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

Alex Hoagland†

December 13, 2022

Abstract

Individuals infer their health risk from observing the health experiences of people around them, particularly family members. I assess how people interpret new information from household health events. When an individual is newly diagnosed with a 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. These responses are most consistent with individual reevaluations of their own health risks in the face of new information. To assess welfare implications, I estimate a structural model of health choices where individuals learn about risk from their own investments and household health events. I find that consumers over-respond to health events by over-weighting their risks *ex-post*, crowding out potential welfare gains from new information for almost half of affected individuals. Over 90% of individuals would benefit from health information were responsiveness limited.

Keywords: Health spillovers, consumer learning, behavioral health economics, discrete choice models, chronic illness

JEL codes: I12, I13, D83, D91, D12

Total Word Count: 12,453

*I am grateful to Kristen Blair, Corinne Andriola, Jason Abaluck, Marshall Drake, Randall Ellis, Keith Ericson, James Feigenbaum, Raymond Fisman, Sarah Gordon, Tal Gross, Jihye Jeon, Timothy Layton, Patrick Power, Marc Rysman, and Paul Shafer for very helpful comments on this proejct. In addition, I am grateful to seminar participants from economics departments at Boston University, Brigham Young University, the Ohio State University, the University of Toronto, and Washington University in St. Louis for useful feedback. I am also grateful to seminar participants from the 10th Annual Conference of the American Society of Health Economists, the 2021 Congress of the International Health Economics Association, the Congressional Budget Office, and the Federal Trade Commission for their comments on the paper.

†University of Toronto, Institute of Health Policy, Management, and Evaluation, 155 College St, Toronto, Ontario, Canada, M5T 3M6. Email: alexander.hoagland@utoronto.ca. 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 information that healthcas shocks casunicate is knowledge about individual health risks, particularly if the shock highlights dimensions of risk which are correlated across family members. Affected individuals may alter their expectations about future health care needs, and even choose to increase take-up of high-value health services after witnessing a family member's health experience. For example, individuals may choose to become vaccinated against COVID–19 after witnessing the infection of a family member (Chen, 2021; Giardinelli, 2021; Salcedo, 2021).

There is evidence that family members react to the health events of their loved ones (Fadlon and Nielsen, 2019; Hodor, 2021), but it remains unclear what drives these reactions. Health events may lead individuals to reassess their health risks, but may also change the expected price of household medical care, alter household preferences for health consumption, or provide knowledge about the availability of health services. Understanding the role that social connections play in altering individual health choices—including the use of both high-return 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 health shocks affecting a household—in particular, the diagnosis with a new chronic condition—change individual assessments of health risks, and how these changes disrupt healthcare consumption. Using claims data for U.S. households with employer-sponsored insurance (ESI) between 2006 and 2018, I show that these health events generate strong informational spillovers among non-diagnosed household members. I find that affected individuals significantly and persistently increase both their overall health care utilization and their investment in preventive care, particularly the use of screenings for the condition just diagnosed in their household.

I then show that individual responses are more consistent with learning about risk than other 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 similarly salient acute hospitalizations; this suggests that households are responding to more than the salience effects of a traumatic health experience (Dalton et al., 2020; Fadlon and Nielsen, 2019). Finally, I show that these responses persist even among individuals who are most familiar with the healthcare system, such as those with current preventive prescriptions for cardiovascular health. This implies that learning about health *systems*, rather than health *risk*, does not drive my results.

In general, one would expect receiving new information about risk to improve decision-making and welfare. Surprisingly, however, the welfare effects of information from household health shocks are ambiguous utilizing only reduced-form evidence. I observe that affected household members increase take-up of "low-value" health care, services that are generally agreed to be cost ineffective due either to their reach (e.g., benefiting few patients) or their average returns (e.g., low levels of benefits relative to costs) (Colla et al., 2015). These responses are most likely to include increased utilization of low-value services that appear,

[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).

from a patient's perspective, closely related to preventive care, including extraneous pre-operative screenings or imaging services. This casts doubt on the extent to which health information improves choice quality.

These findings motivate a structural approach to model the evolution of individual beliefs about risk in the wake of household health shocks. I write and estimate a dynamic model in which forward-looking individuals learn about their health risks in response to investments in preventive screenings as well as diagnoses affecting other household members. The model combines dynamic investments in health production (Grossman, 1972) with static consumption decisions trading off risk protection and health consumption (Cardon and Hendel, 2001). My model allows me to highlight the role of individual learning from new diagnoses separately from the potential moral hazard or salience effects of a new diagnosis in a home. Further, a structural approach also allows me to estimate the welfare effects from individual responses to new information.

Counter to expected thought, information transmitted from household health shocks is not welfare-improving for a large subset of affected individuals. Nearly half of those presented with new risk information would be willing to pay to avoid the resulting change in their beliefs, with welfare losses averaging 90% of baseline expected utility. 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: over 90% of individuals in my sample would find health information welfare-improving were their responses mitigated.

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 educa-

tion 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).[2] I contribute to this discussion in three ways. First, I highlight chronic diagnoses as a new form of risk information, to which individuals respond more than even other acute health events. Second, I show that changes to how individuals assess health risks 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 the potential for individual welfare losses.

Second, I contribute to a literature on non-Bayesian learning in models of health behavior (Barseghyan et al., 2018; Bundorf et al., 2021a). My model combines findings from two distinct threads of the learning literature. First, I emphasize the finding of many previous models that individuals place disproportionate weight on recent, salient events (Kahneman and Tversky, 1973). For example, previous work has found that individuals significantly overweight new evidence when it had a lower *ex-ante* probability of occurring (Epstein et al., 2010; Holt and Smith, 2009). Such disproportionate responses have been found to rationalize individual choices that would otherwise require unreasonably high levels of risk aversion to justify (Ortoleva, 2012; Paserman, 2008; Spinnewijn, 2015). Second, I combine these results with a quickly-evolving literature studying the role of peer signals in learning. In particular, I study how individuals may emphasize peer signals relative to their own when those signals are low-probability; this is related to recent work studying how individuals value signals from sufficiently related (or distinct) distributions (Dasaratha et al., 2022).

[2] 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).

I incorporate findings from these threads of the literature into the first structural model of health risk belief formation. My model incorporates a fully flexible specification for misinterpretation of information (Hauser and Bohren, 2021), and provides a micro-foundation for how individuals form beliefs in a setting of largely small-probability events. My results highlight that even fully forward-looking consumers with rational expectations may over-respond to health shocks *ex-ante* in a quasi-Bayesian framework, resulting in potential welfare penalties from information. 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 risks may arise (Arni et al., 2021).

Finally, my work is relevant to the well-established literature exploring sub-optimal health decisions made by many consumers (Abaluck and Gruber, 2011, 2016a; Baicker et al., 2015; Handel and Kolstad, 2015).[3] This includes discussions 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 overcoming information frictions is not simply a matter of increased access to health information.[4] In particular, individual responses to some forms of health information may not align utilization with service quality, but may simply shift consumers from one type of poor decision-making to another, all while increasing total health spending.

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. Finally, I discuss the relevance of my findings and directions for future work in Section 6.

[3] See Abaluck and Compiani (2020) for a more thorough discussion of these results.

[4] Finkelstein et al. (2022) find a similar result when considering drivers of adherence to medication guidelines for high-return pharmaceutical treatments.

2 Empirical Setting & Data

My primary data on household 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 and firm identifiers from 2006 to 2018. 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.[5]

Table 1 presents summary statistics for the full sample as well as for households in which an individual is affected by a chronic condition. In general, households are comprised of one to two adults and one to two children, with relatively generous insurance coverage. The average (median) household pays out-of-pocket for roughly 18% (16%) of their annual health consumption, and 21% of individuals in the sample do not face any cost-sharing during a year. In addition, roughly a quarter of households are enrolled in plans with no deductibles.[6]

[5]The average household is continuously observed for 7 years in the data. Note that although the results presented in the main text utilize an unbalanced panel—potentially contributing to increased levels of uncertainty at the tails of event study specifications—the results are robust to specifications requiring a fully balanced panel across 6 years.

[6]While insurance contracts are defined by a complicated set of cost-sharing measures—including copayment and coinsurance rates that vary widely across provider specializations, networks, and procedures—the structural model described in Section 4 uses only a family deductible, a simplified non-specialist coinsurance rate, and a family OOP maximum, consistent with prior work (Marone and Sabety, 2022). These measures are constructed using the empirical distributions of payments in the claims data, and described in detail in Appendix A.1 (Zhang et al., 2018). I find that these simplified measures capture a wide degree of variation in my data and harmonize well with measures from earlier work.

Table 1. Household Summary Statistics

	Full Sample	Households Affected by Chronic Events
Panel A: Household Demographics		
Family size	2.84 (0.001)	3.11 (0.004)
Employee age	45.01 (0.007)	43.61 (0.039)
Enrollee age	30.87 (0.008)	29.37 (0.041)
% female employees	41.57 (0.037)	41.04 (0.190)
% female enrollees	50.17 (0.021)	50.11 (0.109)
Risk Score	0.95 (0.001)	1.51 (0.008)
Panel B: Household Medical Utilization		
Total medical spending	\$2,504 [\$680] (4.51)	\$4,546 [\$1,130] (73.13)
OOP medical spending	\$443 [\$110] (0.53)	\$614 [\$175] (4.39)
% enrollees w/ 0 spending	15.39 (0.015)	10.35 (0.067)
% enrollees w/ 0 OOP	21.04 (0.017)	14.68 (0.077)
Household deductible	\$415 (0.619)	\$419 (3.094)
% w/ 0 deductible	28.04 (0.032)	32.47 (0.180)
Panel C: Individual Major Medical Events		
Total cost, Diagnosis	—	\$4,164 [\$1,319] (156.81)
OOP, Diagnosis	—	\$532 [\$212] (27.36)
OOP, Recurring	—	\$489 [\$190] (24.78)
$N_{\text{households}}$	353,403	62,528
$N_{\text{individuals}}$	1,087,353	194,844

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. Column 2 limits the sample to only household-years in which a chronic diagnosis occurred (see Appendix A for diagnostic codes).

2.1 Major Health Events

Major health events, which communicate information about health risk to households, are identified by diagnostic codes in the claims data. I identify these events 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 classifi-

cation of health events to non-pregnancy HCCs that occur with high frequency as discussed in Appendix A.2.[7]

Table 1 shows how households affected by these chronic conditions differ from the full population. The second column of the table limits the sample to only household-years in which a chronic diagnosis occurred. Affected households are riskier, on average—note that this is mechanical, as HCCs are directly used in the calculation of a risk score. In the year of diagnosis, the average (median) household spends about 82% (66%) more on health services than the corresponding household in the full sample. In contrast, affected households look very similar to the full sample in terms of insurance enrollment and plan generosity.

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.[8] 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 $532 ($212) OOP on the associated diagnosis, and then $489 ($190) each year that follows on recurring costs needed to care for chronic conditions.

2.2 Additional Variable Definitions

The rich variation of the data allows me to evaluate the impact of new chronic diagnoses on a wealth of utilization and quality measures. In particular, I define four additional outcome variables which will be useful in identifying the mechanisms by which new health information changes household behavior: preventive health services, acute health events, adherence to prescription medication, and the use of low-value health services. Appendix A contains a

[7] To ensure that I identify new diagnoses, I require that relevant diagnosis codes appear during or after an individual's second observed year.

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

full set of all diagnostic information, procedure codes, and therapeutic classes used in the construction of each of these variables.

First, I define a set of health services typically considered to be preventive in nature, consistent with previous work and federal guidelines (Hoagland and Shafer, 2021; USPTF, 2022). Preventive screenings and wellness visits constitute an important point of entry for the identification of other health concerns (Jiang et al., 2018) and are generally considered to be an important form of high-value care (Tong et al., 2021). For each enrollee, I identified individual preventive services based on commonly used code combinations recommended by the United States Preventive Task Force (USPTF). I include as preventive all billed services that meet these guidelines, including general wellness visits and disease-specific screenings (e.g., for diabetes or depression).

Second, similar to the set of chronic conditions used as the main treatment variable of analysis, I define a set of hospitalizations for *acute* health events. These events capture health events of a similar level of seriousness to new chronic diagnoses, but that are transient in nature. Identification of these events is used to compare how households respond to new health information as opposed to responses to a health event which may change their marginal utility of health consumption (see Section 3.3.1). I identify acute health events as new HCCs within households for conditions which typically do not persist past a year, including hospitalizations for severe viral infections or other non-chronic conditions.

Third, I define adherence to prescription medication. This measure is used to separately identify how changes in new risk information may alter household behaviors independent of other forms of information about the health system as an institution (see Section 3.3.2). I measure adherence to cardiovascular preventive drugs as the proportion of days covered in a year, in keeping with prior literature (Choudhry et al., 2009).[9]

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

Finally, I define categories of medical utilization which are frequently labeled as "low-value" by medical professionals and health officials (Chua et al., 2016; Colla et al., 2015).[10] Low-value services include both those whose cost typically outweighs the benefits to an average patient (e.g., some surgeries, such as arthroscopy) and services which are chronically over utilized in ways that dramatically lower their return (e.g., some imaging services, such as MRI for migraines). I define instances of low-value consumption based on an individual's diagnosis and procedure codes as well as their diagnostic history, based on previous work (Colla et al., 2015). I subdivide these services into 5 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; extraneous screening services for adults, including cardiac testing before low-risk surgeries; and adult surgical procedures.

3 Spillover Effects of Household Health Events

To estimate the causal impact of health shocks on health choices, I use a "stacked" regression design, similar to the approach of Cengiz et al. (2019). This approach avoids potential bias from naive staggered treatment designs in the presence of heterogeneous treatment effects within units over time (Goodman-Bacon et al., 2019; de Chaisemartin and D'Haultfoeuille, 2019).[11] This approach creates event-specific cohorts $r \in \{1, ..., N^1\}$ based on each treatment year, then stacks cohorts into a single estimation with cohort-specific time and group fixed

[10] 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).

[11] My results are robust across a number of alternative specifications, including the standard two-way fixed-effects (TWFE) model and alternative "doubly robust" estimators (Callaway and Sant'Anna, 2018; Sant'Anna and Zhao, 2020). In general, I do not find any evidence that my regression results suffer from concerns of negative weighting (Goodman-Bacon et al., 2019). Robustness checks are reported in Appendix B. In addition, I use a control group of never-treated observations in order separately identify the role of year fixed effects from the dynamic treatment effects, as discussed later in the text (Borusyak and Jaravel, 2016; Sun and Abraham, 2020).

effects, as specified in Equation 1:

$$Y_{tf,r} = \alpha_{f,r} + \tau_{t,r} + \sum_{k=-T}^{T} \gamma_k \mathbb{1}\left\{t - E_{ft,r} = k\right\} + \epsilon_{ft,r}. \tag{1}$$

In Equation 1, f denotes a household, t a year, and r a cohort indicating the first year of treatment (e.g., a new chronic diagnosis in the home); event-by-household and event-by-time fixed effects ($\alpha_{f,r}$ and $\tau_{f,r}$) are included. In my main specifications, $Y_{ft,r}$ includes health utilization outcomes for all family members *except* those who experience the major health event. I measure these outcomes in terms of both raw utilization (e.g., number of visits) and spending (both total and OOP); I adjust for the skewed nature of these distributions by using Poisson regression for count outcomes and the inverse hyperbolic sine transformation for spending outcomes.[12] Throughout, reported regression coefficients can be interpreted as approximate percentage changes in the outcome variable, relative to the year prior to the shock, $t-1$. Standard errors are clustered at the household level.

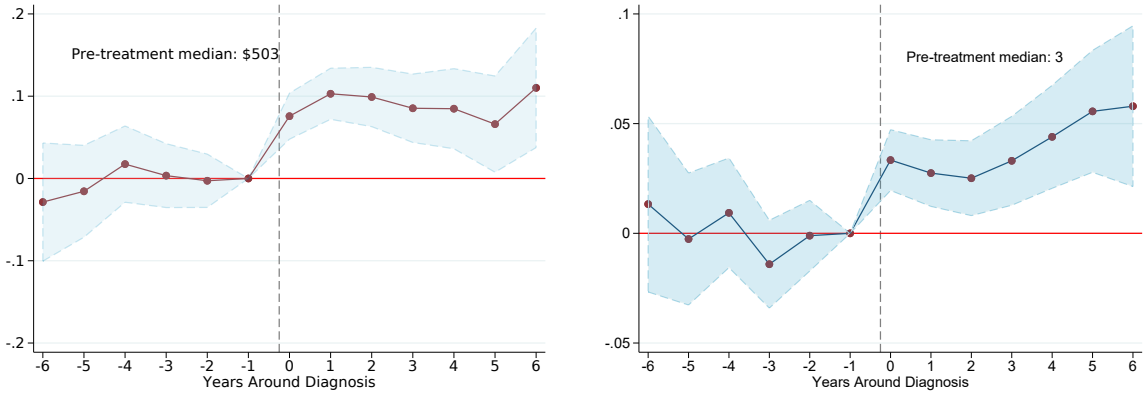
This approach allows me to identify the potentially time-varying effects of health shocks, which might have decaying influence on household choices over time. Identifying the appropriate comparison group is central to identifying these dynamic effects. Previous work has restricted comparisons to only households who have yet to experience a similar diagnosis, to more closely match the treatment group on unobservable characteristics. I include never-treated households in my estimation 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

[12] I use the inverse hyperbolic sine transformation for spending data to accommodate the approximately 15% of individual-years in my data with 0 spending (Harris and Stöcker, 1998; Bellemare and Wichman, 2020). I show in Appendix B that my results are not substantively altered when using alternative transformations, including the more typical $\log(y+1)$ or Poisson regression. I use Poisson regression for outcomes measured as count data, including number of visits.

or particularly life-threatening—concerns about violations of the parallel trends assumption are less plausible in my setting.

Figure 1. Effect of Chronic Diagnoses on Other Household Members' Healthcare Utilization

(a) $\sinh^{-1}$(Total OOP Spending) (b) # of Preventive Visits



Notes: Figures show regression coefficients from "stacked" TWFE regressions, with 95% confidence intervals. Regressions estimate the effect of a new chronic diagnosis on medical utilization of other (non-diagnosed) household members. In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on the number of household preventive services per year using Poisson regression. Coefficients are presented relative to the year prior to diagnosis. Standard errors are clustered at the household level.

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.[13]

While health events may generate spillovers in household spending for many reasons, households responding to the information contained in a diagnosis may be more likely to seek out preventive screenings. The second panel of Figure 1 estimates the effect of chronic diagnoses on the total utilization of preventive services, as defined in Section 2.2. Here, too, I find that new diagnoses in a household are associated with strong responses. Affected,

[13]These results are robust to the measurement of the outcome variable, including total billed spending or number of unique health encounters; see Appendix B.

non-diagnosed household members increase their overall use of wellness visits by about 5% relative to a median of 3 visits annually.[14]

For comparison, I present results in Appendix Figure B1 that illustrate how health events affect outcomes for the diagnosed individuals themselves. This additional analysis allows me to both benchmark the spillover results presented here and to test whether there is evidence that diagnosed individuals have an "anticipation stage" of deteriorating health (and increased spending) prior to a diagnosis. Such anticipation may introduce unobserved pre-trends into the empirical analysis. I find that diagnosed individuals more than double their annual OOP spending in the year of diagnosis (see Table 1), but that spending increases are not as persistent over time; additionally, I find no evidence that diagnosed individuals increase spending prior to diagnosis.[15]

3.1 Changes as Responses to New Health Risk Information

These results suggest a meaningful, persistent change in how a single diagnosis affects health-care utilization of an entire household. Next, I turn to exploring the mechanisms behind such responses. I first show that observed responses are indicative of individuals reassessing their own health risks in response to new health information from a major diagnosis. I then consider alternative explanations in Section 3.3.

I show that households respond to the risk information contained in a diagnosis by estimating the effect of a new diagnosis on the use of *disease-specific* preventive services. 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 (e.g., households which increase their takeup

[14] Appendix B shows that these results are robust to alternative outcome measures, including spending. Here, I show the main outcome measured in number of visits rather than spending to account for the fact that the Affordable Care Act (ACA)'s cost-sharing exclusion took effect in 2010 (or 2012 for certain women's health services), disrupting the costs for preventive services for those with ESI (Hong et al., 2017).

[15] Given that the events in my sample are mainly unexpected diagnoses affecting children (e.g., sudden-onset diabetes), this makes intuitive sense—these events would not typically be associated with a pre-diagnosis "anticipation" stage of longer than a year.

of diabetes screenings differentially more following a diabetes diagnosis than a non-diabetes diagnosis) provides evidence of responsiveness to risk information specifically.[16]

To assess the causal effect of diagnoses on the utilization of disease-specific preventive care, I use a triple-differences approach. This approach separates the disease-specific effect of risk information from the competing general effects which result from experiencing any chronic diagnosis (e.g., salience effects). 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 $chronic_f$ matches 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 β_{DDD} identifies the effect of the specific diagnosis of interest relative to other diagnoses.[17] 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.

[16] Appendix Figure B1 presents simple re-centered time series which validate this intuition in the case of diabetes. Households affected by a diabetes diagnosis increase their screening by about 36% in the first three years after diagnosis, while households affected by other chronic diagnoses do not change their utilization of diabetes screenings.

[17] 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.

The triple difference approach is advantageous because it allows me to compare the causal effect of diagnoses on the use of preventive care across multiple control groups. When the outcome variable of interest is a screening for a specific service (e.g., diabetes), this approach estimates the effect of a corresponding diagnosis relative to all other diagnoses, for which the screening reveals no information. In this context, the identifying assumption for the triple differences approach is the same as the identifying assumption for the simpler difference-in-differences regressions: that spending differences between diagnosed and undiagnosed households would have evolved similarly over time in the absence of treatment.[18]

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.[19]

Finally, to verify my results, I include two "placebo" regressions for cases where health events communicate no useful *risk* information, and hence are expected to change disease-specific screenings little. These include the effect of new diabetes diagnoses on obesity diagnoses (a diagnosis which, while an important risk factor for some types of diabetes, is externally verifiable prior to a household diagnosis), and the effect of a new household mental health condition on screenings for depression. In the second case, while mental health diagnoses may provide meaningful risk information to households, such information may incidentally *reduce* the value of preventive depression screening relative to immediately

[18]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.

[19]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.

seeking treatment. A lack of observed response in these placebo cases underscores the specific role of risk information in changing behavior.

Table 2 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%.[20]

The second panel of Table 2 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 (e.g., new information about the value of preventive care generally) do not appear to drive utilization decisions for preventive care.

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.[21]

[20]Similar to previous work, I also 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 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.

[21]I report additional results in Appendix B, which leverage variation in intra-familial relationships and corresponding risk to show that households are also selective in which members they choose to screen following a diagnosis. For example, 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.

Table 2. 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 six triple-difference regressions for how diagnoses affect household investments in disease-specific preventive care (Equation 2). Outcome variables are binary indicators for the screening in the first column; treatment variables are a binary indicator for the diagnosis in the second column. Difference-in-differences coefficients (β_{DD}) indicate the effect of *any* chronic diagnosis on screenings, while triple differences coefficients (β_{DDD}) indicate the (additive) effect of *specific* diagnoses. Standard errors clustered at the household level shown in parentheses. [1]Due to unavailability of procedure codes, these outcomes are measured using diagnostic codes. [2]Here, the reference group is all acute major health events. $^{*}p < 0.05$,$^{**}p < 0.01$,$^{***}p < 0.001$

3.2 Quality of Induced Spending Changes

Given the high level of household responsiveness to major health events, a natural question is whether new information significantly alters the overall quality of health utilization. While new diagnoses could better align individual care utilization with quality—for example, by pushing individuals to substitute consumption towards high-value preventive services—individuals with updated beliefs about their risks may choose to simply increase consumption across a continuum of services, including some with little or no real risk-mitigating benefit. I examine the role of new diagnoses in household use of low-value care (Section 2.2).[22]

[22]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

I find that new chronic diagnoses are associated with an increase in low-value spending of about 5 percent (Appendix B). However, these results mask significant heterogeneity across different types of services; separating these effects provides intuition towards what type of information households are responding to. Different categories of low-value care may provide different perceived value to an affected household; for example, if a diagnosis communicates risk information, households may increase their take-up of even low return screenings (e.g., imaging services, preoperative screenings). On the other hand, households responding to price changes may be more likely to seek out high-cost, low-return services (e.g., elective surgeries). I explore this heterogeneity in depth by separating low-value services into five categories: low-value pediatric care, as well as non-pediatric low-value care for prescription medications, imaging services, extraneous screening procedures, and surgical operations.

Table 3. 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)
R^2	0.349	0.309	0.293	0.326	0.379

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 3 presents results estimating the effect of a new chronic diagnosis in each of the five categories using a standard difference-in-differences framework.[23] New chronic diagnoses increase household utilization of low-value screenings, imaging services, and pediatric care;

household are associated with marked differences in observed spending behavior, it is still unclear whether these choices are *ex-post* more optimal for affected households. This motivates a more structural approach to quantify the welfare effects of health information (Section 4).

[23]Event study regressions are included in Appendix B.

the effect sizes range from as large as ten percent for screenings to three percent for imaging services. I find no effect on the misuse of prescription drugs among adults.[24]

Taken with the previous results, these findings suggest that households affected by new chronic conditions increase their utilization of a broad set of preventive and "psuedo-preventive" services, without regard to the average return on those services. 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.[25]

3.3 Alternative Explanations for Spending Changes

My results suggest that individuals are highly responsive to new risk information. In particular, the fact that individuals seek out disease-specific preventive care provides strong evidence that individuals are responding to the informational component of a household health shock. However, individuals may be separately responding to other characteristics of new diagnoses, which I explore in this section.

3.3.1 Moral Hazard & Salience

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

[24]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.

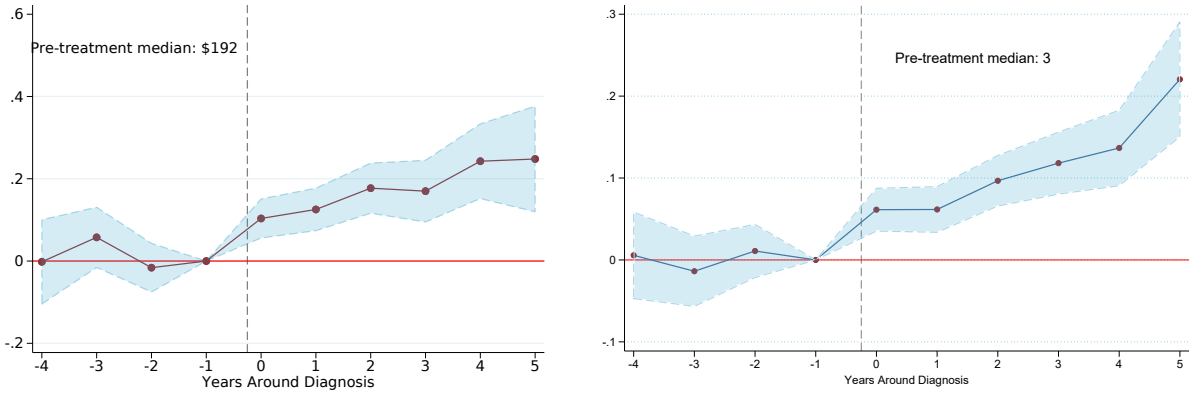
[25]Although I observe households utilizing more of these services—such as preoperative screenings or imaging services—it is unclear whether these are decisions made at the household level or by a physician who knows the family history and hence deems these services as appropriate (Finkelstein et al., 2021).

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 2 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.[26]

Figure 2. Effect of Chronic Diagnoses on Utilization: Households Facing Zero Deductible

(a) $\sinh^{-1}$(Total OOP Spending) (b) # of Preventive Visits



Notes: Figures show regression coefficients (and 95% confidence intervals) from "stacked" TWFE regressions estimating the effect of a new chronic diagnosis on utilization of other (non-diagnosed) household members. The sample is restricted to households enrolled in ESI plans with zero deductible at the time of the event. In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on the number of household preventive services per year using Poisson regression. Coefficients are presented relative to the year prior to diagnosis. Spending is measured in 2020 USD. Standard errors are clustered at the household level.

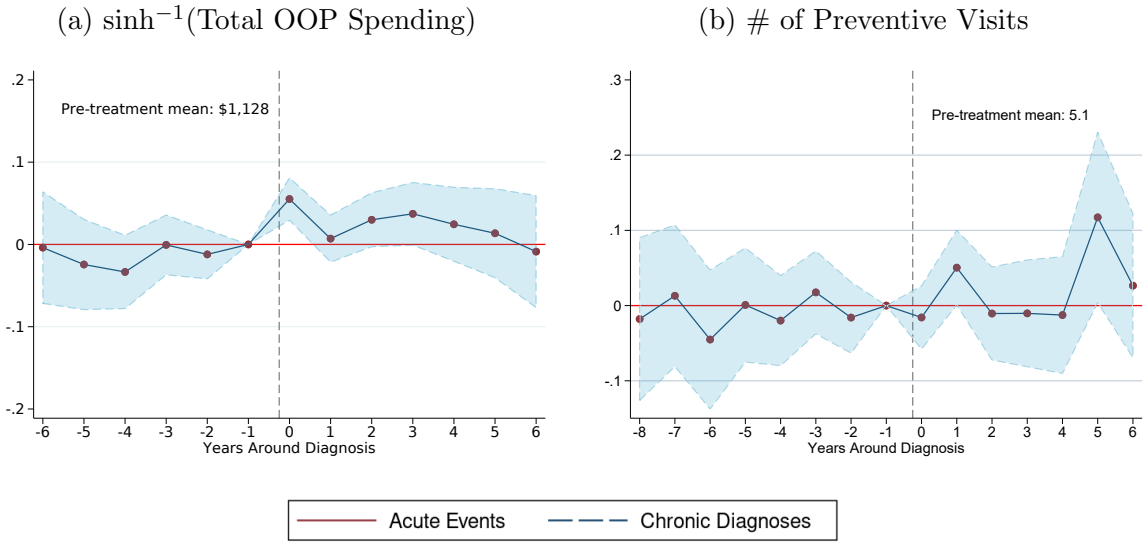
[26]The corresponding figure for plans with nonzero deductibles is included in Appendix B. Additionally, I show that households are not more likely to increase their spending even as they approach a deductible.

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 compare how households respond to new chronic conditions and a complementary set of acute hospitalizations, as defined in Section 2.2. The intuition here is that acute health events—such as hospitalizations for viral infections—provide transient shocks that ought not to meaningfully change individuals' beliefs about risk, but which may alter their relative valuation of health services (e.g., seeking vaccinations at an increased rate in order to reduce the likelihood of future hospitalizations). I can assess the extent to which new health risk information alters behavior beyond salience to the extent that household responses to chronic conditions are larger and more sustained.

Figure 3 presents the results of this comparison. 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 \$1,128) 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. 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.

Figure 3. Effect of Acute Health Events on Other Household Members' Healthcare Utilization



Notes: Figures show regression coefficients from "stacked" TWFE regressions, with 95% confidence intervals. Regressions estimate the effect of an acute hospitalization on medical utilization of other (non-diagnosed) household members (see Section 2.2). In panel (a), the dependent variable is the inverse hyperbolic sine of total OOP spending; panel (b) estimates the effect on the number of household preventive services per year using Poisson regression. Coefficients are presented relative to the year prior to diagnosis. Standard errors are clustered at the household level.

3.3.2 Health Information

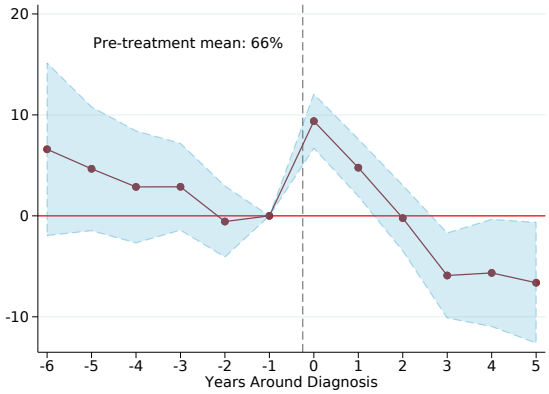
While households appear less responsive to aspects of a diagnosis other than new information, major health events may provide families with institutional knowledge, more than simply risk information. This may include information about the value of medical care, the process of obtaining insurer-covered care, or how to establish provider relationships. In general, it may be difficult to disentangle the effects of this institutional knowledge from new risk information, as the two generally co-move following health events. To separate these, I focus on a particular case where diagnoses provide risk information *without* simultaneously providing key institutional knowledge: the case of adherence to existing prescriptions.

I estimate the effect of a new chronic diagnosis on other household members' adherence to preventative cardiovascular medications (e.g., statins and cholesterol-lowering drugs). These drugs are an extremely common class of medications and are known to be effective in preventing future health problems (O'Connor, 2006). Importantly, those with these prescriptions at

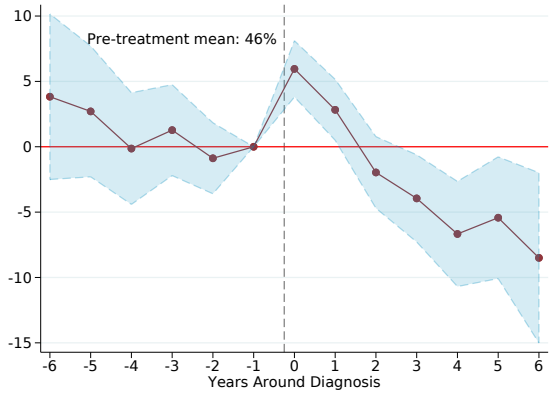
the time of a new diagnosis in their home already have the sufficient institutional knowledge to receive this care; hence, while new diagnoses provide these individuals with new risk information (e.g., the value of adherence), these events are unlikely to provide new knowledge about how to obtain medication. In this exercise, I therefore 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.[27]

Figure 4. Effect of Chronic Diagnoses On Adherence to Existing Preventive Medications

(a) Pr(Refill Old Prescription)



(b) Proportion of Days Covered



Notes: Figures show regression coefficients from "stacked" TWFE regressions, with 95% confidence intervals. Regressions estimate the effect of a new chronic diagnosis on adherence to preventive medications whose prescriptions were first written prior to the major health event. In panel (a), the dependent variable is a binary indicator for whether the prescription was refilled during the year; panel (b) estimates the effect on the proportion of days in a year covered by the medication (see Section 2.2). Standard errors are clustered at the household level.

Figure 4 presents the effects 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.[28] However, diagnoses in the household spur a resurgence in both the likelihood that individuals will fill their prescriptions and the fraction of the year a patient is covered by the medication. Affected individuals are around ten percentage

[27] See Section 2.2 for a description of the measure of prescription adherence.

[28] 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.

points more likely to refill their prescription in the year of a major medical event than in the year before; this translates to an additional eight percentage point increase in overall adherence. The fact that new diagnoses change individual adherence even among a group which has access to and knowledge of specific preventive care illustrates that individuals are learning about more than just how to obtain care. This estimated causal "re-adherence" to prescriptions is consistent with individuals reevaluating the value of their medication given new information about their health risks.[29]

4 Empirical Model of Belief Formation

Based on the empirical results above, I estimate a simple model of belief formation for households learning about health risks. In the model, individual health shocks propagate health information across a household; I then estimate how observed changes in health spending and preventive care takeup can be rationalized by changes to individual beliefs about health risks. The equilibrium model parameters allow me to both identify implied household beliefs based on observed health utilization choices and measure changes in welfare associated with potentially under-informed beliefs.

My model takes into account how major health events in a home may alter both individual investment in health production—e.g., through the use of preventive health services (Grossman, 1972)—as well as other drivers of health utilization, such as the spot price of care (for moral hazard). I write a model that captures both the dynamic returns to investment in health and the tradeoff between risk protection and moral hazard (Cardon and Hendel, 2001).[30] The key parameters of interest—namely, the ways in which major health events

[29]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.

[30]The model presented here is a simplified version that does not take into account household insurance plan choice or changes in risk aversion arising from major health events. The empirical evidence presented in Section 3 suggests that these are second-order concerns; however, I present the details of an enriched two-stage model of consumer choice that takes into account these features in Appendix C. Although static, the results of that model are qualitatively similar to those presented here.

alter beliefs about health risks—are identified based on observed spending and preventive care utilization.

I consider a set of individuals i divided into households, and discretize time t at the annual level.[31] In each period, individuals receive information about their health needs along two dimensions: a transient health shock λ_{it}^{TR} drawn from a stationary distribution $F_{\lambda^{\text{TR}}}(\cdot)$, and the possibility of a chronic health shock λ_{it}^{CH}, which occurs with probability p_{it}. Transient health shocks model the uncertain aspect of demand for healthcare, with individuals with higher λ_{it}^{TR} idiosyncratically demanding greater healthcare consumption in a period. On the other hand, chronic health shocks capture anticipated disruptions in health spending from any chronic diagnoses, whether the expected cost of a new diagnosis or the expected costs of follow-up care for already-diagnosed individuals (when $p_{it} = 1$). These shocks are drawn from separate distributions of empirical diagnostic costs for each HCC in my sample.

In each period, individuals make two choices. First, they choose medical spending m_{it} based on their realized health shocks and individual beliefs; this choice is static in the sense that total health spending does not affect future health states.[32] Second, individuals choose a level of investment in preventive care by selecting a number of preventive screenings s_{it} to consume in a year.[33] This choice is dynamic, as investment in preventive care provides individuals with information about their health risks p_{it}, thereby reducing their uncertainty about expected health shocks in future years. Finally, individuals respond to health shocks occurring in their household, which exogenously affect their beliefs p_{it}.

[31] Hence, as in previous work, I am abstracting away from the timing of claims in a given year (Einav et al., 2013; Marone and Sabety, 2022)

[32] In keeping with previous versions of this model, m_{it} is measured in dollars to capture the value of risk protection and the relative monetary costs associated with moral hazard (Einav et al., 2013; Marone and Sabety, 2022).

[33] I model preventive investments as a discrete choice of number of visits for three reasons: first, this is consistent with the measurements shown in the empirical work in Section 3. Second, this more appropriately models actual individual decision-making, as total billed spending for preventive care is generally unobservable to the individual (this is particularly true for preventive care where OOP spending is near zero). Finally, discretizing the choice set for preventive care aids in making the state space more tractable in estimating the model.

4.1 Medical Spending Decisions in Each Period

I parameterize household utility in each period using a quadratic loss function in the distance between the underlying health shock λ_{it} and chosen spending m_{it}. In addition to being consistent with prior work, this parameterization allows me to directly accommodate any moral hazard effects from major health events in a household, which may affect the marginal costs of an individual's choice of care. Individuals therefore choose m_{it} to maximize each period's utility:

$$u_{it} = (m_{it} - \lambda_{it}) - \frac{1}{2\omega_i}(m_{it} - \lambda_{it})^2 - c_{it}(m_{it}, s_{it}), \tag{3}$$

where $c_{it}(\cdot)$ captures the OOP costs from consuming m_{it} and preventive visits s_{it}. Hence, individuals choose medical spending to approximately match their transient health realization λ_{it}^{TR}, accommodating the associated OOP costs of that spending. Individual heterogeneity in moral hazard—indicating differing levels of demand elasticity for health services—is modeled using the parameter ω_i. As the main goal of this model is to identify the effects of major health events on health risk beliefs while adjusting for individual moral hazard effects, I parameterize these values based on observable demographics using the results of Einav et al. (2013), rather than estimating them directly.[34]

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:

$$1 - \frac{1}{\omega_i}(m_{it} - \lambda_{it}) - c_{ijt} = 0 \Rightarrow m_{it}^* = \max\left[0, \lambda_{it} + \omega_i(1 - c_{ijt})\right]. \tag{4}$$

[34] See Section 4.4 for details on calibration.

That is, medical expenses in each period are chosen so that the marginal utility of those services is equal to the marginal (known) OOP cost. In particular, equation 4 highlights the ways that chronic care costs affect spending decisions through prices—greater OOP spending associated with chronic care costs may drive up non-chronic health utilization to the extent that it changes marginal costs.

4.2 Dynamic Investments in Preventive Care

In addition to choices of health spending, individuals choose a level of investment in preventive services, s_{it}. Preventive visits provide individuals with information on p_{it}, reducing uncertainty associated with underlying health shocks; indirectly, this improves individual expected utility in future periods.[35]

I model household learning about health risks as a Bayesian learning process (in the absence of health events, which are discussed in Section 4.3 below). Households begin with a set of prior beliefs about their underlying health risks centered at $\mu_{p0,i}$, but with uncertainty given by a variance $\sigma^2_{p0,i}$. I model both prior beliefs and future health signals as being normally distributed in log-odds space in order to obtain closed-form solutions for belief updating while maintaining that probabilities are well-defined (e.g., contained in the unit interval). Each preventive visit provides individuals with a noisy signal of their underlying health risks:

$$\log\left(\frac{s_{it}}{1-s_{it}}\right) \sim \mathcal{N}(\overline{p}_i, \sigma^2_s). \tag{5}$$

That is, each preventive visit's signal is distributed around some observable proxy for individual true risk $\overline{p}_i$ with a noise parameter σ^2_s.[36] In estimation, I model σ^2_s as being inversely

[35]Note that in addition to providing individuals with health risk information, preventive care may directly provide health returns, shifting the distribution of F_λ. I do not include these shifts for two reasons: first, they would unnecessarily complicate a model focused on identifying belief formation; second, these positive health returns are unlikely to be realized in the medium run (e.g., within 3 years of consumption).

[36]Hence the evolution of the mean and variance parameters can be written as: $\sigma^2_{pi,t+1} = \frac{\tilde{\sigma}^2_{ift}\sigma^2_{pi0}}{\tilde{\sigma}^2_{ift}+s_{ift}\sigma^2_{pi0}}$ and $\mu_{pi,t+1} = \frac{\tilde{\sigma}^2_{ift}\mu_{pit}+\sigma^2_{pit}\tilde{\mu}_{ift}}{\tilde{\sigma}^2_{ift}+\sigma^2_{pit}}$, where the variable s_{ift} indicates how many health signals an individual has

proportional to an individual's average chosen annual spending on preventive care in order to accommodate the potential stacking of visits. Although I do not have data on actual underlying health risks, I proxy $\overline{p}$ using logistic regressions predicting each individual's probability of a chronic diagnosis as a function of observable demographics, past acute and chronic health events in the household, and family medical history (including pre-existing conditions).[37] This is similar to the type of medical information which might be conveyed by a medical professional in a preventive visit, given that true underlying health risks are not observed perfectly even by a medical professional.

Taken together, individuals choose $\{m_{it}, s_{it}\}_{t=1}^{T}$ in order to solve the dynamic equation

$$V_{it} = \max_{\{m_{it}, s_{it}\}_{t=1}^{T}} \left\{ (m_{it} - \lambda_{it}) - \frac{1}{2\omega_i}(m_{it} - \lambda_{it})^2 - c_{it}(m_{it}, s_{it}) + \delta \mathbb{E}[V_{i,t+1}] \right\}, \tag{6}$$

where households discount future years consumption at a rate of $\delta = 0.95$. A key insight from Equation 6 is that investments in preventive care reduce uncertainty about future medical expenses, and hence may increase the expected utility of future periods. Individuals are therefore forward-looking in their choice of preventive care investments as a central form of health production. Finally, investments in preventive care may help individuals to recenter their beliefs about health risks following discrete changes in beliefs arising from health shocks affecting a household, as discussed in the next section.

4.3 Major Health Events and Belief Formation

In addition to dynamic learning about health risks, health shocks affecting a family f also provide information which exogenously shifts individual health beliefs. I model these re-

received by the end of period t. This is in keeping with previous literature, such as Crawford and Shum (2005).

[37] Specifically, for an individual i and diagnosis d, the underlying risk is the predicted probability from the logistic regression $\mathbb{1}\{d = 1\} = \vec{\beta}(agesex_i) + \gamma_1 \text{Past Acute Event}_i + \gamma_2 \text{Past Chronic Event}_{-i} + \gamma_3 \text{Past Acute Event}_{-i} + \vec{\delta}(familyhistory_i) + \varepsilon$ for a vector of age-sex bins and dummies for pre-existing conditions in a family's medical history. Individual risk probabilities are then pooled across diagnoses with $\overline{p_i}$ set as the maximum probability of a diagnosis. Note that these results are time-varying; in particular, they are conditional on family histories at a given point in time.

sponses as a discrete shift in the mean $\mu_{p,it}$ given by

$$\Delta(\mu_{p,it}) = \pi_1(\text{Chronic Event})_{f,-i} + \pi_2(\text{Acute Event})_{f,-i} + \pi_3(\text{Acute Event})_{f,i}. \quad (7)$$

That is, both chronic and acute health shocks are allowed to propagate across a household. Modeling belief changes as a discrete shift here allows me to abstract away from requiring individuals to respond rationally to health events, including ignoring signals that provide little health risk information (e.g., if $\pi_2 \approx 0$). In addition, this approach allows me to match the results from Figure 1, which suggests a persistent shift in beliefs (and therefore, spending) following a major medical event.[38]

4.4 Parametrization and Estimation

The key parameters of interest in this model are the vector of response coefficients $\vec{\pi}$, which capture how individuals adjust their own beliefs about health risks in response to household signals. These parameters are identified based on observed changes in utilization and preventive investments in the data, as discussed below.

Individuals are characterized by their beliefs about health risks $\{\mu_{p0,i}, \sigma^2_{p0,i}\}$, the underlying distributions about their health states $\{\lambda^{\text{TR}}_{it}, \lambda^{\text{CH}}_{it}\}$, and their individual heterogeneity in price responsiveness ω_i. In addition, key equilibrium parameters include how beliefs are affected by preventive care and health events, $(\sigma^2_s, \pi_1, \pi_2, \pi_3)$.

As the central aim of this model is to provide insight into the evolution of health beliefs based on the empirical results discussed above, I calibrate individual parameters govern-

[38]Note that the extended version of the model in the abstract allows for these effects to decay over time; however, I estimate the size of the parameter governing this decay to be approximately 0. My results are also robust to alternative specifications, including a Bayesian learning framework with signal means as the equilibrium parameters of interest and an adaptive learning framework where individual beliefs are specified as an AR(1) with some dependence $\rho < 1$ on the previous period's beliefs. 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). 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).

ing underlying health shocks and price responsiveness based on previous work. I model λ_{it}^{TR} as a shifted lognormal distribution governed by a mean, variance, and shift parameter $(\mu_{\lambda,i}, \sigma^2_{\lambda,i}, \kappa_i)$. That is, each individual in each period draws λ_{it} from a distribution $F(\mu_i, \sigma_i, \kappa_i)$ such that

$$\log(\lambda_{it} - \kappa_i) \sim \mathcal{N}(\mu_i, \sigma_i^2). \tag{8}$$

This choice of distribution accommodates the skewed nature of spending data while also allowing for a nontrivial fraction of individuals to choose zero spending in a given year (matched by the shift parameter). I calibrate the three hyper-parameters $(\mu_{\lambda,i}, \sigma^2_{\lambda,i}, \kappa_i)$ to match the empirical distribution of annual spending using a matched sample of patients *not* included in the structural estimation, including both individuals enrolled in other firms in Marketscan and spending data for in-sample individuals between 2014 and 2018.[39] Similarly, I calibrate expected spending on chronic conditions λ_{it}^{CH} based on the empirical distributions of disease-specific spending at the HCC level; I modeled separate distributions for the year of diagnosis and followup years to differentiate between diagnostic and maintenance costs.[40]

While I assume that individuals have rational expectations over the distributions of λ, the expected value of a health shock depends explicitly on an individual's belief p_{it} that a chronic shock will be nonzero.[41] Hence, assuming rational expectations makes the model more tractable while also allowing me to separately identify changes in p_{it} from changes to the underlying expected costs of a chronic illness, which are captured by the empirical distribution of costs.

[39] Parameters are calibrated by binning individuals based on including age, sex, risk score quartile, and enrollee relationship, and then constructing moments based on the observed empirical spending distribution. This is done using three properties of a shifted lognormal distribution: $\overline{\lambda} = \exp(\mu + \frac{1}{2}\sigma^2) + \kappa$, $\lambda^M = \exp(\mu) + \kappa$, and $\frac{\text{sd}(\lambda)}{\overline{\lambda}} = \sqrt{\exp(\sigma^2) - 1}$, where λ^M denotes the median. The solution to this system of equations given the moments of the empirical distribution of λ identifies the three hyperparameters μ, σ, κ. In order for shocks to be meaningful, we restrict $\lambda_{\mathcal{I}t} < m_{\mathcal{I}t}$ in each period when drawn.

[40] Disease-specific spending was constructed in collaboration with Rebecca Hughes, MD. See Section 2 for a more detailed discussion.

[41] Although there is some evidence that rational expectations in healthcare utilization depends on how price-aware consumers are when selecting services (Lieber, 2017), this is less concerning when considering both preventive care investments—which generally have low levels of cost-sharing—or with chronic care costs—which are typically stable over time and hence more easily predicted by household members.

Finally, I calibrate heterogeneity in individual price responsiveness using the estimated regression coefficients predicted by Einav et al. (2013). This variation—which captures how moral hazard differs across individuals—is likely a second-order effect in the takeup of preventive care and the development of health risk beliefs, as discussed in Section 3; hence, although I model individual-level heterogeneity, I do not identify individual coefficients ω_i as equilibrium objects in my structural estimation.

These calibrations dramatically reduce the state space of the simplified model presented here, allowing for tractable estimation of the dynamic evolution of beliefs and health investments.[42]. The estimation of the remaining equilibrium parameters takes place in two steps. In the first step, based on a guess of the six initial parameters $(\mu_{p0}, \sigma^2_{p0}, \sigma^2_s, \pi_1, \pi_2, \pi_3)$, I estimate the model via generalized method of moments (Hansen, 1982). Given that the choice of the continuous variable m_{it} is static and the state space of the dynamic variable s_{it} is small, this is done directly using forward induction.[43]

In the second step, the equilibrium parameters are chosen to minimize a set of moments which match the model predictions to observed data. The central moments identifying belief evolution are directly tied to the reduced form results presented in Section 3. The model matches on the predicted levels of overall health spending and preventive care utilization, including the mean, median, and RMSPE of all predictions. In addition, I require the model to capture the observed effects of a chronic diagnosis on both overall health spending and preventive utilization (Figure 1). I do this by matching on regression coefficients for simple difference-in-differences regressions estimating the effect of a chronic diagnosis in the home on spillover health spending and spillover preventive visits.[44]

[42] See Appendix C for the complete model without restrictions on the parameter space

[43] The only restriction made on the state space transition matrix is that individual investments in preventive care cannot *drop* by more than 2 visits from period t to $t+1$. See Appendix C for more details on estimation.

[44] As in Section 3, I appropriately transform the dependent variables of total health spending and preventive care takeup.

5 Structural Results

Table 4 presents the equilibrium parameters resulting from estimation of the model in Section 4 by GMM. Standard errors are calculated as as discussed in Cocci and Plagborg-Møller (2021), and hence represent conservative, "worst-case standard errors" for calibrated structural parameters, that allow for arbitrary correlations across empirical moments.[45]

Table 4. Estimated Structural Parameters

		Dynamic Model	
		Estimate	95% Confidence Interval
Panel A: Initial Beliefs			
μ_{p_0}	Prior mean	0.0236	[0.0183, 0.0288]
$\sigma^2_{p_0}$	Prior variance	2.088	[2.0823, 2.0937]
Panel B: Learning from Preventive Care Investments			
μ_s	Signal mean	0.0228	[0.0227, 0.0229]
σ^2_s	Signal variance	0.934	[0.9115, 0.9565]
Panel C: Learning from Major Health Events			
π_1	Family Chronic Event	0.1028	[0.0867, 0.1189]
π_2	Own Acute Event	0.0261	[0.0189, 0.0333]
π_3	Family Acute Event	0.0279	[0.0243, 0.0315]

Notes: Table presents estimated equilibrium parameters of the model described in Section 4, estimated via GMM on a sample of $N = 387,216$ enrollees observed in 149,938 households between 2006 and 2013. All average parameters are expressed in terms of probabilities, while variances are expressed in log-odds. Standard errors are calculated following Cocci and Plagborg-Møller (2021).

Prior to individual learning about risk—either through household health events or preventive screenings—I estimate that individual beliefs about a future major diagnosis affecting them are centered at about 2.4%. There is considerable heterogeneity in prior beliefs: the 10th percentile of the distribution has prior beliefs of 0.25%, while the 90th percentile has

[45]These standard errors balance the tradeoff of computational feasibility—particularly in a setting where bootstrapping is difficult given that some of the moments are the result of high-dimensional fixed-effect regression estimation—against accurate coverage in conditions where correlations across moments cannot be derived analytically. The 95% confidence intervals presented in this table therefore cover the range of the structural parameters with an asymptotic coverage probability of at least 95%—the coverage probability will be exactly 95% in the "worst-case" correlation structure across parameters, when all structural parameters are perfectly correlated asymptotically.

prior beliefs of 11.7%. However, considering a relative in-sample diagnosis rate of about 2.7%, roughly 62% of the individuals in my sample under-estimate their true risk of a major health event occurring. Furthermore, individuals improve their estimation of health risks through investments in preventive care. The variance of preventive signals is roughly 45% of the variance of prior beliefs, suggesting that these visits provide informative signals improving the beliefs of individuals in the sample.

Most importantly, I estimate that household health events are major drivers in individual health beliefs, consistent with the results presented in Section 3. New chronic diagnoses affecting a household increase individual beliefs by 10.3 percentage points, a more than four-fold increase from the average prior belief. The effect of a chronic diagnosis is estimated to be much larger than the effect of acute events—I estimate that an individual's response to an acute events, regardless of which family member was affected, is just over 2.5 percentage points (Note that the parameters π_2 and π_3 are not significantly different from each other). The results from this simplified dynamic model are qualitatively similar to the results of the full model presented in Appendix C, which also incorporates changes in household risk aversion and plan choice, among other features.[46]

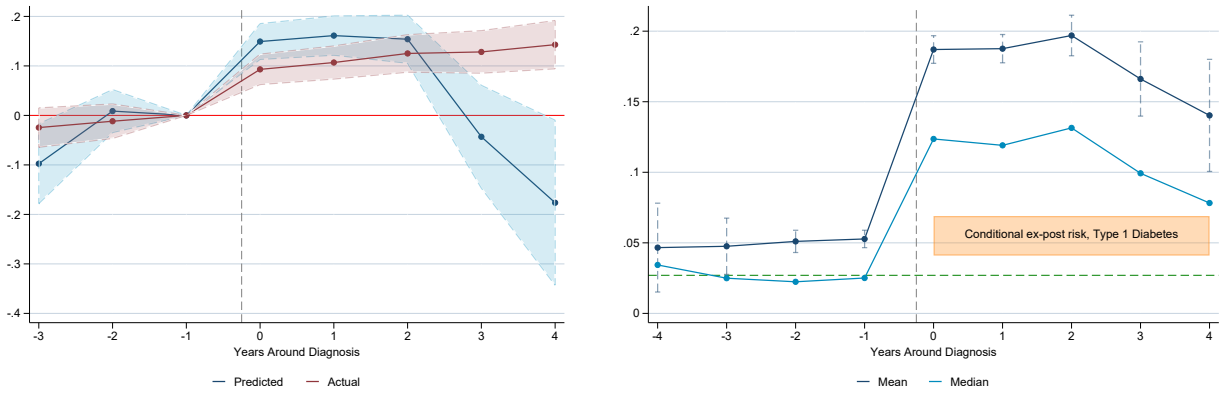
Figure 5 presents the key takeaways from the structural estimation. In the first panel, I highlight the overall match quality of the model, reporting TWFE coefficients for the effect of a new diagnosis on the inverse hyperbolic sine of spillover utilization. I show coefficients for both observed spending and spending predicted by the model. In general, the model captures the increase in spending for the first 3 years following a diagnosis.

The major implication of the equilibrium parameters is presented in the second panel of Figure 5. Here, I present changes in predicted beliefs (in a recentered time series) specifically among individuals affected by a new diagnosis in their home. Prior to diagnosis, individuals

[46]In the more complete model, household risk aversion increases after a diagnosis; however, this does not affect overall plan choice, as discussed in Section 3. I find evidence for heterogeneity in responsiveness to different types of events (e.g., costlier diagnoses), as well as in how risk beliefs co-move with latent individual health shocks. Overall, the model suggests that most of the observed variation in medical utilization can be explained by belief evolution in response to new diagnoses, rather than other individual-level heterogeneity.

Figure 5. Model Predictions: Non-Diagnosed Spending and Beliefs Around a New Diagnosis

(a) TWFE Regression Coefficients, Total Spending (b) Average and Median Health Risk Beliefs



Notes: Figures show key predictions of the structural model described in Section 4. Panel (a) reports TWFE coefficients for the causal effect of a new chronic diagnosis on both observed spillover household utilization and spending predicted by the model (similar to Figure 1; see Equation 1 for the specification). Panel (b) reports recentered time series indicating average and median individual risk beliefs for the same population (individuals affected by a new chronic diagnosis in their home), averaged over draws from individual posterior distributions. The green horizontal line in Panel (b) illustrates the average in-sample rate of diagnosis (∼2.7%); the orange range indicates the estimated *ex-post* risk of a diagnosis of type 1 diabetes following a sibling's diagnosis ([4.1%,6.9%]) (Harjutsalo et al., 2005).

tend to slightly over-estimate their risks, on average. Following the diagnosis, individual beliefs about the probability of a future health shock increase to almost 20% for the average individual (over 10% for the median). These shifts persist following the event, being only slightly mediated by future preventive visits.[47]

Importantly, these changes are much larger than an individual's own *ex-post* conditional probability of a major health event taking into account the clinical information from a household member's diagnosis. To see this, the second panel of Figure 5 includes the estimated range of a sibling's risk of developing type 1 diabetes, conditional on knowing another sibling is already diagnosed (Harjutsalo et al., 2005). This range is an instance where there is large conditional risk-sharing—hence, where a signal is particularly informative—yet even in this case, individual responses to health events well exceed the standard measures of conditional health risk. I explore the welfare implications of these facts in the following section.

[47]Note that I also estimate a version of the model that allows belief responses to major health events to decay linearly with time since the event, and find no evidence of such a decaying response.

5.1 The Welfare Effects of Health Shocks

Based on the estimated structural parameters, I can construct a measure of each individual's expected utility gain from new health risk information. These utility differences constitute a willingness to pay (WTP) measure which allows me to benchmark the welfare effects of providing health risk information, given how such information changes health risk beliefs. It also enables me to consider how welfare might change were information responses altered.

I estimate the value of receiving new health risk information through a household diagnoses by comparing individual expected utility against a state where no information is transmitted. In the counterfactual state, non-diagnosed household members continue to invest in and learn from their own preventive screenings, but do not shift their beliefs in response to acute or chronic health events.[48] Given that the Bellman equation in Section 4 is estimated in dollar terms, individual WTP for information in each period is the difference

$$WTP_{it} = V_{it}(\text{information}) - V_{it}(\text{no information}), \tag{9}$$

for the total valuation V_{it} defined in Equation 6. Given that household beliefs experience a persistent shift at the time of diagnosis and that individuals are forward-looking, I consider differences in WTP in the period where each affected individual experiences a diagnosis in their home.[49] Hence, this welfare criterion measures how much individuals would be willing to pay for shifted beliefs, based on how that change in beliefs would affect their ability to better anticipate health shocks and ultimate spending decisions.

In general, I find that any potential welfare gains from risk information are overshadowed by the interpretation of that information for a substantial fraction of the population.[50] I

[48] Throughout this section, I limit attention to individuals affected by a chronic diagnosis in their household (but not themselves diagnosed). This is to isolate the *spillover* value of health information from chronic health events, and also to ignore the mechanical change in expected utility from experiencing a diagnosis oneself.

[49] Alternatively, I could aggregate WTP at the individual level by summing WTP_{it} across all periods during and after the household diagnosis. The results are qualitatively similar.

[50] Appendix Figure C.1 similarly shows the full distribution of WTP across all affected individuals in the sample, including for counterfactual analyses.

estimate that roughly 48% of the affected individuals in my sample would be willing to pay to *avoid* new information. Among this population, losses are skewed: the average (median) welfare loss is \$100 (\$29), constituting a nearly 89.3% (7.4%) decline from baseline utility. Figure 6 below presents more details on individual WTP for new information.

At first glance, a welfare penalty arising from new information appears counter-intuitive. However, taken together with the fact that individuals may over-update their beliefs about health risks (Figure 5), these results are consistent with the findings in Section 3. To the extent that risk information leads individuals to have priors which overweight adverse health shocks, individual expected utility will be lowered through two mechanisms: first, an over-spending on health services (for example, an over-utilization of low-value health services), and a utility penalty from greater uncertainty in health spending for risk-averse individuals.[51] Utility losses through both channels could be reduced to the extent that individuals place risk information in better context; doing so would result in a mitigated shift in beliefs and, hence, spending.

5.2 The Role of Belief Updating in Welfare Penalties

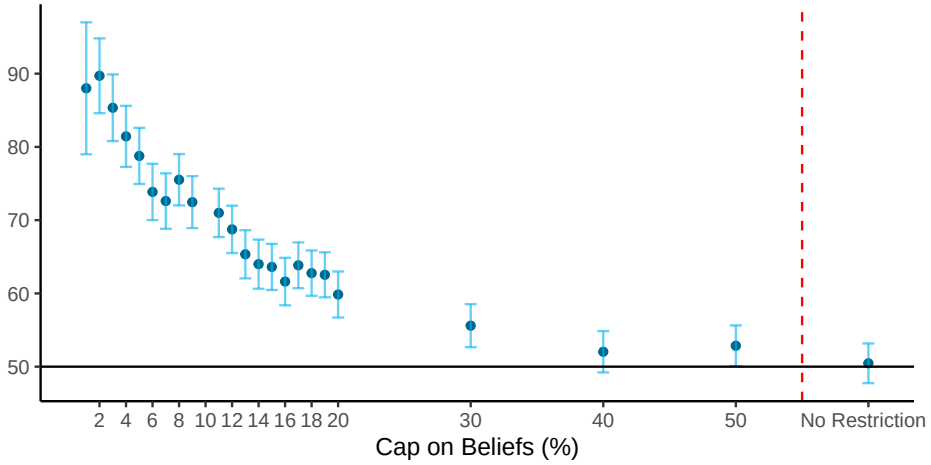
Given the strong belief responses of individuals and potential for resulting welfare penalties, I consider how limiting the magnitude of individual responsiveness to health risk information (captured in the parameter π_1) might change estimated WTP for information. I consider simple counterfactual scenarios which impose arbitrary upper bounds on posterior beliefs; that is, I impose a threshold $\overline{p}$ capping beliefs across all individuals and periods in the sample.[52] Such an exercise intuitively captures the relative value of health risk information in contexts where individual responsiveness is more attuned to true conditional risks.

Figure 6 presents the results, summarizing how individual WTP for information changes as I vary $\overline{p}$. The figures show the results of multiple counterfactual simulations comparing the

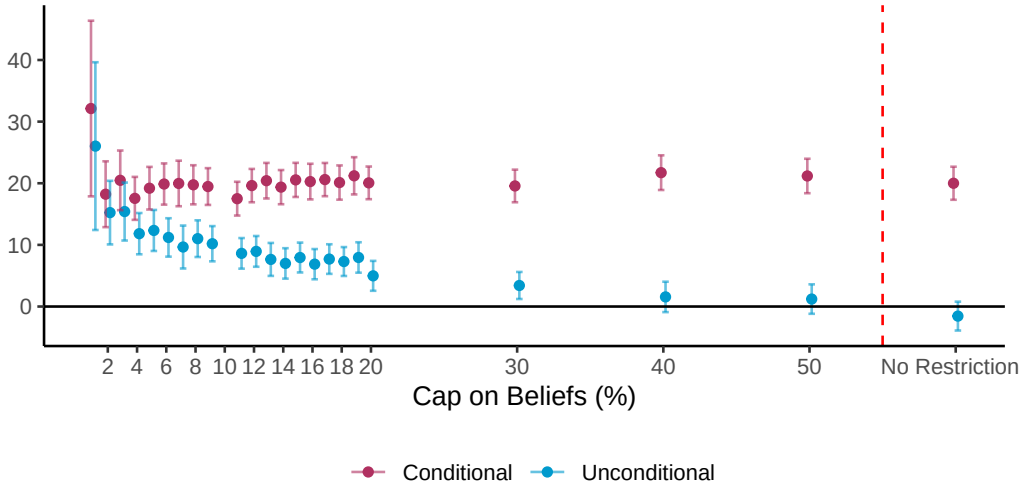
[51] The second mechanism is similar to an anticipatory "disutility" discussed by Oster et al. (2013). Additional results presented in Appendix C highlight that the returns to health information vary with key household characteristics, including household risk levels. Households with high expected health risks prior to a new diagnosis experience lower welfare penalties—this makes intuitive sense, given that these households

Figure 6. Changes in Individual Utility from Bounding Responsiveness to New Diagnoses

(a) % Willing to Pay for Information



(b) Average Percentage Change in Utility



Notes: Figures show results of counterfactual simulations estimating the expected utility gains from new health risk information communicated through household diagnoses. Each point presents estimates from separate simulations, which compare expected utility across states where (i) individuals update risk beliefs according to new information, albeit with some cap $\overline{p}$, and (ii) individuals do not learn from household diagnoses. The value of $\overline{p}$ is shown on the x-axis, with the far right simulation the difference between equilibrium responses and no responsiveness. Panel (a) presents the percentage of individuals whose expected gain from updating is positive (those who would be willing to pay for information); panel (b) presents the unconditional (blue) and conditional (maroon) average percent changes in utility. Models are estimated using the equilibrium parameters presented in Section 4 and Table 4, varying only the belief responses $\vec{\pi}$. Utility differences are measured according to Equation 6 in 2020 USD, and are calculated at the year of diagnosis. Error bars represent 95% confidence intervals.

expected utility gains of individual learning from household diagnoses (albeit with some cap $\overline{p}$) against the state of the world where individuals do not learn from household diagnoses. The value of $\overline{p}$ is shown on the x-axis; on the far right, I include the baseline comparison of utility gains from the observed equilibrium responses with no restrictions on updating.

The first panel highlights that the percentage of individuals who would be willing to pay for health risk information (that is, those whose expected gain from updating is positive) is sharply increasing as individual responsiveness is more restricted. Without meaningful bounds on probability responses, only about half of individuals would be willing to pay for health risk information. However, even restricting household beliefs about the probability of future health events to be 10% or fewer—a relatively generous bound, given true conditional risks generally below 7%—provide a significant utility boost to 10% of individuals. The returns to information continue to increase as bounds become more restrictive; ultimately, roughly 90% of households are willing to pay for risk information that caps their *ex-post* beliefs about risk at 2% or fewer. However, as bounds become too restrictive (in particular, as bounds allow for less updating than a true in-sample diagnostic risk, roughly 2.7%), the fraction of individuals benefiting from information begin to diminish. This makes intuitive sense, given that the value of information lies in its ability to persuade individuals to update their beliefs to a "correct" posterior; it is only when updating is too extreme that welfare penalties begin to appear.

The second panel highlights the relative magnitude of utility gains from imposing restrictions on belief updating. Unconditional utility gains, shown in blue, increase with restrictiveness, in line with the results of the first panel. As in the first panel, gains are maximized at belief caps of 3% or fewer, providing a roughly 15% increase in expected utility on average. I find little evidence of heterogeneity in the value of risk information across those who benefit from it—the conditional gains in utility (shown in maroon) are constant at roughly a

have an already high level of expected spending, meaning that new information changes expected outcomes less.

[52]Results are qualitatively similar when allowing this threshold to vary, for example as a multiple of individual-specific predicted risks. See Appendix C.

20% gain across individuals.[53] This implies that the extent of updating—rather than other dimensions of health information—are driving welfare gains.

Taken together, the results of this simple structural exercise corroborate the main mechanisms and findings presented in Section 3. Individuals appear highly responsive to health shocks affecting their household; their marked increase in both total utilization and preventive care take-up can be explained by large swings in latent beliefs about their health risks. However, this large shift in risk beliefs ultimately incurs welfare penalties which outweigh utility gains from information; the model suggests that individuals would be willing to pay either to moderate their responses to information or else to avoid the information entirely.

6 Conclusion

This paper assesses the extent to which information about one's health risks alters individual decision-making in health care. I demonstrate that an individual diagnosis propagates across household members, who respond to new health risk information by altering their use of both high- and low-return services. These changes in behavior are best explained by individuals reassessing health risks, rather than responding to financial incentives or salience effects. However, 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 an individual's expected utility gains from new health information. I find that for a large subset of the population, gains from information are swamped by extremely large shifts in estimated *ex-post* risks. Bounding the extent to which individuals increase their risk beliefs post-diagnosis makes information welfare-improving for over 90% of individuals.

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

[53]Note that the welfare gains from extremely restrictive caps on beliefs (at 1% or fewer) are estimated to be much more imprecise, presumably because such a restrictive cap induces many of an individual's drawn beliefs to be either zero or extremely small.

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.

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.

Aguirregabiria, V. and Jeon, J. (2020). Firms' beliefs and learning: Models, identification, and empirical evidence. *Review of Industrial Organization*, 56(2):203–235.

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. (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*.

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.

Cengiz, D., Dube, A., Lindner, A., and Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, 134(3):1405–1454.

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

Cocci, M. D. and Plagborg-Møller, M. (2021). Standard Errors for Calibrated Parameters.

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.

Crawford, G. S. and Shum, M. (2005). Uncertainty and learning in pharmaceutical demand. *Econometrica : journal of the Econometric Society*, 73(4):1137–1173.

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.

Dasaratha, K., Golub, B., and Hak, N. (2022). Learning from neighbors about a changing state. *The Review of Economic Studies*, page rdac077.

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.

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

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

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

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.

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., Persson, P., Polyakova, M., and Shapiro, J. (2021). A taste of their own medicine: Guideline adherence and access to expertise. *Working Paper*, page 38.

Finkelstein, A., Persson, P., Polyakova, M., and Shapiro, J. M. (2022). A Taste of Their Own Medicine: Guideline Adherence and Access to Expertise. *American Economic Review: Insights*, 4(4):507–526.

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., 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.

Grossman, M. (1972). On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255.

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 : the journal of the American Medical Association*, 297(10):1073–1084.

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.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the econometric society*, pages 1029–1054.

Harjutsalo, V., Podar, T., and Tuomilehto, J. (2005). Cumulative Incidence of Type 1 Diabetes in 10,168 Siblings of F innish Young-Onset Type 1 Diabetic Patients. *Diabetes*, 54(2):563–569.

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

Hauser, D. and Bohren, J. A. (2021). Learning with heterogeneous misspecified models: Characterization and robustness. *Econometrica : journal of the Econometric Society*, In Press.

Hoagland, A. and Shafer, P. (2021). Out-of-pocket costs for preventive care persist almost a decade after the Affordable Care Act. *Preventive Medicine*, 150:106690.

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.

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.

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.

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. (2022). Should there be vertical choice in health insurance markets? *American Economic Review*, 112(1):304–342.

O'Connor, P. J. (2006). Improving medication adherence: Challenges for physicians, payers, and policy makers. *Archives of Internal Medicine*, 166(17):1802–1804.

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.

Paserman, M. D. (2008). Job search and hyperbolic discounting: Structural estimation and policy evaluation. *The Economic Journal*, 118(531):1418–1452.

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.

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 (Washington, D.C. : 1974)*.

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

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, second edition.

USPTF (2022). Guide to Clinical Preventive Services. Technical report, US Preventive Services Task Force.

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.