such fluctuations serve only to obfuscate the potential benefits of targeting information revelation relative to universal revelation. I therefore use the simplest version of the model that highlights these differences.

scenarios where the information is not revealed to every household indiscriminately, but only to individuals with the highest estimated levels of risk. Each point represents a scenario in which only individuals with the highest $x\%$ risk scores receive health information.

These alternative scenarios highlight that targeting health information along observable demographics may help improve the returns from a policy increasing access to health information. Even under the assumption that this information can be revealed without over-responsiveness, demographics can be used to predict which individuals and households are most likely to benefit from the information.

7 Conclusion

This paper assesses the extent to which information about one's health risks alters individual and household decision-making in health care. I demonstrate that households who receive new information about health risks from a new diagnosis in the household increase their overall levels of spending, including investments in both preventive and low-value services. I illustrate that these changes in behavior are best explained by individual household members reassessing their risks, rather than learning about the value of care or responding to financial incentives or salience effects. However, these risk reassessments do not meaningfully improve the quality of health care choices, with affected household members seeking out a higher rate of health services with expected returns below their costs. The implication of these results is that while access to new health information changes behavior in meaningful ways, it does not necessarily do so in welfare-improving ones.

To explore this further, I present estimates from a structural approach that quantifies a household's willingness to pay for health information, separating out the effects of aggregated health information from moral hazard or salience effects. The model implies low realized returns to health information, perhaps due to individual misinterpretation of health risks following the health event. In particular, bounding the extent to which individual beliefs about health risks may change post-diagnosis substantially improves realized welfare.

The analysis I present could be extended in several meaningful ways. First, future work could relax the assumption that individuals have no control over their chronic care health costs. This would be particularly interesting in non-ESI covered populations, such as those coverd by public insurance programs or without any coverage, for whom chronic diagnoses may impose large financial burdens (Hadley, 2007). Another important consideration left out of the model is how liquidity constraints change *ex-post* spending adjustments as health risks change (Gross et al., 2020). Finally, future work might integrate this model with other costs incurred through living with a chronic condition, including earnings penalties and job

lock (Biasi et al., 2019; Eriksen et al., 2021; Garthwaite et al., 2014).

Increasing an understanding of how consumers interpret new information is at least as vital as improving their access. My results, building on a foundation of documented "mistakes" in choice, suggest that information campaigns, genetic screenings, and other sources of health information may not have nonnegative welfare effects without careful interpretation of results. Individuals and families living with the risk of chronic illness may be better off as they are taught to seek out high-value medical care and tamper high expectations of negative outcomes.

References

Abaluck, J. and Compiani, G. (2020). A Method to Estimate Discrete Choice Models that is Robust to Consumer Search. Technical Report w26849, National Bureau of Economic Research, Cambridge, MA.

Abaluck, J. and Gruber, J. (2011). Choice Inconsistencies among the Elderly: Evidence from Plan Choice in the Medicare Part D Program. *American Economic Review*, 101(4):1180–1210.

Abaluck, J. and Gruber, J. (2016a). Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance. *American Economic Review*, 106(8):2145–2184.

Abaluck, J. and Gruber, J. (2016b). Improving the Quality of Choices in Health Insurance Markets. Technical Report w22917, National Bureau of Economic Research, Cambridge, MA.

Aguirregabiria, V. and Jeon, J. (2020). Firms' Beliefs and Learning: Models, Identification, and Empirical Evidence. *Review of Industrial Organization*, 56(2):203–235.

Agüero, J. M. and Beleche, T. (2017). Health shocks and their long-lasting impact on health behaviors: Evidence from the 2009 H1N1 pandemic in Mexico. *Journal of Health Economics*, 54:40–55.

Alalouf, M., Miller, S., and Wherry, L. (2019). What Difference Does a Diagnosis Make? Evidence from Marginal Patients. Technical Report w26363, National Bureau of Economic Research, Cambridge, MA.

Almond, D., Doyle, Jr., J. J., Kowalski, A. E., and Williams, H. (2010). Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns*. *The Quarterly Journal of Economics*, 125(2):591–634.

Arni, P., Dragone, D., Goette, L., and Ziebarth, N. R. (2021). Biased health perceptions and risky health behaviors—Theory and evidence. *Journal of Health Economics*, 76:102425.

Baicker, K., Mullainathan, S., and Schwartzstein, J. (2015). Behavioral Hazard in Health Insurance. *The Quarterly Journal of Economics*, 130(4):1623–1667.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2013). The Nature of Risk Preferences: Evidence from Insurance Choices. *The American Economic Review*, 103(6):2499–2529. Publisher: American Economic Association.

Barseghyan, L., Molinari, F., O'Donoghue, T., and Teitelbaum, J. C. (2018). Estimating Risk Preferences in the Field. *Journal of Economic Literature*, 56(2):501–564. Publisher: American Economic Association.

Bellemare, M. F. and Wichman, C. J. (2020). Elasticities and the Inverse Hyperbolic Sine Transformation. *Oxford Bulletin of Economics and Statistics*, 82(1):50–61. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12325.

Bhatia, R. S., Levinson, W., Shortt, S., Pendrith, C., Fric-Shamji, E., Kallewaard, M., Peul, W., Veillard, J., Elshaug, A., Forde, I., and Kerr, E. A. (2015). Measuring the effect of Choosing Wisely: an integrated framework to assess campaign impact on low-value care. *BMJ Quality & Safety*, 24(8):523–531.

Biasi, B., Dahl, M. S., and Moser, P. (2019). Career effects of mental health. *Working Paper*.

Bleichrodt, H. and Eeckhoudt, L. (2006). Willingness to pay for reductions in health risks when probabilities are distorted. *Health Economics*, 15(2):211–214. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.1073.

Borusyak, K. and Jaravel, X. (2016). Revisiting Event Study Designs. *SSRN Electronic Journal*.

Bouckaert, N., Gielen, A. C., and Van Ourti, T. (2020). It runs in the family – Influenza vaccination and spillover effects. *Journal of Health Economics*, 74:102386.

Bundorf, K., Polyakova, M., and Tai-Seale, M. (2021a). How do Humans Interact with Algorithms? Experimental Evidence from Health Insurance. Technical Report w25976, National Bureau of Economic Research, Cambridge, MA.

Bundorf, M. K., DeMatteis, J., Miller, G., Polyakova, M., Streeter, J., and Wivagg, J. (2021b). Risk Perceptions and Protective Behaviors: Evidence from COVID-19 Pandemic. Technical Report w28741, National Bureau of Economic Research, Cambridge, MA.

Cairncross, S., Shordt, K., Zacharia, S., and Govindan, B. K. (2005). What causes sustainable changes in hygiene behaviour? A cross-sectional study from Kerala, India. *Social Science & Medicine*, 61(10):2212–2220.

Cardon, J. H. and Hendel, I. (2001). Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey. *The RAND Journal of Economics*, 32(3):408–427. Publisher: [RAND Corporation, Wiley].

Choudhry, N. K., Shrank, W. H., Levin, R. L., Lee, J. L., Jan, S. A., Brookhart, M. A., and Solomon, D. H. (2009). Measuring Concurrent Adherence to Multiple Related Medications. *The American journal of managed care*, 15(7):457–464.

Chua, K.-P., Schwartz, A. L., Volerman, A., Conti, R. M., and Huang, E. S. (2016). Use of Low-Value Pediatric Services Among the Commercially Insured. *Pediatrics*, 138(6). Publisher: American Academy of Pediatrics Section: Article.

Chung, S., Lesser, L. I., Lauderdale, D. S., Johns, N. E., Palaniappan, L. P., and Luft, H. S. (2015). Medicare Annual Preventive Care Visits: Use Increased Among Fee-For-Service Patients, But Many Do Not Participate. *Health Affairs*, 34(1).

Colla, C. H., Morden, N. E., Sequist, T. D., Schpero, W. L., and Rosenthal, M. B. (2015). Choosing Wisely: Prevalence and Correlates of Low-Value Health Care Services in the United States. *Journal of General Internal Medicine*, 30(2):221–228.

Conlon, C. and Gortmaker, J. (2020). Best practices for differentiated products demand estimation with PyBLP. *The RAND Journal of Economics*, 51(4):1108–1161. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1756-2171.12352.

Cutler, D. M. and Zeckhauser, R. (2004). Extending the theory to meet the practice of insurance. Technical report, Brookings Institution Press.

Dalton, A. R. H. and Soljak, M. (2012). The nationwide systematic prevention of cardiovascular disease: The UKs health check programme. *The Journal of Ambulatory Care Management*, 35(3):206–215.

Dalton, C. M., Gowrisankaran, G., and Town, R. J. (2020). Salience, Myopia, and Complex Dynamic Incentives: Evidence from Medicare Part D. *The Review of Economic Studies*, 87(2):822–869.

de Chaisemartin, C. and D'Haultfoeuille, X. (2019). Two-way fixed effects estimators with heterogeneous treatment effects. *NBER Working Paper 25904*.

DesRoches, C. M., Campbell, E. G., Vogeli, C., Zheng, J., Rao, S. R., Shields, A. E., Donelan, K., Rosenbaum, S., Bristol, S. J., and Jha, A. K. (2010). Electronic Health Records' Limited Successes Suggest More Targeted Uses. *Health Affairs*, 29(4):639–646.

Dranove, D. and Sfekas, A. (2008). Start spreading the news: A structural estimate of the effects of New York hospital report cards. *Journal of Health Economics*, 27(5):1201–1207.

Dubin, J. A. and McFadden, D. L. (1984). An Econometric Analysis of Residential Electric Appliance Holdings and Consumption. *Econometrica*, 52(2):345–362. Publisher: [Wiley, Econometric Society].

Dupas, P. (2011). Health Behavior in Developing Countries. *Annual Review of Economics*, 3(1):425–449. _eprint: https://doi.org/10.1146/annurev-economics-111809-125029.

Eichner, M. J. (1997). Medical expenditures and major risk health insurance. *Doctoral Dissertation, Massachusetts Institute of Technology*.

Einav, L., Finkelstein, A., Ryan, S. P., Schrimpf, P., and Cullen, M. R. (2013). Selection on Moral Hazard in Health Insurance. *American Economic Review*, 103(1):178–219.

Ericson, K. M., Kircher, P., Spinnewijn, J., and Starc, A. (2020). Inferring Risk Perceptions and Preferences Using Choice from Insurance Menus: Theory and Evidence. *The Economic Journal*, page ueaa069.

Eriksen, T. L. M., Gaulke, A., Skipper, N., and Svensson, J. (2021). The Impact of Childhood Health Shocks on Parental Labor Supply. *Journal of Health Economics*, page 102486.

Erikstrup, C., Hother, C. E., Pedersen, O. B. V., Mølbak, K., Skov, R. L., Holm, D. K., Sækmose, S., Nilsson, A. C., Brooks, P. T., Boldsen, J. K., Mikkelsen, C., Gybel-Brask, M., Sørensen, E., Dinh, K. M., Mikkelsen, S., Møller, B. K., Haunstrup, T., Harritshøj, L., Jensen, B. A., Hjalgrim, H., Lillevang, S. T., and Ullum, H. (2020). Estimation of SARS-CoV-2 infection fatality rate by real-time antibody screening of blood donors. preprint, Infectious Diseases (except HIV/AIDS).

Fadlon, I. and Nielsen, T. H. (2019). Family Health Behaviors. *American Economic Review*, 109(9):3162–3191.

Finkelstein, A., Luttmer, E. F. P., and Notowidigdo, M. J. (2009). Approaches to Estimating the Health State Dependence of the Utility Function. *The American Economic Review*, 99(2):116–121. Publisher: American Economic Association.

Finkelstein, A., Luttmer, E. F. P., and Notowidigdo, M. J. (2013). What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption. *Journal of the European Economic Association*, 11(suppl_1):221–258.

Finkelstein, A., Persson, P., Polyakova, M., and Shapiro, J. (2021). A Taste of Their Own Medicine: Guideline Adherence and Access to Expertise. *Working Paper*, page 38.

Foundation, K. F. (2020). 2020 Employer Health Benefits Survey. Technical report, Henry J. Kaiser Family Foundation, San Francisco, CA.

Friehe, T. and Pannenberg, M. (2019). Overconfidence over the lifespan: Evidence from Germany. *Journal of Economic Psychology*, 74:102207.

Ganguli, I., Lupo, C., Mainor, A. J., Wang, Q., Orav, E. J., Rosenthal, M. B., Sequist, T. D., and Colla, C. H. (2020). Assessment of Prevalence and Cost of Care Cascades After Routine Testing During the Medicare Annual Wellness Visit. *JAMA Network Open*, 3(12):e2029891.

Garthwaite, C., Gross, T., and Notowidigdo, M. J. (2014). Public Health Insurance, Labor Supply, and Employment Lock. *The Quarterly Journal of Economics*, 129(2):653–696.

Goodman-Bacon, A. (2018). Difference-in-Differences with Variation in Treatment Timing. Technical Report w25018, National Bureau of Economic Research, Cambridge, MA.

Goroll, A. H. (2015). Toward Trusting Therapeutic Relationships — In Favor of the Annual Physical. *New England Journal of Medicine*, 373(16):1487–1489.

Gross, T., Layton, T., and Prinz, D. (2020). The Liquidity Sensitivity of Healthcare Consumption: Evidence from Social Security Payments. Technical Report w27977, National Bureau of Economic Research.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):223–255. Publisher: University of Chicago Press.

Gruber, J., Handel, B., Kina, S., and Kolstad, J. (2020). Managing Intelligence: Skilled Experts and AI in Markets for Complex Products. Technical Report w27038, National Bureau of Economic Research, Cambridge, MA.

Guo, A. and Zhang, J. (2019). What to expect when you are expecting: Are health care consumers forward-looking? *Journal of Health Economics*, 67(102216).

Hackl, F., Halla, M., Hummer, M., and Pruckner, G. J. (2015). The Effectiveness of Health Screening. *Health Economics*, 24(8):913–935. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.3072.

Hadley, J. (2007). Insurance coverage, medical care use, and short-term health changes following an unintentional injury or the onset of a chronic condition. *JAMA*, 297(10):1073–1084.

Handel, B. and Kolstad, J. (2017). Wearable Technologies and Health Behaviors: New Data and New Methods to Understand Population Health. *The American Economic Review*, 107(5):481–485. Publisher: American Economic Association.

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

Handel, B. R. and Kolstad, J. T. (2015). Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare. *The American Economic Review*, 105(8):2449–2500. Publisher: American Economic Association.

Harris, J. W. and Stöcker, H. (1998). *Handbook of Mathematics and Computational Science*. Springer Science & Business Media. Google-Books-ID: DnKLkOb_YfIC.

Hong, Y.-R., Jo, A., and Mainous, A. G. I. (2017). Up-to-Date on Preventive Care Services Under Affordable Care Act. *Medical Care*, 55(8):771–780.

Iizuka, T., Nishiyama, K., Chen, B., and Eggleston, K. (2021). False Alarm? Estimating the Marginal Value of Health Signals. Technical Report 23413, National Bureau of Economic Research, Cambridge, MA.

Imai, K. and Kim, I. S. (2020). On the Use of Two-Way Fixed Effects Regression Models for Causal Inference with Panel Data. *Political Analysis*, pages 1–11.

Jiang, M., Hughes, D. R., and Wang, W. (2018). The effect of Medicare's Annual Wellness Visit on preventive care for the elderly. *Preventive Medicine*, 116:126–133.

Jones, D., Molitor, D., and Reif, J. (2019). What do Workplace Wellness Programs do? Evidence from the Illinois Workplace Wellness Study*. *The Quarterly Journal of Economics*, 134(4):1747–1791.

Kahneman, D. and Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3):430–454.

Kenkel, D. S. (1991). Health Behavior, Health Knowledge, and Schooling. *Journal of Political Economy*, 99(2):287–305. Publisher: University of Chicago Press.

Ketcham, J. D., Lucarelli, C., Miravete, E. J., and Roebuck, M. C. (2012). Sinking, swimming, or learning to swim in Medicare Part D. *American Economic Review*, 102(6):2639–2673.

Kim, H. B., Lee, S. A., and Lim, W. (2019). Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea. *Journal of Health Economics*, 65:1–14.

Kohro, T., Furui, Y., Mitsutake, N., Fujii, R., Morita, H., Oku, S., Ohe, K., and Nagai, R. (2008). The Japanese National Health Screening and Intervention Program Aimed at Preventing Worsening of the Metabolic Syndrome. *International Heart Journal*, 49(2):193–203.

Kolstad, J. T. (2013). Information and Quality When Motivation is Intrinsic: Evidence from Surgeon Report Cards. *American Economic Review*, 103(7):2875–2910.

Kowalski, A. (2016). Censored quantile instrumental variable estimates of the price elasticity of expenditure on medical care. *Journal of Business & Economic Statistics*, 34(1):107–117.

Krummel, T. M. (2019). The Rise of Wearable Technology in Health Care. *JAMA Network Open*, 2(2):e187672.

Layton, T. J. (2017). Imperfect risk adjustment, risk preferences, and sorting in competitive health insurance markets. *Journal of Health Economics*, 56:259–280.

Lee, W.-C. and Lee, S.-Y. (2010). National health screening program of Korea. *Journal of Korean Medical Association*, 53(5):363–370.

Lieber, E. M. (2017). Does It Pay to Know Prices in Health Care? *American Economic Journal: Economic Policy*, 9(1):154–179. Publisher: American Economic Association.

Margolis, J., Hockenberry, J., Grossman, M., and Chou, S.-Y. (2014). Moral Hazard and Less Invasive Medical Treatment for Coronary Artery Disease: The Case of Cigarette Smoking. Technical Report w20373, National Bureau of Economic Research, Cambridge, MA.

Marone, V. and Sabety, A. (2020). Should there be vertical choice in health insurance markets? *Working Paper*.

Mehrotra, A. and Prochazka, A. (2015). Improving Value in Health Care — Against the Annual Physical. *New England Journal of Medicine*, 373(16):1485–1487.

Moore, D. A. and Healy, P. J. (2008). The trouble with overconfidence. *Psychological Review*, 115(2):502–517. Place: US Publisher: American Psychological Association.

Oster, E., Shoulson, I., and Dorsey, E. R. (2013). Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease. *American Economic Review*, 103(2):804–830.

O'Connor, P. J. (2006). Improving Medication Adherence: Challenges for Physicians, Payers, and Policy Makers. *Archives of Internal Medicine*, 166(17):1802–1804.

Paserman, M. D. (2008). Job Search and Hyperbolic Discounting: Structural Estimation and Policy Evaluation. *The Economic Journal*, 118(531):1418–1452.

Pauly, M. V. and Blavin, F. E. (2008). Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance. *Journal of Health Economics*, 27(6):1407–1417.

Posey, L. L. and Thistle, P. D. (2021). Genetic testing and genetic discrimination: Public policy when insurance becomes "too expensive". *Journal of Health Economics*, 77:102441.

Rabideau, B., Eisenberg, M. D., Reid, R., and Sood, N. (2021). Effects of employer-offered high-deductible plans on low-value spending in the privately insured population. *Journal of Health Economics*, 76:102424.

Rabin, M. (2013). Incorporating Limited Rationality into Economics. *Journal of Economic Literature*, 51(2):528–543. Publisher: American Economic Association.

Revelt, D. and Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, 80(4):647–657.

Rubin, R. (2019). Debating Whether Checkups Are Time Wasted or Time Misused. *JAMA*, 322(2):101–102.

Sabety, A. (2020). The value of service sector relationships in health care. *Working Paper*.

Shafer, P. R., Hoagland, A., and Hsu, H. E. (2021). Trends in pediatric wellness visits with out-of-pocket costs before and after the Affordable Care Act. *JAMA Network Open*.

Slejko, J. F., Ho, M., Anderson, H. D., Nair, K. V., Sullivan, P. W., and Campbell, J. D. (2014). Adherence to Statins in Primary Prevention: Yearly Adherence Changes and Outcomes. *Journal of Managed Care Pharmacy*, 20(1):51–57.

Sloan, F. A., Smith, V. K., and Taylor, D. H. (2003). *The Smoking Puzzle: Information, Risk Perception, and Choice*. Harvard University Press. Google-Books-ID: ClkQH7aHyDwC.

Song, Z. (2021). Social Learning and Health Care Utilization: Evidence from Spousal Health Shocks. *Working Paper*.

Song, Z. and Baicker, K. (2019). Effect of a Workplace Wellness Program on Employee Health and Economic Outcomes: A Randomized Clinical Trial. *JAMA*, 321(15):1491–1501.

Song, Z. and Baicker, K. (2021). Health And Economic Outcomes Up To Three Years After A Workplace Wellness Program: A Randomized Controlled Trial. *Health Affairs*, 40(6):951–960. Publisher: Health Affairs.

Sorensen, A. T. (2006). Social learning and health plan choice. *The RAND Journal of Economics*, 37(4):929–945. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1756-2171.2006.tb00064.x.

Spinnewijn, J. (2015). Unemployed but Optimistic: Optimal Insurance Design with Biased Beliefs. *Journal of the European Economic Association*, 13(1):130–167.

Strokoff, S. L. and Grossman, E. G. (2010). Compilation of patient protection and affordable care act. Technical report, Office of the Legislative Counsel, U.S. House of Representatives, Washington DC.

Sun, L. and Abraham, S. (2020). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*, page 53.

Tiefenbeck, V., Goette, L., Degen, K., Tasic, V., Fleisch, E., Lalive, R., and Staake, T. (2016). Overcoming Salience Bias: How Real-Time Feedback Fosters Resource Conservation. *Management Science*, 64(3):1458–1476. Publisher: INFORMS.

Tong, S. T., Webel, B. K., Donahue, E. E., Richards, A., Sabo, R. T., Brooks, E. M., Lail Kashiri, P., Huffstetler, A. N., Santana, S., Harris, L. M., and Krist, A. H. (2021). Understanding the Value of the Wellness Visit: A Descriptive Study. *American Journal of Preventive Medicine*.

Train, K. E. (2009). *Discrete Choice Methods With Simulation*. Cambridge University Press, Cambridge, UK, 2 edition.

Wolfson, D., Santa, J., and Slass, L. (2014). Engaging Physicians and Consumers in Conversations About Treatment Overuse and Waste: A Short History of the Choosing Wisely Campaign. *Academic Medicine*, 89(7):990–995.

Zhang, X., Haviland, A., Mehrotra, A., Huckfeldt, P., Wagner, Z., and Sood, N. (2018). Does Enrollment in High-Deductible Health Plans Encourage Price Shopping? *Health Services Research*, 53(S1):2718–2734. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6773.12784.

A Data Cleaning

This appendix provides more detail on how the sample is constructed for analysis. I discuss the imputing of plan characteristics, identification of major health events, and identification of costs associated with chronic illnesses in detail in each of the sections below.

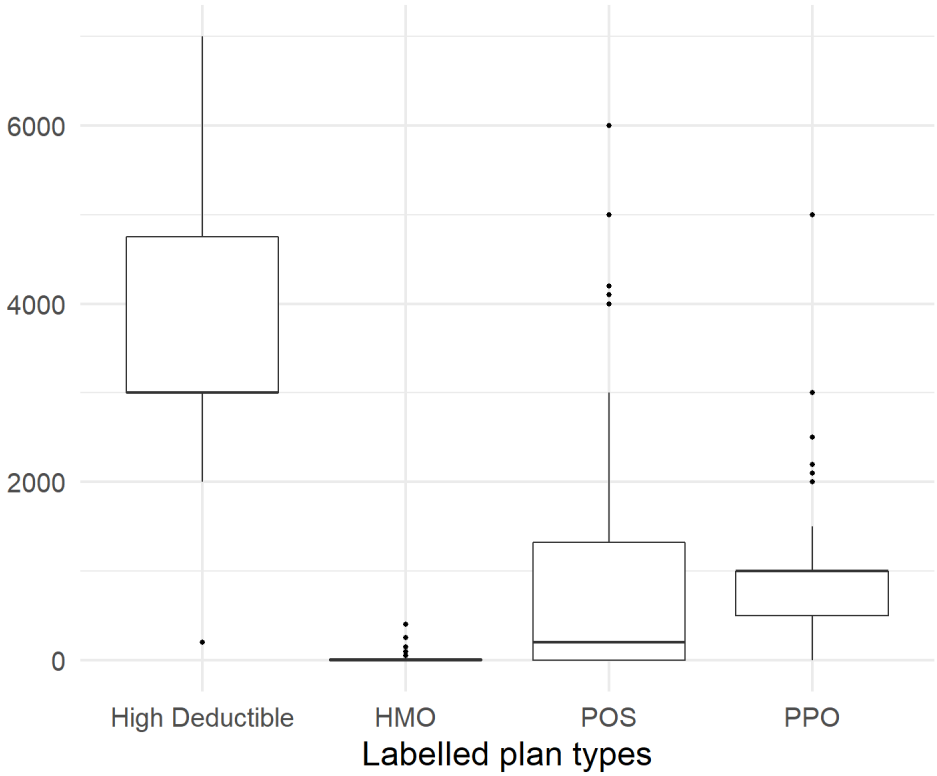
A.1 Identifying Plan Characteristics

I follow the methodology of Zhang et al. (2018) in inferring individual and household deductibles from the empirical distribution of claims. In particular, given the claims for an individual plan-year, I follow these steps:

1. Remove claims that are out-of-network, as well as claims with negative values in any of the total paid, plan paid, deductible, and OOP fields.
2. Limit attention to families that had at least 4 consecutive zero-deductible claims after the last positive deductible claim (to ensure that the deductible has really been reached).
3. Calculate each family's total deductible contribution over the year.
4. Next, estimate the mode and 95th percentile of the deductible within each plan-year.

Figure 13 illustrates the match quality of these assignments by comparing the distribution of imputed plan family deductibles across listed plan types (Rabideau et al., 2021).

Figure 13. Imputed Family Deductibles by Listed Plan Type



Notes: Box and whisker plot summarizing imputed household deductibles for each listed plan type. Each observation is a plan-year. The box in each boxplot extends from the first quartile to the third quartile of all family deductibles, with a line in the middle for the median. Whiskers extend to 1.5 times the interquartile range (the length of the box) if applicable. All plan years with deductibles outside of the whiskers are shown as outlier points.

Once deductibles are estimated, average coinsurance rates and out-of-pocket maxima are estimated using the methodology of Marone and Sabety (2020). These cost-sharing parameters are those which minimizes the sum of squared residuals between predicted and observed out-of-pocket cost, where predicted out-of-pocket costs utilizes the estimated deductible and assumed coinsurance and OOP maximum and observed spending comes directly from the claims data. This estimation is done separately for each plan-year.

Figure 14 illustrates the estimated function used in calculating out-of-pocket costs for a given amount of medical spending in a particular insurance plan and year, compared with the realized distribution of total and out-of-pocket spending for all households enrolled in that plan during the year. Each gray dot represents a household, and gold dots are a binscatter plot of all households, using 100 bins. The estimated features of this plan are a family deductible of $2,000, a coinsurance rate of 13.4%, and a family out-of-pocket maximum of $6,750.41.

Figure 14. Inferred Characteristics for a Sample Insurance Plan

Notes: Data shown for a single plan year. Each gray dot corresponds to a single household's observed total and out-of-pocket spending. Gold dots show averages within 50 bins. Blue line illustrates the estimated piece-wise linear function translating observed billed spending into out-of-pocket spending, determined by a plan deductible, coinsurance rate, and out-of-pocket maximum.

A.2 Identifying Major Health Events

I assign major health events using a set of chronic and acute HCCs, as discussed in Section 2 of the main text. The following table identifies each major health event as well as its corresponding status (acute/chronic) and accompanying diagnosis codes. Prior to October 2015, Marketscan claims data relied on ICD-9-CM diagnosis codes, transitioning to ICD-10-CM diagnosis codes thereafter.

Need to add the following HCCs for acute events:

- 3, 4, 38, 45, 55, 127, 131, 132, 145, 146, 149, 156

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
Adrenal/Pituitary Disorders	Yes	0363, 2510, 25200, 25201, 25202, 25208, 2521, 2528, 2529, 2530, 2531, 2532, 2533, 2534, 2535, 2536, 2537, 2538, 2539, 2540, 2541, 2548, 2549, 2550, 25510, 25511, 25512, 25513, 25514, 2552, 2553, 25541, 25542, 2555, 2556, 2558, 2559, 25801, 25802, 25803, 2581, 2588, 2589, 5881, 58881	A391, E035, E15, E200, E208, E209, E210, E211, E212, E213, E214, E215, E220, E221, E222, E228, E229, E230, E231, E232, E233, E236, E237, E240, E241, E242, E243, E244, E248, E249, E250, E258, E259, E2601, E2602, E2609, E261, E2681, E2689, E269, E270, E271, E272, E273, E2740, E2749, E275, E278, E279, E310, E311, E3120, E3121, E3122, E3123, E318, E319, E320, E321, E328, E329, E344, E892, E893, E896, N251, N2581
Asthma	Yes	49300, 49301, 49302, 49310, 49311, 49312, 49381, 49382, 49390, 49391, 49392	J4520, J4521, J4522, J4530, J4531, J4532, J4540, J4541, J4542, J4550, J4551, J4552, J45901, J45902, J45909, J45990, J45991, J45998
Brain Infections	No	00321, 0065, 01300, 01301, 01302, 01303, 01304, 01305, 01306, 01310, 01311, 01312, 01313, 01314, 01315, 01316, 01320, 01321, 01322, 01323, 01324, 01325, 01326, 01330, 01331, 01332, 01333, 01334, 01335, 01336, 01340, 01341, 01342, 01343, 01344, 01345, 01346, 01350, 01351, 01352, 01353, 01354, 01355, 01356, 01360, 01361, 01362, 01363, 01364, 01365, 01366, 01380, 01381, 01382, 01383, 01384, 01385, 01386, 01390, 01391, 01392, 01393, 01394, 01395, 01396, 0360, 0361, 037, 04500, 04501, 04502, 04503, 04510, 04511, 04512, 04513, 04520, 04521, 04522, 04523, 04590, 04591, 04592, 04593, 0498, 0499, 0520, 0543, 0550, 05601, 05821, 05829, 0620, 0621, 0622, 0623, 0624, 0625, 0628, 0629, 0630, 0631, 0632, 0638, 0639, 064, 0662, 06640, 06641, 06642, 06649, 071, 0722, 09040, 09041, 09042, 09049, 09181, 0940, 0941, 0942, 0943, 09481, 09482, 09483, 09484, 09485, 09486, 09489, 0949, 09882, 10081, 11283, 1142, 11501, 11511, 11591, 3200, 3201, 3202, 3203, 3207, 32081, 32082, 32089, 3209, 3211, 3213, 3214, 3218, 32301, 3231, 3232, 32341, 32351, 32361, 32362, 32381, 3239, 3240, 3241, 3249, 325	A0101, A0221, A066, A170, A171, A1781, A1782, A1783, A1789, A179, A203, A2781, A3211, A3212, A34, A35, A390, A3981, A4281, A4282, A5040, A5041, A5042, A5043, A5044, A5045, A5049, A5141, A5210, A5211, A5212, A5213, A5214, A5215, A5216, A5217, A5219, A522, A523, A5481, A5482, A6921, A800, A801, A802, A8030, A8039, A804, A809, A820, A821, A829, A830, A831, A832, A833, A834, A835, A836, A838, A839, A840, A841, A848, A849, A850, A851, A852, A858, A86, A888, A89, A922, A9230, A9231, A9232, A9239, B004, B0111, B020, B050, B0601, B1001, B1009, B262, B375, B384, B4081, B4281, B431, B5741, B5742, B6011, G000, G001, G002, G003, G008, G009, G01, G02, G0400, G0401, G0402, G042, G0430, G0431, G0432, G0439, G0481, G0490, G053, G060, G061, G062, G07, G08
Breast and Prostate Cancer	Yes	1740, 1741, 1742, 1743, 1744, 1745, 1746, 1748, 1749, 1750, 1759, 179, 1800, 1801, 1808, 1809, 1820, 1821, 1828, 1840, 1841, 1842, 1843, 1844, 1848, 1849, 185, 1880, 1881, 1882, 1883, 1884, 1885, 1886, 1887, 1888, 1889, 1892, 1893, 1894, 1898, 1899, 1900, 1901, 1902, 1903, 1904, 1905, 1906, 1907, 1908, 1909, 1950, 1951, 1952, 1953, 1954, 1955, 1958, 1992, 20100, 20101, 20102, 20103, 20104, 20105, 20106, 20107, 20108, 20110, 20111, 20112, 20113, 20114,	C4A0, C4A10, C4A11, C4A12, C4A20, C4A21, C4A22, C4A30, C4A31, C4A39, C4A4, C4A51, C4A52, C4A59, C4A60, C4A61, C4A62, C4A70, C4A71, C4A72, C4A8, C4A9, C50011, C50012, C50019, C50021, C50022, C50029, C50111, C50112, C50119, C50121, C50122, C50129, C50211, C50212, C50219, C50221, C50222, C50229, C50311, C50312, C50319, C50321, C50322, C50329, C50411, C50412, C50419, C50421, C50422, C50429, C50511, C50512, C50519,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
		20115, 20116, 20117, 20118, 20120, 20121, 20122, 20123, 20124, 20125, 20126, 20127, 20128, 20140, 20141, 20142, 20143, 20144, 20145, 20146, 20147, 20148, 20150, 20151, 20152, 20153, 20154, 20155, 20156, 20157, 20158, 20160, 20161, 20162, 20163, 20164, 20165, 20166, 20167, 20168, 20170, 20171, 20172, 20173, 20174, 20175, 20176, 20177, 20178, 20190, 20191, 20192, 20193, 20194, 20195, 20196, 20197, 20198, 20900, 20901, 20902, 20903, 20910, 20911, 20912, 20913, 20914, 20915, 20916, 20917, 20920, 20921, 20922, 20923, 20924, 20925, 20926, 20927, 20929, 20930, 20931, 20932, 20933, 20934, 20935, 20936, 2250, 2251, 2252, 2253, 2254, 2258, 2259, 2273, 2274, 22802, 2370, 2371, 2373, 2375, 2376, 2379, 2396, 7595, 7596	C50521, C50522, C50529, C50611, C50612, C50619, C50621, C50622, C50629, C50811, C50812, C50819, C50821, C50822, C50829, C50911, C50912, C50919, C50921, C50922, C50929, C510, C511, C512, C518, C519, C52, C530, C531, C538, C539, C540, C541, C542, C543, C548, C549, C55, C577, C578, C579, C61, C661, C662, C669, C670, C671, C672, C673, C674, C675, C676, C677, C678, C679, C680, C681, C688, C689, C6900, C6901, C6902, C6910, C6911, C6912, C6920, C6921, C6922, C6930, C6931, C6932, C6940, C6941, C6942, C6950, C6951, C6952, C6960, C6961, C6962, C6980, C6981, C6982, C6990, C6991, C6992, C760, C761, C762, C763, C7640, C7641, C7642, C7650, C7651, C7652, C768, C7A00, C7A010, C7A011, C7A012, C7A019, C7A020, C7A021, C7A022, C7A023, C7A024, C7A025, C7A026, C7A029, C7A090, C7A091, C7A092, C7A093, C7A094, C7A095, C7A096, C7A098, C7A1, C7A8, C802, C8100, C8101, C8102, C8103, C8104, C8105, C8106, C8107, C8108, C8109, C8110, C8111, C8112, C8113, C8114, C8115, C8116, C8117, C8118, C8119, C8120, C8121, C8122, C8123, C8124, C8125, C8126, C8127, C8128, C8129, C8130, C8131, C8132, C8133, C8134, C8135, C8136, C8137, C8138, C8139, C8140, C8141, C8142, C8143, C8144, C8145, C8146, C8147, C8148, C8149, C8170, C8171, C8172, C8173, C8174, C8175, C8176, C8177, C8178, C8179, C8190, C8191, C8192, C8193, C8194, C8195, C8196, C8197, C8198, C8199, D1802, D320, D321, D329, D330, D331, D332, D333, D334, D337, D339, D352, D353, D354, D420, D421, D429, D430, D431, D432, D433, D434, D438, D439, D443, D444, D445, D446, D447, D496, Q851, Q858, Q859
Cardio-Respiratory Failure	No	42741, 42742, 4275, 514, 5184, 51881, 51882, 51883, 51884, 769, 7703, 7704, 7705, 7707, 77084, 77985, 78550, 78551, 7980, 7981, 7982, 7989, 9584	I462, I468, I469, I4901, I4902, J182, J80, J810, J811, J9600, J9601, J9602, J9610, J9611, J9612, J9620, J9621, J9622, J9690, J9691, J9692, P220, P260, P261, P268, P269, P270, P271, P278, P279, P280, P2810, P2811, P2819, P285, P2981, R570, R579, T794XXA
Chronic Hepatitis	Yes	07022, 07023, 07032, 07033, 07044, 07054, 57140, 57141, 57142, 57149	B180, B181, B182, B188, B189, K730, K731, K732, K738, K739, K754
Chronic Skin Ulcer	Yes	4540, 4542, 45911, 45913, 45931, 45933, 68601, 70710, 70711, 70712, 70713, 70714, 70715, 70719, 7078, 7079	I83001, I83002, I83003, I83004, I83005, I83008, I83009, I83011, I83012, I83013, I83014, I83015, I83018, I83019, I83021, I83022, I83023, I83024, I83025, I83028, I83029, I83201, I83202, I83203, I83204,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
			I83205, I83208, I83209, I83211, I83212, I83213, I83214, I83215, I83218, I83219, I83221, I83222, I83223, I83224, I83225, I83228, I83229, I87011, I87012, I87013, I87019, I87031, I87032, I87033, I87039, I87311, I87312, I87313, I87319, I87331, I87332, I87333, I87339, L88, L97101, L97102, L97103, L97104, L97109, L97111, L97112, L97113, L97114, L97119, L97121, L97122, L97123, L97124, L97129, L97201, L97202, L97203, L97204, L97209, L97211, L97212, L97213, L97214, L97219, L97221, L97222, L97223, L97224, L97229, L97301, L97302, L97303, L97304, L97309, L97311, L97312, L97313, L97314, L97319, L97321, L97322, L97323, L97324, L97329, L97401, L97402, L97403, L97404, L97409, L97411, L97412, L97413, L97414, L97419, L97421, L97422, L97423, L97424, L97429, L97501, L97502, L97503, L97504, L97509, L97511, L97512, L97513, L97514, L97519, L97521, L97522, L97523, L97524, L97529, L97801, L97802, L97803, L97804, L97809, L97811, L97812, L97813, L97814, L97819, L97821, L97822, L97823, L97824, L97829, L97901, L97902, L97903, L97904, L97909, L97911, L97912, L97913, L97914, L97919, L97921, L97922, L97923, L97924, L97929, L98411, L98412, L98413, L98414, L98419, L98421, L98422, L98423, L98424, L98429, L98491, L98492, L98493, L98494, L98499, I70231, I70232, I70233, I70234, I70235, I70238, I70239, I70241, I70242, I70243, I70244, I70245, I70248, I70249, I7025, I70331, I70332, I70333, I70334, I70335, I70338, I70339, I70341, I70342, I70343, I70344, I70345, I70348, I70349, I7035, I70431, I70432, I70433, I70434, I70435, I70438, I70439, I70441, I70442, I70443, I70444, I70445, I70448, I70449, I7045, I70531, I70532, I70533, I70534, I70535, I70538, I70539, I70541, I70542, I70543, I70544, I70545, I70548, I70549, I7055, I70631, I70632, I70633, I70634, I70635, I70638, I70639, I70641, I70642, I70643, I70644, I70645, I70648, I70649, I7065, I70731, I70732, I70733, I70734, I70735, I70738, I70739, I70741, I70742, I70743, I70744, I70745, I70748, I70749, I7075
Congestive Heart Failure	Yes	39891, 40201, 40211, 40291, 40401, 40403, 40411, 40413, 40491, 40493, 4150, 4160, 4161, 4168, 4169, 4170, 4171, 4178, 4179, 4250, 42511, 42518, 4252, 4253, 4254, 4255, 4257, 4258, 4259, 4280, 4281, 42820, 42821, 42822, 42823, 42830, 42831, 42832, 42833,	A3681, B3324, I0981, I110, I130, I132, I2601, I2602, I2609, I270, I271, I272, I2781, I2789, I279, I280, I281, I288, I289, I420, I421, I422, I423, I424, I425, I426, I427, I428, I429, I43, I501, I5020, I5021, I5022, I5023, I5030,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
		42840, 42841, 42842, 42843, 4289, 4290, 4291	I5031, I5032, I5033, I5040, I5041, I5042, I5043, I509, I514, I515
Diabetes w/ Complications	Yes	24940, 24941, 24950, 24951, 24960, 24961, 24970, 24971, 24980, 24981, 24990, 24991, 25040, 25041, 25042, 25043, 25050, 25051, 25052, 25053, 25060, 25061, 25062, 25063, 25070, 25071, 25072, 25073, 25080, 25081, 25082, 25083, 25090, 25091, 25092, 25093, 3572, 36201, 36202, 36203, 36204, 36205, 36206, 36207, 36641	E0821, E0822, E0829, E08311, E08319, E08321, E08329, E08331, E08339, E08341, E08349, E08351, E08359, E0836, E0839, E0840, E0841, E0842, E0843, E0844, E0849, E0851, E0852, E0859, E08610, E08618, E08620, E08621, E08622, E08628, E08630, E08638, E08649, E0865, E0869, E088, E0921, E0922, E0929, E09311, E09319, E09321, E09329, E09331, E09339, E09341, E09349, E09351, E09359, E0936, E0939, E0940, E0941, E0942, E0943, E0944, E0949, E0951, E0952, E0959, E09610, E09618, E09620, E09621, E09622, E09628, E09630, E09638, E09649, E0965, E0969, E098, E1021, E1022, E1029, E10311, E10319, E10321, E10329, E10331, E10339, E10341, E10349, E10351, E10359, E1036, E1039, E1040, E1041, E1042, E1043, E1044, E1049, E1051, E1052, E1059, E10610, E10618, E10620, E10621, E10622, E10628, E10630, E10638, E10649, E1065, E1069, E108, E1121, E1122, E1129, E11311, E11319, E11321, E11329, E11331, E11339, E11341, E11349, E11351, E11359, E1136, E1139, E1140, E1141, E1142, E1143, E1144, E1149, E1151, E1152, E1159, E11610, E11618, E11620, E11621, E11622, E11628, E11630, E11638, E11649, E1165, E1169, E118, E1321, E1322, E1329, E13311, E13319, E13321, E13329, E13331, E13339, E13341, E13349, E13351, E13359, E1336, E1339, E1340, E1341, E1342, E1343, E1344, E1349, E1351, E1352, E1359, E13610, E13618, E13620, E13621, E13622, E13628, E13630, E13638, E13649, E1365, E1369, E138
Diabetes w/o Complications	Yes	24900, 24901, 25000, 25001, 25002, 25003, V5867	E089, E099, E109, E119, E139, Z794
Fibrosis of Lung	Yes	135, 4950, 4951, 4952, 4953, 4954, 4955, 4956, 4957, 4958, 4959, 500, 501, 502, 503, 504, 505, 5060, 5061, 5062, 5063, 5064, 5069, 5080, 5081, 515, 5160, 5161, 5162, 51630, 51631, 51632, 51633, 51634, 51635, 51636, 51637, 5164, 5165, 51661, 51662, 51663, 51664, 51669, 5168, 5169, 5171, 5172, 5178, 5183, 5186	B4481, D860, D862, J60, J61, J620, J628, J630, J631, J632, J633, J634, J635, J636, J64, J65, J660, J661, J662, J668, J670, J671, J672, J673, J674, J675, J676, J677, J678, J679, J680, J681, J682, J683, J684, J688, J689, J700, J701, J82, J8401, J8402, J8403, J8409, J8410, J84111, J84112, J84113, J84114, J84115, J84116, J84117, J8417, J842, J8481, J8482, J8483, J84841, J84842, J84843, J84848, J8489, J849, J99, M3213, M3301, M3311, M3321, M3391, M3481, M3502
Heart Arrhythmias	Yes	4260, 4270, 4271, 4272, 42731, 42732, 42781	I442, I470, I471, I472, I479, I480, I481, I482, I483, I484, I4891, I4892, I492, I495

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
Inflammatory Bowel Disease	Yes	5550, 5551, 5552, 5559, 5560, 5561, 5562, 5563, 5564, 5565, 5566, 5568, 5569	K5000, K50011, K50013, K50014, K50018, K50019, K5010, K50111, K50113, K50114, K50118, K50119, K5080, K50811, K50813, K50814, K50818, K50819, K5090, K50911, K50913, K50914, K50918, K50919, K5100, K51011, K51013, K51014, K51018, K51019, K5120, K51211, K51213, K51214, K51218, K51219, K5130, K51311, K51313, K51314, K51318, K51319, K5140, K51411, K51413, K51414, K51418, K51419, K5150, K51511, K51513, K51514, K51518, K51519, K5180, K51811, K51813, K51814, K51818, K51819, K5190, K51911, K51913, K51914, K51918, K51919, K50012, K50112, K50812, K50912, K51012, K51212, K51312, K51412, K51512, K51812, K51912
Lupus	Yes	0993, 4465, 7100, 7102, 7105, 7108, 7109, 71110, 71111, 71112, 71113, 71114, 71115, 71116, 71117, 71118, 71119, 7144, 71489, 7149, 725	M0230, M02311, M02312, M02319, M02321, M02322, M02329, M02331, M02332, M02339, M02341, M02342, M02349, M02351, M02352, M02359, M02361, M02362, M02369, M02371, M02372, M02379, M0238, M0239, M064, M1200, M12011, M12012, M12019, M12021, M12022, M12029, M12031, M12032, M12039, M12041, M12042, M12049, M12051, M12052, M12059, M12061, M12062, M12069, M12071, M12072, M12079, M1208, M1209, M315, M316, M320, M3210, M3211, M3212, M3213, M3214, M3215, M3219, M328, M329, M3500, M3501, M3502, M3503, M3504, M3509, M351, M353, M355, M358, M359, M368
Major Depressive and Bipolar Disorders	Yes	29600, 29601, 29602, 29603, 29604, 29605, 29606, 29610, 29611, 29612, 29613, 29614, 29615, 29616, 29620, 29621, 29622, 29623, 29624, 29625, 29626, 29630, 29631, 29632, 29633, 29634, 29635, 29636, 29640, 29641, 29642, 29643, 29644, 29645, 29646, 29650, 29651, 29652, 29653, 29654, 29655, 29656, 29660, 29661, 29662, 29663, 29664, 29665, 29666, 2967, 29680, 29681, 29682, 29689, 29690, 29699, E9500, E9501, E9502, E9503, E9504, E9505, E9506, E9507, E9508, E9509, E9510, E9511, E9518, E9520, E9521, E9528, E9529, E9530, E9531, E9538, E9539, E954, E9550, E9551, E9552, E9553, E9554, E9555, E9556, E9557, E9559, E956, E9570, E9571, E9572, E9579, E9580, E9581, E9582, E9583, E9584, E9585, E9586, E9587, E9588, E9589, E959	F3010, F3011, F3012, F3013, F302, F303, F304, F308, F309, F310, F3110, F3111, F3112, F3113, F312, F3130, F3131, F3132, F314, F315, F3160, F3161, F3162, F3163, F3164, F3170, F3171, F3172, F3173, F3174, F3175, F3176, F3177, F3178, F3181, F3189, F319, F322, F323, F332, F333, T1491, T360X2A, T360X2S, T361X2A, T361X2S, T362X2A, T362X2S, T363X2A, T363X2S, T364X2A, T364X2S, T365X2A, T365X2S, T366X2A, T366X2S, T367X2A, T367X2S, T368X2A, T368X2S, T3692XA, T3692XS, T370X2A, T370X2S, T371X2A, T371X2S, T372X2A, T372X2S, T373X2A, T373X2S, T374X2A, T374X2S, T375X2A, T375X2S, T378X2A, T378X2S, T3792XA, T3792XS, T380X2A, T380X2S, T381X2A, T381X2S, T382X2A, T382X2S, T383X2A, T383X2S, T384X2A, T384X2S, T385X2A, T385X2S, T386X2A, T386X2S, T387X2A, T387X2S, T38802A, T38802S, T38812A, T38812S, T38892A, T38892S, T38902A, T38902S, T38992A, T38992S, T39012A, T39012S,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
			T39092A, T39092S, T391X2A, T391X2S, T392X2A, T392X2S, T39312A, T39312S, T39392A, T39392S, T394X2A, T394X2S, T398X2A, T398X2S, T3992XA, T3992XS, T400X2A, T400X2S, T401X2A, T401X2S, T402X2A, T402X2S, T403X2A, T403X2S, T404X2A, T404X2S, T405X2A, T405X2S, T40602A, T40602S, T40692A, T40692S, T407X2A, T407X2S, T408X2A, T408X2S, T40902A, T40902S, T40992A, T40992S, T410X2A, T410X2S, T411X2A, T411X2S, T41202A, T41202S, T41292A, T41292S, T413X2A, T413X2S, T4142XA, T4142XS, T415X2A, T415X2S, T420X2A, T420X2S, T421X2A, T421X2S, T422X2A, T422X2S, T423X2A, T423X2S, T424X2A, T424X2S, T425X2A, T425X2S, T426X2A, T426X2S, T4272XA, T4272XS, T428X2A, T428X2S, T43012A, T43012S, T43022A, T43022S, T431X2A, T431X2S, T43202A, T43202S, T43212A, T43212S, T43222A, T43222S, T43292A, T43292S, T433X2A, T433X2S, T434X2A, T434X2S, T43502A, T43502S, T43592A, T43592S, T43602A, T43602S, T43612A, T43612S, T43622A, T43622S, T43632A, T43632S, T43692A, T43692S, T438X2A, T438X2S, T4392XA, T4392XS, T440X2A, T440X2S, T441X2A, T441X2S, T442X2A, T442X2S, T443X2A, T443X2S, T444X2A, T444X2S, T445X2A, T445X2S, T446X2A, T446X2S, T447X2A, T447X2S, T448X2A, T448X2S, T44902A, T44902S, T44992A, T44992S, T450X2A, T450X2S, T451X2A, T451X2S, T452X2A, T452X2S, T453X2A, T453X2S, T454X2A, T454X2S, T45512A, T45512S, T45522A, T45522S, T45602A, T45602S, T45612A, T45612S, T45622A, T45622S, T45692A, T45692S, T457X2A, T457X2S, T458X2A, T458X2S, T4592XA, T4592XS, T460X2A, T460X2S, T461X2A, T461X2S, T462X2A, T462X2S, T463X2A, T463X2S, T464X2A, T464X2S, T465X2A, T465X2S, T466X2A, T466X2S, T467X2A, T467X2S, T468X2A, T468X2S, T46902A, T46902S, T46992A, T46992S, T470X2A, T470X2S, T471X2A, T471X2S, T472X2A, T472X2S, T473X2A, T473X2S, T474X2A, T474X2S, T475X2A, T475X2S, T476X2A, T476X2S, T477X2A, T477X2S, T478X2A, T478X2S, T4792XA, T4792XS, T480X2A, T480X2S, T481X2A, T481X2S, T48202A, T48202S, T48292A, T48292S, T483X2A, T483X2S, T484X2A, T484X2S, T485X2A, T485X2S, T486X2A, T486X2S, T48902A, T48902S, T48992A, T48992S,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
			T490X2A, T490X2S, T491X2A, T491X2S, T492X2A, T492X2S, T493X2A, T493X2S, T494X2A, T494X2S, T495X2A, T495X2S, T496X2A, T496X2S, T497X2A, T497X2S, T498X2A, T498X2S, T4992XA, T4992XS, T500X2A, T500X2S, T501X2A, T501X2S, T502X2A, T502X2S, T503X2A, T503X2S, T504X2A, T504X2S, T505X2A, T505X2S, T506X2A, T506X2S, T507X2A, T507X2S, T508X2A, T508X2S, T50902A, T50902S, T50992A, T50992S, T50A12A, T50A12S, T50A22A, T50A22S, T50A92A, T50A92S, T50B12A, T50B12S, T50B92A, T50B92S, T50Z12A, T50Z12S, T50Z92A, T50Z92S, T510X2A, T510X2S, T511X2A, T511X2S, T512X2A, T512X2S, T513X2A, T513X2S, T518X2A, T518X2S, T5192XA, T5192XS, T520X2A, T520X2S, T521X2A, T521X2S, T522X2A, T522X2S, T523X2A, T523X2S, T524X2A, T524X2S, T528X2A, T528X2S, T5292XA, T5292XS, T530X2A, T530X2S, T531X2A, T531X2S, T532X2A, T532X2S, T533X2A, T533X2S, T534X2A, T534X2S, T535X2A, T535X2S, T536X2A, T536X2S, T537X2A, T537X2S, T5392XA, T5392XS, T540X2A, T540X2S, T541X2A, T541X2S, T542X2A, T542X2S, T543X2A, T543X2S, T5492XA, T5492XS, T550X2A, T550X2S, T551X2A, T551X2S, T560X2A, T560X2S, T561X2A, T561X2S, T562X2A, T562X2S, T563X2A, T563X2S, T564X2A, T564X2S, T565X2A, T565X2S, T566X2A, T566X2S, T567X2A, T567X2S, T56812A, T56812S, T56892A, T56892S, T5692XA, T5692XS, T570X2A, T570X2S, T571X2A, T571X2S, T572X2A, T572X2S, T573X2A, T573X2S, T578X2A, T578X2S, T5792XA, T5792XS, T5802XA, T5802XS, T5812XA, T5812XS, T582X2A, T582X2S, T588X2A, T588X2S, T5892XA, T5892XS, T590X2A, T590X2S, T591X2A, T591X2S, T592X2A, T592X2S, T593X2A, T593X2S, T594X2A, T594X2S, T595X2A, T595X2S, T596X2A, T596X2S
Multiple Sclerosis	Yes	340, 3410, 3411	G35, G360, G370, G375
Personality Disorder	Yes	30012, 30013, 30014, 30015, 3006, 3010, 30110, 30111, 30112, 30113, 30120, 30121, 30122, 3013, 3014, 30150, 30151, 30159, 3016, 3017, 30181, 30182, 30183, 30184, 30189, 3019	F21, F440, F441, F4481, F481, F600, F601, F602, F603, F604, F605, F606, F607, F6081, F6089, F609
Rheumatoid Arthritis	Yes	1361, 4460, 4461, 44620, 44621, 44629, 4463, 4464, 4466, 4467, 6960, 7101, 7103, 7104, 71120, 71121, 71122, 71123, 71124, 71125, 71126, 71127, 71128, 71129, 7140,	L4050, L4051, L4052, L4053, L4054, L4059, M0500, M05011, M05012, M05019, M05021, M05022, M05029, M05031, M05032, M05039, M05041, M05042, M05049,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
		7141, 7142, 71430, 71431, 71432, 71433, 71481, 7200	M05051, M05052, M05059, M05061, M05062, M05069, M05071, M05072, M05079, M0509, M0510, M05111, M05112, M05119, M05121, M05122, M05129, M05131, M05132, M05139, M05141, M05142, M05149, M05151, M05152, M05159, M05161, M05162, M05169, M05171, M05172, M05179, M0519, M0520, M05211, M05212, M05219, M05221, M05222, M05229, M05231, M05232, M05239, M05241, M05242, M05249, M05251, M05252, M05259, M05261, M05262, M05269, M05271, M05272, M05279, M0529, M0530, M05311, M05312, M05319, M05321, M05322, M05329, M05331, M05332, M05339, M05341, M05342, M05349, M05351, M05352, M05359, M05361, M05362, M05369, M05371, M05372, M05379, M0539, M0540, M05411, M05412, M05419, M05421, M05422, M05429, M05431, M05432, M05439, M05441, M05442, M05449, M05451, M05452, M05459, M05461, M05462, M05469, M05471, M05472, M05479, M0549, M0550, M05511, M05512, M05519, M05521, M05522, M05529, M05531, M05532, M05539, M05541, M05542, M05549, M05551, M05552, M05559, M05561, M05562, M05569, M05571, M05572, M05579, M0559, M0560, M05611, M05612, M05619, M05621, M05622, M05629, M05631, M05632, M05639, M05641, M05642, M05649, M05651, M05652, M05659, M05661, M05662, M05669, M05671, M05672, M05679, M0569, M0570, M05711, M05712, M05719, M05721, M05722, M05729, M05731, M05732, M05739, M05741, M05742, M05749, M05751, M05752, M05759, M05761, M05762, M05769, M05771, M05772, M05779, M0579, M0580, M05811, M05812, M05819, M05821, M05822, M05829, M05831, M05832, M05839, M05841, M05842, M05849, M05851, M05852, M05859, M05861, M05862, M05869, M05871, M05872, M05879, M0589, M059, M0600, M06011, M06012, M06019, M06021, M06022, M06029, M06031, M06032, M06039, M06041, M06042, M06049, M06051, M06052, M06059, M06061, M06062, M06069, M06071, M06072, M06079, M0608, M0609, M061, M0620, M06211, M06212, M06219, M06221, M06222, M06229, M06231, M06232, M06239, M06241,

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
			M06242, M06249, M06251, M06252, M06259, M06261, M06262, M06269, M06271, M06272, M06279, M0628, M0629, M0630, M06311, M06312, M06319, M06321, M06322, M06329, M06331, M06332, M06339, M06341, M06342, M06349, M06351, M06352, M06359, M06361, M06362, M06369, M06371, M06372, M06379, M0638, M0639, M0680, M06811, M06812, M06819, M06821, M06822, M06829, M06831, M06832, M06839, M06841, M06842, M06849, M06851, M06852, M06859, M06861, M06862, M06869, M06871, M06872, M06879, M0688, M0689, M069, M0800, M08011, M08012, M08019, M08021, M08022, M08029, M08031, M08032, M08039, M08041, M08042, M08049, M08051, M08052, M08059, M08061, M08062, M08069, M08071, M08072, M08079, M0808, M0809, M081, M0820, M08211, M08212, M08219, M08221, M08222, M08229, M08231, M08232, M08239, M08241, M08242, M08249, M08251, M08252, M08259, M08261, M08262, M08269, M08271, M08272, M08279, M0828, M0829, M083, M0840, M08411, M08412, M08419, M08421, M08422, M08429, M08431, M08432, M08439, M08441, M08442, M08449, M08451, M08452, M08459, M08461, M08462, M08469, M08471, M08472, M08479, M0848, M0880, M08811, M08812, M08819, M08821, M08822, M08829, M08831, M08832, M08839, M08841, M08842, M08849, M08851, M08852, M08859, M08861, M08862, M08869, M08871, M08872, M08879, M0888, M0889, M0890, M08911, M08912, M08919, M08921, M08922, M08929, M08931, M08932, M08939, M08941, M08942, M08949, M08951, M08952, M08959, M08961, M08962, M08969, M08971, M08972, M08979, M0898, M0899, M300, M301, M302, M303, M308, M310, M311, M312, M3130, M3131, M314, M317, M3300, M3301, M3302, M3309, M3310, M3311, M3312, M3319, M3320, M3321, M3322, M3329, M3390, M3391, M3392, M3399, M340, M341, M342, M3481, M3482, M3483, M3489, M349, M352, M360, M450, M451, M452, M453, M454, M455, M456, M457, M458, M459, M488X1, M488X2, M488X3, M488X4, M488X5, M488X6, M488X7, M488X8, M488X9

Service Type	Chronic?	Diagnosis Codes (ICD-9-CM)	Diagnosis Codes (ICD-10-CM)
Seizures	Yes	1361, 4460, 4461, 44620, 44621, 44629, 4463, 4464, 4466, 4467, 6960, 7101, 7103, 7104, 71120, 71121, 71122, 71123, 71124, 71125, 71126, 71127, 71128, 71129, 7140, 7141, 7142, 71430, 71431, 71432, 71433, 71481, 7200	G40001, G40009, G40011, G40019, G40101, G40109, G40111, G40119, G40201, G40209, G40211, G40219, G40301, G40309, G40311, G40319, G40401, G40409, G40411, G40419, G40501, G40509, G40801, G40802, G40803, G40804, G40811, G40812, G40813, G40814, G40821, G40822, G40823, G40824, G4089, G40901, G40909, G40911, G40919, G40A01, G40A09, G40A11, G40A19, G40B01, G40B09, G40B11, G40B19, P90, R5600, R5601, R561, R569
Sepsis and Shock	No	0031, 0202, 0223, 0362, 0380, 03810, 03811, 03812, 03819, 0382, 0383, 03840, 03841, 03842, 03843, 03844, 03849, 0388, 0389, 04082, 0545, 77181, 78552, 78559, 99590, 99591, 99592, 99593, 99594	A021, A207, A227, A267, A327, A392, A393, A394, A400, A401, A403, A408, A409, A4101, A4102, A411, A412, A413, A414, A4150, A4151, A4152, A4153, A4159, A4181, A4189, A419, A427, A483, A5486, B007, B377, P360, P3610, P3619, P362, P3630, P3639, P364, P365, P368, P369, R571, R578, R6510, R6511, R6520, R6521
Thyroid Cancer	Yes	1720, 1721, 1722, 1723, 1724, 1725, 1726, 1727, 1728, 1729, 1860, 1869, 1871, 1872, 1873, 1874, 1875, 1876, 1877, 1878, 1879, 193, 1941, 1945, 1946, 1948, 1949, 1991, 23770, 23771, 23772, 23773, 23779, 2592	C430, C4310, C4311, C4312, C4320, C4321, C4322, C4330, C4331, C4339, C434, C4351, C4352, C4359, C4360, C4361, C4362, C4370, C4371, C4372, C438, C439, C600, C601, C602, C608, C609, C6200, C6201, C6202, C6210, C6211, C6212, C6290, C6291, C6292, C6300, C6301, C6302, C6310, C6311, C6312, C632, C637, C638, C639, C73, C750, C754, C755, C758, C759, C801, D030, D0310, D0311, D0312, D0320, D0321, D0322, D0330, D0339, D034, D0351, D0352, D0359, D0360, D0361, D0362, D0370, D0371, D0372, D038, D039, E340, Q8500, Q8501, Q8502, Q8503, Q8509
Unstable Angina	No	41000, 41002, 41010, 41012, 41020, 41022, 41030, 41032, 41040, 41042, 41050, 41052, 41060, 41062, 41070, 41072, 41080, 41082, 41090, 41092, 4110, 4111, 41181, 41189	I200, I230, I231, I232, I233, I236, I237, I238, I240, I241, I248, I249, I25110, I25700, I25710, I25720, I25730, I25750, I25760, I25790
Viral Meningitis	No	0470, 0471, 0478, 0479, 048, 0490, 0491, 0530, 05472, 0721, 3212, 3220, 3221, 3222, 3229	A870, A871, A872, A878, A879, A880, B003, B010, B021, B051, B0602, B261, B2702, B2712, B2782, B2792, D8681, G030, G031, G032, G038, G039

When assigning HCCs, I exclude diagnoses associated with the following place of service and procedure codes, due to their high potential for false positive diagnoses:

Excluded ranges		
Place of Service Codes		
	12	Private residence home
	31	Skilled nursing facility
	32	Nursing home
	33	Custodial care
	34	Hospice
	41	Ambulance – land
	42	Ambulance – other
	65	Renal dialysis
	81	Independent lab
	99	Unknown
Procedure Codes		
	36415-36416	Drawing blood
	70000-76999	X-ray and ultrasound
	78000-78999	Imaging
	80000-87999	Lab tests
	88000-88099	Autopsy
	88104-88299	Cytopathology
	88300-88399	Surgical Pathology
	88720-88741	In Vivo
	92551-92569	Hearing tests
	93000-93350	ECG and ultrasound
	99000-99001	Specimen handling
	A0021-A0999	Ambulance
	A4206-A999	Medical and surgical supplies
	B4304-B999	Enteral Supplies
	G0001	Drawing blood
	E0100-E9999	Durable medical equipment
	K0001-K9999	Wheelchairs and accessories
	L0100-L4599	Orthotics
	L5000-L9900	Prosthetics
	P2028-P9999	Pathology and Lab
	R0070-R0076	Radiology

Table 7 identifies additional demographic information, as well as illustrating the balance in my sample across households with and without a chronic condition in the family. The table also shows the frequency of the various chronic conditions utilized in my sample. Households

with chronic conditions are not markedly different in terms of age or sex composition or family size, but do incur significantly higher medical costs in a year. They are not, however, more likely to switch insurance plans from year to year. There is wide variation in the onset of chronic illnesses, the three most common illnesses are asthma, major depressive disorder, and diabetes.

	Full Sample	Households with chronic conditions
Demographics & Utilization		
Enrollee age	30.87 (0.008)	29.61 (0.046)
% female enrollees	50.17 (0.000)	50.46 (0.001)
Mean [median] total spending	\$2,504.41 [\$679.75] (4.510)	\$3,378.17 [\$957.52] (23.752)
Mean [median] OOP spending	\$443.07 [\$109.66] (0.525)	\$531.93 [\$151.18] (3.153)
% switching plans ever	–	
Incidence of chronic illness (per 1,000 individuals)		
Adrenal & pituitary disorders	0.22	7.35
Asthma	2.93	96.08
Breast/prostate cancer	0.35	11.58
Chronic hepatitis	0.10	3.23
Chronic skin condition	0.23	7.46
Congestive heart failure	0.14	4.52
Diabetes with complications	0.39	12.72
Diabetes without complications	1.18	38.57
Fibrosis of lung	0.46	15.10
Heart arrhythmias	0.00	0.00
Inflammatory bowel disease	0.14	4.65
Lupus	0.16	5.20
Major depressive/biploar disorder	1.62	52.76
Multiple sclerosis	1.10	36.17
Personality disorder	0.09	2.81
Rheumatoid arthritis	0.17	5.70
Seizures	0.30	9.82
Thyroid cancer	0.14	4.69
N_{families}	353,403	52,747
$N_{\text{individuals}}$	1,087,353	165,694

Table 7. Relative Incidence of Chronic Conditions

A.3 Identifying Chronic Care Costs

An important component of my model is that chronic illnesses correspond to annual diagnostic and maintenance costs that are not strictly choice variables, in the sense that certain health utilization is more or less required. I identify the costs associated with these illnesses from the claims data as procedures which have the major diagnosis listed on that line item. Additionally, in conjunction with Rebecca Hughes, MD, I identify specific therapeutic classes for prescription medications that are associated with treating each chronic condition, shown in the table below. Empirical distributions of these estimated diagnostic and maintenance costs for each major health event are available upon request.

Major Health Event	Therapeutic Classes
Breast and Prostate Cancer	Antineoplastic Agents Hematopoietic Agents Antiemetics Adrenals Androgens Immunosuppresants Antiinf S/MM, Antibiotics & Comb Antiinf S/MM, Antivirals & Comb Antiinf S/MM, Antifungals & Comb Antiinf S/MM, Scabic/Pediculic Antiinf S/MM, Antiinf Local Misc Antineoplastics S/MM Phosphodiesterase Inhibitors Hormone-Modifying Therapy Molecular Targeted Therapy
Thyroid Cancer	Antineoplastic Agents Antihyperlipidemic Drugs Antiemetics Adrenals Androgens Thy/Antithy, Thyroid Hormones Thy/Anithy, Anithyroid Hormones Immunosuppresants Antineoplastics S/MM Phosphodiesterase Inhibitors
Diabetes (w/ or w/o Complications)	Diabetes Mell/Diab Supply, NEC Antidiabetic Agents, Insulin Antidiabetic Ag, Sulfonylureas Antidiabetic Agents, Misc Antidiabetic Ag, Meglitinides Antidiabetic Ag, SGLT Inhibit Antidiabetic Ag, TZD
Adrenal/Pituitary Disorders	Adrenals Androgens Estrogens & Comb Parathyroid Hormones Pituitary Hormones Progestins Thy/Antithy, Thyroid Hormones Thy/Antithy, Antithyroid Hormones
Chronic Hepatitis	Antivirals Adrenals
Inflammatory Bowel Disease	Antineoplastic Agents Gastrointestinal Drugs Misc. Adrenals Immunosuppresants

Major Health Event	Therapeutic Classes
Rheumatoid Arthritis	Antineoplastic Agents Adrenals Immunosuppresants
Lupus	Quinolones, NEC Adrenals
MDD and Bipolar	Stimulant, Non-Amphetamine Anticonvulsants, Misc. Psychotherapeutics, Antidepressants Psychotherapeutics, Tranq/Antipsychotics Stimulant, Amphetamine Stimulant, Non-Amphetamine ASH, Barbiturates ASH, Benzodiazepines Anxiolytic/Sedative/Hypnotic Antimaniac Agents
Personality Disorder	Stimulant, Non-Amphetamine Anticonvulsants, Misc. Psychotherapeutics, Antidepressants Psychotherapeutics, Tranq/Antipsychotics Stimulant, Amphetamine Stimulant, Non-Amphetamine ASH, Barbiturates ASH, Benzodiazepines Anxiolytic/Sedative/Hypnotic Antimaniac Agents
Multiple Sclerosis	Sympathomimetic Agents Stimulant, Amphetamine Stimulant, Non-Amphetamine Adrenals
Seizures	Anticonvulsants, Misc.
Congestive Heart Failure	Cardiac Drugs, NEC Cardiac, ACE Inhibitors Cardiac, Alpha-Beta Blockers Cardiac, Beta Blockers Cardiac, Calcium Channel Antihyperlipidemic Drugs Repl Preps, Potassium Supp Diuretics, Loop Diuretics Diuretics, Misc. Diuretics, Osmotic Diuretics, Potassium-Sparing Diuretics, Thiazides Diuretics, Carb Anhydrase Inhib
Heart Arrhythmias	Coag/Anticoag, Anticoagulants Coag/Anticoag, Antiheparin Agents Coag/Anticoag, Hemostatics

Major Health Event	Therapeutic Classes
	Cardiac Drugs, NEC Cardiac, ACE Inhibitors Cardiac, Alpha-Beta Blockers Cardiac, Beta Blockers Cardiac, Calcium Channel Antihyperlipidemic Drugs
Asthma	Sympathomimetic Agents Adrenals Leukotriene Modifiers
Fibrosis of Lung	Sympathomimetic Agents Adrenals
Chronic Skin Ulcer	Adrenals

A.4 Identifying Cardiovascular Preventive Medications

Cardiovascular preventive medications are identified using the following set of therapeutic classes

Therapeutic Class	Example Medications
Angiotensin-converting-enzyme (ACE) Inhibitors	benazepril (Lotensin), zofenopril, perindopril
Anticoagulants	warfarin (Coumadin), heparin
Antihyperlipidemic Agents	atorvastatin (Lipitor), fluvastatin, lovastatin
Beta Blockers	propranolol (Inderal), pronethalol
Hypotensive Agents	midodrine (Amatine), norepenephrine

Table 8. Therapeutic Classes Used in Identifying Cardiovascular Preventive Medications

A.5 Identifying Low-Value Services

Low value services are identified at the procedure level using CPT codes for medical procedures and therapeutic classes for prescription medications. I aggregate these services into five broad categories, as illustrated in the following table.

Category	Service	CPT Codes / Therapeutic Classes	Additional restrictions (age/sex restrictions, diagnosis or procedure codes)
All Pediatric	Vitamin D Screening	82306,82652	Age < 18
All Pediatric	Cervical Cancer Screening	87620,87621,87622, 87623, 87624, 87625, 88141, 88142, 88143, 88147, 88148, 88150, 88152, 88153, 88154, 88155, 88164, 88165,88166, 88167, 88174, 88175, G0123, G0124, G0141, G0143, G0144, G0145, G0147, G0148, P3000, P3001, Q0091	Age < 18, age >= 14, female
All Pediatric	Head imaging for headache	70450,70460,70470,70551,70552,70553	Age < 18, Diagnosis codes: 3390, 3391, 3460, 3461, 3462, 3464, 3465, 3467, 3468, 3469, 7840, 3393, G440, G441, G442, G444, G430, G431, G435, G437, G438, G439, 30781,33983, 33984, 33985, R51, R510, R519, G4483, G4484, G4485
All Pediatric	Antibiotics for upper respiratory infections	Antibiotics (multiple classes)	Diagnosis codes: 460,465, J00, J06, H65, H60, H61, H62, 3810, 3814
All Pediatric	Antibiotics for bronchiolitis	Antibiotics (multiple classes)	Diagnosis codes: 46611,46619, J210, J218
All Pediatric	Cough or cold medicine	Antitussives, Expectorants, Mucolytics, Cough/Cold Combinations	Age < 6
Adult Drugs	Opioids to treat migraines	Opiate Agonists, Opiate Part Agonists, Opiate Antagonists	Diagnosis codes: 346**, G43**
Adult Imaging	Head imaging for headache	70450,70460,70470,70551,70552,70553	Diagnosis codes: 3390, 3391, 3460, 3461, 3462, 3464, 3465, 3467, 3468, 3469, 7840, 3393, G440, G441, G442, G444, G430, G431, G435, G437, G438, G439, 30781,33983, 33984, 33985, R51, R510, R519, G4483, G4484, G4485
Adult Imaging	Imaging for lower-back pain	72010, 72020,72052, 72100, 72110, 72114,72120, 72200, 72202, 72220, 72131, 72132, 72133, 72141, 72142, 72146, 72147, 72148,72149, 72156, 72157, 72158	Diagnosis codes: 7213, 7226, 7242, 7243, 7244,7245, 7246,7385, 7393,7394, 8460, 8461, 8462, 8463, 8468, 8469, 8472, M432, M512, M513, M518, M533, M545, M541, M543, M998, 72190, 72210, 72252, 72293, 72402,72470, 72471, 72479, M47817, M532X7, M9903, M9904, S338XXA, S336XXA, S339XXA, S335XXA, M47819, M4647, M4806, M532X8

Category	Service	CPT Codes / Therapeutic Classes	Additional restrictions (age/sex restrictions, diagnosis or procedure codes)
Adult Imaging	Screening for carotid artery disease	36222, 36223, 36224, 70498, 70547, 70548,70549, 93880, 93882, 3100F	Diagnosis codes: 430, 431, 434,436,781, I63, I66, R25, R26, R27, R29, R47, G45, H34, R55, R20, 4350, 4351, 4353, 4358, 359,3623, 7802, 7820, I609, I619, 43301, 43311, 43321, 43331,43381, 43391, 99702, V1254, 36284, 78451, 78452, 78459, I6789, I67848, I97811, I97821, Z8673, H3582
Adult Imaging	Cardiac imaging	0144T, 0145T, 0146T, 0147T, 0148T, 0149T, 0150T, 75552, 75553, 75554, 75555, 75556, 75557, 75558, 75559, 75561, 75562, 75565, 75571, 75572, 75573, 75574, 78451, 78452, 78453, 78454, 78460, 78461, 78464, 78465, 78478, 78480, 78459, 78481, 78483, 78491, 78492, 78494, 78496, 78499	
Adult Screening	Vitamin D Screening	82306,82652	
Adult Screening	Cardiac testing for low-risk patients	93015, 93016, 93017, 93018, 93350, 93351,78451, 78452, 78453, 78454, 78460, 78461,78464, 78465, 78472, 78473, 78481, 78483,78491, 78492, 93303, 93304, 93306, 93307, 93308, 93312,93315, 93318, 3120F, 93000, 93005, 93010, G0366, G0367, G0368, G0403, G0404, G0405	
Adult Screening	Pre-operative testing before low-risk surgery	71010, 71015, 71020, 71021, 71022, 71023, 71030, 71034, 71035, 93303, 93304, 93306, 93307, 93308, 93312, 93315, 93318, 94010, 78451, 78452, 78453, 78454, 78460, 78461, 78464, 78465, 78472, 78473, 78481, 78483, 78491, 78492, 93015, 93016, 93017, 93018, 93350, 93351	Procedure codes for surgery: 19120, 19125, 47562, 47563, 49560, 58558
Adult Surgery	Arthroscopic surgery for knee osteoarthritis	29877, 29879, G0289	Diagnosis codes: 8360, 8361, 8362, 7170, S832, 71741, M23202, M23205

B Additional Reduced Form Results

B.1 Robustness of Results to Transformations

Table 10 demonstrates that results are robust to two standard transformations for skewed spending variables: the inverse hyperbolic sine transform, as reported in the main text, and the $\log(y+1)$ transformation.

B.2 Robustness of Results to Event Study Specification

Table 9 shows the standard difference-in-differences coefficients for each of the main event study regressions performed in the main text.

Outcome Variable	Treated$_f \times$ Post$_t$	Adusted R^2	N
OOP, chronic, full sample	0.09***	0.51	1,538,162
	(0.012)		
OOP, chronic, zero-deductible plans	0.13***	0.55	390,335
	(0.020)		
OOP, acute, full sample	0.42***	0.50	1,374,481
	(0.031)		
OOP, acute, zero-deductible plans	0.39***	0.54	358,860
	(0.063)		
Billed spending, wellness visits, full sample	0.13***	0.43	1,538,162
	(0.013)		
Billed spending, wellness, zero-deductible plans	0.18***	0.40	390,335
	(0.027)		
Cardiovascular Prescriptions, Prob(fill scrip)	2.56	0.42	439,542
	(1.501)		
Cardiovascular Prescriptions, PDC	1.46	0.48	439,542
	(1.142)		
Billed Spending, Low Value Services	0.06***	0.20	1,538,162
	(0.011)		
Utilization, Low Value Services	0.03***	0.20	1,538,162
	(0.008)		

Table 9. Difference in Differences Coefficients, Main Regressions

Table 11 implements the robust alternative event study estimator described by de Chaisemartin and D'Haultfoeuille (2019).

	OOP, chronic diagnosis		OOP, acute diagnosis		Wellness spending		Low-value spending	
	$sinh^{-1}(y)$	$log(y+1)$	$sinh^{-1}(y)$	$log(y+1)$	$sinh^{-1}(y)$	$log(y+1)$	$sinh^{-1}(y)$	$log(y+1)$
$t-5$	-0.02	-0.02	-0.11	-0.10	-0.09**	-0.08**	-0.06*	-0.05*
	(0.028)	(0.026)	(0.070)	(0.064)	(0.031)	(0.028)	(0.033)	(0.03)
$t-4$	0.02	0.01	-0.11	-0.10	-0.03	-0.03	-0.04	-0.03
	(0.024)	(0.022)	(0.059)	(0.055)	(0.026)	(0.024)	(0.028)	(0.024)
$t-3$	0.00	0.00	-0.02	-0.02	-0.02	-0.02	-0.03	-0.02
	(0.020)	(0.018)	(0.052)	(0.048)	(0.022)	(0.020)	(0.023)	(0.021)
$t-2$	-0.00	-0.00	-0.07	-0.06	-0.03	-0.03	-0.01	-0.01
	(0.017)	(0.015)	(0.045)	(0.042)	(0.019)	(0.017)	(0.020)	(0.018)
$t-1$	–	–	–	–	–	–	–	–
t	0.08***	0.07***	-0.01	-0.01	0.12***	0.11***	0.05*	0.04*
	(0.014)	(0.013)	(0.041)	(0.037)	(0.016)	(0.015)	(0.018)	(0.016)
$t+1$	0.10***	0.10***	0.10*	0.09*	0.09***	0.08***	0.05**	0.04**
	(0.016)	(0.014)	(0.047)	(0.043)	(0.017)	(0.016)	(0.019)	(0.017)
$t+2$	0.10***	0.09***	0.06	0.07	0.10***	0.10***	0.05*	0.04*
	(0.018)	(0.017)	(0.055)	(0.050)	(0.020)	(0.018)	(0.021)	(0.019)
$t+3$	0.09***	0.08***	0.10	0.09	0.11***	0.10***	0.04	0.04
	(0.018)	(0.019)	(0.062)	(0.057)	(0.022)	(0.020)	(0.024)	(0.021)
$t+4$	0.08***	0.08***	0.14	0.13	0.13***	0.12***	0.09**	0.08**
	(0.025)	(0.022)	(0.074)	(0.068)	(0.025)	(0.023)	(0.028)	(0.024)
$t+5$	0.07***	0.06*	0.12	0.12	0.10***	0.09***	0.12***	0.11***
	(0.030)	(0.028)	(0.088)	(0.081)	(0.030)	(0.027)	(0.033)	(0.029)
R^2	0.51	0.52	0.50	0.51	0.43	0.44	0.20	0.20
N	1,538,161	1,538,161	1,374,359	1,374,359	1,538,161	1,538,161	1,538,161	1,538,161

Notes: This table presents estimates for the main event study regression results reported in the paper. The first column of each pair of results are the results shown graphically in the text, while the second column uses the log transformation. Standard errors are clustered at the household level.

Table 10. Robustness: Inverse Hyperbolic Sine & Log Transformations

	OOP spending, chronic		OOP spending, acute		Billed spending, wellness		Prescriptions, PDC		Billed Spending, low-value care	
	Standard	Robust	Standard	Robust	Standard	Robust	Standard	Robust	Standard	Robust
t	0.08***	0.06***	-0.01	-0.02	0.12***	0.11***	5.96***	1.33	0.05*	0.04**
	(0.014)	(0.013)	(0.041)	(0.039)	(0.016)	(0.02)	(1.100)	(1.013)	(0.018)	(0.016)
$t+1$	0.10***	0.08***	0.10*	0.06	0.09***	0.07***	2.82**	2.24**	0.05**	0.04**
	(0.016)	(0.016)	(0.047)	(0.043)	(0.017)	(0.018)	(1.189)	(1.137)	(0.019)	(0.016)
$t+2$	0.10***	0.06***	0.06	0.03	0.10***	0.07***	-1.97	0.12	0.05*	0.04*
	(0.018)	(0.018)	(0.055)	(0.055)	(0.020)	(0.019)	(1.397)	(1.444)	(0.021)	(0.021)
$t+3$	0.09***	0.04**	0.10	0.04	0.11***	0.06***	-3.95*	0.46	0.04	0.02
	(0.018)	(0.021)	(0.062)	(0.063)	(0.022)	(0.021)	(1.701)	(1.949)	(0.024)	(0.022)
$t+4$	0.08***	0.02	0.14	0.11	0.13***	0.07**	-6.67***	-2.05	0.09**	0.06**
	(0.025)	(0.025)	(0.074)	(0.09)	(0.025)	(0.021)	(2.051)	(2.685)	(0.028)	(0.028)
$t+5$	0.07***	-0.02	0.12	0.12	0.10***	0.02	-5.42*	-0.20	0.12***	0.11***
	(0.030)	(0.031)	(0.088)	(0.107)	(0.030)	(0.034)	(2.375)	(3.461)	(0.033)	(0.038)
N	1,538,161	1,538,161	1,374,359	1,374,359	1,538,161	1,538,161	439,542	439,542	1,538,161	

Notes: This table compares regression results from the typical two-way fixed effects event study regression and the robust alternative estimator proposed by de Chaisemartin and D'Haultfoeuille (2019). Estimation is performed using the did_multiplegt package in Stata (note that pre-trends are not estimated using this command, hence not reported). Standard errors clustered at the household level are reported in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

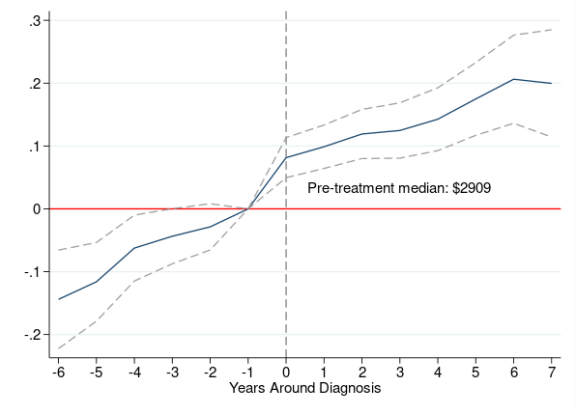
Table 11. Model Comparison: Event Studies and DID_M

B.3 Household Response to Major Medical Events

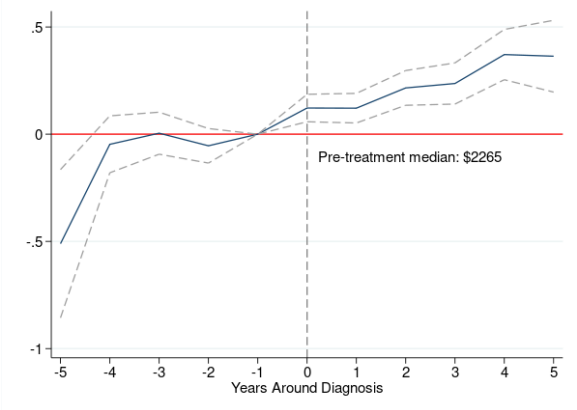
In this section, I include additional results from a suite of two-way fixed effects models estimating the causal effect of major medical events on health behaviors. Figures 15 and 16 illustrate the estimated effect on billed spending for both chronic and acute medical events.

Figure 15. Estimated Effect of a Chronic Diagnosis on Billed Non-Diagnosed Spending

(a) Full sample



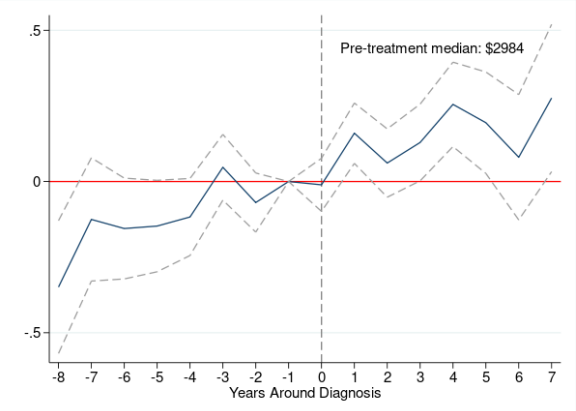
(b) Households facing zero deductible



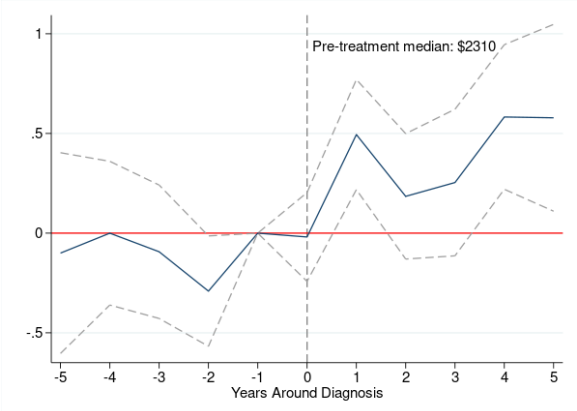
Note: Dependent variable is the inverse hyperbolic sine of total billed spending for all non-diagnosed individuals in a household. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

Figure 16. Estimated Effect of an Acute Health Event on Billed Non-Diagnosed Spending

(a) Full sample



(b) Households facing zero deductible



Note: Dependent variable is the inverse hyperbolic sine of total billed spending for all non-diagnosed individuals in a household. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

I also explore the effect of acute health events on household out-of-pocket spending,

similar to Figure 1 in the text. In general, acute events do not generate the same household response that chronic diagnoses do.

To explore the role that these conditional price changes have on the observed spending responses, I first examine the potentially heterogeneous effects of major medical events by families' typical pre-diagnosis deductible contributions. Figure 17 illustrates various difference-in-difference estimates for the effect of a major medical event on billed spending, estimated on the sample of families who contributed up to a certain fraction of their deductible on average prior to diagnosis. For this approach, I examine billed spending instead of OOP spending because OOP spending will mechanically rise more for those who tend to have a larger portion of their deductible to pay off, as the deductible is typically the largest contributor to OOP expenses.

The figure shows much larger utilization effects among families that typically spent less than a quarter of their deductible OOP. In fact, families that spent 10% or less of their deductible on average prior to diagnosis are estimated to increase their utilization by about 50%. These large effects decay as more of the sample is included, and I find that even families spending 50% of their deductible may not increase their health utilization following major medical events. Taken together, these results suggest that the families who experience the largest price reductions in care are not the families increasing their utilization the most, suggesting that demand responses are not the major driver of health behavior changes.

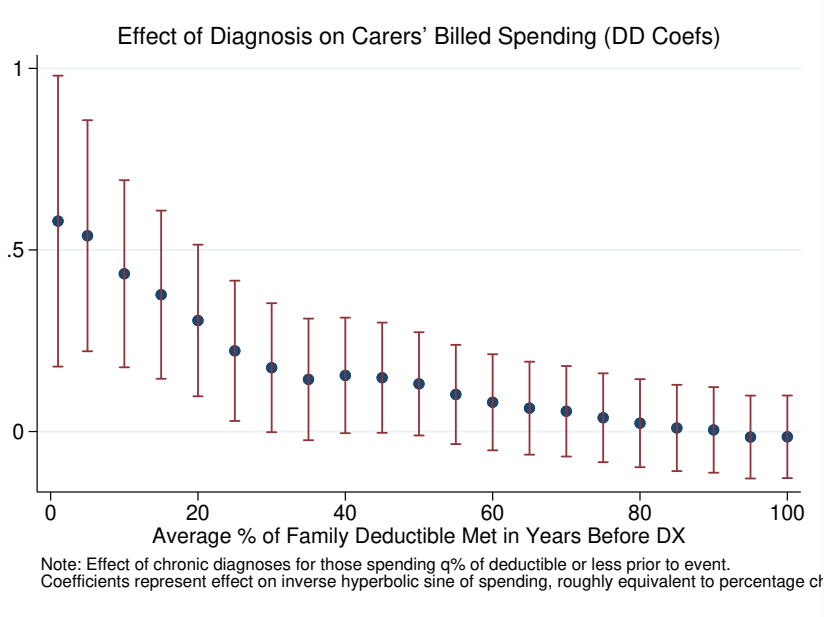


Figure 17. Heterogeneous Utilization Responses based on Pre-Diagnosis Deductible Contributions

Finally, I find a strong extensive margin response among household members who ex-

perience major medical events in their families. Table 12 shows that individuals are more likely to spend any positive amount (billed and OOP) on medical care, use any outpatient visits or preventive care, or fill any prescriptions. This effect is strongest in the year of the diagnosis and decays slightly over time, but remains significant for five years following the health event.

	Year 0	Years 1–5 (average)
Any Billed Spending	1.54***	0.60***
	(0.08)	(0.13)
Any OOP Spending	2.62***	1.41***
	(0.11)	(0.18)
Any Outpatient Visits	2.20***	0.65***
	(0.09)	(0.15)
Any Preventive Care	3.23***	0.90***
	(0.15)	(0.22)
Any Prescription Fills	4.74***	2.45***
	(0.41)	(0.53)

Table 12. Estimated Extensive Margin Health Effects of Family Diagnosis

B.4 Intra-Familial Relationships

For example, while a diabetes diagnosis is most likely to affect adult household members with similar lifestyles to the original diagnosed individual,[33] a mental health diagnosis may have a stronger genetic component. Hence, households where an adult was diagnosed with diabetes may choose to screen other adults, such as spouses, while households where someone received a mental health diagnosis may choose to screen children or siblings of the affected individual.

To assess these potentially heterogeneous effects, I utilize a simple difference-in-differences framework. In Table 13, I present estimation results for the same six diagnosis/outcome pairs shown in Figure 5. The dependent variable—either a screening or a new diagnosis—is shown in the top row, with the treatment variable—the chronic illness affecting the household—below in italics. I explore the potentially heterogeneous responses for four family relationships: parents, spouses, siblings, and children of the affected individual, with children as the

[33]The vast majority of diabetes diagnoses in my sample are for Type 2 Diabetes Mellitus, which generally affects adults and risk of which is increased or decreased based on specific lifestyle choices, such as diet and exercise. The same is not as true for Type 1 DM diagnoses.

Screening *Diagnosis*	Hypertension *Any Chronic*	Diabetes *Diabetes*	Cholesterol *Diabetes*	High BMI *Diabetes*	Cancer *Cancer*	Depression *MDD/Bipolar*
$\text{Post}_t \times \text{Diagnosis}_f$	0.39***	-0.85***	-2.20***	-0.38**	2.55***	0.30**
	(0.03)	(0.21)	(0.29)	(0.12)	(0.43)	(0.10)
$\text{Post}_t \times \text{Diagnosis}_f \times \text{Parent}_j$	-0.34**	3.49*	3.73	1.73*	-1.90	-0.93***
	(0.11)	(1.71)	(2.26)	(0.70)	(2.49)	(0.13)
$\text{Post}_t \times \text{Diagnosis}_f \times \text{Spouse}_j$	-0.74***	2.54***	5.15***	1.03***	-3.33***	-0.62***
	(0.13)	(0.45)	(0.60)	(0.20)	(0.81)	(0.11)
$\text{Post}_t \times \text{Diagnosis}_f \times \text{Sibling}_j$	0.09	0.76	2.89	0.16	1.56	0.68*
	(0.04)	(1.09)	(1.86)	(0.69)	(1.55)	(0.32)
Observations	4,039,602	3,680,725	3,680,725	3,680,725	3,671,064	3,724,608
Adjusted R^2	0.024	0.217	0.388	-0.025	0.473	0.117

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Table shows results of a difference-in-differences estimation strategy highlighting the potentially differential effects of chronic illnesses on preventive care utilization by household relationships. The primary outcome variable in each column is a screening or new diagnosis, shown in the top row. The specific chronic illness used as the Diagnosis_f dummy is shown in the second row. Standard errors are clustered at the household level.

Table 13. DDD Estimates: Disease-Specific Spending

reference group.

Throughout, I find consistent evidence that households respond by not only selecting screenings associated with the health events they experienced, but also selecting which individuals to screen based on their associated risk. New hypertension diagnoses following a chronic event are concentrated among children rather than parents and spouses, suggesting that households are identifying previously ignored risks among the previously lower-risk members of their household. Additionally, households affected with diabetes focus screenings on spouses more than on children, consistent with the lifestyle factors that affect diabetes risk. In contrast, households affected with chronic illnesses that communicate a greater level of genetic risk—cancer and mental health conditions—choose instead to screen children and siblings (in the case of mental health conditions) more than parents or spouses.

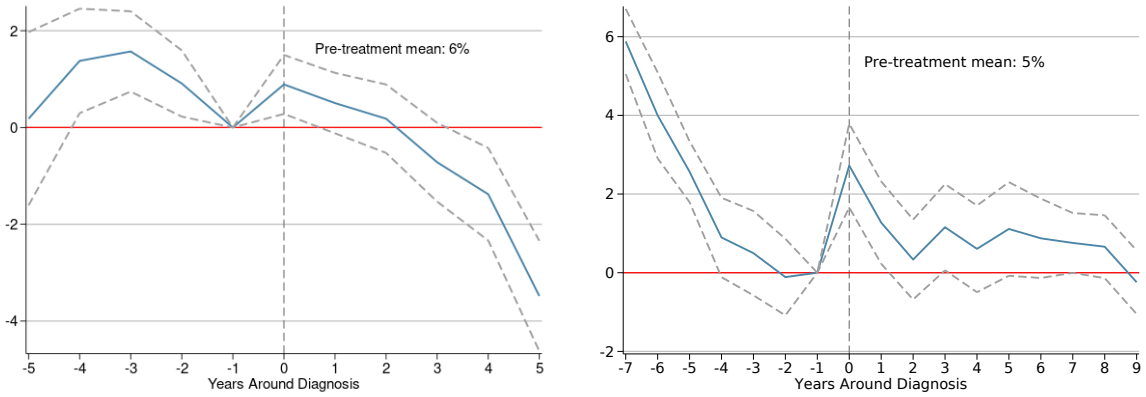
B.5 Plan Choices

Finally, using the portion of my sample with identifiable plan choice information, I estimate the effect of chronic health events on household decisions to switch plans. Figure 18 illustrates that affected households are less likely to switch insurance plans following their major health events relative to the general population. I observe both that plan switches do not become more likely overall (Panel (a)), and that even among active choosers, plan switches do not

become higher-quality (proxied by the use of zero-deductible plans; see Panel (b)).

Figure 18. Effect of Chronic Diagnoses on Health Plan Switching

(a) Pr(Any Health Plan Switch) (b) Pr(Switch to 0-Deductible—Any Switch)



Note: These figures assess the impact of major health events on plan switches. The outcome variables are a binary indicator for whether the household switched plans in the first panel, and whether they switched plans to a plan with zero-deductible in the second panel. The second panel restricts the sample to those who ever made an active plan choice. Standard errors are clustered at the household level.

C Additional Modeling Notes

C.1 Solving the Utility Maximization Problem

In the final choice stage of the model, households choose medical spending m_{it}^* based on the realization of their acute shocks $\{\lambda_{it}, m_{ft}^{\text{CH}}\}$ and their type parameters $\{p_{it}, \omega\}$. Their expected utility is given by

$$\begin{aligned} u_{it}(m_{it}) = p\left[(\alpha_1 m_{it} + \alpha_2 m_{ft}^{\text{CH}} - \lambda_{it}) - \frac{1}{2\omega}(\alpha_1 m_{it} + \alpha_2 m_{ft}^{\text{CH}} - \lambda_{it})^2 - c_j(m_{it})\right] \\ + (1-p)\left[(m_{it} - \lambda_{it}) - \frac{1}{2\omega}(m_{it} - \lambda_{it})^2 - c_j(m_{it})\right] + \varepsilon_{ijt}. \end{aligned} \tag{18}$$

Ignoring the idiosyncratic shock ε_{ijt}, the first order condition for utility maximization implies that optimal spending is given by:

$$m_{it}^* = \frac{1}{1 + p_{it}(\alpha_1 - 1)}\left[\lambda_{it} + \omega(1 - c_j'(m_{it}) + p_{it}\left((\alpha_1 - 1)\omega - \alpha_2 m_{ft}^{\text{CH}}\right)\right]. \tag{19}$$

Without the expected utility framework or allowing for state-dependent utility across states, this reduces to the typical solution of $m^*_{it} = \lambda_{it} + \omega(1 - c'_j(m_{it})$. Here, $c'_j(m_{it})$ depends on the optimal level of spending, with $c' = 1$ when households choose a level of spending below the deductible, and then declining to $c' = c < 1$ when OOP spending is between the deductible and the OOP max, and $c' = 0$ otherwise. The piecewise linear structure of the cost-sharing scheme does not yield a closed form solution for m^*_{it}, but rather implies a discrete set of possible solutions that must be evaluated.

C.2 Alternate Interpretations of p_{it}

The evidence presented in Section 3 suggests that health events generate spending responses as household members reevaluate their health risks. This leads to the simple interpretation of the dynamic learning parameter p_{it} as a probability of an adverse health event occurring. However, to the extent that other informational effects affect spending choices in ways that are separate from health risk information, moral hazard effects, or salience effects, these effects may "load" onto the estimated p_{it} parameter, affecting its interpretation. These informational effects may include physician relationship building, increased comfort obtaining care covered by an insurer, or other, more general health information effects, which alter consumer *preferences* for health care rather than their *beliefs* about risk.

The transition probability parameter p_{it} can therefore be interpreted, in part, as an adjustment to consumer preferences for care in addition to risk beliefs. Consider equation 18. If we assume that $\alpha_1 \approx 1$, as estimated in Section 5, the equation reduces to:

$$u_{it}(m_{it}) = m_{it} - \lambda_{it} - c_j(m_{it}) + p_{it}\alpha_2 m^{\text{CH}}_{ft} - \frac{p_{it}}{2\omega}(m_{it} + \alpha_2 m^{\text{CH}}_{ft} - \lambda_{it})^2 - \frac{1 - p_{it}}{2\omega}(m_{it} - \lambda_{it})^2. \quad (20)$$

Hence, p_{it} can be construed, together with the estimated parameter α_2, to be representative of the preference weight individuals place on chronic care, relative to all non-chronic care. In this setting, the informational effect of health shocks increases individual preferences for chronic care.

D Additional Structural Results

D.1 Estimation Algorithm

I estimate the model described in Section 4 using a maximum likelihood approach similar to Train (2009) and Revelt and Train (1998), with the appropriate extension to a dis-

crete/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 (2020). I estimate the parameter values θ that maximize the probability density of households' observed total healthcare spending conditional on their plan choices.

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.

D.1.1 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.

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 19 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:

$$f_m(m_{it}|c_{jt}, \beta_{fts}, \theta) = \begin{cases} \mathbf{\Phi}\left(\frac{\log(\kappa_i)-\mu_{\lambda,i}}{\sigma_{\lambda,i}}\right) & m_{it} = 0 \\ \mathbf{\Phi}'\left(\frac{\log(\lambda_{its})-\mu_{\lambda,i}}{\sigma_{\lambda,i}}\right) & m_{it} > 0, \end{cases} \tag{21}$$

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 (2020).

D.1.2 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:

$$\lambda_{itsd} = \exp\left(\mu_{is} + \sigma_{is} Z_d\right) + \kappa_{is}, \tag{22}$$

where Z_d is the appropriate Gaussian quadrature vector of points (with corresponding weights W_d). The utility maximization framework discussed above (Equation 19) 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 11 and summed (with weights) over all 100 points.[34] Choice probabilities for a plan j are then given by the standard logit formula

$$L_{ftjs} = \frac{\exp(U_{ftjs}/\sigma_\epsilon)}{\sum_{i \in \mathcal{J}_{ft}} \exp(U_{ftis}/\sigma_\epsilon)}. \tag{23}$$

D.1.3 Likelihood Function

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

$$LL_f = \sum_{j=1}^{J} d_{fjt} \sum_{s=1}^{S} W_s \prod_{t=1}^{T} f_m(m_{it}|c_{jt}, \beta_{fts}, \theta) L_{ftjs}, \tag{24}$$

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 to be maximized is therefore the sum over households:

$$LL(\theta) = \sum_{f=1}^{F} \log(LL_f). \tag{25}$$

D.2 Additional Parameters

Describe table of mean shifters for preferred specification:

Include versions of this for other specifications?

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

	p_0	λ	κ	ψ_0
Intercept	0.089	0.190	-0.105	0.112
Age	0.084	-0.088	-0.097	
Age^2	0.115	-0.006	-0.087	
Female	0.102	0.219	-0.117	
Individual risk score	0.100			
Any PE condition in family	0.107			
Type		0.152		
Family size				0.107
Average family age				0.052
Average family risk score				0.140

Table 14. Estimated Type Mean Shifting Parameters

Notes: Values reported for preferred specification only.

D.3 Additional Model Fit

D.4 Additional Simulations