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

Alex Hoagland†

September 30, 2021

Click here for the latest version

Abstract

Individuals infer their health risk from observing the health experiences of people around them, particularly family members. I assess whether people correctly interpret new information from household health events and analyze resulting welfare implications. When an individual is diagnosed with a new chronic condition, unaffected family members increase their healthcare spending by over 10 percent. Informational spillovers are associated with increased use of both high- and low-return care, including takeup of new services and renewed adherence to extant ones. I show these responses are consistent with individual reevaluations of health risk and reject other mechanisms. To assess welfare implications, I estimate a structural model of health choices in which individuals learn about risk after health events reveal information. I find that consumers over-respond to recent, salient health events by over-weighting their risks *ex-post*. This leads to annual welfare losses of $2,788 per family on average; suppressing responsiveness results in net gains for 86 percent of households. Revealing health risk information can be optimally targeted on household demographics to improve social welfare gains.

Keywords: Health spillovers, consumer 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, for expert advice that greatly benefited this research. I am also grateful to Kristen Blair, Corinne Andriola, Jason Abaluck, Marshall Drake, Keith Ericson, James Feigenbaum, Raymond Fisman, Sarah Gordon, Timothy Layton, Patrick Power, and Paul Shafer for very helpful comments and support. Finally, I also thank seminar participants at the 10th Annual Conference of the American Society of Health Economists, the 2021 Congress of the International Health Economics Association, and Boston University 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

Social networks provide important information for consumers making health care choices. Through connections with family, friends, and neighbors, individuals form expectations of their own health risks, learn about the value of specific medical practices, and identify how or from whom to receive care. Family relationships provide particularly influential sources of health information due to their close proximity and the high relevance of their health experiences, as both shared genetic profiles and lifestyle choices influence expected health care consumption. Understanding how individual health experiences shape family health behaviors is essential for policies aiming to improve public health, such as those attempting to address high levels of health care spending or the takeup of high-value health services.

One especially salient dimension of health information individuals may learn from family health experiences is knowledge about health risks, including both current and expected future health care needs. Individuals may choose to seek out high-value, life-saving care after witnessing a family member's health experience, particularly if that experience reveals their own risk. For instance, some may choose to become vaccinated against COVID–19 once a family member becomes infected (Chen, 2021; Giardinelli, 2021; Salcedo, 2021).

There exists evidence that family members react to the health events of their loved ones (Fadlon and Nielsen, 2019; Hodor, 2021), but it remains unclear what information drives these reactions. Health events may lead individuals to reassess their specific health risks, but individuals may also respond to other features of an event. These include changes to the expected price of medical care, preferences for health consumption, or knowledge about the availability of health services. Understanding the role that social connections play in both the utilization of high-return medical care and the propagation of low-value services relies on separating these competing effects. In particular, assessing the welfare effects of transmitting new health information requires both understanding whether individuals respond to information itself and the extent to which they update their beliefs correctly.

In this paper, I examine how consumers who receive health risk signals through witnessing a major health event within their household—such as a diagnosis with a new chronic condition—modify their assessments of their own risks and alter their choices accordingly. I study households with employer-sponsored insurance (ESI) obtained through large employers between 2006 and 2018. Highly-detailed claims data provides insight into how individuals respond to quasi-random health events in their family, including overall responses in plan choices and health spending as well as decisions about the use of specific services. Additionally, these data include rich variation in coverage generosity and plan characteristics among

enrollees, an important fact I leverage to separate changes in household beliefs about risk from other, potentially confounding, effects of health events.

I show that major health events generate strong informational spillovers among non-diagnosed household members. Those exposed to new health information significantly and persistently increase both their overall health care utilization and their investments in preventive care, particularly for services that are specific to the condition just diagnosed in their household. I show that these spillover effects are more consistent with individual learning than other potential mechanisms. The magnitude of these increases is constant across insurance plan designs—including plans without deductibles—suggesting that moral hazard concerns are not driving changes.[1] Additionally, chronic events induce stronger and more persistent behavior changes than acute health events, suggesting that salience effects arising from a traumatic health experience do not fully explain observed results (Dalton et al., 2020; Fadlon and Nielsen, 2019). Finally, I show that even individuals who are most familiar with the health care system—such as those taking regular preventive medications for cardiovascular health—are responsive to major health events, implying that learning about health *systems*, rather than health *risk*, is not the main driver of observed results.[2]

In general, one would expect receiving new information about one's risk to lead to improvements in decision-making and welfare. Surprisingly, however, I demonstrate that the welfare effects of new information are not obvious from reduced-form analysis alone. Affected household members increase their use of "low-value" health services, procedures 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 services that appear, from their perspective, closely related to preventive care, including cardiac screenings before low-risk surgeries or imaging services for lower back pain. In addition, households do not alter their insurance plan choices even after large expected increases in health costs from managing chronic conditions. Both of these findings cast doubt on the extent to which health information improves choice quality.

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 model in which households form beliefs about their health risks over time. In my model, households first make decisions about their insurance coverage

[1] 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. For a more in-depth discussion of this abuse of notation, see Einav et al. (2013).

[2] This general learning may include systematic learning about health care organizations, the process of receiving insurance coverage for care, or building physician relationships (Sabety, 2020).

prior to receiving information about their health state in a period; once this information is realized, households choose health spending (Cardon and Hendel, 2001; Einav et al., 2013; Marone and Sabety, 2021). Novel to my model, health shocks take two forms: major health events and non-chronic health fluctuations. Major health events occurring in a household induce other members to update beliefs about their health risks, but also affect consumer choices by potentially lowering the conditional cost of non-chronic care and increasing risk aversion. A structural approach allows me to separately identify these competing effects, yielding clear estimation of the welfare effects from receiving health information.

A key challenge in my model is identifying changes in an individual's beliefs about their health risks separate from these alternative explanations. I use multiple sources of variation in the data to decompose the effects of household health events. First, I use a broad set of health events which vary in their expected treatment costs to identify the effects of price changes on spending decisions. More expensive conditions (e.g., cancers) are associated with stronger price effects than cheaper ones (e.g., asthma) and therefore are expected to induce stronger moral hazard responses. Second, I 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 health 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 vary (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. Finally, I use both acute and chronic health events to assess the extent to which individuals learn more generally about the health care system, rather than the causal effect of new information about health risks.

Counter to expected thought, the new information gained from health events is not welfare-improving for many households. In fact, new health risk information lowers expected household utility by an average of $2,788 per year. The central insight of the model is that there is a tension between the seriousness of a major health event and the appropriate level to which individuals should update their beliefs: new diagnoses in a household spur overly 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: 86% of the households in my sample would find health information welfare-improving were their responses mitigated. Finally, I demonstrate that the societal value of communicating health information can be improved by selectively revealing it to specific groups, such as those with higher *ex-ante* risk. This suggests that population health information campaigns—

including genetic testing programs and screening practices for important conditions such as COVID-19—can benefit from targeting specific groups.

My analysis contributes to a burgeoning discussion on the causal spillover effects of health information within social networks, particularly the family. The importance of family relationships in economic decision-making has been well-documented in labor supply and education choices (Browning et al., 2014; Altmejd et al., 2021), but the role of these relationships in forming health behaviors is not as well understood. Previous work has suggested that an individual's social network informs their decision-making following acute health events (Bouckaert et al., 2020; Hodor, 2021; Song, 2021), health trials (Archibong and Annan, 2021), and infectious disease outbreaks (Agüero and Beleche, 2017).[3]

I contribute to this discussion in three ways. First, I highlight a new type of health information to which individuals are highly responsive: household chronic conditions. I provide evidence that individuals are even more responsive to these chronic diagnoses than to household acute health events. Second, I explore the mechanisms behind these responses, showing that changes to how individuals assess their health risks appear to drive observed spending changes. Finally, I provide evidence that while health events increase investments in high-value care, they are also associated with large errors in risk assessments and the takeup of low-value care, resulting in welfare losses for households on average.

I also contribute to a growing literature that incorporates learning and preferences in structural models of health behavior (Barseghyan et al., 2018; Bundorf et al., 2021a). I incorporate the findings of this literature into the first structural model addressing the value of health information spillovers, and highlight the particular behaviors—such as information misinterpretation—that dampen potential welfare gains. My model encompasses previous work highlighting the role major health events play in inducing demand responses by changing spot prices for other care (Eichner, 1997; Kowalski, 2016). Additionally, I make use of previous identification results to simultaneously estimate weighted probabilities and standard risk aversion parameters in a nonlinear framework (Ericson et al., 2020).[4]

Related to this, I also contribute to a literature on non-Bayesian learning, which emphasizes the disproportionate weight put on recent, and particularly salient, events (Kahneman and Tversky, 1973). This literature emphasizes the role of individual over- and under-reactions to new signals, and how this affects the ultimate convergence of individual beliefs

[3]A rich literature has highlighted how individuals respond to information about their own health risks, including their own diagnosis. 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); however, examinations of other diagnoses revealed a lack of noticeable responses (Dupas, 2011; Kim et al., 2019).

[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.

(Epstein et al., 2010). Models that incorporate such ideas include Holt and Smith (2009), who find in an experimental setting that individuals significantly overweight new evidence (relative to typical Bayesian predictions) when it had a lower *ex-ante* probability of occurring. Other important models draw attention to biased beliefs in models of consumer choices, including their role in rationalizing choices that would otherwise require unreasonably high levels of risk aversion (Ortoleva, 2012; Paserman, 2008; Spinnewijn, 2015).

My model highlights that over- or under-reactions can be accommodated *ex-ante* in a quasi-Bayesian framework by varying the timing of belief updating. In addition, I simultaneously estimate biased beliefs and risk preferences, providing a micro-foundation of how individuals form beliefs in a setting of largely small-probability events. My model provides additional insight into the development of subjective health beliefs; in particular, I provide new evidence that explains why consumers may be better at predicting their relative risk rather than their absolute risk (Bundorf et al., 2021b), and how biases in assessing their own health risks may arise (Arni et al., 2021).

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; Baicker et al., 2015; Handel, 2013; Handel and Kolstad, 2015; Iizuka et al., 2021; Ketcham et al., 2012). This literature includes an ongoing discussion about the extent to which improving health information generally may improve decision-making (Abaluck and Gruber, 2016b; Cutler and Zeckhauser, 2004; Gruber et al., 2020). 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.

I present my empirical setting and data in Section 2. Following a discussion of major health events, I provide evidence of their spillover effects and 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 its results 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 enrolled in ESI through large U.S. firms which 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 Consumer Price Index for All Urban Consumers series.

My final sample includes households with two or more members observed for two or more years 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. It is important to ensure that the two samples are relatively balanced given that I use only the plan-identified sample in my structural estimation (Section 4). 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 out-of-pocket (OOP) costs than the full sample, suggesting that they possess more generous insurance coverage on average. However, this is likely due to time trends arising from the fact that the plan-identified sample runs only through 2013. 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 are low, consistent with prior work (Handel, 2013).

2.1 Major medical events

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 high frequency 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

Table 1. Household Summary Statistics

	Full Sample	Plan-Identified Sample
Panel A: Household demographics		
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)
Panel B: Medical spending & plan choices		
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)
Panel C: Major medical events		
% 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	\$1,082.05 [\$464.69] (11.59)	\$854.62 [\$329.90] (17.72)
Diagnosis OOP, acute	\$2,494.42 [\$1,419.91] (68.05)	\$2,107.09 [\$964.62] (122.50)
Recurring OOP, chronic	\$983.03 [\$521.39] (17.32)	\$683.60 [\$446.69] (19.20)
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.

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 Rebecca Hughes, MD, to identify a set of disease-specific procedures and prescriptions associated with each health condition in my sample.[5] 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. As reported in Table 1, the average (median) household in my sample spends $683.60 ($446.69) out-of-pocket on recurring costs needed to care for chronic conditions.

2.2 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.[6] 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 sum of squared residuals between predicted and observed OOP spending for households within each plan year (Marone and Sabety, 2021). Appendix A.1 describes this methodology in more detail and evaluates the quality of these inferences. I find that these simplified measures capture a wide degree of variation in my data and harmonize well with measures from earlier work. 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 contribu-

[5] Appendix A.3 lists the relevant codes used for each diagnosis.

[6] My data does not distinguish whether there exist plan "tiers" within firms (for example, a university that offers one set of plans to its faculty and a different set to its graduate students). These unobserved barriers may cause measurement error in the plan choice sets used in the structural model in Section 4; however, such error would not affect any of my primary results, which focus on how new health information alters spending choices conditional on the choice of plan.

tions are consistent with the national averages for household coverage (on average about 28% of the household premium; KFF (2020)).

There is substantial variation across firms, regions, and years in the generosity of coverage offered to employees, which I describe in Table 2. As I describe in Section 4.3.2, such variation provides an intuitively useful means of attributing household behaviors to changes in risk *preferences* versus risk *beliefs*; households who are more risk averse tend to minimize their overall variation in *ex-ante* expenditures by choosing more generous health plans, while households who are less risk averse but believe they are at higher risk for major health events may choose less-generous plans overall that instead provide more targeted coverage. The average household has between two and four plans to choose from in a given year, with a wide degree of variation in the average family deductible. This variation is comprised of both 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.

Table 2. Average Plan Characteristics, 2006–2013

	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
Spending/Enrollee ($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.

3 Spillover Effects of Household Health Events

This section presents my main reduced-form empirical results. I first show that after experiencing a chronic major health event, households increase their overall medical utilization by about 10% annually, as well as increasing their investment in billed spending on preventive care. I illustrate that the observed responses are consistent with a reevaluation of one's own risk by showing that households are more likely to invest in preventive care that is specific to the illness their family member experienced. I then consider other potential mechanisms, including financial incentives, salience effects, and general learning about the health care system. Finally, I show that household members increase their utilization of "pseudo-

preventive" low-value services—such as extraneous screenings and imaging services—showing that while health events generate strong spending responses, these responses are not necessarily targeted at high-return services.

3.1 Induced Spending Changes

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

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

The variable y_{ft} represents a spending outcome for a household f in year t; in my main specification, this outcome is annual OOP payments made by all family members *except* those who experience the major health event. I adjust for highly-skewed distributions of spending variables by using the inverse hyperbolic sine transformation.[7] An 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. The coefficients on these indicator variables, $\{\gamma_k\}$, are the objects 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 TWFE estimators can be difficult to interpret without strong modeling assumptions (Callaway and Sant'Anna, 2018). In particular, coefficients estimated by TWFE models represent the weighted average of many two-by-two comparisons. When treatment effects are heterogeneous across groups—and hence, these comparisons—some comparisons may be assigned negative weights (de Chaisemartin and D'Haultfoeuille, 2019; Goodman-Bacon, 2018). This makes the interpretation of estimated treatment effects—static or dynamic—difficult to interpret. Furthermore, when estimating dynamic treatment effects, researchers must take care that dynamic parameters of interest

[7] I use the inverse hyperbolic sine transformation to accommodate the approximately 15% of individual-years in my data with 0 spending (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 that my results are not substantively altered when using the logarithm transformation.

(including both pre-trends and estimated time-varying treatment effects) are separately identified from time fixed-effects included in the regression (Borusyak and Jaravel, 2016; Sun and Abraham, 2020). Without including a control group of observations which are never treated, separate identification of time fixed effects and dynamic treatment effects is impossible.

I demonstrate that my analysis is robust to both concerns. First, I show that my coefficients of interest do not suffer from problems of negative weighting by considering a number of additional specifications in Appendix B. These include both robust estimators proposed by de Chaisemartin and D'Haultfoeuille (2019) and Sant'Anna and Zhao (2020), as well as simple recentered time series graphs and standard difference-in-differences coefficients.[8] This provides evidence that my results are not idiosyncratic to my estimation method; rather, my results appear even in the raw data.

Second, I utilize a large control group in my sample, allowing me to separately identify the time-varying treatment effects from yearly fixed effects. Previous work examining health spillovers within families has restricted the control group to only those who experience a similar diagnosis in the future in order to utilize a control group that more closely matches the treatment group on unobservable characteristics. I include never-treated households in my sample in order to identify dynamic treatment effects. The central tradeoff in doing so lies in the validity of the parallel trends assumption: namely, that in the absence of major health events, the treated and control groups would continue to have similar spending and utilization trajectories. Given that my setting spans a large range of chronic conditions—many of which are neither directly related to health behaviors or particularly life-threatening—concerns about violations of the parallel trends assumption are less plausible in my setting.

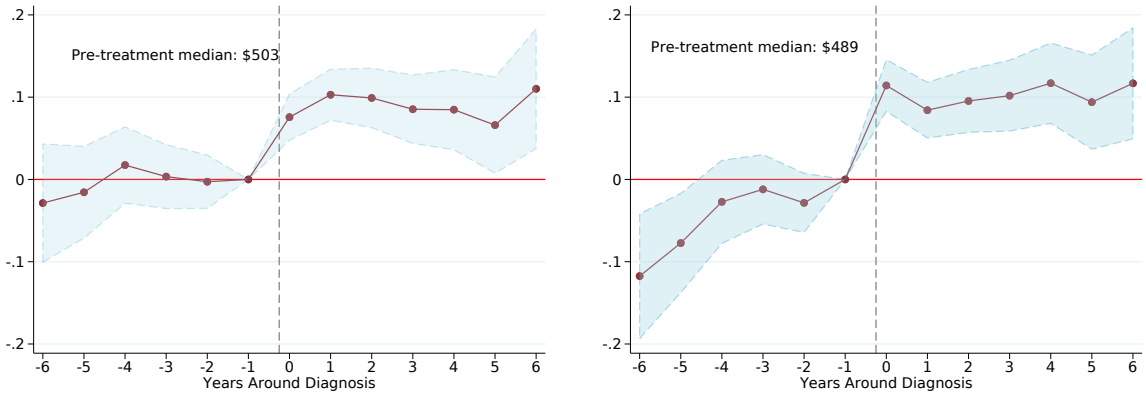
Figure 1 presents the time-varying causal effect of a health shock on household OOP spending for all non-diagnosed individuals. The first panel illustrates that non-diagnosed household members increase their annual OOP spending by about 10% relative to the year prior to the event. For the median (average) household, this corresponds to an increase of about \$50 (\$115) annually. This effect begins in the year of the health event and persists following the diagnosis. Additional results in Appendix B corroborate this finding with other outcome variables including total billed spending or visit frequencies.

Importantly, this increase in utilization encompasses an increased investment in preventive care. The second panel of Figure 1 illustrates this by limiting the scope of the analysis to household spending only on wellness visits. Wellness visits are non-problem-based visits with a family or primary care physician that are generally recommended about once a year;

[8]Using the Bacon decomposition reveals that the estimands in my primary specification are not constructed using negative weights (Goodman-Bacon et al., 2019). However, I present these additional robustness results for completeness.

Figure 1. Effect of Chronic Diagnoses on Non-Diagnosed Household Members' Spending

(a) $\sinh^{-1}$(Total OOP Spending) (b) $\sinh^{-1}$(Billed Spending on Wellness Visits)



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new chronic diagnosis on medical spending. In both panels, the sample includes spending for all household members without major health events. In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on total spending (insurer spending + OOP spending) on wellness visits only. Coefficients are presented relative to the year prior to diagnosis. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

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). Here, too, I find that new diagnoses in a household are associated with strong responses. Affected, non-diagnosed household members increase their overall spending on wellness visits by about 10%, matching the increase in overall utilization.[9]

3.2 Changes as Responses to New Health Risk Information

These results suggest a meaningful, persistent change in how non-diagnosed household members engage with the health care system. I first show that these responses are indicative of household members updating their beliefs about their own health risks following the receipt

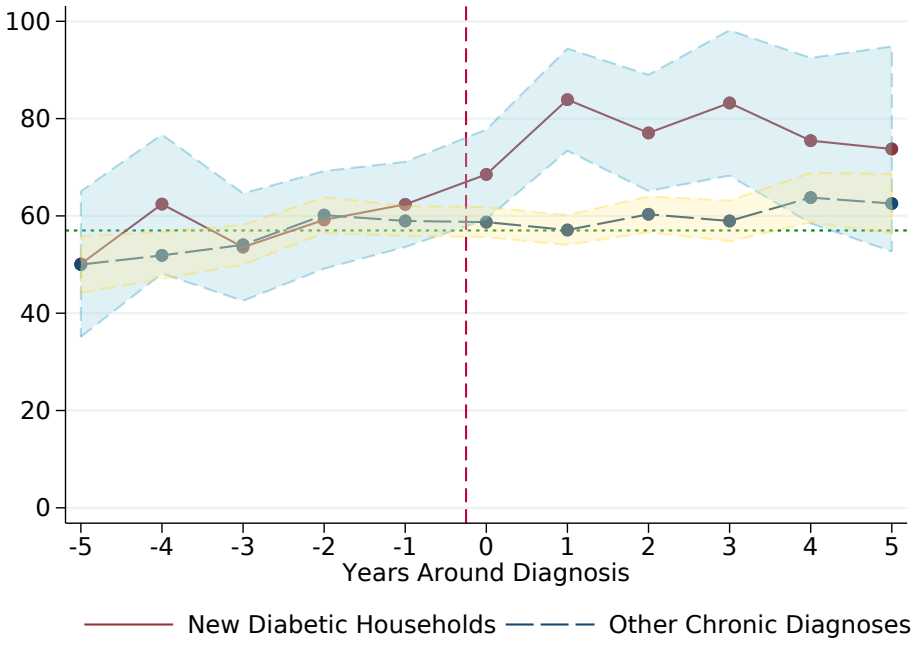
[9]Even before the Affordable Care Act (ACA)'s cost-sharing exclusion took effect in 2010 (or 2012 for certain women's health services), OOP costs for preventive care were steadily declining for those with ESI (Hong et al., 2017). Once the ACA took effect, the majority of wellness visits should be free to enrollees in my sample (Shafer et al., 2021), a feature I observe in the data. Although time fixed effects in the regression specification should absorb these trends for both pre- and post-ACA trends, I use billed spending rather than OOP spending as my outcome variable of interest. Note that in my data set, billed spending represents the sum of individual OOP payments and insurer payments to the provider; it does not reflect any price negotiations or other discounts that were provided at the time of service, and therefore does not reflect the listed prices of services.

of health information from a major event. Such observed responses could also be driven by factors beyond changes in a household's assessment of their health risks, including changes in the price of care, salience effects, overall exposure to the health care system, or improved physician relationships. I explore these alternative mechanisms in Section 3.3.

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 2. Rate of Diabetes Screenings Around Time of Diagnoses

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



Notes: Figure shows average utilization rates of diabetes screenings for non-diagnosed household members 18 years of age and older, measured in rates per 1,000 adults. Point estimates and 95% confidence intervals are presented. The top (solid maroon) line indicates average rates for households who experience a diabetes diagnosis, and the bottom (dashed navy) line indicates rates for those affected by other chronic diagnoses. The horizontal, dotted green line indicates the average utilization rate for all other households in the sample who do not experience a diagnosis, about 59 screenings per 1,000 adults.

For example, individuals who have learned that they are at increased risk for developing diabetes may have a higher likelihood of seeking out screenings for abnormal blood glucose levels than individuals who have learned that they are at increased risk for another chronic condition. Figure 2 plots re-centered time series that depict the associations between house-

hold diagnoses and the takeup of diabetes screenings for adults within a household. The figure plots average utilization rates of diabetes screenings for two groups: those who are exposed to a diabetes diagnosis in their home and those who are exposed to a different chronic diagnosis. Individuals whose family members are diagnosed with conditions other than diabetes do not appear to significantly alter their screening behaviors from unaffected households (whose average is depicted in the horizontal, dotted green line). On the other hand, household members of those diagnosed with diabetes increase screenings in the first three years following the diagnosis, being about 36% more likely to be screened for diabetes than unaffected individuals.

To assess the causal effect of multiple diagnoses simultaneously on the utilization of disease-specific preventive care, I use a triple-differences approach. This approach disentangles two competing effects: those arising from experiencing any chronic illness (e.g., salience effects) and a 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 periods following a diagnosis. The triple interaction variable includes an additional constraint that the chronic diagnosis $chronic_f$ match the specific diagnosis d (e.g., a diabetes diagnosis when the outcome variable is a diabetes screening). Hence, β_{DD} identifies the effect of any chronic diagnosis on screening, while the triple interaction β_{DDD} identifies the effect for the specific diagnosis of interest relative to other diagnoses.[10] For example, using this approach I can estimate the impact of a diabetes diagnosis on diabetes screenings as $\beta_{\text{DD}} + \beta_{\text{DDD}}$, where β_{DD} indicates the impact of experiencing any chronic diagnosis in the household on diabetes screenings and β_{DDD} indicates the specific differential effect of a new diabetes diagnosis occurring in the household.

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

[10]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 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.[11]

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. 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.[12]

Finally, to verify my results, I estimate this model for screenings for which health events communicate little useful information, and hence are expected to change behavior little. This might be because a diagnosis doesn't require a doctor's visit to diagnose (e.g., obesity) or doesn't require preventive screening prior to seeking treatment (e.g., mental health conditions, such as major depressive disorder). Hence, observing a lack of response among these types of preventive services serves to underscore the role that health information, specifically, plays in altering individual behavior. I include "placebo" regressions for the effect of new diabetes diagnoses on obesity diagnoses and the effect of new mental health disorder diagnoses on screenings for depression.

Table 3 presents the estimation results from these six regressions in two panels. First, I highlight that new chronic diagnoses alter specific preventive behaviors in cases where they transmit important information about health risk. The occurrence of any chronic diagnoses in a household is associated with a 19.4% increase in the rate of hypertension diagnoses among other affected household members. Furthermore, specific diagnoses such as cancer and diabetes increase the likelihood that a non-diagnosed household member will seek out screening by 13.2% and 21.1%, respectively. Finally, diabetes diagnoses are associated with an increase in cholesterol screenings of 7.2%. Similar to previous work, I find evidence that new diagnoses reduce the rate of other, unrelated screenings (Fadlon and Nielsen, 2019); for example, a non-diabetes chronic diagnosis is associated with a 7.4% *decline* in the rate of

[11] 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 includes standard difference-in-differences regression results that corroborate the findings reported here.

[12] 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. Coding practices reduce my ability to test this finding for each individual diagnosis in my sample; for example, there are no diagnostic or procedure codes used exclusively for asthma screenings.

diabetes screenings among non-diagnosed household members. These effects, however, are typically smaller than the estimated increases in disease-specific screenings, suggesting that this crowding out is not necessarily one-to-one.

The second panel of Table 3 reports results for placebo regressions including obesity diagnoses and depression screenings. Here, I find no strong evidence that health events alter screenings. This is consistent with the notion that individuals respond by altering their use of preventive care only when the major health event communicates health risk information that necessitates preventive care utilization. Other dimensions of a health event—such as learning about the role of preventive care in medical maintenance overall—do not appear to drive individual behavior changes, at least in the use of preventive services.

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

Own Screening (Dependent Variable)	Household Diagnosis	Pre-Diagnosis Average	Effect of Any Diagnosis (β_{DD})	Effect of Specified Diagnosis (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. $^{*}p < 0.05,^{**}p < 0.01,^{***}p < 0.001$

I report additional results in Appendix B. I find that in addition to selecting screenings based on the health risk information they receive, households are selective in which members

they choose to screen. I utilize variation in intrafamilial relationships and corresponding risk to show that households screen those who are most affected by the new health information. When households are affected by a chronic illness with a strong genetic component, such as type 1 diabetes, children and siblings of the affected individual are more likely to be screened than other household members. On the other hand, diagnoses such as type 2 diabetes—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 Alternative Explanations for Spending Changes

Although individuals appear highly responsive to new information about their own risk, additional factors could separately cause or exacerbate observed changes in health spending, including moral hazard effects, salience effects, and learning about the health care system. In this section, I explore each of these potential competing explanations and show that they are each insufficient to explain my observed results.

3.3.1 Moral Hazard

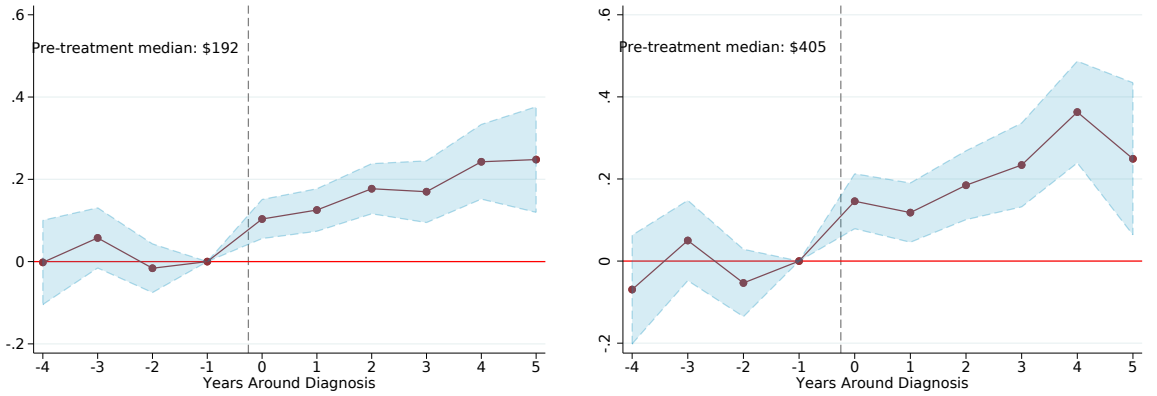
A natural response to observing the phenomenon illustrated in Figure 1 is to conclude that the spending increase is driven by induced demand responses among the non-diagnosed individuals. A chronic diagnosis—such as diabetes— implies consistent, predictable costs on a household—such as through insulin 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. These induced-demand responses have been studied within families experiencing sudden acute health shocks that unexpectedly meet their household deductible (Eichner, 1998; Kowalski, 2016).

Two features of the results suggest that these induced-demand responses are unlikely to be the principal driver of the 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 diagnose 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 closer to the diagnostic event, and more muted in following years. Figure 1 does not show this to be true, either for overall utilization or the use of wellness visits specifically. Second, Figure 3 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.[13]

Figure 3. Effect of Chronic Diagnoses on Spending: Households Facing Zero Deductible

(a) $\sinh^{-1}$(Total OOP Spending) (b) $\sinh^{-1}$(Billed Spending on Wellness Visits)



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new chronic diagnosis on medical spending. This figure uses a limited sample of only households enrolled in health insurance plans with zero deductible at the time of the event. In both panels, the sample includes spending for all household members without major health events. In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on total spending (insurer spending + OOP spending) on wellness visits only. Coefficients are presented relative to the year prior to diagnosis. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

3.3.2 The Effect of Salience

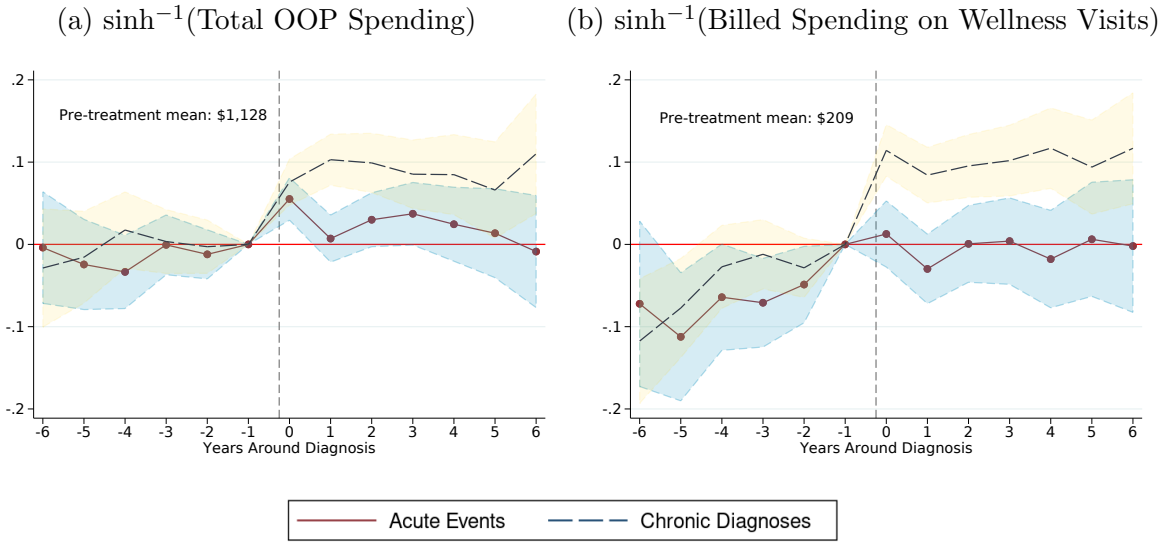
It may also be that the intensity of major health events realigns household preferences to prioritize medical care. Individuals who experience the hospitalization of a household member may (over-)respond to the trauma of the event itself, changing their health consumption behaviors in order to avoid future hospitalizations. The critical difference is that when individuals respond to this health trauma, health events alter an household's risk *preferences* by affecting their marginal utility of medical care, rather than affecting risk *beliefs*.

To examine the impacts of these salience effects relative to risk reassessments, I analyze the responses of individuals who experience acute, rather than chronic, health events in their households. These include hospitalizations for family members who experience severe viral

[13] A corresponding result for the subset plans with nonzero deductibles is included in Appendix B. Additional results in this Appendix show 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 meaningfully change family cost-sharing rates.

infections or other serious conditions unrelated to chronic disease. I use health events that are still assigned HCCs to capture health events of a similar level of seriousness to new chronic diagnoses; however, these events do *not* communicate any information to household members about health risks. Comparing observed household responses to these acute events against responses to chronic diagnoses allows me to assess the extent to which new health risk information alters behavior beyond salience.

Figure 4. Effect of Acute Health Events on Non-Diagnosed Household Members' Spending



Notes: These figures show estimated coefficients and 95% confidence intervals for the effect of a new acute hospitalization on medical spending. The solid maroon line indicates estimates from an acute event; the dashed navy line presents estimated results from Figure 1 as a reference. In both panels, the sample includes spending for all household members without major health events. In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on total spending (insurer spending + OOP spending) on wellness visits only. Coefficients are presented relative to the year prior to diagnosis. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

Figure 4 presents the results. I find that, unlike new chronic diagnoses, acute hospitalizations spur few changes in health behaviors among other household members. Acute hospitalizations are associated with a short-term increase in spending of about five percent (from a baseline of about $1,100) in the year of the diagnosis, but these effects do not persist across time. Acute health events are also not associated with increased investments in preventive care for other household members. In particular, Figure 4 compares these regression coefficients to those estimated in response to new chronic diagnoses (Figure 1). I find that chronic health events are associated with overall spending responses almost twice as large as for acute hospitalizations, differences which are significant at the 95% confidence level for

the first three years following diagnosis. Furthermore, chronic diagnoses induce significantly more investment in preventive services for the first five years following a diagnosis.

Given that acute hospitalizations make health care at least as salient—if not more so—than chronic diagnoses, these findings suggest that changes in risk preferences arising from a "health scare" are insufficient to entirely explain changes in behavior. Rather, new health risk information, such as about one's inherent genetic risk for a chronic condition, appear to drive observed changes.

3.3.3 Health Information

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 risks and this more systematic learning imply similar responses among affected individuals, making their effects difficult to disentangle.

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. Cardiovascular preventive drugs, including statins and other cholesterol-lowering drugs, are an extremely common class of medications and are known to be effective in preventing future health problems when used appropriately (O'Connor, 2006).[14] 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.

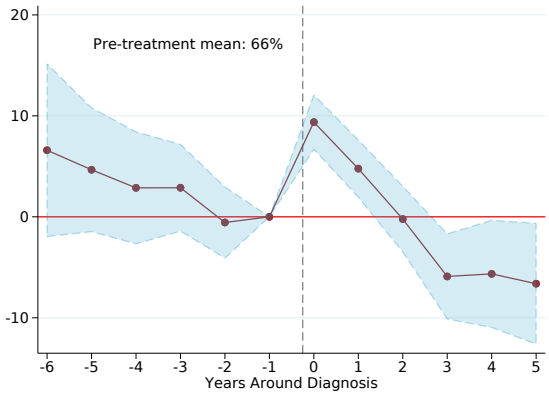
This setting provides a unique environment in which to disentangle the effects of general learning about health systems and learning about one's own health risk. Individuals with existing prescriptions already have sufficient knowledge about the health care system to receive this care from their provider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

I show, however, that new diagnoses alter adherence to these prescriptions. I estimate the effect of a chronic diagnosis on both the likelihood of any medication use and overall adherence, measured as the proportion of days covered in a year (Choudhry et al., 2009). This measure is standard in the literature on adherence, and corresponds to the fraction of the year after a patient's first prescription fill for which the patient has a supply of

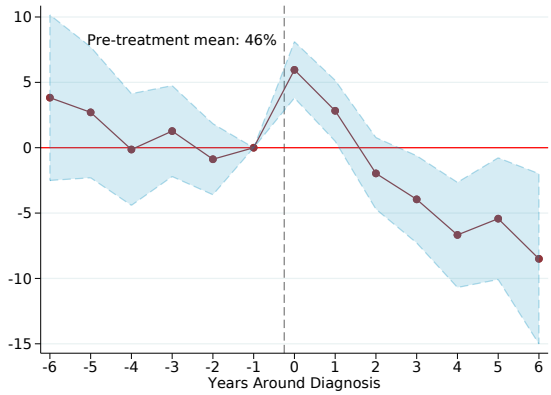
[14] Appendix Table A.5 contains a detailed list of the therapeutic classes used in my sample.

Figure 5. Effect of Chronic Diagnoses On Adherence to 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). Coefficients are presented relative to the year prior to diagnosis. Standard errors are clustered at the household level.

the medication. One concern in identifying the 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 may occur at different rates for different individuals both within and across households, meaning that these trends would not be accounted for using only household and year fixed effects. I therefore add a variable controlling for the number of years an individual has been in the sample to Equation 1.

Figure 5 presents the estimated dynamic treatment effect of a new chronic diagnosis on adherence to existing preventive prescriptions. As expected, in the absence of new health information, individuals become less adherent to prescriptions over time. However, diagnoses in the household 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. 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. The estimated causal "re-adherence" to

prescriptions is consistent with individuals reevaluating the value of their medication given new information about their health risks.[15]

3.4 Quality of Induced Spending Changes

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).[16] Low-value services include both services whose cost typically outweighs any benefits to an average patient (e.g., unnecessary surgeries such as arthroscopies) as well as services which are chronically over utilized in ways that dramatically lower their return (e.g., imaging services such as MRI services for chronic migraines). 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.

I find that new chronic diagnoses are associated with an increase in overall low-value spending of about 5 percent (Appendix B). However, these results mask significant heterogeneity across different types of low-value services. Low-value services may differ in their perceived value to an affected household depending on the ways in which health events induce behavior changes. For example, if a chronic diagnosis communicates new risk information to a household, they may find low-value screening services—such as imaging services and preoperative visits—to be more attractive. On the other hand, households that respond 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.

[15] These effects are likewise observed in the sample of households with zero deductible, suggesting that this re-adherence is also not exclusively driven by moral hazard responses.

[16] These health services are based on recommendations made with the Choosing Wisely initiative, directed by the American Board of Internal Medicine 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.

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

Population	Pediatric	Adult Services			
Service Category	All Services	Prescriptions	Imaging	Screening	Surgery
$\text{Post}_t \times \text{Diagnosis}_f$	0.051*	-0.004	0.029***	0.103***	-0.096***
	(0.017)	(0.000)	(0.013)	(0.014)	(0.012)
Adjusted R^2	0.192	0.143	0.123	0.163	0.230

Notes: Table shows estimated difference-in-difference regression coefficients for the effect of a new chronic diagnosis (N=1,538,161). Outcome variables are the inverse hyperbolic sine of billed spending in each category. See Appendix A 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 presents results estimating the effect of a new chronic diagnosis in each of the five categories using a standard difference-in-differences framework (event study regressions are included in Appendix B). New chronic diagnoses shift households spending and utilization into low-value service categories comprised of screening services, pediatric care, and imaging services. The effect sizes range from an increase as large as ten percent for low-value screenings to three percent for imaging services.[17] 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 understood by health professionals as being low return. 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. 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. This is in keeping with 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 (Appendix B). 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

[17]The results also provide preliminary evidence that major health events provide a deterrent from low-value elective surgeries. However, Appendix Table B.3 highlights the strong presence of pre-trends in these models, which obfuscates the true causal effect of the diagnosis.

optimal for affected households. This motivates a more structural approach to quantify the welfare effects of health information.

4 Empirical Model of Belief Formation

In this section, 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). In the first stage, 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 stage, individuals within the household choose their spending and utilization based on realized health shocks and their chosen health plan's features.

I extend the existing model in two important ways. First, I allow consumers' types to be adaptive in response to health experiences. In my model, individuals learn about their probability of adverse health events; in addition, health events may alter household risk aversion to capture potential salience effects. Second, I explicitly model the differences between acute and chronic health shocks, as chronic health shocks impose recurring costs on a family, thereby altering conditional OOP prices for non-chronic care and inducing moral hazard effects within a household.

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. Time is discrete and indexed by t; I am thus abstracting away from the timing of health spending within a year. Households and individuals are characterized by three main variables: individual beliefs about health risks (p_{ift}), household risk aversion (ψ_{ft}), and the distributions of their health shocks (discussed below). New health events—including both new chronic diagnoses and acute hospitalizations—cause all household members to update their beliefs about their health risks, as well as potentially altering household risk aversion and OOP prices.

In each period, two types of shocks are realized. Following typical convention, each individual has an acute health realization λ_{ift} drawn from an individual-specific distribution $F_{\lambda_{ift}}(\cdot)$. Acute health realizations model the uncertain aspect of demand for healthcare, with individuals with higher λ_{ift} being sicker and hence demanding greater healthcare con-

sumption.[18] Second, households in each period receive a chronic health shock, m_{ft}^{CH}. This shock represents the disruptions in health *spending* affecting the household that arise from any new chronic diagnoses affecting an individual in the family; for households without a new diagnosis, this amounts to the expected cost of a new diagnosis. For households with pre-existing chronic conditions, these shocks are the health costs associated with maintaining health for those affected by the conditions.[19]

4.2 Model Stages

Families make two choices during each period. First, families choose their insurance coverage; then, acute and chronic health shocks are realized; finally, individuals choose their yearly health spending. These choices are static, in the sense that both households choose plans and individuals make spending decisions on the basis of the current period's utility and type parameters only (including their beliefs about health risks). The model is static, in the sense that household decisions in period t do not affect outcomes in period $t+1$. I can therefore ignore forward-looking behavior.[20] However, individual and household type parameters—including beliefs and risk aversion—are responsive to exogenous shocks, including major health events. These parameters adjust at the end of each model period, following individual utilization choices. I model the evolutions of these parameters using a Bayesian framework.

In the following sections, I outline the stages of the model in reverse—that is, I first present details of the individual spending choices in Section 4.2.1, followed by a discussion of household plan choices in Section 4.2.2. I then discuss how type parameters respond to exogenous health shocks in Section 4.2.3.

4.2.1 Utilization Choice

After choosing a health plan $j \in \mathcal{J}$ and realizing the vector of acute and chronic health shocks $(\vec{\lambda}_{ift}, m_{ft}^{\text{CH}})$, each individual in the household chooses their medical spending on non-chronic medical care, m_{ift}^{*}. In this stage, individuals make decisions independently to maximize their personal welfare; in the first (plan choice) stage, households make a collective decision.

[18] 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.

[19] Many related models incorporate heterogeneity in individual demand elasticities in order to accommodate heterogeneity in moral hazard effects (Einav et al., 2013; Marone and Sabety, 2021). As my model is concerned with disentangling only moral hazard events induced by major health events, I restrict the demand elasticity parameter ω in my model to be homogeneous across individuals and periods.

[20] Households are, however, forward-looking within a period, as they anticipate second-stage outcomes as part of their first-stage choices. See equation 9.

Given the flexibility in health states, which vary across individuals, households will ultimately distribute health spending so that the least healthy members receive the most care, as would be expected. Hence, this assumption makes the model more tractable without imposing restrictions on household behavior.

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, which they believe occur with probability p_{ift}.[21] Individuals then choose m_{ift} in order to maximize their expected utility over states:

$$m^*_{ift} \equiv \text{argmax}_{m_{ift}} EU(m_{ift}; p_{ift}) = p_{ift} u_{ift,\text{C}} + (1 - p_{ift}) u_{ift,\text{H}}, \tag{3}$$

where $u_{it,\text{C}}$ and $u_{it,\text{H}}$ represent individual utilities when diagnosed with a chronic illness and when not diagnosed, respectively. Note that Equation 3 nests the case where an individual has already been diagnosed with a chronic illness, in which case $p_{ift} = 1$. I assume that each individual's utility function is separable in health and wealth for both chronic and healthy individuals:

$$u_{ift,\text{H}}(m_{ift}; \lambda_{ift}, m^{\text{CH}}_{ft}) = h_1(m_{ift}; \lambda_{ift}, m^{\text{CH}}_{ft}) + y_{ift}(m_{ift}; m^{\text{CH}}_{ft}) + \varepsilon_1 \tag{4}$$

$$u_{ift,\text{C}}(m_{ift}; \lambda_{ift}, m^{\text{CH}}_{ft}) = h_2(m_{ift}; \lambda_{ift}, m^{\text{CH}}_{ft}) + g(m^{\text{CH}}_{ft}; \lambda_{ift}) + y_{ift}(m_{ift}; m^{\text{CH}}_{ft}) + \varepsilon_2. \tag{5}$$

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_{ift} alter the optimal level of utilization m^*_{ift}. Remaining annual income is denoted by $y_{ift}(m_{ift}; m^{\text{CH}}_{ft})$. $\varepsilon_1(\cdot)$ 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, but allow for a potentially state-dependent utility function in which health status potentially alters the marginal utility of medical spending.[22] Individuals without chronic conditions face the typical utility function:

$$u_{ift,\text{H}}(m_{ift}; \lambda_{ift}, m^{\text{CH}}_{ft}, j) = (m_{ift} - \lambda_{ift}) - \frac{1}{2\omega}(m_{ift} - \lambda_{ift})^2 - c_j(m_{ift}). \tag{6}$$

[21] 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 have 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_{ift}.

[22] Previous work discuss and provide evidence for state-dependence in the utility of *non-medical* consumption (Finkelstein et al., 2013, 2009); this model introduces suggestive evidence for the state-dependence of non-chronic medical consumption as well.

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

On the other hand, individuals in the state of chronic illness face a utility function that depends on both acute and chronic health shocks, with potentially differing preference parameters. Their utility, which depends on the same model primitives as Equation 6, is given by:

$$u_{ift,\mathrm{C}} = (\alpha_1 m_{ift} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{ift}) - \frac{1}{2\omega}(\alpha_1 m_{ift} + \alpha_2 m_{ft}^{\mathrm{CH}} - \lambda_{ift})^2 - c_j(m_{ift}). \tag{7}$$

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; however, as the marginal OOP cost changes based on where it is evaluated, the solution depends on which "region" of OOP costs an individual finds themselves in conditional on their health shocks (see Appendix C for details). If the realized acute health shock is negative (or sufficiently small relative to the shift parameter), individuals will choose $m_{ift}^* = 0$ as spending is required to be non-negative; otherwise, optimal spending follows the condition:

$$m_{ift}^* = \frac{1}{1 + p_{ift}(\alpha_1 - 1)} \left(\lambda_{ift} + \omega(1 + p_{ift}(\alpha_1 - 1) - c_j'(m_{ift}; m_{ft}^{\mathrm{CH}})) - p_{ift}\alpha_2 m_{ft}^{\mathrm{CH}}\right). \tag{8}$$

The interpretation of Equation 8 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 households mismeasure p_{ift} 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_{ift}, potentially explaining the dramatic and persistent shifts observed in Section 3.

Equation 8 also highlights the ways that chronic care costs affect spending decisions through prices. The OOP cost function $c_j(m_{ift}; m_{ft}^{\mathrm{CH}})$ 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 so-

lutions 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.

4.2.2 Plan Choice

In the first stage of the model, households choose an insurance plan to maximize their *ex-ante* expected utilities without knowing their realization of individual health shocks λ_{ift} or major health costs m_{ft}^{CH}. This expected utility depends on the distributions of both health shocks as well as a household risk aversion parameter, which depends flexibly on household demographics and is allowed to evolve over time to capture the salience effects associated with health events, as discussed in Section 4.3.1. The household expected utility function for a given plan j is therefore:

$$U_{fjt} = -\sum_{i \in \mathcal{I}_f} \left[\int \int \frac{1}{\psi_{ft}(x_{ft})} \exp\{-\psi_{ft}(x_{ft}) u_{ift}^*\} \, dF_{\lambda_i} dG_{m^{\text{CH}}} \right] - c_j(m_{ft}^{\text{CH}}) - \pi_{fj} - \eta \mathbb{1}_{fj,t-1}, \quad (9)$$

where u_{ift}^* represents the optimal payoff to individual i in period t given the realization of acute and chronic health states.[23] In addition to each individual's realized OOP costs for non-chronic medical spending, households face OOP costs for chronic care represented by $c_j(m_{ft}^{\text{CH}})$. Households also face plan premiums π_j and a perceived monetary cost η for switching plans ($\mathbb{1}_{fj,t-1}$ is an indicator for whether the family chose plan j in year $t-1$).[24]

4.2.3 Parameter Updating

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

[23] 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_{ift}; hence, this modeling choice does not give rise to families allocating all of their care to a single individual. An alternative approach is to use a CES function for utilities; however, this introduces more nuisance parameters into the estimation framework. Finally, I assume that the von Neumann Morgenstern (vNM) utility index for this decision possesses a constant coefficient of absolute risk aversion, a common choice for these models as it implies no wealth effects.

[24] 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 28% of that value (KFF, 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.

update their risk aversion parameters (ψ_{ft}) according to an adaptive framework; I discuss this further in Section 4.3.1.

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

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

$$\sigma^2_{pi,t+1} = \frac{\tilde{\sigma}^2_{ift}\sigma^2_{pi0}}{\tilde{\sigma}^2_{ift} + s_{ift}\sigma^2_{pi0}} \tag{10}$$

$$\mu_{pi,t+1} = \frac{\tilde{\sigma}^2_{ift}\mu_{pit} + \sigma^2_{pit}\tilde{\mu}_{ift}}{\tilde{\sigma}^2_{ift} + \sigma^2_{pit}}, \tag{11}$$

where the variable s_{ift} 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 chronic health shocks occur, updating of probabilities after each period would result in posterior beliefs that are tightly centered around the initial mean, varying little with new information. In such a regime, individuals would have to perceive health shocks as being impossibly likely (e.g., $\tilde{\mu}_{ift}$ much greater than 1) in order for health shocks to meaningfully change health beliefs. This is inconsistent with the analysis I have presented previously, which shows that individuals are highly responsive to chronic health shocks.[25]

[25]In addition to the analysis presented here, I also find that older individuals have stronger responses to chronic health events in their household than younger individuals, even after conditioning for risk score (not shown). If individuals behaved as though they updated their health beliefs in each period—regardless of if a signal or health event occurred—then older individuals would have belief distributions more tightly centered around their mean, hence their posterior distributions following a realized health signal would shift *less* than younger individuals with more flexible priors. I do not observe this to be the case.

I address this inconsistency in my preferred specification by assuming that households update their beliefs *conditional* on a health event occurring. This reduces the number of uninformative signals individuals process, and hence avoids problems of weight degeneracy, and is consistent with 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 their health risk beliefs (e.g., after a diagnosis has occurred within the household), they do so in a completely standard way, including updating beliefs in all following years without major health events.

Such an approach is an intuitively appealing way to deal with the issue of Bayesian updating when signals are infrequent. However, my results are robust to alternative specifications, including (i) an adaptive learning framework where individual beliefs are specified as an AR(1) with some dependence $\rho < 1$ on the previous period's beliefs, and (ii) a more traditional setup where individuals update their beliefs with some probability $p > 0$ in the absence of health events.[26] 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), or where individuals over-weight "good news" relative to "bad news" (Eil and Rao, 2011).

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 faces a choice of plans that varies at the firm-year-state level.[27] Households are characterized by their unobserved type variables $\{p_{ift}, \lambda_{ift}, \psi_{ft}\}_{i \in \mathcal{I}_f}$. I allow the initial parameters $(p_{if0}, \lambda_i ft, \psi_{f0})$ to be arbitrarily correlated, and link them to observable data by assuming that they are drawn from a multivariate normal distribution which depends on observed demographics:

$$\begin{bmatrix} p_{if0} \\ \mu_{\lambda if} \\ \log(\psi_{f0}) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \beta_p \boldsymbol{X}^p \\ \beta_\lambda \boldsymbol{X}^\lambda \\ \beta_\psi \boldsymbol{X}^\psi \end{bmatrix}, \begin{bmatrix} \sigma_p^2 & & \\ \sigma_{p,\lambda} & \sigma_\mu^2 & \\ \sigma_{p,\psi} & \sigma_{\lambda,\psi} & \sigma_\psi^2 \end{bmatrix} \right). \tag{12}$$

[26] 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).

[27] 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.

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

Individual beliefs evolve in response to signals about their health risks as discussed in section 4.2.3. I assume that these signals y_{ift} are normally distributed with variance σ_π^2 (to be estimated) and a mean given by the logit regression:

$$y_{ift} = \pi_1 \mathbb{1}\{\text{chronic}\}_{f,-i} + \pi_2 \mathbb{1}\{\text{acute}\}_{f,-i} + \pi_3 \mathbb{1}\{\text{acute}\}_{f,i} + \pi_4 x_{ift}, \tag{13}$$

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

To parameterize the distribution of acute health shocks, I assume that $F_\lambda(\cdot)$ is a shifted lognormal distribution. This is a natural parameterization as the distribution of annual health expenditures is highly skewed (Mitchell, 2020). 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 model an individual's (correct) beliefs about their transient health shocks by

$$\ln(\lambda_{ift} - \kappa_{if}) \sim \mathcal{N}(\mu_{\lambda,if}, \sigma^2_{\lambda,if}). \tag{14}$$

When κ_{if} is sufficiently large (and negative), small and negative values of λ_{ift} may lead to zero spending being the utility-maximizing solution for an individual.[28]

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

In contrast, I directly use empirical distributions of chronic care costs from my data in household expected utility. I assume that individuals have rational expectations over the distributions of their chronic health care costs, which change when they experience major

[28]Previous work has allowed the distributions of these shocks to evolve over time. 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_{ift} already present.

health events. This is a simplifying assumption employed for tractability, as my model already allows for the identification of rich heterogeneity governing individual expectations about health shocks. However, although there is evidence that consumers do not fully know the price of health care before selecting services (Lieber, 2017), this is less concerning with chronic care costs, which are typically stable over time and hence more easily predicted by household members. The empirical distributions are similarly assumed to be stable across years, but I use a separate distribution in the year of diagnosis to accommodate potentially higher costs in that year (e.g., for unexpected hospitalizations).

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

$$\psi_{ft} = \gamma_0 \psi_{f,t-1} + \gamma_1 \left\{ \text{Post}_t \times m_{f0}^{\text{CH}} \right\} + \gamma_2 \left\{ \text{Post}_t \times c_j(m_{f0}^{\text{CH}}) \right\} + \gamma_3 \left\{ \text{Post}_t \times \text{Hosp}_{f0} \right\} + \zeta_{ft}, \quad (15)$$

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

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

I estimate the model via maximum likelihood with the appropriate adaptation for modeling a discrete choice followed by a continuous one (Dubin and McFadden, 1984; Revelt and Train, 1998; Train, 2009). For a given household, likelihood functions are constructed as the density of their observed health spending conditional on their observed plan choices. I provide additional estimation details in Appendix D.

4.3.2 Identification & Interpretation

My model utilizes multiple sources of variation to separate multiple effects arising from major medical events. In addition to any changes in individual risk beliefs, health events may alter health behaviors by changing the price of non-chronic care, increasing the salience of health consumption, providing experiential learning about how to obtain high-quality health care, or altering preferences for medical care in other ways. 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.

I use a rich set of major health events that vary in their expected costs, both in the year of diagnosis and in following years. This variation in the expected costs needed to maintain health for someone with a chronic condition changes the extent to which a specific chronic condition significantly alters the expected prices for other, non-chronic medical care. This variation, coupled with variation in plan spending characteristics, allows me to separate moral hazard effects from other drivers of behavior.

To separate risk aversion from beliefs, 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 wide range of cost-sharing parameters (Table 2). 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, highly risk-averse households will gravitate towards the plans in their choice sets that most limit high expenses (e.g., low-deductible plans). Finally, I use data on the circumstances of major medical events—including the resulting costs and whether a hospitalization occurred—to incorporate the role of salience associated with health trauma in changing household risk aversion.

The principal estimated structural parameters of interest in my model are those governing the evolution of the transition probabilities p_{ift}. 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_{ift} responds to new diagnoses. I therefore interpret changes in p_{ift} as resulting from an aggregate informational effect, rather than from moral hazard or salience effects.[29]

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

[29] Appendix C discusses an alternative interpretation of p_{ift} as a preference weighting across states rather than explicitly health beliefs.

information—including incidental parameters such as switching costs and individual mean-shifting regression coefficients—can be found in Appendix Table D.1.

I consistently find strong effects on non-diagnosed beliefs associated with household chronic diagnoses. New chronic diagnoses are associated with an average increase in an individual's belief of a major health event of 33 percentage points, an effect which is far larger than those estimated for acute events for either the individual or their family members, which are estimated to only increase risk beliefs by five and six percentage points, respectively. These increases are persistent, with little evidence that risk beliefs decrease over time (the estimated time trend coefficient is only one percentage point each year). The estimated variance for the unobserved dimension of belief changes is low, indicating that unobserved events are not contributing to large changes in risk assessments.

Table 5 also presents parameters illustrating how the effects of new chronic illnesses alter behaviors in other meaningful ways. Major health events—both acute and chronic—are associated with strong salience effects that increase household risk aversion. On average, experiencing a major health event increases the coefficient of household risk aversion by 0.61, a 34.9% increase over the pre-diagnosis average coefficient of 1.75.[30] 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.

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 weakly 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.

Models 1 and 2 of Table 5 illustrate simplifications of the model that help validate the estimated parameters and build intuition. In Model 1, I estimate a version of the model with no heterogeneity in acute health shocks or changes in household risk aversion. That

[30]To put these numbers into context, I follow the results of Cohen and Einav (2007) and consider the amount \$X that would make the average household in my sample indifferent between a sure payoff of \$0 and an equal-odds gamble between winning \$100 and losing \$X. Prior to a diagnosis, the average value of \$X is roughly \$85.08; after diagnosis, this value changes to \$80.85. These results are comparable with previous estimates of household risk aversion for health insurance (Einav et al., 2013; Marone and Sabety, 2021)—however, as mentioned in Einav et al. (2013), the coefficients from models incorporating both health and financial risk do not compare to those of models with pure financial risk (Cohen and Einav, 2007; Handel, 2013).

Table 5. Estimated Structural Parameters of Interest

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

Notes: This table presents estimates for selected parameters of the structural model of health choice; Appendix Table D.1 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.

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. A key difference between Model 1 and my preferred specification is that the estimated impact of chronic health shocks on risk belief distributions is much higher when I do not accommodate heterogeneity either in period-level health shocks or salience effects. This result is intuitive, as the absence of this heterogeneity leads to the inaccurate "loading" of belief changes onto specific events.[31] This loading is observed on a comparable scale for coefficients for acute major health events as well; however, note that these effects are associated with higher overall variance in belief evolution, presumably because the simplified model attempts to explain multiple sources of variation through a single channel.

Column 2 adds variation in acute health status to the model while continuing to hold household risk aversion constant over time. Accounting for this heterogeneity explains a substantial portion of the belief evolution pattern suggested by the most simplified model, decreasing the size of the effect of all major health events by about two-thirds and the variance of unobserved belief shocks (σ_π) even more drastically. Similarly, including acute health shocks in each period reduces the estimated variation in initial coefficients of risk aversion and the correlation between risk aversion and beliefs, suggesting that including that accounting for variation across health states is important in estimating both health learning and salience effects. A key difference between column 2 and column 3 is that after incorporating the explicit modeling of salience effects, the estimated effect of major health events on belief changes is almost double. Notice that there is a strong negative correlation between household beliefs and risk aversion; this means that when estimated together, salience effects may have muted the estimated effect of belief changes. Hence, it is to be expected that separating salience effects from belief changes increases the estimated effect of events on beliefs.

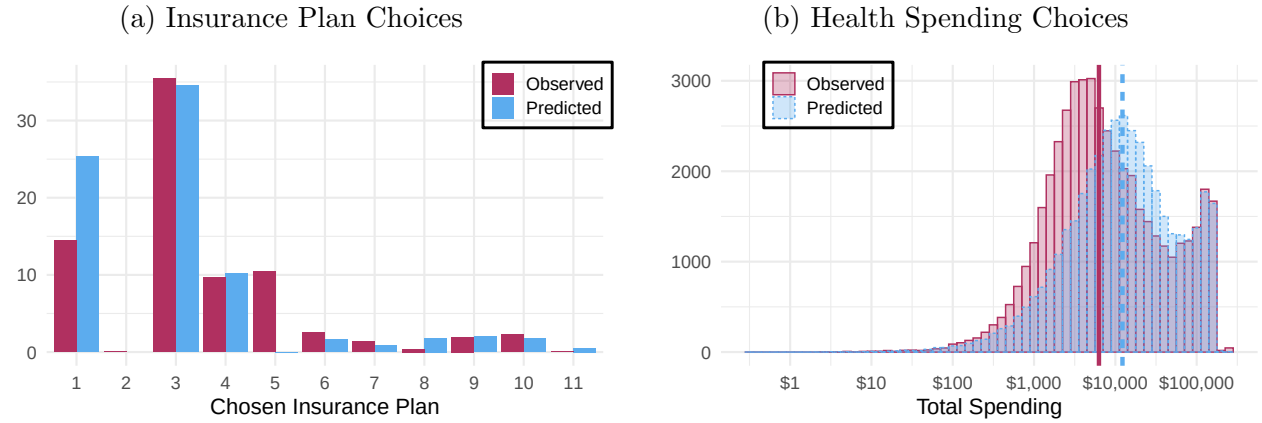
5.1 Model Fit

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 6. Predicted choice probabilities are influenced by premiums, inertia, and household expectations of their acute and chronic health shocks, valued based on their level of risk aversion. At the level of household spending, I compare observed household spending distributions to those predicted by the model. As spending decisions are made after the realization of two random variables (acute and chronic health

[31] This is exacerbated by the fact that acute health states and chronic diagnoses are correlated, as presented in Panel B of Table 5.

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 6. 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 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.

Figure 6 presents the results. The first panel shows the observed and predicted market shares for enrollment in plans offered in the largest firm in my sample. Overall, predicted shares are closely matched. The panel on the right presents observed and estimated spending conditional on a plan choice. Here, the model predicts slightly higher levels of billed spending than are typically observed, with a difference of about $1,000 between the means of the two distributions. The model appropriately predicts the extensive margin of spending, appropriately capturing the fraction of individuals who choose zero medical spending in a given year.

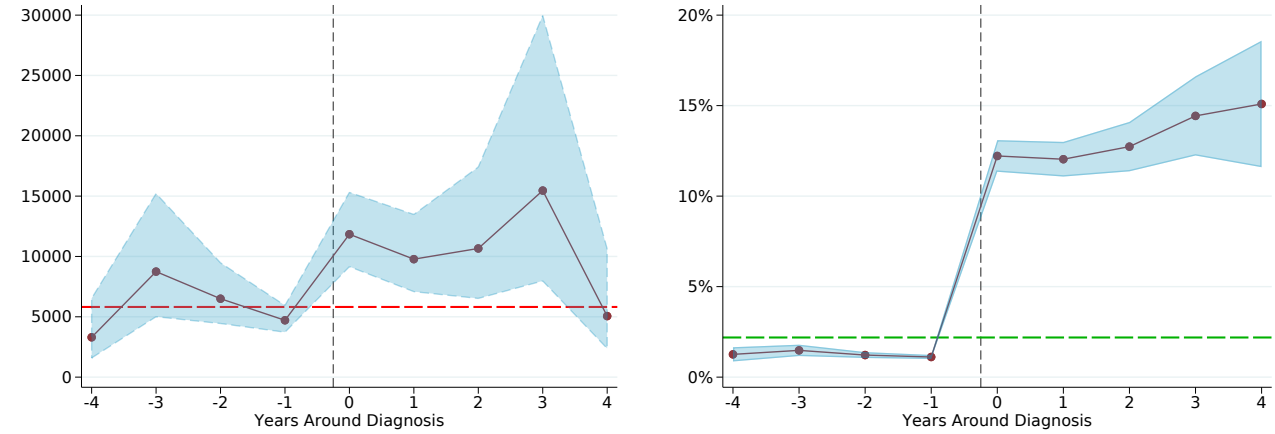
5.2 Spending Response to Major Health Events

Figure 7 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). In my model, household diagnoses are

associated large increases in OOP spending (about 20%, a difference which is statistically indistinguishable from the 10% reported earlier).

Figure 7. Model Predictions: Non-Diagnosed Spending and Beliefs Around a New Diagnosis

(a) Effect on Non-Diagnosed OOP Spending (b) Effect on Health Risk Beliefs



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

Importantly, I predict large accompanying changes in individual health risk beliefs following a new chronic diagnosis in the family. The horizontal green line in the Panel (b) of Figure 7 depicts the pooled average risk of diagnosis within my sample, which is roughly 2.5%. Prior to health events, individuals tend to underweight their health risks by about 58%; however, following a diagnosis, individuals move to *over-weighting* their risks by over *six* times the true in-sample rates of diagnosis. Instead, these households make choices as though they perceived their risk of a chronic diagnosis to be greater than one in ten. 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.

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 par-

ticular 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. My model allows me to estimate the spillover value of new health information for non-diagnosed household members by comparing these expected payoffs in the observed data—where household members use information to alter choices—and a counterfactual regime where the information is not revealed. In this counterfactual state, non-diagnosed household members experience the observed sequence of acute health shocks without any of the changes to p_{ift}, ψ_{ft}, or $c_j(\cdot)$ that would arise from a chronic event in the household.[32]

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

$$CE_{fjt} = -\psi_{ft}^{-1}\log(-U_{fjt}), \tag{16}$$

where U_{fjt} is the total *ex-ante* expected utility family f expects when enrolling in plan j at time t, as defined in equation 9. I assume that conditional on the estimated parameters, households are fully rational and enroll in the plan that gives the highest expected utility at the time of choice.[33] 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.

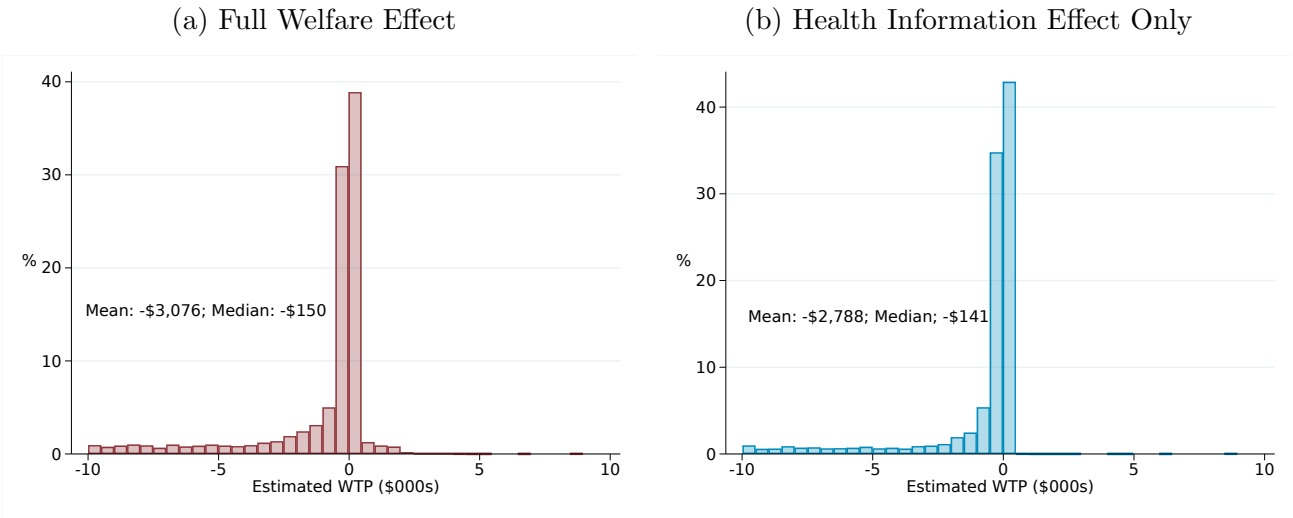
The utility-maximizing decision in my model is one where agents choose an appropriate level of spending relative to an uncertain multi-dimensional health shock; new health risk information changes the relative weight agents place on the dimensions of that shock when making their decisions. Hence, this welfare criterion measures how much households would be willing to pay for the information, based on their resulting changes in utilization choices during that period. My model does not allow me to measure the welfare effects of information in terms of long-term health production, for example from an increased investment in preventive health services. Such welfare effects are interesting particularly in conjunction

[32] I limit attention to non-diagnosed household members in order to estimate the *spillover* value of new health information, as well as to ignore the mechanical changes in household expected welfare that arise when a household member's health state is permanently altered, as with a new chronic diagnosis.

[33] 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.

with feasible health policies that jointly reveal information about health risk *and* the relative quality of health services. However, these returns would take more years to be realized than my sample permits me to analyze.

Figure 8. Variation in Welfare Effects Associated with Health Events and Health Information



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 8 depicts variation in household willingness to pay for health information in the year of the new chronic diagnosis.[34] Household members who are exposed to a new chronic diagnosis experience a welfare penalty 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 8 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 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 beliefs. The average household experiences a welfare penalty of $2,788 associated with changes to how they evaluate their

[34]These welfare effects are stable in the first few years following the diagnosis; hence, for ease of interpretation, I only focus on the year of diagnosis itself.

risk of developing a chronic condition. This corresponds to an average decrease in welfare of about 11.6% (Appendix Figure D.1).

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 (Appendix Figure D.2). 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 to an already high level of expected spending, meaning new health events change outcomes (in percentage terms) less.

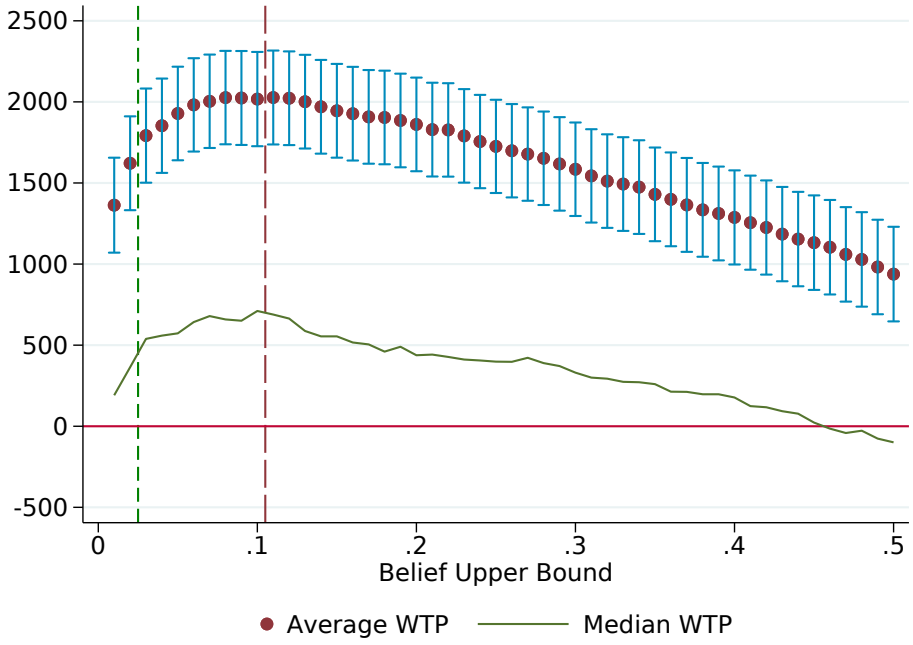
6.2 Evaluating Household Over-Responsiveness to Information

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 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 upper bounds on an individual's beliefs about their own health risks; that is, imposing that any predicted value p_{ift} in the model be no greater than some threshold $\overline{p}$. This exercise illustrates 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.

Figure 9. WTP ($) for Health Information After Bounding Responsiveness



Notes: Figure depicts estimated household willingness to pay for new health information across multiple counterfactual scenarios in which post-event health beliefs are capped at $\overline{p}$. Each point represents a distinct scenario with $\overline{p}$ indicated along the x-axis. Average and median household WTP for new information are depicted as the maroon scatter plot (with 95% confidence intervals) and the smoothed blue-gray line, respectively. The vertical dashed green line represents the in-sample rate of diagnosis (about 2.5%), while the long-dashed maroon line represents the upper bound at which welfare is maximized (about 10%).

Figure 9 presents the results. The figure summarizes household WTP for information across multiple scenarios, each with a varying degree of restrictiveness on $\overline{p}$. Average and median welfare gains are plotted; notice that the distribution of welfare gains is skewed as suggested in Figure 8. As opposed to a scenario with no restrictions—where the median

household's informational WTP was -$141—the median household would be willing to pay a positive amount for information whenever $\overline{p}$ is less than 45%. Welfare gains continue to improve as this bound becomes more restrictive until $\overline{p}$ is about 11% (shown in the Figure as the marroon long-dashed line). At this point, the average (median) household's welfare is estimated to be $2,027 ($711); in addition, about 86% of households receive welfare benefits from information, compared to 0.2% in the baseline scenario.

As the upper bound moves past this point, average household welfare gains begin to diminish. The belief upper bound which achieves an average WTP maximum is larger than the true in-sample risk of diagnosis (shown in the Figure as the green dashed line); this is because declines in consumer welfare following this point represent heterogeneous returns to new health information. Although the generic household in the model prefers, *ceteris paribus*, to have beliefs matching their true risk of chronic diagnosis onset (due to the state-dependence of preferences for non-chronic care), these risks vary across households. For some, these risks skew much higher than the average rate of illness onset, meaning that arbitrary bounds such as $\overline{p}$ risk harming households for whom information *does*, in fact, reveal large changes to beliefs.

To examine this further, 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. The predicted probabilities are small and match in-sample diagnostic risks.

When I impose these predicted probabilities as individual-specific upper bounds, I find that the average household would be willing to pay $2,385 for information, an 18% increase in average returns over the welfare-maximizing point in Figure 9. This underscores that exploiting individual risk characteristics to further refine household responsiveness can increase welfare. Importantly, accommodating for these heterogeneous returns to information explains the average differences between the welfare-maximizing upper-bound $\overline{p}$ predicted by the model and the in-sample rate of diagnosis demonstrated by the data. I explore methods to harness these heterogeneous returns to maximize social welfare of information-revealing social policies in the following section.

6.2.2 Targeting Information to Maximize Gains

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 (Appendix Figure D.2). Based on these results, I consider the effects of targeting information revelation based on observable characteristics, such as individual risk scores.[35] 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_{ift}. When individuals receive this information, this probability is adjusted to be equivalent to their predicted risk probability $\hat{p}_{ift}$ defined above. I assume that following this information, individual beliefs are constant at their predicted risk level, with no residual uncertainty or updating across periods.[36] As before, I assume away salience and moral hazard effects.

This scenario therefore mirrors a hypothetical transmission of health information that informs consumers of their health risks as perfectly as population-data allows.[37] I present results of the individual and social value of this revelation based on 50,000 households in my sample which do not experience major health events. These individuals may still have erroneous beliefs about their health risks and may benefit from new health information. Furthermore, the estimated welfare effects of this policy validates the results presented earlier, documenting the value of information transmitted in a more quasi-random setting.

Figure 10 presents the results, showing both average welfare gains and the fraction of targeted households benefitting from the information. Each point represents a scenario in which only individuals with risk scores falling in the top $x\%$ of the sample receive health information. The average household in the full sample would be willing to pay approxi-

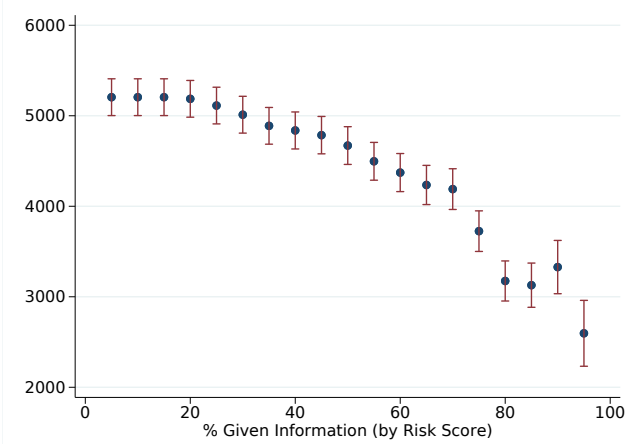
[35] Appendix Figure D.2 illustrates that other, less-easily observable characteristics (e.g., household risk aversion) may also be beneficial.

[36] I ignore residual uncertainty that would arise from individuals treating information revelation as a single signal, rather than true information. For the purposes of this exercise, such fluctuations would serve only to obfuscate the potential benefits of targeting information revelation relative to universal revelation.

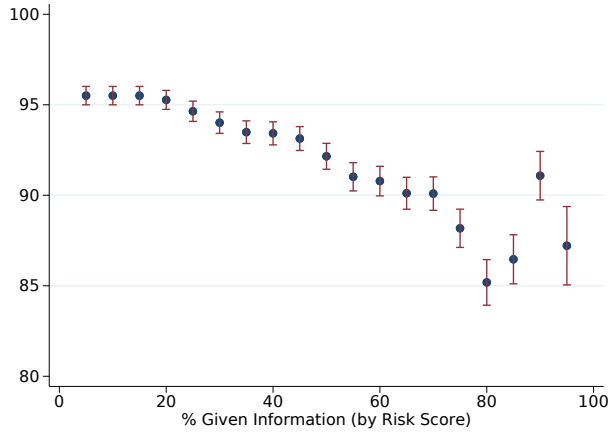
[37] Note that $\hat{p}_{ift}$ is not equivalent to one's *true* risk as private information (e.g., underlying health status) is not incorporated.

Figure 10. Changes in Welfare Gains From Targeted Revelation of Information

(a) Average Welfare Gains (WTP, $)



(b) Households with Welfare Gains (%)



Notes: Figures show estimated welfare gains from revelation of health information. Individuals are organized by their average risk scores, from highest to lowest. Each point in both panels represents a different counterfactual scenario, where individuals with risk scores in the top x% of the sample are given information about their predicted health risks, $\hat{p}$, as described in the text. Returns to health information are presented as (a) average expected welfare changes, measured as willingness to pay in 2020 USD, and (b) the percentage of households with non-negative welfare gains.

mately $2,500 per year for updated health information (the right-most point in Panel (a)); this information benefits roughly 85% of households (the right-most point in Panel (b)).[38] In contrast, revealing information only to higher-risk individuals improves welfare gains: revealing information only to individuals within the top quartile of the risk score distribution increases average welfare gains to over $5,000 per household per year, benefiting more than 95% of households.

Hence, even policies that are capable of revealing information that closely matches individuals' true risks without inducing salience responses, moral hazard effects, or over-responsiveness may still benefit from using demographic information to identify households that are most likely to benefit from the policy. For example, policies such as universal genetic screening programs—such as common programs in the U.S. providing risk information to newborns in many developed countries—may incur private welfare costs to specific households, even as they improve societal welfare more generally.

[38]Not every household in the sample benefits from information about predicted risk. There are two reasons why even such high-quality information may make a household worse off. First, the household may have private information regarding their true risks, making public information counter-productive. Second, highly risk averse households may benefit from placing smaller weights on the adverse state of the world than are objectively accurate; this is similar to an "optimal expectations" model where individuals do not benefit from information when it lowers utility in an anticipation period (Oster et al., 2013). Overall, this highlights a central tension inherent in the dissemination of health information: even high-quality information can incur individual welfare costs based on how households value health care across states.

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. These changes in behavior are best explained by individual household members reassessing their risks, rather than responding to financial incentives or salience effects. However, these reassessments do not meaningfully improve the quality of their health care choices. 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 use a structural approach to quantify a household's willingness to pay for health information, isolating the specific effects of new health information from other mechanisms. The model implies low realized returns to health information, most likely due to individual misinterpretation of their health risks following the health event. Bounding the extent to which individuals increase their beliefs about risks post-diagnosis substantially improves realized welfare. Finally, my analysis illustrates that information revelation is privately most optimal for individuals with high *ex-ante* risk and those with low risk aversion.

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 covered 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. Family health experiences are powerful forces in shaping individual behaviors and decisions; however, witnessing these experiences may lead individuals to "over-react" when making future consumption decisions. 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 temper 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.

Altmejd, A., Barrios-Fernandez, A., Drljie, M., Goodman, J., Hurwitz, M., Kovac, D., Mulhern, C., Nielson, C., and Smith, J. (2021). O brother, where start thou? Sibling spillovers on college and major choice in four countries. *The Quarterly Journal of Economics*, 136(3):1831–1886.

Archibong, B. and Annan, F. (2021). 'We are not guinea pigs': The effects of negative news on vaccine compliance. *SSRN Electronic Journal*.

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.

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.

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.

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.

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.

Browning, M., Chiappori, P.-A., and Weiss, Y. (2014). *Economics of the Family*. Cambridge University Press.

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 the COVID-19 pandemic. Technical Report w28741, National Bureau of Economic Research, Cambridge, MA.

Callaway, B. and Sant'Anna, P. (2018). Difference-in-differences with multiple time periods and an application on the minimum wage and employment. *SSRN Electronic Journal*.

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.

Chaisemartin, C. d., D'Haultfoeuille, X., and Guyonvarch, Y. (2021). DID_multiplegt: Stata module to estimate sharp Difference-in-Difference designs with multiple groups and periods.

Chen, M. (2021). Family of San Diego COVID-19 victim makes emotional vaccine plea. *ABC10 News*. Section: Local News.

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.

Cohen, A. and Einav, L. (2007). Estimating risk preferences from deductible choice. *American Economic Review*, 97(3):745–788.

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.

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

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. Technical Report Technical Report w25904, National Bureau of Economic Research, Cambridge, MA.

Dubin, J. A. and McFadden, D. L. (1984). An econometric analysis of residential electric appliance holdings and consumption. *Econometrica*, 52(2):345–362.

Dupas, P. (2011). Health behavior in developing countries. *Annual Review of Economics*, 3(1):425–449.

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

Eichner, M. J. (1998). The Demand for Medical Care: What People Pay Does Matter. *The American Economic Review*, 88(2):117–121.

Eil, D. and Rao, J. M. (2011). The good news-bad news effect: Asymmetric processing of objective information about yourself. *American Economic Journal: Microeconomics*, 3(2):114–138.

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.

Epstein, L. G., Noor, J., and Sandroni, A. (2010). Non-Bayesian Learning. *The B.E. Journal of Theoretical Economics*, 10(1):0000102202193517041623.

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*.

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*.

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.

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.

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.

Giardinelli, C. (2021). 'Get the vaccine:' Oregon man pleads as 23-year-old wife fights for her life. *KATU ABC2*.

Goodman-Bacon, A. (2018). Difference-in-differences with variation in treatment timing. Technical Report w25018, National Bureau of Economic Research, Cambridge, MA.

Goodman-Bacon, A., Goldring, T., and Nichols, A. (2019). BACONDECOMP: Stata module to perform a Bacon decomposition of difference-in-differences estimation.

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.

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.

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. 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.

Harris, J. W. and Stöcker, H. (1998). *Handbook of Mathematics and Computational Science*. Springer Science & Business Media.

Hodor, M. (2021). Family health spillovers: Evidence from the RAND health insurance experiment. *Journal of Health Economics*, 79:102505.

Holt, C. A. and Smith, A. M. (2009). An update on Bayesian updating. *Journal of Economic Behavior & Organization*, 69(2):125–134.

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.

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.

Kahneman, D. and Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4):237–251.

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.

KFF (2020). 2020 Employer health benefits survey. Technical report, Henry J. Kaiser Family Foundation, San Francisco, CA.

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.

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.

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

Lieber, E. M. (2017). Does it pay to know prices in health care? *American Economic Journal: Economic Policy*, 9(1):154–179.

Marone, V. and Sabety, A. (2021). Should there be vertical choice in health insurance markets? *American Economic Review*, In Press.

Mitchell, E. M. (2020). *Concentration of Health Expenditures and Selected Characteristics of High Spenders, U.S. Civilian Noninstitutionalized Population, 2017.* Statistical Brief #521. Agency for Healthcare Research and Quality (US), Rockville, MD.

Ortoleva, P. (2012). Modeling the change of paradigm: Non-Bayesian reactions to unexpected news. *American Economic Review*, 102(6):2410–2436.

Oster, E., Shoulson, I., and Dorsey, E. R. (2013). Optimal expectations and limited medical testing: Evidence from Huntington's 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.

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.

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.

Rabin, M. (2013). Incorporating limited rationality into economics. *Journal of Economic Literature*, 51(2):528–543.

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.

Rios-Avila, F. and Naqvi, A. (2021). DRDID: Doubly-robust differences estimators for stata.

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

Salcedo, A. (2021). He put off getting vaccinated. Now, he's in the ICU pleading for others to avoid his mistake: 'I messed up'. *Washington Post*.

Sant'Anna, P. H. C. and Zhao, J. (2020). Doubly robust difference-in-differences estimators. *Journal of Econometrics*, 219(1):101–122.

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.

Song, Z. (2021). Social learning and health care utilization: Evidence from spousal health shocks. *Working Paper*.

Spinnewijn, J. (2015). Unemployed but optimistic: Optimal insurance design with biased beliefs. *Journal of the European Economic Association*, 13(1):130–167.

Sun, L. and Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, page 53.

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.