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

Alex Hoagland†

November 19, 2024

Abstract

I examine how individuals learn about health risks from household health shocks using US administrative data. When a family member is diagnosed with a chronic condition, relatives increase healthcare spending by 10%, a response that would require price declines as large as 50% to justify on demand alone. I quantify the mechanisms behind these effects, showing they are most consistent with individuals updating their beliefs about health risks. I evaluate the welfare and efficiency implications of this learning using a structural approach. I find that the majority of individuals overreact to diagnoses, over-weighting their ex-post risks and offsetting potential welfare gains from informed decision-making.

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

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

*I am grateful to Kristen Blair, Corinne Andriola, Jason Abaluck, Michael Darden, Randall Ellis, Keith Ericson, Itzik Fadlon, James Feigenbaum, Raymond Fisman, Sarah Gordon, Tal Gross, Jihye Jeon, Timothy Layton, Marc Rysman, and Paul Shafer for very helpful comments on this project. 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 ASSA 2023 Annual Meeting, 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 healthcare 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 pathways to receiving care. Family health experiences provide particularly influential sources of health information due to their proximity and relevance. Understanding how individual health experiences shape family health behaviors is essential for policies aiming to improve public health, including addressing high levels of healthcare spending or incentivizing the takeup of high-value health services.

One especially salient feature of a health shock is the information it communicates about risks for future shocks, particularly when risks are correlated within a household. However, while health events may lead individuals to update their beliefs about health risks, shocks may also drive changes in the expected prices of medical care (Anderson et al., 2024), household preferences for health consumption (Finkelstein et al., 2009), or knowledge about the availability of health services (Dwyer and Liu, 2013). Hence, decomposing and quantifying the mechanisms underlying these spillover effects is critical to understanding the efficiency and welfare implications of household responses to shocks.

More generally, identifying how beliefs about health risk evolve sheds light on the value of health information amid potential costs from imperfect information transmission. Theoretically, informative health events such as health shocks or preventive screening outcomes should improve individual risk beliefs, allowing for earlier detection of health conditions and improved health in the long run. Conceptually, however, there is a tension between the seriousness of a health event and appropriate belief updating: households responding to health shocks may either be slow to update their beliefs—incurring potentially large health and financial costs in the meantime—or may be overly responsive to uninformative events, placing large ex-post weights on low-probability health events even after conditioning on household genetic risk. These over-reactions may lead to over-consumption of low-value care for the individual, potentially at the cost of crowding out other, more valuable health services. Assessing the welfare effects of new health information requires understanding whether, how quickly, and how precisely individuals respond to it.

In this paper, I examine these questions using new administrative data in the US, a novel setting in which to study household responses to health shocks. I use claims data for US households insured through their employers to examine how a new diagnosis of a chronic condition (for example, a cancer diagnosis for a parent or a child's diagnosis with type 1 diabetes) affects household health behaviors. I first provide clear reduced-form

evidence that these health shocks generate spillover effects, and that the effects appear to be driven by responses to health *information* rather than other possible explanations. The observed spending increases are large, on the order of 10% for households; to put this into context, using the RAND Health Insurance Experiment's estimate of the price elasticity of demand for medical care at -0.2, a 50% decline in prices would be required to rationalize these spending increases using price changes alone (Newhouse, 1993). These spillover effects include significant and persistent increases in both overall utilization and investments in disease-specific preventive care, novel evidence of responsiveness to new risk information.

This work leverages rich administrative data in the United States, allowing me to study household health spillovers in a novel environment. The unique healthcare environment of the US is a particularly fruitful one for studying spillover effects and learning about health risk, as households may also be subject to financial considerations and liquidity constraints (Gross et al., 2022). I test for these financial incentives, and show that spillover effects are unaffected by changes in the marginal price of future care, suggesting ex-post moral hazard responses do not drive the results.[1] The chronic diagnoses I study also induce larger changes in spending than similarly intensive, but uninformative, acute health events, which have been the focus of previous papers. Importantly, this suggests that households respond to genetic risk information in addition to any salience effects. Finally, I show these effects are not driven by learning about the healthcare system generally. I do this by evaluating adherence to existing preventive prescriptions as a way of isolating spillover effects independent of changes to institutional knowledge of US healthcare.

In general, one would expect that if households are responding to new risk information, this information should improve individual decision-making and, subsequently, welfare. Surprisingly, the welfare effects of chronic health shocks are ambiguous using only reduced-form evidence. I observe that affected household members increase take-up of "low-value" health care, including services that do not typically benefit the marginal patient (Colla et al., 2015). These responses are most likely to include increased utilization of low-value services that appear, from a patient's perspective, closely related to preventive care, including extraneous pre-operative screenings or imaging services. Importantly, this suggests individuals may have trouble interpreting risk information signals from a chronic diagnosis or may not appropriately choose high-quality services conditional on their specific medical histories.

[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 (Einav et al., 2013).

These findings motivate a structural approach to model how risk beliefs evolve following a health shock and the welfare implications of these effects. I present a dynamic model where forward-looking individuals learn about health risks through preventive screenings and household diagnoses. The model combines dynamic investments in health production with consumption decisions trading off risk protection and consumption; this provides identifying variation pinning down individual learning from new diagnoses separately from dynamic effects, moral hazard, or salience. The model therefore provides a simple framework to estimate the elasticity of health risk beliefs to household health events. I estimate the model using observed demand data and compare its predictions to counterfactual scenarios with more targeted risk signals. This allows me to isolate households' willingness to pay for new health risk information and present novel welfare calculations.

Counter to expected thought, information about future health risks 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 40% of baseline expected utility. This represents a significant utility loss; the average household would need to reduce healthcare consumption by between 7% and 10% to experience the same reduction in utility. The model's central finding is that new diagnoses in a household spur large changes in individual assessments of health risk, resulting in average posterior beliefs that are well above true diagnostic risk, even conditional on household genetic risk. Bounding this updating can be welfare improving for over 60% of households previously unwilling to pay for health risk information.

My analysis contributes to a well-established literature on the spillover effects of health shocks within a household. Family relationships provide important information for economic decisions, and economic shocks in a family affect the health of its members (Fontes et al., 2024; Altmejd et al., 2021); acute family health shocks have similarly been shown to induce spillover demand changes (Fadlon and Nielsen, 2019; Hodor, 2021; Arteaga et al., 2024; Fadlon et al., 2024).[2] I contribute to this literature in three ways: First, I provide novel evidence for household spillovers in the US, where responses to health shocks are

[2]A rich literature has also highlighted how individuals respond to health risk information contained in their *own* diagnosis; see Alalouf et al. (2019). Some previous work has demonstrated that certain diagnoses can have dramatic impacts (Almond et al., 2010), while others have no noticeable responses (Dupas, 2011; Kim et al., 2019). In addition, previous work has suggested that acute health events within an individual's broader social network may also shape behavior (Bouckaert et al., 2020; Song, 2021), including economic decisions (Yi et al., 2015). There is also a rich literature identifying how health shocks affect family members' employment and labor supply outcomes (Maestas et al., 2024; Fadlon and Nielsen, 2021; Frimmel et al., 2020; Arrieta and Li, 2023).

potentially very different than in other healthcare settings. Thus far, previous work has been limited to administrative data in publicly-funded health insurance systems or historical data, despite the relevance of studying household health behaviors in the complex US healthcare system. Second, I provide clear reduced-form evidence disentangling learning about health risks from other relevant drivers of the observed effects. Finally, motivated by these results, I present a general framework to estimate the effect of a health shock on individual beliefs about health risk and the relative value of seeking care.

I also contribute to the literature on non-Bayesian learning in models of health behavior, combining two distinct threads of the learning literature (Barseghyan et al., 2018; Bundorf et al., 2021). First, I emphasize the role of disproportionate weight individuals place on high-cost, low-probability events, which rationalize individual choices that would otherwise require unreasonably high levels of risk aversion to justify (Goldstein et al., 2023; Ortoleva, 2012; Spinnewijn, 2015). I combine these results with the quickly-evolving literature studying the role of peer signals in learning, highlighting that individuals may over-emphasize high-risk peer signals relative to their own (Dasaratha et al., 2022; Hauser and Bohren, 2021). I incorporate these disparate findings into a novel structural approach modeling the evolution of beliefs about health risk. This work is related to other models of learning in healthcare settings (Darden, 2017; Chan and Hamilton, 2006; Crawford and Shum, 2005), but distinct in that by identifying the evolution of individuals' own health beliefs, I am able to comment on the efficiency of responses to household health shocks.

Finally, my work is relevant to the well-established literature exploring many suboptimal health decisions made by consumers (Abaluck and Gruber, 2016; Baicker et al., 2015; Handel and Kolstad, 2015).[3] This includes discussions about whether improving health information generally would improve decision-making (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. Rather, individual responses to some information may not improve the match quality between patients and services, but simply shift consumers from one type of poor decision-making to another, all while increasing health spending.[4]

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. I then present the details of my model in Section 4 and its

[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.

results in Section 6. Finally, I discuss the relevance of my findings and directions for future work in Section 7.

2 Empirical Setting & Data

Data on household health shocks and utilization come from the Merative (formerly IBM Truven) Marketscan *Commercial Claims and Encounters* Data from 2006 to 2018. These data contain detailed inpatient, outpatient, and pharmaceutical claims for households enrolled in an employer-sponsored insurance (ESI) plan provided through participating insurance carriers to several large U.S. firms.

Households are defined as all enrollees covered under a single insurance contract, including the primary employee and any covered dependents.[5] I limit the sample to only households with two or more members observed for two or more years and insured with one of eight large firms for whom plan benefit information is readily available. Households with any gaps in their enrollment or eligibility were dropped from the sample (however, households who change insurance plans within their observed window are included). My final sample consists of 353,403 households and 5,439,482 individual-year observations.[6]

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.[7]

Table 1 also shows how households affected by these chronic conditions (defined below) differ from the full population. Column 2 limits the sample to only household-years in

[5] Households may include dependent children living away from home, and may exclude family members such as spouses covered under their own ESI contract.

[6] Households did not need to be observed for all of 2006–2018 to be included in the sample; the average household is observed for 7 years. The main results are based on an unbalanced panel, but are robust to fully balancing the panel across 6 years. Finally, I drop households where the diagnosed individual is not observed for at least a full year after diagnosis, excluding less than 1% of households. This avoids biasing my results to capture only household reactions to extremely rare fatal diagnoses (note that these diagnoses are rarely fatal in the short run).

[7] While insurance contracts are defined by a complicated set of cost-sharing measure across provider specializations, networks, and procedures, Section 4 uses only a family deductible, a simplified non-specialist coinsurance rate, and a family OOP maximum in the model, as in prior work (Marone and Sabety, 2022). These measures are constructed using empirical distributions in the claims data, and described in detail in Appendix Section A.1 (Zhang et al., 2018). These simplified measures capture a wide degree of variation in the data and harmonize well with 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: Enrollees are employees plus their covered dependents. Spending values are reported in 2020 USD. Standard errors are reported in parentheses and sample medians are in brackets. Panel B includes unconditional spending averages for the entire household (summing across all household members) while Panel C includes conditional spending averages for affected individuals only. Column 2 limits the sample to only household-years in which a chronic diagnosis occurred.

which a chronic diagnosis occurred. Affected households are riskier, on average—note that this is mechanical, as chronic conditions directly affect risk scores. 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.

2.1 Major Health Events

Major health events, which communicate information about health risk to households, are identified by diagnostic codes in the data, following the Department of Health and Human Services' Hierarchical Condition Categories (HCC) diagnostic codes. HCCs are

typically used in risk adjustment models and identify a basic set of chronic illnesses that alter overall health utilization and spending. I limit my classification of health events to non-pregnancy HCCs that occur with high frequency.[8]

Health Shock Category	Conditions/Diagnostic Groups
Cancers	Breast cancer, prostate cancer, thyroid cancer
Cardiovascular Conditions	Congestive heart failure, heart arrhythmias
Autoimmune Conditions	Type 1 diabetes, lupus, multiple sclerosis, rheumatoid arthritis
Mental Health Conditions	Major depressive disorder, bipolar disorder, personality disorder
Others	Asthma, inflammatory bowel disease, seizures, type 2 diabetes

Table 2. Sample Chronic Condition Health Shocks

Notes: Table summarizes the most common chronic conditions used as health shocks throughout the paper. For a complete list and relevant diagnostic and procedure codes, see Appendix A.

Table 2 illustrates the most common major health events used as focal health shocks in my analysis. These include common conditions affecting family members across a number of disease categories, including cancers, chronic cardiovascular conditions, autoimmune conditions (e.g., type 1 diabetes), and mental health conditions. I also include additional chronic health shocks commonly observed in the sample, such as new diagnoses with type 2 diabetes.

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. I identified a set of disease-specific procedures and prescriptions associated with each health condition in my sample.[9] 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

[8]See Appendix Section A.2 for details. To ensure that I identify new diagnoses, I require that relevant diagnosis codes appear during or after an individual's second observed year.

[9]These procedures and prescriptions were identified in collaboration with Rebecca Hughes, MD. Appendix Section A.3 lists the relevant codes used for each diagnosis.

information changes household behavior: preventive health services, acute health events, adherence to prescription medication, and the use of low-value health services. Appendices A.3 through A.6 contain a full set of all diagnostic information, procedure codes, and therapeutic classes used in the construction of each of these variables.

Preventive health services. 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).

Acute health events. Second, I define a set of *acute* health events, to capture health shocks of similar severity to new chronic diagnoses, but which are transient in nature and do not communicate intra-household health risk information (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. Appendix Section A.4 compares acute and chronic events by pre-event spending, event cost, and hospitalization incidence, finding the two groups to be comparable.

Prescription adherence. Third, I define adherence to prescription medication; this is used in Section 3.3.2 to separately identify learning about the health system following a health shock. 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).[10]

Low-value health services. 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).[11] 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

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

[11] 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).

these services into five categories: pediatric services, including imaging services and the early use of medications such as antibiotics; adult prescription drugs, such as the use of opiates to treat migraines; unnecessary imaging services for adults; 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 local projections difference in differences (LP-DID) estimator (Dube et al., 2023). This estimator performs a "stacked" regression of treated units combined with their clean controls to estimate treatment effects without bias from naive staggered adoption designs with heterogeneous treatment effects (Roth et al., 2023). The regression uses local projections methods to restrict the estimation sample so that previously-treated observations (which may be experiencing time-varying or heterogeneous treatment effects post-adoption) are not included in the control group, eliminating bias. The LP-DID regression performs similarly to other approaches in this context, including weighted stacked DID regressions (Wing et al., 2024; Cengiz et al., 2019) and imputation estimators (Sun and Abraham, 2020; Callaway and Sant'Anna, 2021). Formally, for a household f and $h = 6$ years pre- and post-treatment, I estimate the equation

$$y_{f,t+h} - y_{f,t-1} = \beta_h^{\text{LP-DID}} \Delta D_{ft} + \alpha_f + \tau_t + \varepsilon_{ft}^h, \tag{1}$$

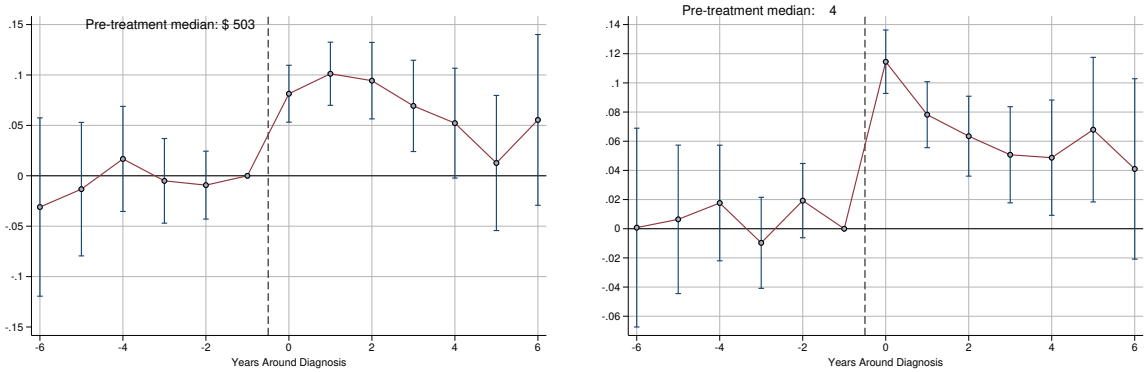
where the sample is restricted to newly treated households ($\Delta D_{ft} = 1$) or clean controls ($\Delta D_{f,t+h} = 0$), with effects estimated relative to the year prior to diagnosis, $t - 1$. To assess spillover responses, my main outcomes Y_{ft} aggregate utilization across a household *excluding* those who experience the major health event. I measure these outcomes both in counts (e.g., number of visits) and spending (both total and OOP), and adjust for the skewed nature of these distributions by using the inverse hyperbolic sine transformation for spending outcomes.[12] Throughout, reported coefficients can be interpreted as approx-

[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 Section B.1 that my results are not substantively altered when using alternative transformations, such as log transformations or Poisson regression. I also discuss the choice of control group in more detail in Section B.2. Effects were estimated using the LPDID package in Stata (Busch and Girardi, 2023).

imate percentage changes in the outcome variable, relative to the year before the shock, $t-1$. Standard errors are clustered at the household level.

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

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



Notes: Figures show LP-DID regression coefficients and 95% confidence intervals. Regressions estimate the effect of a new chronic diagnosis on medical utilization of other (non-diagnosed) household members, measured as (a) the inverse hyperbolic sine of total OOP spending, and (b) number of household preventive services per year. Standard errors are clustered at the household level.

Figure 1 presents the time-varying causal effect of a health shock on household utilization 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 before 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 for at least three years following the diagnosis.

This constitutes a sizable increase in health spending. For reference, given the range of estimates for the price elasticity of demand for healthcare services from -0.2 (Newhouse, 1993) to -1.5 (Kowalski, 2016), this magnitude of change in health spending would require between a 7% and 50% decline in prices to be induced as a pure demand effect. Recent work has argued that demand for preventive care is even less price sensitive than demand for other medical services, suggesting that this might be a conservative range (Haviland et al., 2011).

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 (Section 2.2). Here, too, I find that new diagnoses in a household are associated with strong responses. Affected, non-diagnosed household members increase their overall use of wellness visits by about

3% relative to a median of 4 visits annually. These effects persist for longer than overall spending increases, with statistically significant increases observed even five years after the shock.[13]

A principal identification concern in these regressions is that anticipation of a diagnosis—through, for example, deteriorating health—may introduce unobserved pre-trends into the empirical analysis, even for adjacent household members. For comparison, I present results in Appendix Section B.3 illustrating how diagnoses affect the focal individuals, which allows me to directly observe anticipation effects. Diagnosed individuals more than double their annual OOP spending in the year of diagnosis but not thereafter; additionally, I do not observe spending increases pre-diagnosis. This is intuitive given the types of health shocks being studied, which typically affect children and have a rapid onset.

3.1 Changes as Responses to New Health Risk Information

These results suggest that a new chronic diagnosis induces a meaningful, persistent change in household demand for healthcare. Next, I turn to exploring the mechanisms behind these effects: I first show that effects are indicative of individuals reassessing their health risks given new health information, and consider alternative explanations in Section 3.3.

To explore household responsiveness to risk information contained in a diagnosis, I estimate spillover demand for *disease-specific* preventive services following a diagnosis. The intuition I rely on is that a chronic health shock provides targeted information about a household member's risk for a certain disease, and hence induces demand particularly for screenings or other services relevant to that disease. For example, households with a newly diagnosed diabetic may use diabetes screenings more post-diagnosis than households affected by a non-diabetes diagnosis.[14]

I assess the causal effect of diagnoses on the utilization of disease-specific preventive care using a triple-differences modification of Equation 1. This approach separates the disease-specific effect of risk information from more general diagnosis effects by comparing household responses to any chronic event with a "third difference" that stratifies chronic health shocks by the specific condition affecting a household. That is, I estimate the effect

[13]Results in both panels are robust to alternative measurement of the outcome variables, including total billed spending, number of unique health encounters, or spending on preventive care. Panel B measures preventive care in visits rather than spending to account for the fact that the Affordable Care Act (ACA)'s cost-sharing exclusion disrupted the costs for preventive services for those with ESI (Hong et al., 2017).

[14]This is exactly what is shown in the raw data in Appendix Figure B.3.

of a new chronic diagnosis on a household f's decision to screen for a specific diagnosis d during time t:

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

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 "first difference" effect—the effect of a chronic diagnosis generally—while β_{DDD} identifies the "second difference" effect of the specific diagnosis of interest, compared to other diagnoses.[15] For example, when the outcome of interest is diabetes screenings, the first difference β_{DD} identifies how any chronic diagnosis changes diabetes screenings, while β_{DDD} identifies the specific effect of a diabetes diagnosis compared to any other chronic diagnosis in a household.

The triple difference approach is advantageous because it allows me to compare the causal effect of diagnoses on the use of preventive care across multiple control groups. When the outcome variable is a screening for a specific service (e.g., diabetes), Equation 2 estimates the effect of a "matching" diagnosis on screening, relative to other diagnoses where a screening is not informative. The identifying assumption for the triple differences approach is the same as for the earlier regressions: that spending differences between diagnosed and undiagnosed households would have evolved similarly over time in the absence of treatment.

I estimate several versions of Equation 2 for different diagnoses-screening pairs. These include 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.[16]

[15] These triple differences results are robust when considering the LP-DID estimates of post-treatment effects separately across the third difference (e.g., for a diabetes diagnosis vs. a non-diabetes diagnosis, as is shown in the raw time series in Appendix B.3).

[16] 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.

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

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

Notes: Table presents six DDD regressions estimating the effect of a chronic condition on spillover household investments in disease-specific preventive care (Equation 2). Outcome variables are binary indicators for a screening (column 1); the specific diagnosis d is listed in column 2. DD coefficients (β_{DD}) indicate the effect of *any* chronic diagnosis on screenings, while DDD coefficients (β_{DDD}) indicate the (additive) effect of *specific* diagnoses. [1]Outcome is measured using diagnostic codes. [2]Here, the reference group is acute events. $^{*}p < 0.05$,$^{**}p < 0.01$,$^{***}p < 0.001$

Table 3 presents the results 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. This suggests that household members responding to an informative (chronic) health shock are differentially more likely to seek out information about their health risks than households responding to a less-informative (acute) health shock. Furthermore, specific diagnoses such as cancer and diabetes increase the likelihood that a non-diagnosed household member will seek out specific screenings by 13.2% and 21.1%, respectively. Diabetes diagnoses are also associated with an increase in cholesterol screenings of 7.2%.[17]

[17]I also find evidence that new diagnoses reduce the rate of unrelated screenings; 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.

The second panel of Table 3 reports results for two "placebo" regressions for cases where health events communicate little *risk* information, and hence would be expected not to meaningfully change screening behaviors. These include the effect of new diabetes diagnoses on obesity screenings/diagnoses and the effect of a new household mental health condition on depression screenings. In both cases, while the focal event may increase the salience of highly relevant risk factors or symptoms—including obesity and depression symptoms—the shock is unlikely to change the marginal value of screenings in these instances. For example, while obesity is an important risk factor for chronic conditions such as diabetes, it is typically externally verifiable prior to a physician's diagnosis, limiting the value of the screening even for at-risk household members. I find no evidence that health shocks affect these screenings.[18]

3.2 Quality of Induced Spending Changes

Given observed household responses to chronic diagnoses, a natural question is whether new information improves overall quality of care. While new diagnoses could feasibly lead to substitution of healthcare towards high-value preventive services, affected individuals may increase overall consumption, with limited regard for a service's underlying risk-mitigating value. I examine how diagnoses affect household consumption of low-value care (Section 2.2).[19]

New chronic diagnoses are estimated to increase total low-value spending by about 5% (Appendix Table B.4). However, pooled results mask significant heterogeneity across services that yields useful intuition as to the type of information households are reacting to. Households may seek out different types of care if they are responding to new risk information—by demanding low-value screenings such as preoperative screenings—or responding to marginal price changes following a diagnosis—by demanding elective surgeries.

Table 4 presents results estimating the effect of a new chronic diagnosis in each of five categories using the LP-DID and traditional DD estimators. New chronic diagnoses

[18] Appendix Section B.5 contains further results leveraging within-family variation in relationships and corresponding risk to show that households are selective in which members receive screenings 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.

[19] In addition to the utilization of low-value care, I explore other ways health events alter the quality of consumers' health care decisions, including their plan choices. In general, I do not find that major health events prompt households to switch their health insurance plans.

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

Population	Pediatric	Adult Services			
Service Category	All Services	Prescriptions	Imaging	Screening	Surgery
DD Coefficient	0.048***	-0.001	0.027***	0.095***	-0.098***
	(0.006)	(0.001)	(0.007)	(0.009)	(0.007)
LP-DID Pooled Effect	0.046*	-0.006	0.025	0.077**	-0.086***
	(0.020)	(0.004)	(0.020)	(0.026)	(0.014)
R^2	0.349	0.309	0.293	0.326	0.379

Notes: Table shows DD regression coefficients and LP-DID pooled post-treatment effects for the effect of a new chronic diagnosis. Outcome variables are the inverse hyperbolic sine of billed spending in each category (Section 2.2). Standard errors clustered at the household level.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

increase household utilization of low-value screenings (9.5%), pediatric care (4.8%), and imaging services (2.7%), while decreasing demand for elective surgeries by 9.8%. I find no effect on the misuse of prescription drugs among adults. Taken with the previous results, these findings suggest affected households increase utilization of a broad set of preventive and "psuedo-preventive" services, with less distinction between the average return on those services.[20]

3.3 Alternative Explanations for Spending Changes

My results suggest household health shocks provide important health risk information that changes behavior. However, individuals may be responding to other features inherent in a health shock; I explore this in this section.

3.3.1 Moral Hazard & Salience

A natural response to 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

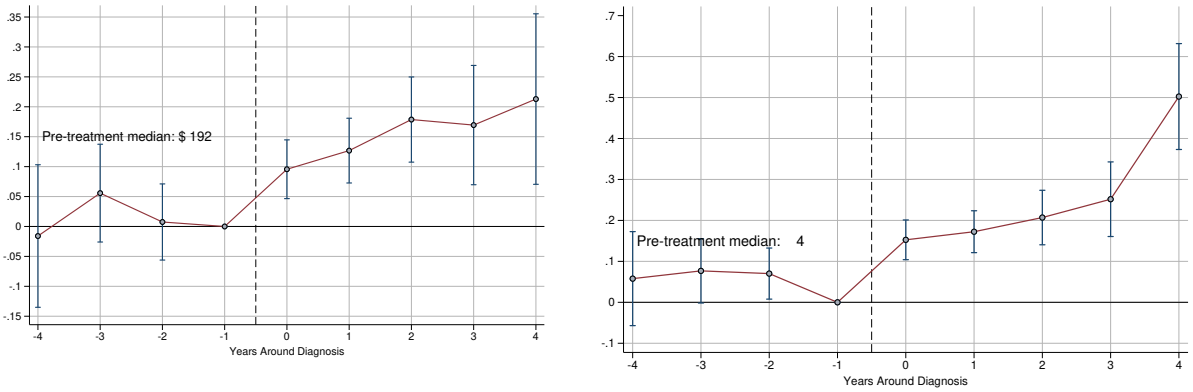
[20] Although I observe households utilizing more of these services—such as preoperative screenings or imaging services—it is unclear whether these are decisions made at the household level or by a physician who knows the family history and hence deems these services as appropriate (Finkelstein et al., 2021). This is particularly relevant in the context of referred services, such as pre-operative screenings (Hoagland et al., 2023).

fixed for the individual, effectively reduce cost-sharing for the rest of the household, lowering future spot prices of (non-chronic) health care (Eichner, 1998; Kowalski, 2016).

Two features of the results suggest that induced-demand responses are insufficient to explain the results. First, the costs of a chronic diagnosis are typically larger in the year of diagnosis than in future years, especially when a diagnosis requires hospitalization. This would lead observed spillover effects to be much larger closer to the diagnostic event and 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, in Figure 2, I examine households enrolled in plans without a deductible, for whom a diagnosis does not meaningfully change household prices; even in this group, health shocks continue to promote strong spillover responses.[21] Were moral hazard the principal mechanism, households in plans with no deductibles would have no incentive to adjust spending decisions (Anderson et al., 2024).

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 LP-DID regression coefficients and 95% confidence intervals. The sample is restricted to households enrolled in ESI plans with zero deductible at the time of the event. Regressions estimate the effect of a new chronic diagnosis on medical utilization of other (non-diagnosed) household members, measured as (a) the inverse hyperbolic sine of total OOP spending, and (b) number of household preventive services per year. Standard errors are clustered at the household level.

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. These "salience effects" are different from effects capturing the risk information contained in a specific condition

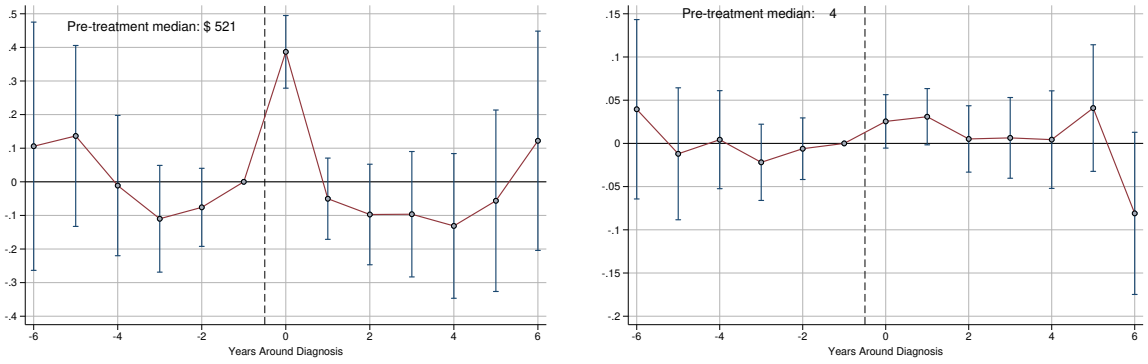
[21] Note that the regression coefficients in Figure 1 and Figure 2 are statistically indistinguishable from each other in both panels, except for in year 4 on panel (b).

(Section 4.2). Rather, the critical difference here is that these general salience effects capture individual responses to overall negative health experiences that alter future risk *preferences*—by affecting the marginal utility of care—rather than risk *beliefs*.

To examine the impacts of salience effects in the absence of new health risk, I compare household responses to the acute events (Section 2.2). The intuition is that these events—for example, a hospitalization for a viral infection—provide transient shocks that don't change individuals' risk beliefs, but may alter the relative valuation of health services (e.g., seeking vaccinations to reduce the likelihood of future hospitalizations). Hence, differences in responses across these two events provide a sense of how health risk information—over and above salience effects—drives the observed spillover responses.[22]

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

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



Notes: Figures show LP-DID regression coefficients and 95% confidence intervals. Regressions estimate the effect of an acute hospitalization (Section 2.2) on medical utilization of other (non-diagnosed) household members, measured as (a) the inverse hyperbolic sine of total OOP spending, and (b) number of household preventive services per year.

Figure 3 presents the results. Unlike in Figure 1, acute hospitalizations spur few changes in health behaviors among other household members, particularly in years following a diagnosis. Acute hospitalizations are associated with a transitory spending increase in the year of diagnosis, potentially capturing moral hazard effects or intra-household correlations in health status. However, these events do not cause increased spillover investments in preventive care. Given that acute hospitalizations make healthcare at least as salient—if not more so—than chronic diagnoses, the findings suggest changes in risk

[22]Recall that acute events are comparable to chronic events as discussed in Appendix A. In fact, acute hospitalizations tend to be slightly more expensive and require longer hospital stays than chronic diagnoses; hence, were salience effects alone driving household responses, I should observe stronger household responses among the sample affected by acute diagnoses. I do not observe this.

preferences alone are insufficient to explain overall changes in behavior. Rather, new health risk information, such as about one's inherent genetic risk for a chronic condition, plays an important role.

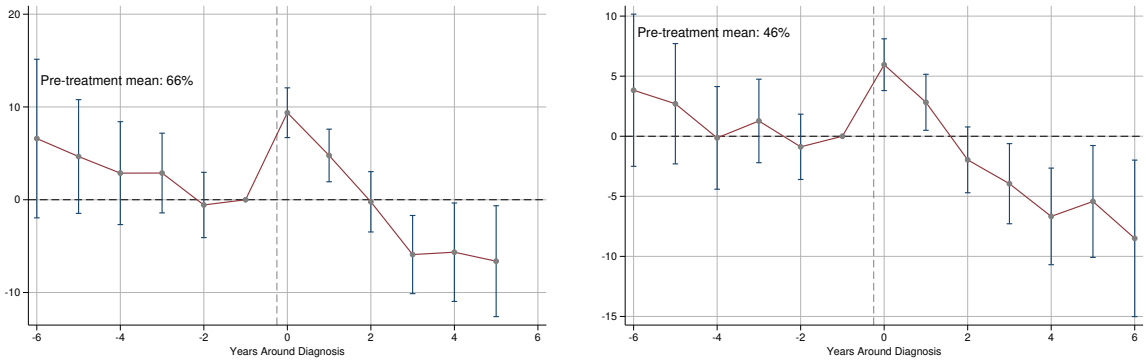
3.3.2 Health Information

In addition to new health risk information, a diagnosis may give families institutional knowledge, such as about the value of medical care, the process of obtaining insurer-covered care, or how to establish provider relationships. It may be difficult to disentangle effects of this knowledge from new risk information, as the two generally co-move.

I separate these by focusing on a unique case where a diagnosis provides information about risk but not institutions: adherence to existing prescriptions, including preventative cardiovascular medications such as statins. These drugs are extremely common and are known to be effective in preventing future health problems (O'Connor, 2006). Importantly, individuals with active prescriptions already have sufficient institutional knowledge to receive this care; hence, when affected by a new household diagnosis, individuals may update risk beliefs without gaining new knowledge about *how* to obtain medication.

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 and 95% confidence intervals. Regressions estimate the effect of a new chronic diagnosis on adherence to preventive medications, measured as (a) a binary indicator for whether the prescription was refilled during the year, and (b) the proportion of days in a year covered by the medication (Section 2.2). Standard errors are clustered at the household level.

Figure 4 summarizes the effects of a chronic diagnosis on adherence to existing preventive prescriptions (Section 2.2).[23] As expected, in the absence of new health information,

[23]The sample includes individuals with preventive prescriptions for at least two years pre-diagnosis. An identification concern in these regressions is that adherence decays over time in response to financial

individuals become less adherent to prescriptions over time. However, diagnoses in the household spur a resurgence in multiple measures of adherence, with affected individuals around ten percentage points more likely to fill a prescription immediately after a household health event. This illustrates that individuals respond to updated risk beliefs communicated by a diagnosis, not simply new information about the logistics of obtaining care.

4 Empirical Model of Belief Formation

Based on the results above, I estimate a dynamic model of belief formation for households learning about health risks. In the model, one individual's health shock propagates health information across a household, leading each member to update their belief about subsequent health risks. The goal of the model is to identify implied health expectations based on observed health utilization choices—separate from other potential mechanisms—and measure changes in welfare associated with potentially under-informed beliefs.

The central contributions of the model—relative to the reduced-form evidence—are threefold. First, the model isolates the effect of health risk information on shaping behavior, as it allows for major health events to affect demand through both updated beliefs and changes to marginal prices of care. Second, the model identifies both the level of perceived health risks and how those beliefs are implied to change following health shocks, given the observed choice data. Finally, the model allows for welfare calculations, and considering counterfactual simulations for how different responses to health shocks (primarily in belief updating) would affect estimated consumer choices and welfare.

Formally, consider a household comprised of individuals i and their corresponding medical demand in a year t. In the model, individuals form beliefs about their future health risks, and in each period receive a draw of their health needs, λ_{it}. In order to maximize expected lifetime utility given their beliefs, individuals then choose two things: medical spending, m_{it}, to meet their transitory health needs; and preventive care investments, s_{it}, to learn about health risks and possibly reduce future health costs. Chronic health shocks, when they occur, exogenously change individual beliefs about health risk.

or knowledge barriers (Slejko et al., 2014). As this may occur at different rates across individuals, these trends are not accounted for by household and year fixed effects. I therefore include additional controls for the number of years an individual has been in the sample.

4.1 Decision 1: Static Spending

In each period, individuals choose healthcare spending to match their expected health needs (Cardon and Hendel, 2001). Annually, individual health needs λ_{it} are given by transient health costs λ_{it}^{TR} incurred with a probability of 1, and chronic health costs λ_{it}^{CH} incurred with some endogenous probability p_{it}. Both costs are drawn from stationary distributions; $F_{\text{TR}}(\cdot)$ represents year-to-year fluctuations in health needs for acute conditions (e.g., illnesses and injuries), while $F_{\text{CH}}(\cdot)$ indicates the underlying costs of chronic conditions, should they occur. An individual's *ex-ante* expected spending, given beliefs p_{it}, is therefore

$$\mathbb{E}[\lambda_{it}] = \mathbb{E}[\lambda_{it}^{\text{TR}}] + p_{it}\mathbb{E}[\lambda_{it}^{\text{CH}}]. \tag{3}$$

Individuals make static decisions about total health spending m_{it} to match their realized health needs λ_{it} given a quadratic loss function that takes into account the OOP costs $c(m_{it}, s_{it})$ of spending,[24]

$$u_{it} = (m_{it} - \lambda_{it}) - \frac{1}{2\omega_i}(m_{it} - \lambda_{it})^2 - c_{it}(m_{it}, s_{it}). \tag{4}$$

That is, absent cost-sharing, individuals would choose medical consumption m_{it}^* to exactly match their health needs λ_{it}. Individual heterogeneity in moral hazard is modeled using the parameter ω_i; those with larger values for ω_i have more elastic demand for care, and therefore increase m_{it} by more given a fixed decline in OOP costs.

Solving the expected-utility maximization problem is straightforward; as marginal OOP costs change based on where consumers are in their insurance contract, the solution depends on which "region" of OOP costs an individual finds themselves in given λ_{it}. For negative or sufficiently small values of λ_{it}, individuals will choose $m_{it}^* = 0$; otherwise, optimal spending follows the condition:

$$m_{it}^* = \max\left[0, \lambda_{it} + \omega_i(1 - c_{ijt})\right]. \tag{5}$$

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. Equation 5 highlights two important

[24] m_{it} measures total spending in dollars—rather than OOP spending—in keeping with earlier literature using this model. This allows a more consistent measurement of utilization regardless of plan characteristics and marginal prices. s_{it} is measured as a discrete number of annual visit, to be consistent with Section 3 and to more closely capture individual decision-making (total billed spending for preventive care is generally unobservable to the individual, especially given preventive care is typically cost-sharing exempt). As in previous work, I abstract away from within-year claim timing (Einav et al., 2013; Marone and Sabety, 2022).

features of the model: first, new chronic diagnoses may generate moral hazard effects among price-sensitive individuals by exogenously lowering marginal costs of care for other household members. Second, despite the apparent symmetry of the quadratic loss function, the features of marginal cost-sharing in insurance contracts ensure that individuals will always choose to over-spend—rather than under-spend—relative to a realized health shock λ_{it}.[25]

4.2 Decision 2: Preventive Care Investments

While the choice of spending (m_{it}) is a static one, individuals' choices of preventive services (s_{it}) influence future shocks λ_{it} and provide information on p_{it}. That is, individuals make dynamic investment decisions to maximize future returns on health and health spending (Grossman, 1972). Both of these effects are specified in the model.

First, preventive care may directly affect an individual's future expected health spending, by both reducing the likelihood of future expensive procedures and ensuring chronic diagnoses are discovered in a primary care setting (avoiding costly hospitalizations). This return on investment is modeled by linking expectations for future shocks $\lambda_{i,t+1}$ to decisions today:

$$\mathbb{E}\left[\lambda_{i,t+1}\right] = \gamma\left(\sum_{\tau<t} s_{i\tau}\right) \cdot \left(\mathbb{E}\left[\lambda^{\text{TR}}_{i,t+1}\right] + p_{i,t+1}\mathbb{E}\left[\lambda^{\text{CH}}_{i,t+1}\right]\right), \tag{6}$$

where $\gamma(\cdot)$ is a function relating total preventive care consumption to reductions in expected health needs. Although one could consider allowing preventive care to affect the *distributions* of health needs instead, to capture the intuition that preventive care reduces one's underlying risk of a health shock, this is not consistent with the types of chronic health shocks studied here, where preventive care prioritizes risk monitoring and early diagnosis (e.g., reducing costs) over mitigating risk (e.g., shifting distributions). This is particularly true given the heritability and rapid onset of the conditions in my sample.

Preventive care also provides individuals with information about their health risks, p_{it}. This is modeled by a Bayesian process, with each individual's prior specified as a normal distribution (in log-odds space) with mean $\mu_{p0,i}$ and variance $\sigma^2_{p0,i}$. Future health signals

[25] The restriction that consumer marginal costs are in the interval $[0, 1]$ implies that individuals prefer a constant amount of over-spending (e.g., $\lambda_{it} + \delta$) than a constant amount of under-spending (e.g., $\lambda_{it} - \delta$). This is consistent with the notion that all else equal, an individual would prefer to spend slightly more than their health shock than slightly less. This is a desirable feature in models of health spending with imperfect information, where the utility costs of over-spending might be systematically lower than under-investing in health care. Appendix Section C.1.1 contains a proof of this result and additional discussion.

are also normally distributed in log-odds space in order to obtain closed-form solutions for belief updating in the unit interval:

$$\log\left(\frac{s_{it}}{1-s_{it}}\right) \sim \mathcal{N}(\overline{p}_i, \sigma_s^2). \tag{7}$$

That is, each preventive visit's signal is distributed around some observable proxy for individual true risk $\overline{p}_i$ with noise σ_s^2.[26] Absent data on actual underlying health risks, I proxy $\overline{p}$ using logistic regressions predicting each individual's probability of a new chronic diagnosis in a year as a function of observable demographics, past acute and chronic health events in the household, and family medical history (including pre-existing conditions). Although imperfect, these proxies are similar to the information a medical professional might convey in a preventive visit, as true health risks are not observed perfectly in that setting either.[27]

4.3 Beliefs & the Role of Major Health Events

Individual beliefs p_{it} capture the relative likelihood that they will incur additional health costs λ_{it}^{CH} associated with a new diagnosis. In the absence of major health events, these beliefs are identified based on two key moments linking the observed choice data and the model. First, given distributions of health needs $(F_{\text{TR}}(\cdot), F_{\text{CH}}(\cdot))$, beliefs link expected health needs λ_{it} to observed health choices in the data. Second, investments in preventive care and their future returns on health spending identify changes in beliefs over time. Both of these use the full sample—including households unaffected by chronic health shocks—for identification (See Section 5.1 for more details).

When a chronic health shock affects an individual, that shock propagates through a household and provides each member with new information about p_{it}. Given the quasi-randomness of these diagnoses, I model these responses as a discrete shift to p_{it}:

$$\Delta(p_{it}) = \pi_1(\text{Chronic Event})_{f,-i}, \tag{8}$$

[26] In expectation, therefore, average individual beliefs are a linear combination of $\mu_{p0,i}$ and $\overline{p}$. The evolution of the mean and variance parameters can be written as: $\sigma^2_{pi,t+1} = \frac{\tilde{\sigma}^2_{it}\sigma^2_{pi0}}{\tilde{\sigma}^2_{it}+s_{it}\sigma^2_{pi0}}$ and $\mu_{pi,t+1} = \frac{\tilde{\sigma}^2_{it}\mu_{pit}+\sigma^2_{pit}\tilde{\mu}_{it}}{\tilde{\sigma}^2_{it}+\sigma^2_{pit}}$, where the variable s_{it} indicates how many health signals an individual has received by the end of period t (Crawford and Shum, 2005).

[27] Specifically, for an individual i and diagnosis d, the underlying risk is the predicted probability from the logistic regression $\mathbb{1}\{d=1\} = \vec{\delta}(agesex_i) + \gamma_1\text{Past Acute Event}_i + \gamma_2\text{Past Chronic Event}_{-i} + \gamma_3\text{Past Acute Event}_{-i} + \vec{\delta}(familyhistory_i) + \varepsilon$ 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.

where f denotes household members affected by i's diagnosis. The parameter π_1 can be identified as the change in beliefs needed to rationalize the LP-DID results from Equation 1 for total spending and preventive care, after adjusting for household-specific moral hazard or salience effects. That is, by including these coefficients as moments in estimation and using information about changes to household marginal OOP costs following health shocks (which are known), one can back out an implied value for π_1.

Identification of both baseline beliefs and the shift implied by a health shock is important for identifying welfare effects. However, estimated equilibrium beliefs on their own are not informative about the correctness or precision of belief updating following a major health event. To define a notion of "over" or "under" reaction, I compare implied beliefs to two important quantities: an individual's own predicted risk for health shocks given demographic information ($\overline{p}_i$), and additional estimates from the epidemiological literature quantifying the genetic risk of diagnoses given family histories (Figure 5).

4.4 Summary

Taken together, individuals respond to their own beliefs p_{it} and both individual and household shocks by choosing $\{m_{it}, s_{it}\}_{t=1}^{T}$ to maximize the sum of per-period discounted expected utilities:

$$\sum_{\tau=0}^{\infty} \delta^{\tau} \mathbb{E}\left[U(m_{i,t+\tau}, s_{i,t+\tau}) | p_{it}\right], \tag{9}$$

where households discount future years consumption at a rate of $\delta = 0.95$.[28] Given this maximization problem, our quadratic loss function, and the evolution of p_{it} defined above, individuals solve the Bellman equation given by

$$V(p_{it}; \lambda_{it}) = (m_{it} - \lambda_{it}) - \frac{1}{2\omega_i}(m_{it} - \lambda_{it})^2 - c_{it}(m_{it}, s_{it}) + \delta \mathbb{E}\left[V(p_{i,t+1}; \lambda_{i,t+1}) | m_{it}, s_{it}, p_{it}\right] \tag{10}$$

Equation 10 highlights that individuals are forward-looking in choosing preventive care investments as a central form of health production, underscoring two consequences of family health shocks. First, exogenous changes to p_{it} may change medical spending m_{it}^* via increased anticipation of negative utility shocks from chronic conditions (in addition to any moral hazard effects). Second, households may respond to this new information by consuming additional preventive care, as this provides accurate information about p_{it}. Both of these results are consistent with the evidence presented in Section 3.

[28] Underlying changes to λ as an evolving state variable are suppressed for ease of notation.

The model's equilibrium parameters include those governing initial beliefs $\{\mu_{p0,i}, \sigma^2_{p0,i}\}$; governing changes to beliefs from preventive care (σ^2_s) or health shocks (π_1); and those governing the return to preventive investments (γ) and demand for health services (e.g., ω_i). Importantly—given the relatively simple state space for preventive care investments and the fact that m_{it} is a static choice in each period—the model can be solved directly for a given guess of these parameters and choice data. Given a starting guess for a sequence of preventive care choices $\{s_{it}\}_t$, I can directly determine the expected sequences of $\{\mathbb{E}[p_{it}]\}_t$ and, therefore, $\{\mathbb{E}[\lambda_{it}]\}_t$. Combined with an individual's plan features, this pins down estimated choices $\{m^*_{it}\}_t$. Finally, selecting the sequence $\{s^*_{it}\}_t$ that maximizes an individual's expected lifetime utility given these implied choices yields the equilibrium choices.[29]

5 Parameterization & Estimation

In addition to the key parameters governing belief formation and evolution, additional parameters include those governing the distributions for health shocks $\{\lambda^{\text{TR}}_{it}, \lambda^{\text{CH}}_{it}\}$, and heterogeneity in price responsiveness, ω_i. As I focus here on the evolution of health beliefs, I calibrate these parameters based on previous work. I model λ^{TR}_{it} 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^{TR}_{it} from a distribution $F(\mu_i, \sigma_i, \kappa_i)$ such that

$$\log(\lambda_{it} - \kappa_i) \sim \mathcal{N}(\mu_i, \sigma^2_i). \tag{11}$$

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 κ_i). I calibrate the three hyper-parameters $(\mu_{\lambda,i}, \sigma^2_{\lambda,i}, \kappa_i)$ using the empirical distribution of annual spending of 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 2014–2018.[30] Similarly, I calibrate expected spending on chronic conditions λ^{CH}_{it} based on the empirical distributions of

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

[30] 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.

disease-specific spending at the HCC level; I modeled separate distributions for the year of diagnosis and follow-up years to differentiate between diagnostic and maintenance costs.

Finally, I calibrate ω_i following regression approaches discussed in Einav et al. (2013). Variation in price sensitivity across individuals is likely a second-order effect in the takeup of preventive care and risk belief updating, particularly as noted in Section 3. The model continues to adjust for individual-level variation in price sensitivity, even though individual coefficients are not equilibrium objects in my structural estimation.

I estimate the model via GMM, using moments that match the model predictions to observed data, including predicted levels of overall health spending and preventive care utilization (mean, median, and RMSPE of predictions) and LP-DID regression coefficients following Equation 1 for health spending and preventive visits (via indirect inference).

5.1 Identification

The main identification challenge is to pin down the elasticity of beliefs to information, π_1; this is identified using two sources of variation. Differences in spending decisions for household members affected by a diagnosis—compared to not-yet-treated and never-treated household members—generate the needed variation to identify the effect of a diagnosis. Empirically, I use the regression coefficients in Section 3 to leverage this variation. Second, to separate the evolution of risk beliefs from alternative mechanisms such as price and salience effects, the model leverages variation across households enrolled in different plans (e.g., households facing large marginal price changes compared to others with no relevant changes) and within-household variation in diagnostic risk (e.g., differences between affected members with the appropriate intrafamilial relationship for a risk signal to be relevant). Finally, π_1 and returns to preventive investments (γ) are separately identified, as preventive care choices made by households unaffected by diagnoses identifies γ while changes over time within affected households helps to pin down π_1.

Despite this identifying variation, residual concerns may remain that there are unobserved mechanisms that causally affect spending after a diagnosis and inappropriately load onto the estimated belief elasticities. Although addressing all concerns is impossible, I present evidence in Section 6 and Appendix C that the estimated values for $\vec{\pi}$ and the subsequent welfare calculations are robust to a suite of modeling choices.

5.2 Robustness

The model presented here is constructed to intuitively identify how individual belief parameters evolve in response to chronic health shocks, and how changes in beliefs affect

consumer demand. However, the principal model results—particularly with regards to belief formation and evolution—are robust to a number of alternative modeling choices, which I outline here. A subset of these model alternatives are discussed and presented in Appendix C.

First, the model is robust to parameterizing additional features of the model, such as arbitrary correlations between λ^{TR} and λ^{CH}; a more flexible functional form for $\gamma(\cdot)$; alternative specifications for the Bayesian learning framework in Equation 7; and belief responsiveness to additional health shocks in Equation 8 (as well as allowing beliefs to revert to baseline over time). Second, model results are robust to estimating calibrated parameters—such as ω_i—directly as structural objects. Finally, the key results do not change if additional modeling complexity is added, such as considering household—rather than individual-level—maximization, or incorporating insurance plan choices as an additional stage in the model.

6 Structural Results

Table 5 presents the equilibrium model parameters estimated by GMM. Standard errors are calculated as discussed in Cocci and Plagborg-Møller (2021), and hence represent conservative, "worst-case standard errors" for calibrated structural parameters.[31]

Prior to individual learning about risk—either through household health events or preventive screenings—I estimate that individual beliefs about a future health shock are, on average, 3.1%. There is considerable heterogeneity in prior beliefs: the 25th percentile of the distribution has prior beliefs of 0.25%, while the 75th percentile has prior beliefs of 28.9%. Considering a relative in-sample diagnosis rate of about 2.7%, roughly 48.5% of the individuals in my sample under-estimate their true risk of a major health event. Preventive care is therefore an important vehicle for individual learning: the variance of preventive signals is roughly 40% of the variance of prior beliefs, suggesting that preventive care improves the beliefs of the representative individual. Finally, chronic diagnoses increase individual risk beliefs by 18.5 percentage points, a more than six-fold increase from the average prior belief.

Figure 5 presents key takeaways from estimation. The first panel highlights the overall match rate between predicted and observed spending, using the LP-DID specification for

[31] These standard errors allow for arbitrary correlations across empirical moments, and 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.

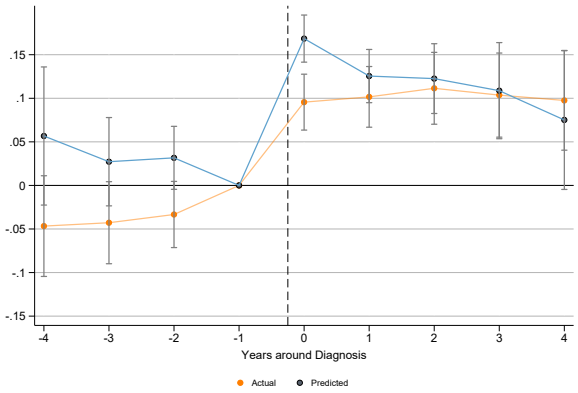
Table 5. Estimated Structural Parameters

		Dynamic Model	
		Estimate	95% Confidence Interval
Panel A: Initial Beliefs			
μ_{p_0}	Prior Mean	0.031	[0.012, 0.050]
$\sigma^2_{p_0}$	Prior Variance	14.16	[8.20, 20.12]
Panel B: Learning from Preventive Care Investments			
μ_s	Signal Mean	0.023	—
σ^2_s	Signal Variance	0.982	[0.880, 1.083]
γ	Health Returns from Prevention	0.996	[0.992, 1.001]
Panel C: Learning from Major Health Events			
π_1	Family Chronic Event	0.185	[0.096, 0.274]

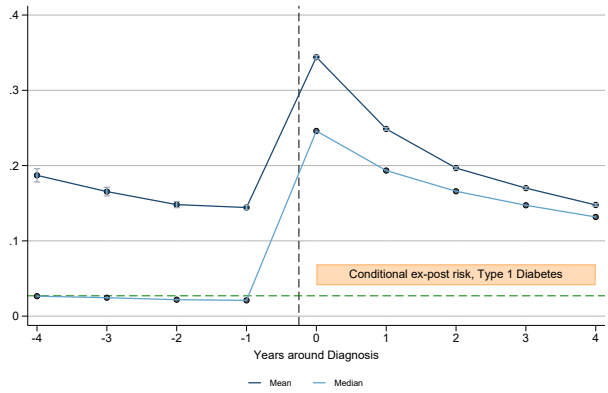
Notes: Table presents estimated equilibrium parameters of the model estimated via GMM on a sample of $N = 472,521$ enrollees in $172,598$ households between 2006 and 2013. All average parameters are expressed in terms of probabilities, while variances are expressed in log-odds. Signal mean μ_s is not estimated via GMM, but rather through individual-level risk predictions. Standard errors are calculated following Cocci and Plagborg-Møller (2021).

Figure 5. Model Predictions: Spillover 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 LP-DID coefficients (Equation 1) for the causal effect of a new chronic diagnosis on both observed and predicted total spending (similar to Figure 1; note that the relationship for OOP spending is mechanical). 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).

the inverse hyperbolic sine of total spending for both groups as in Equation 1.[32] The model captures the key reduced-form results, highlighting a roughly 10% increase in spending following a household member's chronic diagnosis (Figure 1).

The key result of the model—highlighted in the right panel of Figure 5—is that household chronic diagnoses are major drivers in individual health beliefs. The figure illustrates changes in predicted beliefs for households affected by a health shock, illustrating changes in both average and median risk beliefs. Prior to diagnosis, the median individual's beliefs are close to the unconditional probability of a chronic event (roughly 2.7%, shown in green); the skewed distribution of beliefs leads the average to be considerably higher. Following a chronic health shock, risk beliefs increase sharply to over 20% (30%) for the median (average) individual. Investments in preventive care—which increase after the chronic health shock—assist households in correcting these shifts, but beliefs remain higher than 10% for both groups even four years after the health shock.

On their own, these estimated beliefs are uninformative about the value of household reactions, as large changes in beliefs could be warranted by the conditional risk distributions individuals face given a family member's diagnosis. I therefore place these changes in context by comparing them to expected *ex-post* conditional probabilities of a diagnosis, based on external clinical estimation. The second panel of Figure 5 shows this estimated range in orange for a sibling's conditional risk of developing type 1 diabetes given another sibling's diagnosis (Harjutsalo et al., 2005). This is an example where there is large conditional risk-sharing, so the signal is particularly informative. However, estimated changes in beliefs well exceed this standard measure of conditional health risk even four years post-diagnosis.

6.1 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 and consider how welfare changes as information evolves.

Throughout this section, I focus only on individuals affected by a chronic diagnosis in their household. WTP comparisons are made against a counterfactual state in which

[32]Note that as the model used LP-DID coefficients on OOP spending as a moment in estimation, the match between predicted and observed *OOP* spending—rather than total spending—is mechanical. This is also true for estimated effects on preventive visits.

household members continue to invest in and learn from their own preventive screenings, but have no discrete shift in their beliefs following a major health event. As Equation 10 is estimated in dollar terms, this comparison is then simply the difference:

$$WTP_{it} = V_{it}(\pi_1 = \pi_1^*) - V_{it}(\pi_1 = 0). \tag{12}$$

Equation 12 therefore measures individual willingness-to-pay for the risk information contained in a diagnosis, given how they interpret that information. I present results measuring differences only in the diagnosis year, to eliminate concerns about differential timing of health shocks; including all post-diagnosis utilities does not qualitatively change results.

For a substantial fraction of the population, any potential welfare gains from new information are overshadowed by the interpretation of that information: 43% of affected individuals in my sample would be willing to pay to *avoid* new information.[33] Among this population, losses are highly skewed: the average (median) welfare loss is $3,984 ($145), a roughly 40% (24%) decline from baseline utility. To put this into context, the average (median) household would incur this utility loss if they reduced their health consumption by 9% (7%) relative to a fixed health shock λ)

At first glance, a welfare penalty associated with new information is counter-intuitive. This result, however, highlights the discrepancy between how informative a signal actually is and how informative an individual interprets it to be. Given the estimated equilibrium shift π_1 is substantially larger than external benchmarks for true conditional risk (Figure 5), the welfare results are consistent with the findings presented in Section 3. Over-updating reduces individual expected utility through two mechanisms: first, increased expected spending with lower expected health returns (e.g., over-utilization of low-value services), and second, greater uncertainty in spending, especially for more risk-averse individuals.[34] Placing information into context could mitigate utility losses through both channels.

6.2 The Role of Belief Updating in Welfare Penalties

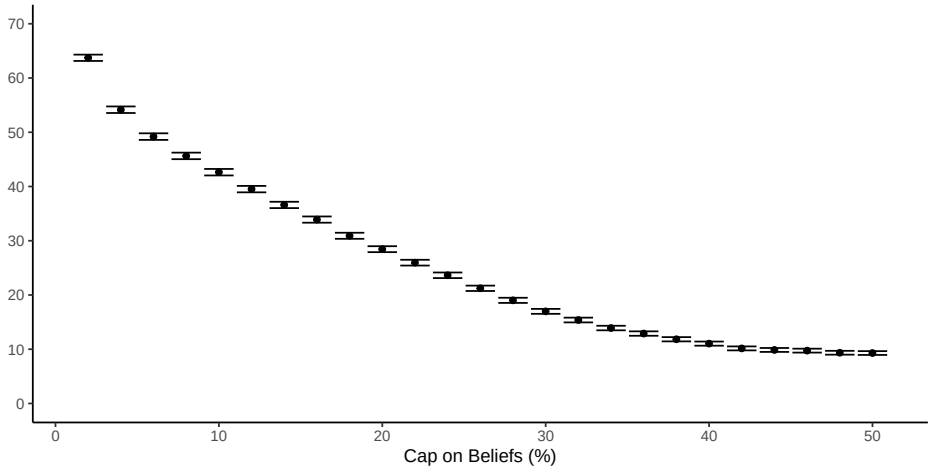
Given these results, I consider how limiting responsiveness to chronic risk information (π_1) might change its value. I perform simple counterfactual scenarios imposing arbi-

[33] Appendix Figure D.1 shows the full distribution of WTP across the affected individuals in the sample.

[34] The second mechanism is similar to an anticipatory "disutility" discussed by Oster et al. (2013). Appendix Figure D.2 highlights the correlation between welfare results and household characteristics, such as pre-diagnosis risk and risk aversion.

trary upper bounds $\overline{p}$ on post-diagnosis beliefs.[35] This exercise intuitively illustrates how valuable one might expect health information to be in contexts where individuals more appropriately interpret information based on their conditional risks.

Figure 6. Bounding π_1 Increases the % of Individuals Valuing Health Risk Information



Notes: Figure shows results from a counterfactual simulation bounding the extent to which beliefs update following a chronic diagnosis. Each point represents the result of a different simulations where beliefs are capped at the indicated value, $\overline{p}$; the $y-$axis shows the fraction of individuals affected by a chronic diagnosis who have a nonnegative WTP for (capped) learning from the diagnosis. Sample is restricted to those with negative WTP for uncapped health risk information, and evaluated at the first year following diagnosis, relative to the valuation under no learning. Utility gains are shown in Appendix Figure D.4, measured according to Equation 10 in 2020 USD.

Figure 6 presents the results for individuals who, in equilibrium, incurred welfare penalties from new health risk information. Each point plots the fraction of these individuals with a nonnegative WTP for information as the exogenous cap on updating, $\overline{p}$, varies. This share increases as $\overline{p}$ declines, prohibiting large *ex-post* swings in underlying beliefs. Without restrictions, just over half of individuals have a positive WTP for information; however, even restricting beliefs of future chronic risk to be 10% or fewer—a relatively generous bound, given true conditional risks are generally below 7%—expand this share to be over 75%. Of those unwilling to pay for health information without a cap on belief updating, over 60% value information when beliefs are capped at 3% or fewer.

6.3 Additional Policy Simulations

The structural model estimated here allows us to move beyond welfare estimation into policy simulations comparing alternative approaches to improve the value of risk infor-

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

mation. Health-related spillovers—especially within families—can be leveraged as policy tools to improve screening and public health, and have therefore attracted recent research (Acosta et al., 2021). However, welfare implications of differing approaches are not obvious, particularly given limited understanding of individual reactions to information; using the estimated model allows for careful comparisons of these options.

My model results are consistent with the empirical findings presented above and additional work studying the role of selection into screenings (Einav et al., 2020). In particular, I find that the value of new risk information—without corrections to belief updating rules—is greatest for individuals with low *ex-ante* risk scores, who may have large existing errors between underlying and perceived risk.[36] Similarly, the model predicts that medical histories are particularly valuable for those with salient familial relationships; as discussed in Section 3, individual WTP for information exhibits strong within-family variation, with some relationships (e.g., siblings for Type 1 diabetes and spouses for Type 2 diabetes) exhibiting greater returns for new health risk information. Together, this suggests value in policies leveraging medical histories and machine learning, among other techniques, to construct a more targeted approach to screenings and the transmission of new risk information.

However, the model suggests that interventions seeking to improve information *interpretation*, rather than simply information *access*, may be more valuable and effective. Health literacy programs that either improve the precision of risk signals or more clearly underscore the value of specific health services for a risk condition may improve patient welfare while reducing overall health spending. This may include improving the return on primary care investments as a way to correct inappropriate health beliefs or to limit the use of pseudo-preventive low-value services.

The model results could be extended in several meaningful ways to improve its use for policy evaluation. First, future work could relax the assumption that individuals have no control over their chronic care health costs. This is particularly interesting in non-ESI populations, including uninsured or Medicaid-enrolled individuals for whom chronic diagnoses may impose large financial burdens (Hadley, 2007). This is related to the effect of liquidity constraints on spending adjustments (Gross et al., 2020), another important consideration to be integrated into the model. 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).

[36] Appendix Figure D.3 shows the results of a simulation using the structural model to estimate the optimal revelation of health risk information based on observable demographics and taking into account individual belief responses.

7 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 leave individuals better off.

I use a structural approach to quantify the welfare effects of new health information. I find that for nearly one-third of affected individuals, information gains are swamped by overly large shifts in estimated *ex-post* risks. Bounding how much individuals increase their risk beliefs post-diagnosis makes information welfare-improving for over 80% of individuals. The findings of the model are robust to multiple specifications, and can be meaningfully used to consider policies leveraging health information to improve social welfare.

Increasing 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 induce over-corrections in 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. (2016). Evolving Choice Inconsistencies in Choice of Prescription Drug Insurance. *American Economic Review*, 106(8):2145–2184.

Acosta, A., Camilleri, M., Abu Dayyeh, B., Calderon, G., Gonzalez, D., McRae, A., Rossini, W., Singh, S., Burton, D., and Clark, M. M. (2021). Selection of antiobesity medications based on phenotypes enhances weight loss: a pragmatic trial in an obesity clinic. *Obesity*, 29(4):662–671.

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.

Anderson, D. M., Hoagland, A., and Zhu, E. (2024). Medical bill shock and imperfect moral hazard. *Journal of Public Economics*, 236:105152.

Arrieta, G. R. and Li, G. (2023). Caring to work or working to care: The intra-family dynamics of health shocks. *American Journal of Health Economics*, 9(2):175–204.

Arteaga, C., Vigezzi, N., and Garcia-Gomez, P. (2024). In sickness and in health: The broad impact of spousal health shocks.

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*.

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

Bundorf, K., Polyakova, M., and Tai-Seale, M. (2021). How do humans interact with algorithms? Experimental evidence from health insurance. Technical Report w25976, National Bureau of Economic Research, Cambridge, MA.

Busch, A. and Girardi, D. (2023). Lpdid: Stata module implementing local projections difference-in-differences (lp-did) estimator.

Callaway, B. and Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

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.

Chan, T. Y. and Hamilton, B. H. (2006). Learning, private information, and the economic evaluation of randomized experiments. *Journal of Political Economy*, 114(6):997–1040.

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.

Darden, M. (2017). Smoking, expectations, and health: A dynamic stochastic model of lifetime smoking behavior. *Journal of Political Economy*, 125(5):1465–1522.

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.

Dube, A., Girardi, D., Jorda, O., and Taylor, A. M. (2023). A local projections approach to difference-in-differences event studies. Technical report, National Bureau of Economic Research.

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.

Dwyer, D. S. and Liu, H. (2013). The impact of consumer health information on the demand for health services. *The Quarterly Review of Economics and Finance*, 53(1):1–11.

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

Einav, L., Finkelstein, A., Oostrom, T., Ostriker, A., and Williams, H. (2020). Screening and Selection: The Case of Mammograms. *American Economic Review*, 110(12):3836–3870.

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.

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.

Fadlon, I. and Nielsen, T. H. (2021). Family labor supply responses to severe health shocks: Evidence from danish administrative records. *American Economic Journal: Applied Economics*, 13(3):1–30.

Fadlon, N., Gross, T., Hoagland, A., and Layton, T. J. (2024). The protective effects of a healthy spouse: Medicare as the family member of last resort.

Finkelstein, A., Luttmer, E. F. P., and Notowidigdo, M. J. (2009). Approaches to estimating the health state dependence of the utility function. *American Economic Review*, 99(2):116–121.

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.

Fontes, L. F., Mrejen, M., Rache, B., and Rocha, R. (2024). Economic distress and children's mental health: Evidence from the brazilian high-risk cohort study for mental conditions. *The Economic Journal*, 134(660):1701–1718.

Frimmel, W., Halla, M., Paetzold, J., and Schmieder, J. (2020). Health of elderly parents, their children's labor supply, and the role of migrant care workers.

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.

Goldstein, Y., Chodick, G., and Shurtz, I. (2023). Realization of Low Probability Clinical Risks and Physician Behavior: Evidence from Primary Care Physicians. *American Journal of Health Economics*.

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.

Gross, T., Layton, T. J., and Prinz, D. (2022). The liquidity sensitivity of healthcare consumption: Evidence from social security payments. *American Economic Review: Insights*, 4(2):175–190.

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.

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.

Haviland, A., McDevitt, R., Sood, N., et al. (2011). Healthcare spending and preventive care in high-deductible and consumer-directed health plans. *The American journal of managed care*, 17(3):222–230.

Hoagland, A., Kantarevic, J., Stutely, J., and Wijeysundera, H. C. (2023). Importance of direct exposure in continuing medical education: Primary care physician learning through patients with transcatheter aortic valve implantation. *The Canadian Journal of Cardiology*, pages S0828–282X.

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.

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.

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.

Maestas, N., Messel, M., and Truskinovsky, Y. (2024). Caregiving and labor supply: New evidence from administrative data. *Journal of Labor Economics*, 42(S1):S183–S218.

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

Newhouse, J. P. (1993). *Free for all?: lessons from the RAND health insurance experiment*. Harvard University Press.

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.

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.

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.

Roth, J., Sant'Anna, P. H. C., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature - ScienceDirect. *Journal of Econometrics*, 235(2):2218–2244.

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, 2 edition.

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

Wing, C., Freedman, S. M., and Hollingsworth, A. (2024). Stacked Difference-in-Differences. Technical report, National Bureau of Economic Research.

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.

Yi, J., Heckman, J. J., Zhang, J., and Conti, G. (2015). Early health shocks, intra-household resource allocation and child outcomes. *The Economic Journal*, 125(588):F347–F371.

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.