Medical Bill Shock and Imperfect Moral Hazard*

David M. Anderson† Alex Hoagland‡ Ed Zhu§

January 16, 2023

Abstract

A central problem in the optimal design of health insurance is medical price sensitivity among consumers. However, delays in when pricing information is communicated to consumers may create distortions in consumption choices. We study spillover household responses to scheduled medical services before and after a medical bill arrives, leveraging variation in the time an insurer takes to process a claim. Immediately after services, non-diagnosed household spending increases by roughly 60%; however, a bill's arrival causes a reduction in spending by 8.5%, nearly 15% of the increase. Importantly, responses are not entirely comprised of strategic delays in care; a bill also affects *where* consumers seek care even for non-delayable services, such as hospital care for respiratory infections. Our results suggest households learn pricing information from a bill: our effects are much larger when the bill communicates that a household is just shy of meeting a deductible. We model how households form beliefs about marginal prices, and find households overestimate their expenditures by 10% prior to a bill. This leads to over-consumption of $842.80 ($480.59) for the average (median) affected household member. Policy simulations show that novel plan designs—such as shortening deductible periods—may stabilize consumption trajectories.

Keywords: *Ex-post* moral hazard, price transparency, learning, low-value care
JEL codes: I12, I13, D01, D90

*The authors are grateful to Laura Derksen, Tal Gross, Ryan McDevitt, Petra Rasmussen, Paul Shafer, and Jonathan Zhang, for very helpful comments on this project. In addition, we acknowledge participants in the APPAM 2022 Research Conference, as well as seminar participants at Brigham Young University, Duke University, and the University of Toronto for useful feedback.

†Duke University, Duke-Margolis Center for Health Policy, 100 Fuqua Drive, Durham, NC. dma34@duke.edu.

‡Corresponding author. University of Toronto, Institute of Health Policy, Management, and Evaluation, 155 College St, Toronto, Ontario, Canada. alexander.hoagland@utoronto.ca. Website: alex-hoagland.github.io.

§Boston University, Department of Economics, 270 Bay State Road, Boston, MA 02215. edzhu1@bu.edu.

1 Introduction

The provision of health insurance plays a vital role in protecting consumers against the risk of volatile, unpredictable health shocks. However, incomplete information plagues health insurance markets, ultimately leading both public institutions (e.g., governments) and private organizations (e.g., insurers) to provide sub-optimal coverage (Einav and Finkelstein, 2018; Dave and Kaestner, 2009).

A primary information friction in healthcare markets is "*ex-post* moral hazard," or the extent to which healthcare consumption decisions are price sensitive.[1] This price sensitivity ultimately justifies exposing consumers to out-of-pocket (OOP) cost-sharing for health services (Chandra et al., 2010; Goldman and Philipson, 2007), for example, through increasing enrollment in high-deductible health plans (HDHPs) (Geyman, 2012). While this exposure is nominally to limit potential over-consumption of low-return health services, price pressures reducing consumption may harm households who either delay or forego necessary medical care[2] or who reduce consumption of even high-value health services, such as preventive care.[3]

While consumers are responsive to prices when making healthcare consumption decisions, there is ongoing debate about the extent to which consumers actually have access to information about the marginal costs of care (Lieber, 2017). Much of this debate has focused on consumer knowledge of the *ex-ante* OOP price of a service, such as how consumers search across multiple medical providers offering the same service at different prices (Brown, 2017). In contrast, in this paper we highlight an overlooked feature of medical demand under price uncertainty with significant implications for models of *ex-post* moral hazard and cost-sharing: lack of timely *ex-post* pricing information.

Consumers are rarely, if ever, given information about the total, negotiated, or final prices of a health service at the point of its consumption, including their own expected OOP contribution.[4] Although patients value price transparency and would like to know

[1]The use of "moral hazard" to refer to elastic demand for medical care is an abuse of notation that is now widely used in this literature, beginning with Arrow (1963). We use moral hazard to refer to how, conditional on health status, individuals adapt consumption to the price of care (Pauly and Blavin, 2008; Cutler and Zeckhauser, 2000). Previous work has also underscored the role of this price sensitivity in decision-making (Kowalski, 2016a; Duarte, 2012; Dunn, 2016).

[2]Consumers exposed to higher rates of cost-sharing are more likely to report delaying medical care, a finding that is exacerbated among households with low income (Kullgren et al., 2010) or high-cost chronic conditions (Fu et al., 2021; Gaffney et al., 2020).

[3]While value-based insurance designs—where certain high-value services are carved out of cost-sharing obligations for consumers—have become more prevalent (Chernew et al., 2007), confusion about insurance contracts may still lead to reduced take-up (Hoagland and Shafer, 2021; Shafer et al., 2021).

[4]Notably, health services are characterized by a total amount billed by physicians (a "sticker" price); a negotiated total price approved by the patient's insurer; and the relative fraction of that negotiated price that is the patient's OOP responsibility. Importantly, price transparency for medical pricing must take into account these various prices, including the relative lack of information contained in the sticker price of a

their OOP costs before agreeing to a service (Henrikson et al., 2017), consumers in our sample wait an average of 4.1 weeks before any pricing information arrives, either as an Explanation of Benefits (EOB) from their insurer or a bill from their physician. During the waiting period, consumers must form expectations about their already realized expenses when making future care decisions, a nontrivial task given substantial variation in the price of even basic health services (Gruber, 2022; Cooper et al., 2019).[5] Given that health insurance contracts aggregate spending across household members using a multipart tariff, residual uncertainty about realized spending affects future marginal costs for care.

We isolate the causal impact of receiving information about realized spending on household spillovers in healthcare consumption. We study how households with employer-sponsored insurance (ESI) in the US make collective spending decisions after one household member incurs a significant health expenditure.[6] Our identification is based on exogenous variation in the time it takes insurers to receive and process bills, which affects the length of the household's "interim period" between a service and a bill.

Using a triple-differences regression design, we estimate how scheduled healthcare consumption generates distinct household consumption spillovers before and after pricing information arrives. We find that, in the interim period between the service and its bill, household members increase their total health spending by about 60% (roughly \$72 per person per week). However, once the bill arrives, consumption drops significantly by 8.5%, almost 15% of the initial increase.

We claim that these effects are primarily driven by the pricing information contained in the bill—rather than eliminating other information frictions—using three supporting analyses. First, the effects of a bill are largest when the bill is most informative about prices, including households with lower initial consumption and those whom the bill leaves just shy of meeting a deductible. Second, we show that a bill's arrival changes how households seek care for even unforeseen consumption, such as for respiratory infections. Bills both reduce a household's overall use of services for these infections as well as shift where households seek care (e.g., from a hospital to an outpatient clinic). Finally, we explore bill effects across a spectrum of services and find households are most likely to respond in areas of healthcare which are more elastic, such as preventable hospitalizations and general practitioner visits.

Our results provide evidence that household medical decision-making may be influenced

medical service.

[5] Appendix Figure A.7 illustrates some of the variation in prices for common services in our sample.

[6] Specifically, we assess spillover household responses following the use of a health service classified as "shoppable" by the Centers for Medicare & Medicaid Services (CMS) (CMS, 2019); see Section 2 for details. We exclude the household member who received the service in order to estimate spillover responses among the unaffected household members, and identify the causal effect of a bill's arrival in changing these responses.

by incorrect beliefs about their spending histories, given the uncertainty associated with marginal prices. We therefore develop and estimate a model of imperfect moral hazard in which households form beliefs about marginal prices in an environment where information is delayed. We use the exogenous variation in our data to identify both household beliefs about prices as well as learning over time. Consistent with our reduced-form results, our model predicts that households are under-informed about OOP spending prior to a bill; average household beliefs about realized OOP spending are 11% higher than actual spending. As a result, 10.5% of households spend more than they would were pricing information immediately available, with the average (median) over-consuming household spending \$843 (\$481) more per household member per plan year.[7] We also find strong evidence of consumer learning: at the beginning of a plan year, households over-estimate the OOP costs of a medical service to be as high as 180% of the truth, an over-estimate which is quickly corrected as bills for medical services provide new information.

We present the first model of healthcare demand under price uncertainty and highlight its implications for consumption and welfare; hence, our work makes several important contributions. First, we contribute to an ongoing literature on dynamic responses to cost-sharing, including the strategic delay of some services such as dental care (Cabral, 2017), and models of "forward-looking" moral hazard (Aron-Dine et al., 2015; Baicker et al., 2015). Recent work has found that patients will defer care to take advantage of future changes in prices, perhaps more so than they will increase within-period utilization as prices fall (Hettinger, 2022; Johansson et al., 2023). Although there is strong evidence for the role of dynamic moral hazard in healthcare (Klein et al., 2022; Diaz-Campo, 2022), our results highlight that information about *ex-post* prices changes even real-time decisions about both when and from whom to receive care. In particular, we show that even in cases where care cannot be strategically delayed (e.g., for respiratory infections), households respond to pricing information by changing where they seek care.

Our findings also fit into a larger discussion of the usefulness of price transparency policies in mitigating large levels of healthcare consumption in the United States (Muir et al., 2012; Zhang et al., 2020). In contrast to previous work—which highlighted how the availability of price information may change the strategic decisions of patients shopping for a service (Gondi et al., 2021; Reed et al., 2005)—we highlight a new mechanism through which price transparency may affect *future* care decisions across entire households. Our findings provide strong evidence that reducing price uncertainty in the weeks or months after a service may

[7]This over-consumption occurs most often in our model because households believe erroneously that they have already met their deductible, leading to lower marginal costs of care. Although this leads to greater total spending, this is not to say that increased consumption is normatively of lower value to the household.

have snowball effects reducing the utilization of unnecessary care for a greater number of people over a longer duration of time. Policies which shorten the length of a delay for medical cost estimates would reduce variance between expected and actual cost-sharing, as would real time claims adjudication, similar to prescription drug claims adjudication (Hartzema et al., 2011).

Finally, we provide the first estimate of household beliefs about health expenditures in a delayed learning setting. These estimates are particularly policy-relevant as they allow us to estimate the effect of under-information on household medical consumption trajectories, including simulating how care coordination decisions may change as insurance plan characteristics evolve. Our model allows us to compare alternative plan designs that may reduce pricing uncertainty faced by households, including plans with deductibles that reset more frequently than once per year (Korenstein et al., 2012; Elshaug et al., 2017).[8] Our findings are therefore related to a broader discussion of how consumers respond to nonlinear health insurance contracts (Brot-Goldberg et al., 2017; Stockley, 2016) and belief formation in healthcare (Hoagland, 2022). We find that consumers are responsive to pricing information across a broad spectrum of services, fluctuations which could be smoothed by targeted health policies (Shafer et al., 2022).

Our setting of healthcare consumption is not the only industry where marginal price uncertainty affects consumption decisions. "Bill shock" is common in other domains of consumption, including household utilities (e.g., electricity), cell phone service, and even college financing decisions (Grubb and Osborne, 2015). Our work furthers models of demand under marginal price uncertainty by providing a tractable estimation of consumer beliefs, as well as providing insight into the welfare effects of various policies aimed at reducing information frictions. Hence, our model is related to those focused on learning about prices, including uncertain prices of financial assets and agricultural goods (Ngangoue, 2021; Boyd and Bellemare, 2020). In contrast to other models, our model does not rely on consumer inattentiveness to past consumption, but rather underscores the role of uncertainty arising from complex contracts involving multiple parties (insurers, physicians, and patients) (Grubb, 2015).[9] Estimating the effects in the healthcare consumption has the added advantage that

[8] Although the relevance of price uncertainty in healthcare for developing countries has been noted briefly (Knowles, 1995), our work is the first to formalize this and directly discuss policy implications. Our model is useful in the context of countries where consumers face demand-side cost-sharing for health care, even in countries with universal health care (e.g., Australia, Germany, and the Netherlands, among others) (Globerman, 2016).

[9] Our work is also related to a literature on learning models with delays in belief updating (Karlsson et al., 2009; Peng, 2005). However, in these models, delays typically arise endogenously as consumers either choose to delay learning or have limited information processing abilities. In contrast, our model exploits exogenous variation in the delayed *arrival* of information outside the consumer's control, but which still affects the marginal utility and costs associated with choices retroactively.

we avoid concerns about endogenous price setting, given that bill shock arises as a disconnect between insurers and physicians, rather than a single organization such as a cell phone service provider.[10] Finally, studying bill shock in healthcare is particularly salient given that privately-provided healthcare comprises roughly 6% of US GDP.

We discuss the setting of shoppable services and the data in Section 2. We then present our methods and identifying assumptions in Section 3, followed by our empirical results in Section 4. We incorporate these findings into a model of imperfect moral hazard in Section 5, with estimated results and insights in Section 6. Finally, Section 7 highlights the relevance of these findings for optimal design of insurance contracts.

2 Setting and Data

2.1 Data

Our primary data on household health utilization come from the IBM/Truven Marketscan *Commercial Claims and Encounters* Data, spanning from 2006 to 2018. These data contain detailed inpatient, outpatient, and pharmaceutical claims for a sample of households enrolled in ESI through large U.S. firms. Each observation includes diagnostic, procedural, and payment information, including the date of service and the corresponding date on which the insurer paid their portion of the claim. In addition, the data includes household, firm, and insurance plan identifiers.[11]

We limit our analytical sample to enrollees in one of eight large firms with plan identifiers available.[12] Our final sample includes 386,240 households with two or more members, full eligibility, and continuous enrollment across their window of observation. Throughout, spending data has been normalized to 2022 USD using the Consumer Price Index for All Urban Consumers series.

Table 1 presents summary statistics for the full sample as well as the subset of the sample with insurance plan identifiers. Households tend to be young and relatively low-risk, with an average age of 31.7 years and between 3 and 4 household members. Insurance coverage is more generous than average, although the conditional average deductible is over $1,000, and household members who select into shoppable services typically spend close to a full

[10]This is in contrast to endogenous price setting in the context of *ex-ante* prices for specific medical services, as discussed in Brown (2017).

[11]Note that insurance plan identifiers are only available through 2013, which will affect the analytical sample used in the structural exercise.

[12]These firms are selected randomly from a larger sample of firms with plan identifiers available, and do not otherwise differ meaningfully from the full Marketscan data. Note also that all plans have a start date of January 1 in all observed years.

	Full Sample	Plan-Identified Sample
Panel A: Demographics		
Age (individual)	31.67 (0.000)	31.15 (0.000)
% female (individual)	0.51 (0.000)	0.51 (0.000)
Risk Score	0.29 (0.000)	0.29 (0.000)
Family Size	3.08 (0.000)	3.10 (0.000)
Panel B: Medical Utilization		
Total medical spending (individual)	$4,764 [$975] (0.002)	$4,406 [$887] (0.002)
% of individuals with no spending	0.17 (0.000)	0.20 (0.000)
OOP medical spending (individual)	$650 [$198] (0.000)	$562 [$167] (0.000)
Household deductible | deductible > 0	—	$1,040.24 (0.001)
% Households with zero deductible	—	0.26 (0.000)
Household coinsurance rate	—	0.29 (0.000)
% individuals with shoppable services	0.06 (0.000)	0.06 (0.000)
Total cost, shoppable service	$5,572 [$3,721] (0.011)	$5,645 [$3,814] (0.015)
OOP, shoppable service	$691 [$388] (0.002)	$574 [$290] (0.002)
Years	2006–2018	2006–2013
N_{families}	368,237	367,445
$N_{\text{individuals}}$	1,357,392	1,311,554

Notes: Enrollees include employees and their covered dependents. Risk scores are calculated using the CMS-HCC 2014 community model. Household plan characteristics are calculated as discussed in Section 2. Spending values are reported in 2022 USD. Standard errors are reported in parentheses; medians (when reported) are in brackets.

Table 1. Household Summary Statistics

year's OOP costs on that service alone. Note that the sub-sample with plan identifiers does not appear substantially different from the full sample, an important fact given that we use the plan-identified sample in our structural approach (Section 6). Households in the plan-identified sample incur slightly lower OOP costs than the full sample; however, this is likely indicative of decreasing insurance coverage generosity over time, given that the latest 5 years of data are excluded in this sub-sample.

2.2 CMS Shoppable Services

Our goal is to assess how pricing information contained in a medical bill alters household utilization patterns. In our primary specifications, we analyzed the impact of medical bills for individual health services that are expected to generate a significant—but unknown—amount of OOP spending for the household. We identified the utilization of 30 CMS "shoppable services," which correspond to frequently billed healthcare services that patients can schedule in advance and for which there exists substantial variation in charges across providers(CMS, 2019; White and Eguchi, 2014).[13] In particular, CMS shoppable services constituted 16% of overall OOP spending for individuals on ESI in 2017 (Bloschichak et al., 2020).

The complete list of services is available in Appendix Table A.7.[14] In general, our services are divided into three broad categories: pathology services (e.g., diagnostic biopsies), radiology services (e.g., electrocardiograms), and surgical services (e.g., spinal fusion or removal of cancerous growths).[15] Our choice of services is not based on the relative quality or value of a service, in contrast to other sets of high-frequency health events (for example, urgent or non-urgent hospitalizations as discussed by Card et al. (2009)). Instead, we assess how households affected by these relatively costly medical procedures make decisions about potentially non-urgent or low-value services as a result of the exposure to pricing information (see Table 5).

Our focus on shoppable services provides a tractable means to assess the influence of price uncertainty on future household consumption. By focusing on commonly-billed services with both a relatively high average cost and a sizable variance across providers, we are able to cleanly identify the effects of pricing information on consumption across many households in a large dataset. In addition, simplifying the set of treatment events enhances the tractability of reduced-form regressions, given that many households consume far fewer of these services than more general services (and 94% of households do not consume any of these services in a year). The tradeoff associated with this limited focus, however, is that our results may not apply to simpler (and generally cheaper) services, such as general wellness visits. However, in the structural model in Section 5, we generalize our setting to include all medical services consumed in a plan year, significantly widening the scope of our analysis.

2.3 Bill dates & Plan characteristics.

One limitation of our data is that we are not able to view the exact date on which consumers first received a bill from their provider for the services rendered. Instead, we observe the date the insurance plan paid the provider their portion of the pill. As this is the first possible date at which a patient will receive their Explanation of Benefits (EOB), it is the earliest

[13] We identified these services in the claims data using Current Procedural Technology (CPT) codes for outpatient and inpatient services and Diagnostic Related Groups (DRGs) for inpatient hospitalizations.

[14] Effective January 1, 2021, hospitals must publish standard charges for these services online, including negotiated rates. This does not affect our analytical sample (which goes through 2018). Prior to implementing this rule, there has been little empirical evidence found that patients engage in price shopping for these procedure ahead of time (Mehrotra et al., 2017).

[15] The final list of CMS shoppable services includes commonly used hospital evaluation and management (E&M) codes; we did not include these as medical events in our sample due to the substantially lower average cost of these services compared to other categories.

date that definitive OOP cost information becomes available to a patient. Hence, we use this date as a proxy for patient bill information.

While this is a noisy proxy, the effects of any measurement error here are expected only to attenuate our findings. Since our proxy measures the earliest possible date at which households have access to pricing information, noise in our context always leads to a misclassification of the post-bill indicator to be 1 when it should be 0, rather than the other way around. Hence, any contamination bias arising from misclassification only operates in one direction, meaning the resulting coefficient on the post-bill indicator will be a weighted average of true post-bill effects and contamination from the interim period for any misclassified treatment dates. Hence, as long as the effects of a bill's arrival are of opposite sign than the effects of service (for example, if spillover household consumption increases following a service but then declines after the bill arrives), any contamination bias will attenuate the correction parameter towards zero.

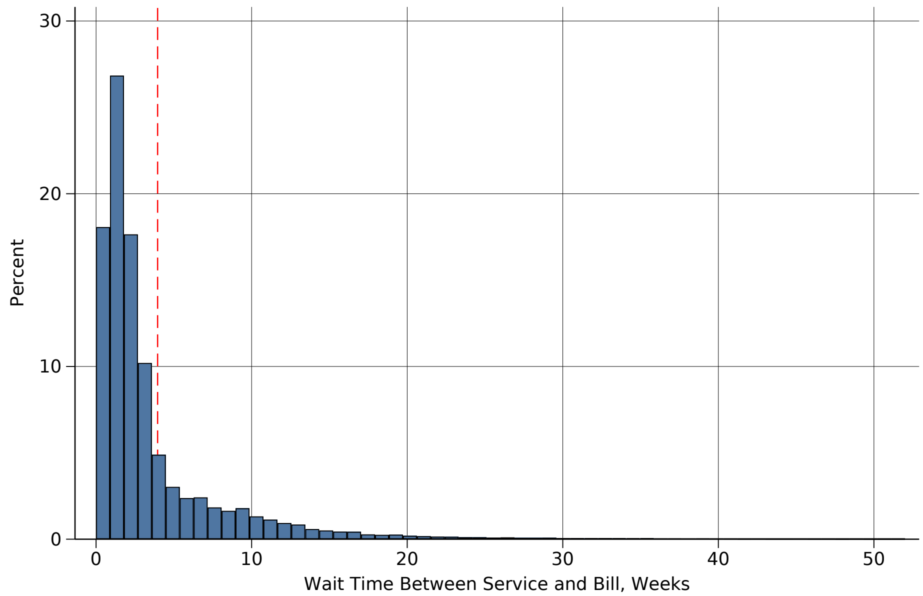
Figure 1 presents the distribution of wait times (in weeks) between the date a shoppable service was received and the date the insurance plan paid their portion of the service bill. Note that there is substantial variation in this wait time, with roughly 60% of bills being paid by insurers within the first four weeks, and the rest taking longer than a month for payment to be settled.

We claim that the length of this waiting period for a bill to arrive is exogenous at the household level, thereby allowing us to identify the causal impact of receiving information about OOP expectations on household health utilization. Appendix Figure A.8 illustrates the substantial variation in how long it takes an insurer to receive and process a claim, both within and across years. The average household waits 4.1 weeks for a bill from a shoppable service; however, waiting times tend to be higher at the beginning of a calendar year and the first month of each quarter, when insurers have billing changes and new policies to incorporate into their processing algorithms.[16] Waiting times are also affected by other time-varying features of the healthcare system that are exogenous to the household, including the rate at which physicians submit claims to insurers for reimbursement. The exact variation in bill waiting times is therefore the result of interactions between an insurer—typically chosen at the employer level in our context, rather than the household level—and specific physicians or hospitals. Even if households attempted to choose general practice providers based on the relative efficiency of billing with their specific insurer, this is unlikely to be a driver in household choice of physicians and hospitals from whom they receive the shoppable services

[16]Waiting times are also affected by more general health policies, such as the national transition to the International Classification of Diseases, 10th Revision, Clinical Modification (ICD-10-CM), in October 2015. This transition increased billing complexity by roughly five times and, subsequently, the rate of administrative frictions in processing billing information (Caskey et al., 2014).

in our data (e.g., the surgeon who performs a mastectomy). Hence, the variation in the length of time a household waits for their bills is both unpredictable and exogenous at the consumer level.

Figure 1. Variation in Wait Times Between Service Date and Bills' Arrival



Notes: Figure depicts the distribution of wait times between the date the service was provided and the date the insurance company paid their portion of the service bill to the provider, measured in weeks. Only services included as shoppable health events in our analytical sample are shown here. Vertical dashed red line indicates the average duration of the waiting period, approximately 4.5 weeks.

In addition to data on individual health services, we utilize data on insurance plan characteristics to estimate how households respond during the period when OOP prices remain unknown to them. In constructing measures for these characteristics, we follow previous literature (Hoagland, 2022; Marone and Sabety, 2022).[17]

3 Methods

Household consumption of shoppable health services—particularly those with large expected OOP costs relative to a deductible—may generate strategic responses in household health spending (Cabral, 2017). When exact OOP spending information is not immediately transmitted, responses take place in two stages: first, households respond to the event itself, based

[17]For tractability, our model assumes cost-sharing contracts are comprised of: a family deductible, a single non-specialist coinsurance rate, and a family OOP maximum. Rates are constructed using the empirical distribution of payments in the data (Zhang et al., 2018; Marone and Sabety, 2022). See Hoagland (2022) Appendix A for a detailed description of this methodology and an evaluation of the quality of these inferences.

on their expectations of spending; second, and only after the bill arrives, households make decisions with full information in hand. Hence, we leverage these two distinct response periods in a triple-differences regression framework to estimate spillover responses to scheduled healthcare consumption *separately* for the periods before and after a bill's arrival.

We estimate the causal impact of a bill on spillover spending (e.g., for all household members excluding the original consumer of the shoppable service). There is strong evidence that individual health events generate spillovers affecting the utilization decisions of other household members (Hoagland, 2022; Fadlon and Nielsen, 2019).[18] We therefore estimate a bill's effect on total spillover health spending (measured per week per household-member) in household i at week t of year y as given by Equation 1:

$$\mathbb{E}[\text{spend}_{ity}] = \exp\left\{\beta_1 \mathbb{1}(\text{post\_service}_{ity}) + \beta_2 \mathbb{1}(\text{post\_bill}_{ity}) + \gamma \vec{X}_{ity} + \alpha_{\mathcal{I}} + \tau_t + \delta_y + \xi_{\text{MD}}\right\}, \tag{1}$$

where the two main regressors are dummy variables indicating whether the shoppable service had already been performed and if the bill for the service had arrived by week t, respectively. We also control for linear time trends before and after the service, as contained in the vector $\vec{X}_{ity}$.[19] Finally, we consider the robustness of our estimation approach to controlling for various time-invariant fixed effects, including those for individual households, years, relative week of the year (to account for within-year seasonality in health spending), and provider fixed effects (for the providers offering the shoppable service).[20]

We use Poisson regression to estimate multiplicative effects on spending. A Poisson regression model is advantageous as it allows us to deal with the skewed nature of our (non-negative) spending data while appropriately including weeks with zero spending and avoiding a complete specification of the dependent variable's distribution (Manning and Mullahy, 2001). Our estimator will be consistent as long as the conditional mean of the dependent variable is correctly specified, as is the case in ordinary least squares (OLS) regression Gourieroux et al. (1984). Additionally, Poisson regression allows us to avoid the inconsistency of regression coefficients induced by heteroskedasticity in a log-linear transformed model (Santos Silva and Tenreyro, 2006) or conerns associated with nonlinear transformations of the dependent variable (Mullahy and Norton, 2022).[21]

[18]There may be situations where information is not shared fully across a household (e.g., young adults in the household still covered on a plan but no longer living at home). This would tend to bias our estimated results towards zero, a problem discussed in other work (Kowalski, 2016b; Fadlon and Nielsen, 2019).

[19]In our preferred specification, X_{ity} includes two controls for separate linear time trends before and after the shoppable service. Our results are robust to more flexible specifications, including allowing "dynamic treatment effects" of the shoppable service—independent of the bill—in a two-way fixed-effects framework.

[20]Our results are robust to including a procedure-specific fixed effect as well, allowing for potential differences in household behavior following different types of shoppable services.

[21]Poisson regressions were estimated in Stata using the "ppmlhdfe" command to handle high-dimensional

A critical assumption for the identification of our parameter of interest ($\beta_{\text{post\_bill}}$ in Equation 1) is that the arrival of spending information through the bill is exogenous to the affected households' strategic spending decisions. Previous work has highlighted the endogeneity inherent in estimating demand elasticities to major health events, especially when those events are planned or scheduled in advance (Duarte, 2012). In our estimation, as we are not attributing spending responses to changes in price (e.g., $\beta_{\text{post\_service}}$ is not a demand elasticity), there are no potential endogeneity concerns in estimation, as we are including both strategic and non-strategic responses collectively in $\beta_{\text{post\_service}}$. As long as bill arrival times are exogenous to the household, $\beta_{\text{post\_bill}}$ will represent a causal estimate of household "corrections" in response to pricing information. However, should bills take systematically longer for higher-risk patients (who might have families who are more likely to respond to health spending in the first place), our estimates would be inconsistent.

There is strong evidence that variation in the time households wait for a bill to arrive is exogenous to the household and uncorrelated with underlying patient risk or procedure severity, as discussed in Section 2. Bill wait times are highly seasonal within a year for a given payer, depending on both the total volume of claims they are processing (Appendix Figure A.8) and administrative frictions of initiating new enrollees and groups to new benefits. In addition, factors such as new risk-adjustment policies, transitions in billing systems (such as the 2014-2015 shift to ICD-10-CM), or even the COVID-19 pandemic can overwhelm payer processing of claims, occasionally even drastically increasing the wait time before patients learn with certainty their ultimate OOP costs (Snowbeck, 2022).

Still, bill wait times may be associated with underlying patient risk, potentially introducing selection concerns affecting causal inference. If payers have an incentive to slow down payments for highly-expensive procedures, or if a higher-complexity patient takes longer to process, bill times could be systematically longer for the most at-risk patients in our sample. We test these claims directly in our sample, by comparing differences in the average total cost of shoppable services based on the length of the bill wait time in days.

Table 2 presents the results. For each potential service, the hypothesis that bills that took longer to arrive ($d \geq 30$ days) are associated with more (or less) expensive procedures is tested. We test these hypotheses for both unadjusted means and averages adjusted for provider-specific trends. We find that the large volume of procedures in each group lends itself to statistically significant differences, but not economically meaningful ones. The average difference across services constitutes only $220, 6.1% of total payments. In addition, the estimated value of the differences varies widely, with almost a quarter of the included procedures estimated to have *shorter* wait times for more expensive instances of the proce-

fixed-effects (Correia et al., 2020).

	Average Spending		**Difference**			**Sample Size**	
Procedure	$d \leq 30$	$d > 30$	Unadjusted	Adjusted	p-value	$d \leq 30$	$d > 30$
Removal, prostate	$21,834	$25,362	$3,528	$1,260	0.41	917	403
Removal, knee cartilage	$7,619	$8,021	$402	$697	0.00	46,937	15,606
Removal, breast growth	$4,887	$5,173	$286	$674	0.00	10,550	3,916
Injection, anesthetic	$3,258	$3,537	$279	$484	0.00	49,604	16,667
Biopsy, esophagus/stomach	$3,317	$3,238	-$79	$406	0.00	245,411	65,603
Removal, tonsils (age < 12)	$4,578	$4,871	$293	$342	0.00	21,503	4,962
Shaving, shoulder bone	$12,262	$12,040	-$222	$233	0.07	27,952	11,410
Biopsy, prostate	$2,653	$2,377	-$276	$124	0.01	23,172	6397
Removal, gallbladder	$9,217	$9,794	$577	$96	0.38	36,756	13,252
Hernia repair	$6,753	$6,724	-$28	$28	0.83	14,314	5,215
Removal, cataract (no insertion)	$1,408	$1,198	-$210	-$179	0.05	11,776	2,388
Vaginal delivery	$7,789	$7,927	$139	-$344	0.00	82,968	36,068
Removal, cataract (lens insertion)	$6,114	$5,958	-$156	-$346	0.00	43,129	9,266
Vaginal delivery, prior C-section	$8,429	$8,634	$205	-$912	0.01	1298	503

Notes: Table shows differences in means for total spending (patient + insurer payments) by procedure category for the shoppable services included in our analytical sample. Services are divided into groups if (1) the bill took 30 days or fewer to arrive following the service ($d \leq 30$) or (2) the bill took more than 30 days to arrive ($d > 30$). Differences in means are presented both in raw, unadjusted terms, as well as adjusted for provider-specific fixed-effects. p-values are the results of two-way difference in means testing on the adjusted differences.

Table 2. Bill Balance Table (Unadjusted and Adjusted for Provider Fixed Effects)

dure. Taken together, we find little evidence that bill wait times may be endogenous at the household level.

4 Empirical Results

4.1 Effect of Bills on Spending

Table 3 presents the regression results from estimating Equation 1. We find robust evidence that although spillover household spending increases following the utilization of a shoppable service, the bills arrival causally affects these responses. Without conditioning on the bill's arrival, the overall estimated spending increase is roughly 71.5% of average weekly per-person household spending, an increase of roughly \$85 per person-week. However, this conflates a period prior to the plan's payment of the claim where spending is estimated to increase by 60%, only to decline by 8.5% once the bill arrives. This decline (roughly 14% of the overall change in spending) is consistently estimated across our specifications. The correction amounts to approximately \$10 per person per week for the average household in the sample; based on the timing of bills in our sample, this amounts to roughly \$244 in per-person annual spending.

These findings are consistent with a model in which consumers over-estimate their actual OOP contributions prior to receiving the definitive pricing information of a medical bill. Due to this over-estimation, consumers enrolled in insurance plans with piecewise-linear cost-sharing insurance plan designs (e.g., a nonzero deductible) may incorrectly assume that the marginal cost associated with additional services has declined discontinuously (e.g., by meeting a deductible). Once the bill arrives correcting any errors in perception, however, individuals curtail their spending increases in response.

Given that we are using the dates insurance plans pay providers as a proxy for bill arrival, it is possible that we are artificially splitting the post-service period into two essentially random periods and attaching significance to a spurious difference between the two periods. To test this possibility, we conducted placebo tests, running regressions on artificial data that randomly assigned consumers new wait times for their bill based on the empirical distribution of wait times.[22] The results of 1,000 such placebo simulations are reported in Appendix Figure A.9; placebo regression coefficients are centered around zero and generally statistically indistinguishable from a null effect. Taken with the results from Table 3, this suggests that it is unlikely that our results are spurious correlations from a convenient semi-

[22]For each shoppable service, we fixed the service date and artificially varied the bill arrival date as the service date plus a random draw of a wait time, drawn from the empirical distribution of waits (Figure 1).

	Main Models		Alternative Specifications		
Post Service	0.715***	0.599***	0.754***	0.626***	0.625***
	(0.0026)	(0.0034)	(0.0033)	(0.0034)	(0.0034)
Post Bill		-0.085***	-0.075***	-0.082***	-0.084***
		(0.0030)	(0.0033)	(0.0032)	(0.0033)
Weeks Prior to Service	0.012***	0.014***	0.019***	0.014***	0.015***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
Weeks Following Service	-0.002***	-0.002***	-0.001***	-0.001***	-0.002***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
$\overline{\text{spend}}_{it}$	\$120.49	\$120.49	\$120.49	\$120.49	\$120.49
Household FEs	X	X	X	X	X
Year FEs	X	X		X	X
Week of Year FEs	X	X			X
Provider FEs	X	X			
Observations	61,860,735	61,860,735	61,860,735	61,860,735	61,860,735

Notes: Table presents results from triple-difference Poisson regressions highlighting the role of a bill's arrival on health spending of affected household members. Each column in this table estimates the impact of a single household member's shoppable health service—and accompanying bill—on health spending for all other household members. Regression coefficients displayed illustrate the expected change in log household spending (measured per person-week) associated with the service date and bill arrival (both measured as dummy variables). Throughout, standard errors were clustered at the household level.
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Table 3. Estimated Impact of Bill Arrival on Household Health Spending

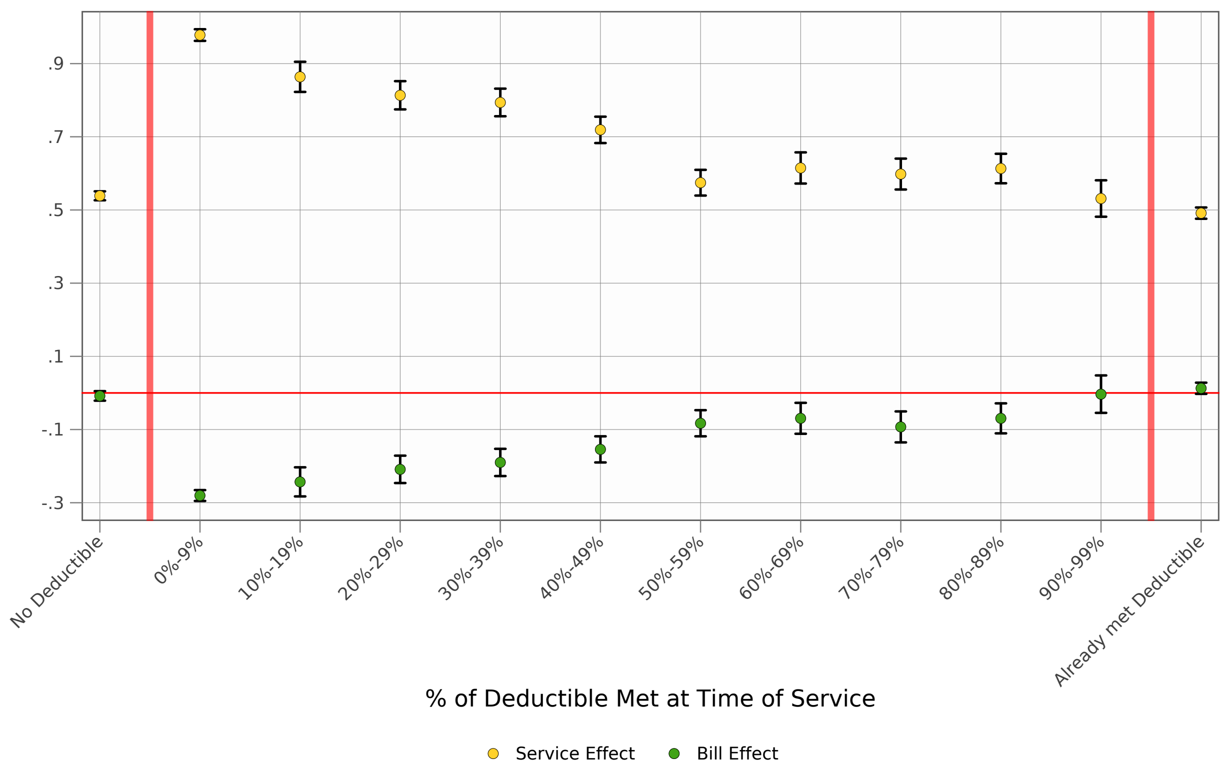
random splitting of the post-service period.

4.2 Heterogeneity by Deductible Spending

There are several ways households might find bills informative enough to alter their health-care utilization patterns. First, households may learn about the overall prices of the services they received, particularly their own OOP burden for their care. This may include more detailed information about the percentage of a household deductible that has now been met as a result of the service. In particular, a bill that informs consumers that they are still short of meeting a deductible may generate the corrective action observed in Section 4.1. Second, households may learn about the extent to which their insurance does or does not cover certain procedures. In this sense, bills inform households not about the overall prices of services, but correct a misunderstanding of the fraction of services they will have to cover OOP. Finally, a bill may reveal discrepancies between a patient's understanding of a service and the provider's billing, including up-coding practices. This information may also alter future healthcare spending to the extent that it erodes household trust in the healthcare system (Webb Hooper et al., 2019).

In order to understand the mechanisms behind household responses to bills, we assessed how responses differed based on household plan structure and spending histories prior to the shoppable service. Here, the intuition is that variation in household pre-event spending provides useful variation in the relevance of the bill (e.g., if the bill provides information about marginal costs, such as deductible spending) as well as variation in the amount of engagement with the health system in a particular year.

Figure 2. Heterogeneous Bill Effects Across Household Deductible Status at Time of Service



Notes: Figure shows estimated coefficients and 95% confidence intervals for $\mathbb{1}\{\text{Post\_Service}_{it}\}$ and $\mathbb{1}\{\text{Post\_Bill}_{it}\}$ in Equation 1 by decile of household deductible spending prior to the event. Standard errors are clustered at the household level. Deductibles are imputed based on previous literature (Hoagland, 2022; Zhang et al., 2018).

Figure 2 presents results stratified by decile of household deductible spending prior to the event. The spending responses for both the interim period between the service and the bill (yellow) and the post-bill correction (green) are shown. Both spending responses and corrections are largest for households who have spent little towards their deductible before the event: post-service spending increases are estimated to be over 90% for households with less than 10% of their deductible met, and fall to just under 50% for households close to their deductible. Correspondingly, post-bill corrections are estimated to be as high as 30% (roughly one-third of the post-spending increase) for the low-spending group, and converging to zero for the high-spending group. In both cases, spending responses for the group closest to meeting their deductible are statistically indistinguishable from households who met their deductible prior to the shoppable service. Finally, households who don't face changes to their marginal cost of care from the bill (e.g., those in zero-deductible households and those which have already met their deductible) exhibit no spending responses to the bill.

Taken together, these results suggest that households respond, at least in part, to a bill's information about OOP expenditures. This information appears to be especially relevant to households who have yet to contribute much to their deductible, while households without a deductible or who have already met it exhibit negligible responses to the bill.

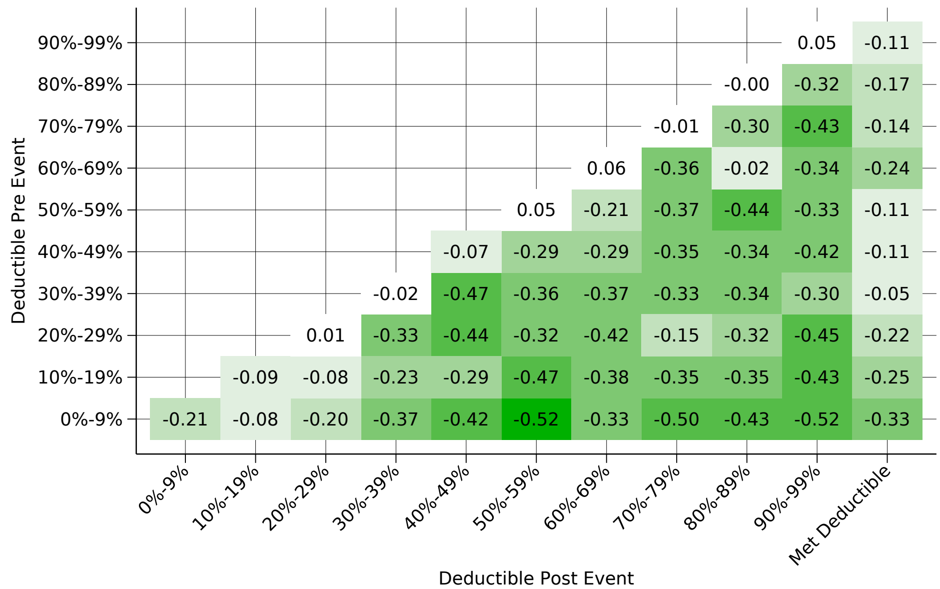
While variation in pre-event spending provides useful information, further insight can be gained by leveraging a second dimension of variation: the relative cost of the shoppable service itself. The intuition for this exercise is that high-cost events are more likely to ultimately alter a household's marginal cost of services, while lower-cost events may have similar price uncertainty without any economically meaningful costs to that uncertainty. Given that price information is most valuable to households when it communicates whether or not households have crossed the threshold of their deductible, we identify the extent of household learning about *prices*—separate from other forms of learning—by comparing household responses to high- and low-cost events. To do this, we explore two-way heterogeneity in the effects of household responses considering both pre-event deductible contributions *and* the resulting change in deductible spending after the scheduled health consumption.

Figure 3 presents the results. We restrict our attention to households enrolled in plans with a non-zero, unmet deductible at the time of service. We then separately estimate Equation 1 across cells of households who have similar deductible spending both before and immediately following the shoppable service.[23] The figure depicts a two-way heatmap of estimated bill responses across cells. Consistent with Figure 2, we find that households starting at lower levels of their deductible exhibit greater sensitivity to their bill. In addition, we find that households appear considerably less responsive to low-cost services; coefficients are estimated to be much closer to zero when households do not move across deciles of spending, and weakly increasing as the OOP costs of services become more expensive.

Comparing household responses across the discontinuous threshold of meeting the de-

[23]For each regression, the control group is households who did not consume a shoppable service over the course of the year.

Figure 3. Heterogeneous Bill Effects By Household Deductibles and Service Cost



Notes: Figure depicts estimated coefficients for $\mathbb{1}\{\text{Post\_Bill}_{it}\}$ in Equation 1 across deciles of household deductible spending prior to *and* following an event. Here, sample is restricted to individuals in a non-zero deductible plan who have not yet met their deductible at the time of service. Each row indicates a different decile of deductible spending prior to the event, while each column indicates deciles following the event. Standard errors are clustered at the household level. Deductibles are imputed based on previous literature (Hoagland, 2022; Zhang et al., 2018).

ductible in Figure 3 is particularly informative about mechanisms. Across all levels of pre-event spending, estimated coefficients are at least 41% higher when a bill left households just short of meeting their deductible, rather than services that pushed households into a lower marginal-cost region of their contract. The average effect of receiving a bill for the shoppable service declines by 45% from 0.39 to 0.18. Note that coefficients are still highly significant even after crossing the deductible threshold for two reasons: first, the extremely large size of our data lends itself to an increased likelihood of statistical significance even for economically insignificant changes; and second, large residual cost-sharing for a shoppable service may persist even after a deductible is met (e.g., large coinsurance charges, out-of-network expenses), which may still change household behavior. Overall, these findings illustrate that households are much more responsive to a bill when it contains important information about future marginal costs, consistent with price sensitivity driving household responses.

Taken together, accounting for heterogeneity in household responsiveness to bills—including both spending histories and the cost of services—suggests that households are most responsive to bills when their pricing information is particularly salient. When households have little information about their deductible, or when the bill indicates that a household's marginal cost of additional care has not changed, the arrival of pricing information curbs overly large responses in spillover spending following a shoppable health service.

4.3 What Services Are Affected?

4.3.1 Do Bills Only Affect Strategic Delaying of Services?

Our results suggest strong evidence that households respond to price information contained in a bill by contracting their total consumption of medical care after the bill arrives. A remaining question is whether these contractions merely represent a strategic delay in the use of medical care or a more fundamental change in the quantity and type of medical care households seek. For example, households may strategically delay the use of some services from the end of a plan year with a higher effective end-of-year deductible to one with a lower expected end-of-year-deductible.

While there is strong evidence for such dynamic moral hazard concerns, we find that these dynamic effects are insufficient to explain our results. That is, find evidence that households alter the level and type of care they choose even among services for which strategic delays are infeasible. To see this, we study how household responses to the same shoppable services discussed above affect household spending on mild acute respiratory infections, for which care cannot be delayed far into the future (Hwee et al., 2018). We investigate how a bill's arrival affects overall spending on these infections, as well as stratify our results by the place of service to investigate where households seek care (see Appendix Table A.8 for a list of relevant diagnosis and place of service codes).

Table 4 presents the results. Overall, we continue to find that a bill's arrival significantly alters total household spending even when limiting attention to only respiratory infections. We find that households reduce their spending on respiratory infections by 35% (roughly $2.81 in the unconditional average) after a bill arrives. Given that this care cannot be strategically delayed to a new plan year, this contraction in spending must be a change in household decisions about the level of care to seek for a stochastically-realized infection. In contrast to the overall spending results in Table 3, this correction almost entirely eclipses the post-service increase in spending; this further supports the hypothesis that bills affect extensive margin decisions about whether or not to seek care for mild respiratory infections. That is, the results are consistent with a model where households have a lower threshold at which they seek care for an infection when they believe (in the absence of information) that marginal prices of doing so are lower than they really will be.

Importantly, our results represent both an overall level change in spending on infections (representing an extensive margin effect) and a change in *where* households go for that care.

	Coefficient (SE)	% of Service Effect	Conditional Mean
Total Spending			
Total Bill Effect	-0.085*** (0.0030)	14.2%	$378
Respiratory Infection Spending			
Total Bill Effect	-0.353*** (0.0135)	93.4%	$190
Physician Office	0.050*** (0.0070)	38.8%	$139
Urgent Care	0.081* (0.0397)	36.4%	$184
Emergency Department	-0.044 (0.0375)	102.8%	$815
Hospital Campus (incl. outpatient)	-0.639*** (0.0220)	82.9%	$1,001

Notes: Table presents results from triple-difference Poisson regressions ($N = 61,860,735$). Only the regression coefficient on *post_bill* is shown; we also report the coefficient as a percentage of the *post_service* coefficient and as the approximate change in weekly spending at the person level. Respiratory infections and place of service were identified using the methodology of Hwee et al. (2018) (see Appendix Table A.3). All models include fixed effects for households, years, relative week of year, and providers; standard errors were clustered at the household level.
$^{*}p < 0.05, ^{**}p < 0.01, ^{***}p < 0.001$

Table 4. Bill Effects on Care for Respiratory Infections

Households respond to bills by dramatically decreasing the rate at which they seek care for infections in a hospital setting (including both inpatient care and on-campus outpatient clinics). In fact, our estimates suggest that responses along this dimension entirely explain the overall 35% drop in expenditures following a bill's arrival. In contrast, households actually increase spending on respiratory infections at physician's (non-hospital) offices and urgent care clinics. Finally, we see little change in the rate at which consumers seek care at the emergency department; we might expect this to the extent that such visits are the most inelastic form of health consumption (see Section 4.3.2).

Taken together, our results suggest that households are not simply responding to bills by rearranging the date at which they seek care. Reductions in total expenditures persist even among services which cannot be strategically delayed; these reductions appear to be driven by households being more selective about when to seek hospital-based care for infections, and increasing substitution towards cheaper points of service, including physicians' offices and urgent care centers. Overall, price information changes both the level and type of healthcare that households seek out.

4.3.2 Heterogeneous Effects by Type of Care

Finally, we assess whether household responsiveness to price information varies across broad categories of medical services, including hospital care, specific types of outpatient services, and pharmaceutical spending. This decomposition allows us to examine whether household spending responses—or the extent to responses are corrected after a bill arrives—varies with any measure of perceived or real quality of care. Particularly, we examine how bills affect future utilization of typically high-value health services (e.g., preventive screenings, behavioral health services) as well as typically low-value care (e.g., unnecessary pre-operative screenings, imaging services, or surgeries; see Table A.10).

Table 5 presents the results. We separately estimated the coefficients for Equation 1 with each sub-category of spending as its own dependent variable.[24] Overall, we find that households respond to shoppable services across a spectrum of services, increasing their consumption in the short-run after a service is performed, and then reducing that consumption significantly once the bill arrives.

After an individual household member receives a significant health services, other household members are more likely to seek hospital care, including a 10.8% increase in emergency department visits and a 37.1% increase in visits for potentially preventable hospitalizations (e.g., admissions to treat dehydration) (Agency for Healthcare Research and Quality, 2007). Following the receipt of the bill, however, use of inpatient care for preventable hospitalizations falls by roughly 52.3% of the increase (a 19.4% decrease). These large swings can be explained as the result of perceived changes in the cost of accessing care, which may have particularly large effects for more expensive services (e.g., hospitalizations after a deductible is perceived to be met).[25]

Some outpatient services, including those for behavioral health (e.g., psychotherapy) and chiropractic care (e.g., physical therapy), are affected neither by the consumption of a shoppable service nor its accompanying bill; this is presumably because these services have more inelastic demand and lower rates of cost-sharing generally. However, we find that when households increase demand for a type of outpatient service following a major health event, they tend to over-increase spending. Bill arrivals cause households to correct spending increases by between 32% and 90%. Households increase their utilization of general practice visits (e.g., E&M visits, lab work, and preventive screenings) the most, followed by specialist visits (e.g., dermatology). While household demand for prescription drugs increases slightly following a health event in the home (by 1.8%), we do not observe a corresponding reduction in demand following the bill's arrival. This could be because of the already high levels of pharmaceutical spending relative to other medical consumption.

Somewhat surprisingly, we do not observe that households reduce their utilization of low-

[24] Appendix Table A.9 includes detailed descriptions of the construction of each of these variables.

[25] Whether this increase is an over-utilization of unnecessary care or simply increased access to relevant hospital services—particularly considering the "layperson standard" for hospital care—is an open question which warrants future research (Siegfried et al., 2019).

	Regression Coefficients		Pre-Treatment Averages	
	Post Service	Post Bill	% ≥ 0	Conditional Mean
Hospital Care				
Emergency Department	0.108***	-0.021	0.67%	$929.98
	(0.0129)	(0.0133)		
Preventable Hospitalizations	0.371***	-0.194*	0.04%	$19,979.89
	(0.0829)	(0.0848)		
Outpatient Care				
Behavioral Health	-0.020	0.018	1.19%	$119.47
	(0.0132)	(0.0134)		
Chiropractic Care	-0.005	0.027	1.86%	$133.39
	(0.0147)	(0.0151)		
Evaluation & Management	1.440***	-0.518***	1.05%	$121.45
	(0.0066)	(0.0062)		
Imaging	0.098***	-0.037***	2.55%	$265.52
	(0.0108)	(0.0111)		
Lab Services	0.198***	-0.178***	3.96%	$62.14
	(0.0113)	(0.0119)		
Low-Value Services	0.084***	0.028**	6.58%	$148.61
	(0.0094)	(0.0097)		
Preventive Care	0.345***	-0.249***	11.68%	$120.47
	(0.0036)	(0.0037)		
Specialist Care	0.550***	-0.181***	0.57%	$114.70
	(0.0192)	(0.0198)		
Prescriptions	0.018***	-0.006	18.30%	$147.14
	(0.0047)	(0.0048)		

Notes: Table shows coefficients from triple-difference regressions capturing service-specific effects of pricing information ($N = 59,177,995$). Columns (1) and (2) present regression coefficients; column (3) indicates the fraction of pre-treatment weeks when spending was positive; and column (4) presents pre-treatment weekly averages, conditional on positive spending. See Appendix Table A.2 for a complete list of the CPT codes for each of the outpatient categories. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 5. Estimated Impact of Bill Arrival on Service-Specific Spending

value care following a bill's arrival. These services, which include services such as imaging for lower-back pain, misuse of prescription medications to manage migraines and bacterial infections, or unnecessary pre-operative screenings, are determined based on the recommendations of the Choosing Wisely campaign (Colla et al., 2015). We find that households increase their use of low-value care by 8.4% following a major service, and then further by another 2.8% once the bill arrives. This may be a result of a "cascade of care" effect associated with increased consumption of general medical care, which in turn prompts downstream increases in physician ordering of low-value services (Ganguli et al., 2020). Physicians typically retain control over when low-value services are performed, in order to reduce their own uncertainty, liability, or "just to be safe" (Colla and Mainor, 2017).[26]

5 Model

Based on the empirical findings from our reduced-form analysis, we propose a model of imperfect moral hazard, in which consumers make medical care choices based on beliefs about realized spending. Central to the model is the delayed nature of pricing information, which may lag behind consumption by weeks or months while still affecting the spot prices of care in ways that are unknown to the consumer before the bill arrives. As a result, consumers must form expectations about realized OOP spending and the implied marginal cost of care in each period. We first consider a case where consumer beliefs are static (e.g., where there is no learning) before introducing a learning component to the model in Section 5.2.

In each period t, an individual i receives a health shock λ_{it}, which represents a combination of both acute fluctuations in health status and persistent health needs. Patients then choose an appropriate level of medical spending m_{it}—measured in the dollar value of the services—in response to λ_{it}, spending histories, and individual preferences.[27] Following Einav et al. (2013), we calibrate individual patient utility as a quadratic loss function in the distance between selected health spending and the unobserved health shock:

$$u_{it} = (m_{it} - \lambda_{it}) - \frac{1}{2\omega_i}(m_{it} - \lambda_{it})^2 - c_{ijt}(m_{it}; M_{\mathcal{I}t}). \tag{2}$$

[26] Our findings are consistent with prior work, but warrants future exploration as to whether it is physician- or patient-driven (Hoagland, 2022). In particular, increases in low-value service take-up could be initiated at a (more elastic) general practitioner visit and simply continue into the future despite the bill's arrival.

[27] We make the simplifying assumption that shocks can be measured in dollars to make comparable health production and OOP spending, consistent with previous versions of this model (Einav et al., 2013). The parameterization is useful because it is both tractable and incorporates rational responses to nonlinear pricing schemes; for example, individuals close to a deductible will choose to slightly increase their consumption, anticipating the approaching nonlinear change in marginal costs Marone and Sabety (2022). To be consistent with Section 4, we model spending choices at the weekly level and normalize spending by household size.

Here, ω_i is an individual time-invariant "moral hazard" parameter capturing individual heterogeneity in demand responsiveness to the price of services.[28] In addition, $c_{ijt}(m_{it}; M_{\mathcal{I}t})$ denotes the OOP costs associated with m_{it}, which depends on the piecewise-linear cost-sharing contract of individual i's chosen insurance plan, j, as well as the OOP spending to date at the household level, $M_{\mathcal{I}t} = \sum_{i\in\mathcal{I}} \sum_{s=1}^{t-1} m_{is}$. Note that an individual's OOP costs for services are weakly decreasing in $M_{\mathcal{I},t}$.

Under full information, patients know both the value of $M_{\mathcal{I}_t}$ and how it affects $c_{ij}(\cdot)$. Furthermore, in the case where cost-sharing is linear at all stages of the contract, a patient's marginal OOP cost is given by $c_{ijt} \in [0, 1]$, where $c = 1$ applies to all services before a deductible has been met and $c = 0$ applies for all services after an OOP-max has been met. Between the deductible and the OOP-max, c is typically in the open interval $(0, 1)$. With full information about prices, the static choice of m_{it} in each period is simply the solution to the first order condition:

$$1 - \frac{1}{\omega_i}(m_{it} - \lambda_{it}) - c_{ijt} = 0 \Rightarrow m_{it}^* = \max\left[0, \lambda_{it} + \omega_i(1 - c_{ijt})\right]. \tag{3}$$

That is, medical expenses in each period are chosen so that the marginal utility of those services is equal to the marginal (known) OOP cost. In particular, as c changes from 1 to $c < 1$ as households meet their deductible, household members will have a discontinuous increase in their medical consumption in future periods.

However, based on the discussion in Section 4, we suppose that $M_{\mathcal{I},t}$ is not known with certainty at the time a service is performed. Rather, household spending can be divided into two components: spending for services whose bills have already arrived (e.g., where prices are known), and spending for services without pricing information yet available. For ease of notation, suppose that each bill takes τ weeks to arrive, so that a bill for a service procured in week t would arrive in week $t + \tau$.[29] Based on these components, households respond to a signal of their spending θ:

$$\theta_{it} = \underbrace{\sum_{s=0}^{t-\tau} \sum_{i\in\mathcal{I}} c_{ij}(m_{is})}_{\text{known spending}} + \underbrace{\sum_{s=t-\tau+1}^{t} \sum_{i\in\mathcal{I}} s_i(m_{is}|x_{is})}_{\tilde{\theta}_{it}=\text{unknown spending}}, \tag{4}$$

where $s_i(m_{is}|x_{is})$ represents service-specific signals of spending, which may depend on individual, household, and service level characteristics.

[28] The individual-specific moral hazard parameter ω_i has a helpful interpretation as the incremental spending induced by a move from no insurance to full insurance (Einav et al., 2013).

[29] Note that in the empirical estimation of the model, the length of time between a service and bill's arrival is allowed to vary across services; this assumption is only made in this section for ease of exposition.

Hence, the timing of the model in each period t is as follows:

1. Individuals form expectations about their spending histories M_{it}, based on θ_{it}.
2. Individual health shocks λ_{it} are realized.
3. Spending decisions m^*_{it} are made based on realized health shocks and perceived spending histories, which govern the perceived marginal cost of additional units of care $\hat{c}_{it}$.
4. A new signal of spending $s_{it}(m^*_{it})$ is received for the individual and all household members enrolled in the same plan. Household members update their expectations of M_{it} based on the signal and any learning. Then we proceed to period $t+1$.

5.1 Simple Case: Constant Under-information

In the simplest case, we suppose that signals do not vary across services, but rather assume that cost signals are a constant multiple of true costs:

$$s_i(m_{is}|x_{is}) = \beta \cdot c_{ijs}(m_{is}). \tag{5}$$

That is, before a bill arrives, patients inflate (or deflate) their true OOP spending by a constant parameter β, which does not vary across services or individuals.[30] In the simplest version of the model, we also assume that there is no learning about β over time; we introduce learning in Section 5.2. Based on these assumptions, a household's signal of their OOP spending (and hence, of their marginal cost for additional spending) can be simplified as

$$\theta_{it} = \sum_{i \in \mathcal{I}} \sum_{s=0}^{t} (1 - D_{is})\beta c_{ijs}(m_{is}) + D_{is} c_{ijs}(m_{is}), \tag{6}$$

where D_{is} is a binary variable indicating if the bill for services performed in week s has arrived ($D_{is} = 1$) or not ($D_{is} = 0$). Based on the household's value of θ_{it} in each period, the signal for the marginal cost of future expenditures in the simple piecewise-linear insurance contract setting is given by

$$\hat{c}_{it} = \begin{cases} 1 & \theta_{it} < d \\ c & \theta_{it} \geq d, \end{cases} \tag{7}$$

[30] Note that allowing β to be a random coefficient varying across individuals is a simple extension of the model; for the present purposes, however, we focus on an average of β across the population of interest.

where $c < 1$ in general.[31]

In this simplified case, the central parameter of interest is β, the rate at which households systematically over- or under-inflate their true levels of OOP spending prior to the arrival of the pricing information contained in a bill. Additional unobservable parameters in the model, which threaten identification, include heterogeneity in moral hazard ω_i and individual health shocks λ_{it}. Separate identification of β relies on being able to credibly identify the hyper-parameters governing the distributions of these characteristics.

When estimating the model, we calibrate these nuisance hyper-parameters to match moments predicted by (a) previous research and (b) training data not used in the structural estimation. We use the estimated regression coefficients predicted by Einav et al. (2013) in order to capture variation in moral hazard parameters across households.[32] We model individual-level health shocks as draws from an individual-specific shifted lognormal distribution; this distribution captures both the skewed nature of the observed spending data and the nonzero probability of an individual choosing zero spending in a period. That is, each individual in each period draws λ_{it} from a distribution $F(\mu_i, \sigma_i, \kappa_i)$ such that

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

We calibrate the three hyper-parameters $(\mu_i, \sigma_i, \kappa_i)$ to match the empirical distribution of weekly spending using the individuals in our analytical sample who are *not* included in the structural estimation. These include individuals enrolled in plans with no deductible, as well as patients enrolled in any type of plan between 2014 to 2018. Individuals in this sample are grouped into cells based on patient demographics—including age, sex, risk score quartile, and relationship to the main employee—and the empirical distribution in each cell is matched to the shifted lognormal moments.[33] Once these parameters are identified, individual-period shocks are drawn for each member of the model sample and then summed to the household-period level.[34]

Given these calibrations, identification of the main parameter of interest β comes centrally from exogenous variation (at the household level) in the length of time required for a bill to arrive after different health services. This variation may exist across services as well as across

[31]Note that in practice, we estimate the model on the sample of individuals enrolled in plans with non-zero deductibles. This is to cleanly capture the ways in which misperception of OOP spending may affect discontinuous changes in the marginal cost of spending across thresholds of the linear insurance contract.

[32]Note that these regression models result in individual-level predictions for ω_i; in estimation, we aggregate these to the household level by taking the mean of $\log(\omega_i)$ across all members $i \in \mathcal{I}$.

[33]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 $\frac{\mathrm{sd}(\lambda)}{\overline{\lambda}} = \sqrt{\exp(\sigma^2) - 1}$, 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.

[34]In order for shocks to be meaningful, we restrict $\lambda_{\mathcal{I}t} < m_{\mathcal{I}t}$ in each period.

households; importantly, underlying variation in $\tilde{\theta}_{it}$ which artificially moves households above or below their deductible is central to identifying how β changes household estimates of $\hat{c}$ in ways that most closely fit the observed choice data.

5.2 Learning

Once beliefs about OOP costs can be reasonably calibrated in the model, a natural question is whether consumers correct their beliefs with repeated exposure to health information. Households with frequent interactions with the health system, particularly within a plan year, may have beliefs about their bills which converge to the truth over time.

To assess this question, we incorporate household learning about the calibration parameter β. We model each bill's arrival as a signal from which consumers can learn.[35] Households are assumed to have prior beliefs about β which follow a normal distribution with a mean $\mu_{\beta,0}$ and variance $\sigma^2_{\beta,0}$:

$$\hat{\beta}_{i0} \sim \mathcal{N}(\mu_{\beta,0}, \sigma^2_{\beta,0}). \tag{9}$$

When a bill arrives for a household conveying information about the prices of medical services, it in essence communicates that $\beta = 1$. Hence, we model each signal s_{it} as being drawn from a normal distribution centered at 1 and with a signal variance σ^2_s:

$$s_{it} \sim \mathcal{N}(1, \sigma^2_s). \tag{10}$$

We assume that households update their prior beliefs in accordance with Bayes' Rule, conditional on their observed signal. Assuming normal distributions for both the prior and posterior distributions allows for closed-form solutions for household beliefs about β at each period, and is consistent with previous learning models (Crawford and Shum, 2005). This expanded version of the model therefore has three parameters of interest. First, the average prior mean $\mu_{\beta,0}$ captures the extent to which households are uninformed about the relative costs of their medical care at the start of an enrollment period. The dispersion of this lack of information across households is captured in the variance of the average prior, $\sigma^2_{\beta,0}$. Finally, the variance of the signal, σ^2_s, reflects how precise the information communicated by each medical bill is, and subsequently how rapidly household beliefs converge to the true parameter of 1.

Estimating household learning allows deeper insight into the spread of household beliefs about their expenses both across households in the sample and over the relative course of

[35] For now, we model each signal as having equal impact; future extensions of this model could flexibly model heterogeneous signals based on the total cost of a service or by different service types.

an enrollment period. In particular, the speed with which beliefs converge informs the rate of over-consumption of medical care relative to fully-informed households. As in the simpler case of the model, identification of the three learning parameters $(\mu_{\beta,0}, \sigma^2_{\beta,0}, \sigma^2_s)$ stems from exogenous variation in bill timing. When the parameter space is expanded, identification comes from various sources. Within-household variation in expenditures relative to pending (hence, unknown) OOP expenditures serves to identify both the starting point of household beliefs (the prior mean) and the rate of convergence (governed, in this case, by the signal variance). Similarly, variation in choices across households identifies the spread of beliefs, summarized in the prior variance; this parameter informs both the spread of households' starting beliefs as well as how that spread evolves over time.

6 Model Results

6.1 The Case of Constant Over-/Under-Estimation

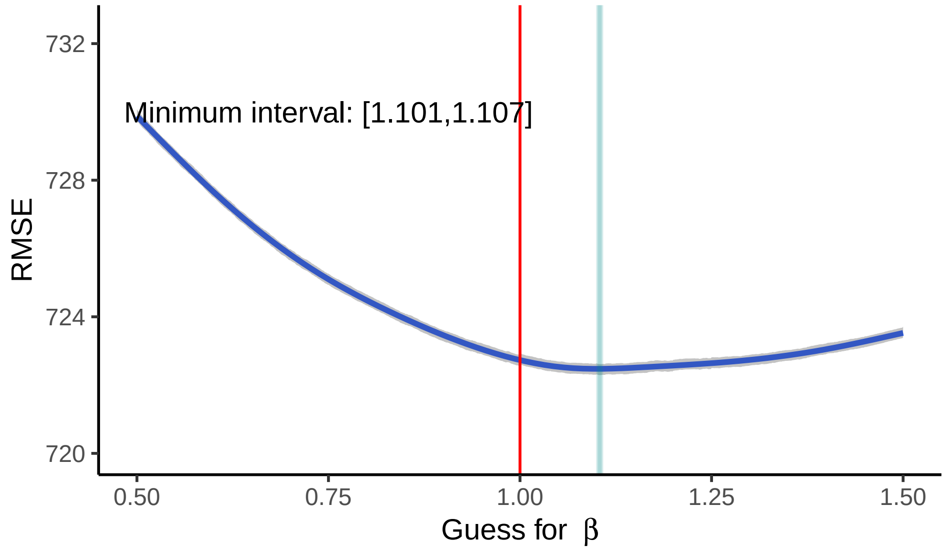
We estimate the model presented in Section 5 for 240,111 households in our analytical sample enrolled in plans with nonzero deductibles from 2006 to 2013. For each household-week, we simulate household health shocks and draw idiosyncratic moral hazard parameters; then, for different values of β, we estimate household signals of underlying OOP spending and the marginal cost of incremental spending, $\hat{c}_{it}$. Taken together, these estimates produce a prediction of spending $m_{it}(\beta)$, which differs as β changes. Our primary measure of model fit is the root mean squared error (RMSE) between observed and predicted levels of weekly spending at the household level.

Figure 4 presents the estimated relationship between β and model fit, based on 50 simulations with different individual health shocks. The median RMSE for each value of β, as well as a confidence band of two standard deviations, are shown. Increasing the guess of β reduces the RMSE until the function reaches a minimum at $\beta = 1.104$ (in the median simulation), after which RMSE increases. The blue band in the figure shows the estimated 95% confidence region for the minimizing value of β, $[1.101, 1.107]$.

The model estimates, as illustrated, suggest that households over-estimate the OOP spending of services prior to the arrival of price information by between 10.1% and 10.7%. This is consistent with the findings of Section 4, which similarly illustrated a "correction" in implicit marginal costs following the arrival of the bill.

We conduct a simple counterfactual analysis to compare how spending predictions differ given this inflation, against a counterfactual world where β is correctly perceived to be 1 for all household-weeks (see Appendix Figure A.10 for details on the simulation). We find that

Figure 4. Estimating Household Responsiveness β to Spending Before Bills' Arrival



Notes: Figure depicts the relationship between chosen level of household pre-bill discounting parameter β and the mean squared error (MSE) of the model presented in Section 5. MSE is measured as the mean squared error between observed and predicted household spending at the weekly level. For each value of β, the median result of 50 simulations with independently drawn health shocks is shown in the black line; the confidence band illustrates one standard deviation above and below the median. The blue band denotes the full range of observed $\min_\beta MSE(\beta)$.

over-estimating the costs of medical services leads to over-spending for 11.9% of households, with the average (median) affected household spending \$856 (\$486) more per household member in medical services that they would not have selected had they been correctly informed of their true OOP costs. This corresponds to an over-spending of 34.8% (33.1%) for the average (median) affected household (see Section 6.3 below for more details).[36]

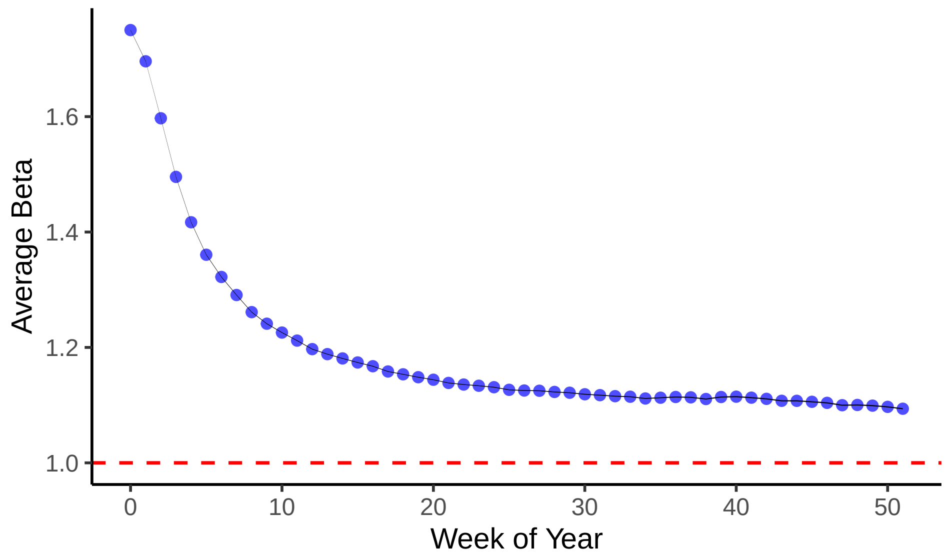
6.2 Learning

Finally, we incorporate the possibility of household learning into our estimates. We estimate that the median household's prior for β is roughly 1.75, indicating an 75% over-estimate of OOP costs (95% bootstrapped confidence interval: [1.702,1.798]). There is relatively little variation across households, as captured by the estimated prior variance parameter $\sigma^2_{\beta,0} = 0.011$ (95% CI: [0.002, 0.020]). Put into context, we estimate that roughly 95.5% of households (two standard deviations to either side of the mean) have prior beliefs that fall in the interval $(1.54, 1.96)$, indicating high levels of misinformation.[37]

Figure 5 shows how beliefs evolve in response to medical spending. We estimate that bills provide extremely precise information, with the signal variance term estimated to be

[36]Note that these percentage changes are measured relative to the counterfactual predicted spending; that

Figure 5. Evolution of Beliefs about β Across Plan Year



Notes: Figure depicts average value of simulated β across the relative week of a plan year for the full sample, with 95% confidence intervals shaded in black. Simulations are performed based on the median equilibrium parameters of the model discussed in Section 5.2.

0.0002 (95% CI: [-8e-5,.0005]). That is, roughly 95.5% of household signals for β fall in the interval $(0.97, 1.03)$. This leads to rapid convergence of beliefs as the year progresses, as illustrated in the figure in the blue curve, which indicates the average value of $\hat{\beta}$ across the sample by week of year. Within the first quarter of the year, average household inflation for OOP costs has converged to below 20%. Following this rapid convergence, however, belief convergence stalls—it isn't until week 33 that the average household's value of β crosses the upper bound of the 95% confidence interval for β estimated in the non-learning model (1.107), and average beliefs don't dip below 110% until week 45. The average household does not have sufficiently many medical encounters for their beliefs to converge completely; by the end of the plan-year, the average household value of β is estimated to be about 1.094, just outside of the confidence interval for the β in the non-learning model.

Figure A.11 in the Appendix presents results which further illustrate the heterogeneity in household beliefs across the year. The figure shows the fraction of households in the sample with simulated β greater than or equal to some threshold $\beta_{\min}$ for various thresholds. In general, extreme beliefs are rare after the first quarter of the year (fewer than 5% of households have an estimated $\beta \geq 1.5$ after week 10); however, many households stall in the convergence of their beliefs, with over a quarter of households estimating their OOP costs

is, as $(Actual - Counterfactual)/Counterfactual$.

[37]Note that given the estimated mean and standard deviation of prior beliefs, fewer than 5e-11% of households would be expected to have beliefs of $\beta \leq 1$.

as at least 25% more expensive than the truth for the entire plan-year.

6.3 Counterfactual Exercises

The full model with household learning permits the same counterfactual simulations as above. We estimate predicted spending differences between the observed data (using the equilibrium model parameters) and three counterfactual states of the world: one where consumers are fully informed about their OOP costs at the time of service (e.g., where $\beta = 1$ for all household-weeks); one where priors are re-centered around the truth (e.g., where $\mu_{\beta,0} = 1$, but the other learning parameters remain unchanged); and one where deductibles reset more frequently than at the yearly level. While the first counterfactual exercise assumes full information, the second exercise allows for idiosyncratic differences across signals, but centered around the truth. That is, in each household-week, beliefs are centered around $\beta = 1$ but drawn randomly, with decreasing uncertainty over time. Finally, the exercises in which deductibles reset more frequently capture changes in both how often household uncertainty affects the estimated marginal costs of services and the length of time for which uncertainty is allowed to persist.

Table 6 presents the results.[38] Accounting for greater heterogeneity in household beliefs about OOP spending results in a greater fraction of individuals being affected by changes to the learning parameters. This makes intuitive sense, given that households restricted to no learning in the simpler model may have been estimated to have reasonably correct beliefs for the full plan year, when in fact they experienced a period of rapid learning early in the year. Compared to a state where β is fixed at 1, roughly 21% of households over-consume care, with the average (median) affected household consuming $1,051 ($575) more per household member. This corresponds to an over-spending of 44.4% (40.8%) for the average (median) household relative to the predicted spending under full information. This over-spending can be nearly entirely attributed to high household priors, as can be seen in row 3 of Panel B of the table. Re-centering household priors—without completely eliminating residual uncertainty around prices for each unique medical event—accounts for more than 95% of the over-spending for both the average and median household.

6.3.1 Optimal Plan Design with Under-information

In addition to counterfactual exercises with different belief structures or learning mechanisms, we also conduct policy simulations exploring the role of plan design in potentially mitigating over-consumption arising from a lack of plan pricing information. This over-consumption

[38]See also Appendix Figure A.10 for a distribution of estimated spending differences.

	Spending/Person			Conditional Differences	
	Average	Median	% Diff > 0	Average	Median
Panel A: Non-Learning Model					
Observed Data	\$2,001	\$1,123	–	–	–
Full Information ($\beta = 1$)	\$1,899	\$1,066	11.9%	\$856	\$486
Deductible Resets Quarterly	\$3,499	\$2,040	12.7%	\$751	\$424
Deductible Resets Monthly	\$3,745	\$2,272	8.5%	\$657	\$361
Panel B: Learning Model					
Observed Data	\$2,119	\$1,181	–	–	–
Full Information ($\beta = 1$)	\$1,899	\$1,068	21.0%	\$1,051	\$575
Re-centered Priors ($\mu_{\beta,0} = 1$)	\$1,896	\$1,065	21.7%	\$1,031	\$562
Deductible Resets Quarterly	\$3,662	\$2,135	29.9%	\$863	\$477
Deductible Resets Monthly	\$3,851	\$2,335	22.7%	\$714	\$404

Notes: Table presents average and median spending per household member predicted by the models outlined in Section 5 under different assumptions of the underlying structural parameters. Panel A uses the model described in Section 5.1 without learning, and Panel B uses the model outlined in Section 5.2. The first row in each panel indicates predicted spending using the observed choice data and the estimated equilibrium parameters presented in Section 6. The subsequent rows impose arbitrary assumptions on the parameter space to capture salient features of counterfactual scenarios, including full information without learning (row 2), learning with re-centered priors (row 3, panel B only), and policies shortening the length of a deductible. Shortened deductibles are calculated at actuarially fair rates while holding premiums constant (in our case, the quarterly deductible is 30.75% of the annual deductible, while the monthly deductible is 10.25%)—see Hong and Mommaerts (2022) for a discussion. All currencies are reported in 2022 USD.

Table 6. Comparison of Predicted Spending Across Counterfactual States of the World

is most prevalent in the case where households incorrectly behave as though they have met their deductible (e.g., their perceived marginal cost of care is below 1) prior to a bill's arrival. Hence, we first evaluate the potential tradeoffs associated with policies which shorten the length of a deductible's effective period. The intuition for this exercise is shown in Appendix Figure A.12: limiting the length of time a deductible applies to future spending mechanically shortens any periods of with waiting for a bill. This is one reason why shorter deductible periods have been proposed by health policy experts recently (Shafer et al., 2022).

Shortening the length of a deductible trades off a reduced intensity of uncertainty—through both shorter periods of uncertainty and overall lower levels of a deductible—against the possibility of more frequent periods of uncertainty. That is, as deductibles reset more frequently (and are therefore lower), patients at the threshold of these lower deductibles may be induced to change their spending behaviors even while waiting for bills for low-cost medical procedures.[39] In addition, a lower threshold for changes in marginal costs may induce greater levels of *ex-post* moral hazard, particularly in short periods of time where medical care can be highly concentrated (e.g., after shoppable health services).

We model these tradeoffs in policy simulations which consider how patients and affected household members may change their consumption as deductibles are applied annually, quarterly, and monthly. Table 6 reports results for both the learning and non-learning models. We find that as deductibles reset more frequently, consumers are expected to nearly double their chosen levels of consumption: the average (median) household increases their consumption by 81% (92%) per-person in the non-learning model (Panel A) and 77% (89%) in the learning model (Panel B). Note that this arises principally due to moral hazard concerns concentrated in the relatively short lengths of time following a large health shock (either expected or scheduled); hence, we do not observe large differences in predicted spending increases as deductibles move from quarterly to monthly, or after incorporating learning into our model.

While increasing the frequency of deductibles increases overall levels of health spending, we observe limited changes in over-consumption arising from under-information regarding pricing. In the absence of learning, the rate of over-consumption drops to 8.5% (nearly a 30% decrease from the baseline of 11.9%), with the average (median) household's over-consumption dropping by 24% (26%). However, when household beliefs about prices improve over the plan-year, we do not observe a reduction in the frequency of over-spending arising from poor pricing information; instead, resetting deductibles more frequently may increase the likelihood of over-spending along the extensive margin (such as is suggested in Figure A.12). However, we do find evidence that this over-consumption is more limited in scope: when deductibles reset monthly, the average (median) over-consumption per person drops to \$714 (\$404), down about 32% (20%) from the baseline.

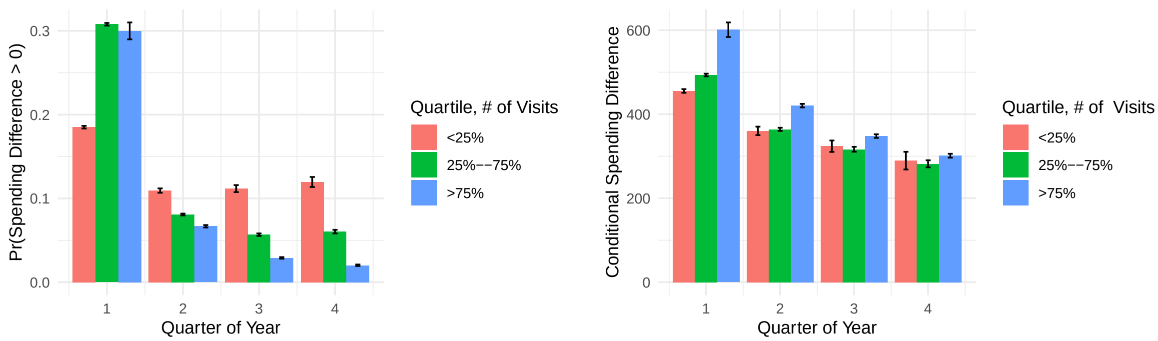
However, these results mask significant time heterogeneity across households, given that some learn more rapidly about β as they have more exposure to medical billing. The results presented in Table 6 conflate households with limited exposure to health information in the observed plan year (and hence, large degrees of under-information about prices) and other households with more refined knowledge. We report how our policy simulations affect these groups differently by stratifying households based on the total number of health encounters across the plan year.

Figure 6 presents the results. The figure illustrates how predicted over-spending varies

[39] When shortening the length of the deductible period by a factor $c < 1$, actuarially fair deductibles decrease by a factor $c' \in (c, 1)$; that is, deductibles do not decrease at a one-to-one replacement rate. We calculate new dedudcitbles at actuarially fair rates while holding premiums constant (Hong and Mommaerts, 2022). In our case, the quarterly deductible is 30.75% of the annual, while the monthly deductible is 10.25%.

Figure 6. Heterogeneous Effects of Resetting Deductibles Quarterly on Over-Spending

(a) Extensive Margin (b) Intensive Margin



Notes: Figure shows estimated the differences in predicted per-person health spending between two models where deductibles reset quarterly: the equilibrium parameters estimated in the learning model (Section 5.2), and a counterfactual model where $\beta = 1$ across all individuals and periods. Quarterly deductibles are 30.75% of the annual deductible, calculated at actuarially fair rates while holding premiums constant. Spending is stratified by the number of health encounters a household has had within a plan-year—those in the lowest quartile of number of medical visits are shown in red (group 1; 2 or fewer visits), while those in the highest quartile are in blue (group 3; 9 or more visits). Group 2 (blue) indicates the middle two quartiles of the distribution (between 3 and 8 visits). Panel (a) shows results for the probability of any over-spending, while panel (b) shows the conditional median level of over-spending. 95% confidence intervals are shown in error bars.

across the plan year when the deductible resets quarterly.[40] Predicted levels of over-spending are stratified based on how many health visits a household has had over the plan-year, indicating the extent to which the household has updated their beliefs about β: households in the lowest quartile of visits are shown in red (group 1; 2 or fewer visits), while those in the highest quartile are in blue (group 3; 9 or more visits). Group 2 (shown in blue) indicates the middle two quartiles of the distribution (between 3 and 8 visits).

Panel (a) shows the fraction of households with any over-spending predicted in the quarter. Given that over-spending in the model arises from households incorrectly assuming they have met their deductible, this panel is informative of the extent to which under-informed beliefs about β drive perceived changes in marginal costs. In the first quarter, when households have similar beliefs about β, households with more visits (and hence, higher levels of spending) are the most likely to over-spend, with the top 75% of the distribution over-spending roughly 30% of the time. However, as these groups proceed through the plan year, their increasingly informed beliefs about β reduce their rates of over-spending, down to as little as 2% in the fourth quarter. On the other hand, households who have few medical encounters over the year (and hence, do not update their beliefs even into the fourth quarter of the year)

[40]Here, over-spending is measured as the difference in predicted spending between the model using the equilibrium learning parameters and a counterfactual model where β is restricted to always be equal to 1

are more consistent in their likelihood of over-spending; over 10% of these households are still predicted to spend more when under-informed by the end of the plan-year, consistent with the results of the non-learning model.

Panel (b) presents the conditional median level of over-spending among affected households. When households are more under-informed about β in the first quarter, those with more visits (and hence, more uncertainty about their health spending) are estimated to spend more than those with fewer visits; in particular, households who have more than 8 visits in the first quarter of the year (the first blue bar) have levels of predicted over-spending as high as \$600. Given that this over-spending is independent of health shocks, this can be entirely explained by high levels of under-information about β, compounded by the number of visits with bills still pending. As households become more informed about β, this difference vanishes; by the end of the plan-year, the median amount of over-spending is similar across all affected households, roughly \$300.

Hence, while our pooled estimates suggest little effects of deductibles resetting more frequently, households with more information about β appear more likely to respond to shorter deductible periods. This is driven mainly by a large reduction in the incidence of over-consumption, as households have shorter periods time when they are unsure if they met their deductibles. Our policy simulations suggest that by the last quarter of the year (when the average value of β has fallen to about 1.1, only 4% of households would spend more in the absence of full information.

Taken together, our modeling exercise and counterfactual simulations corroborate the findings of the empirical strategy: households tend to over-estimate their actual OOP expenditures relative to their deductibles, which leads a significant fraction to elect for greater levels of spending. These over-estimates are worse when households have little information with which to form expectations, such as early in a plan-year or for households with few medical encounters. More frequent signals, shorter waiting periods for pricing information, and more frequent deductible reset periods could all help reduce the effects of incorrect beliefs on over-consumption of care.

7 Discussion & Conclusion

This paper assesses how households respond to pricing information in their strategic decisions for future care consumption. We show that although households increase their spending following health events which may reduce the future (spot) marginal cost of care, they do so based on misinformed over-estimates of actual spending. When a bill arrives with meaningful and accurate price information, households curtail their spending increases. This delayed pricing information meaningfully contributes to over-consumption of medical care, particularly general medical services which may spark downstream cascades of care.

We encapsulate our findings into a model of imperfect *ex-post* moral hazard with delayed learning from prices. Our model, just as our reduced-form evidence suggests, indicates that consumers over-inflate expectations of OOP spending before they receive bills, particularly early on in a plan year. Our model allows us to consider alternative plan designs—including more frequent deductible resetting—that might curtail the associated over-expenditures of such under-information.

Our analysis provides several important contributions to models of price uncertainty and household moral hazard in healthcare; however, our results should be viewed in the context of their limitations. In particular, by limiting our analysis to households enrolled in group ESI plans, we are unable to determine how price uncertainty affects consumption decisions for other populations, such as couples on Medicare or low-income households covered by Medicaid. Examining other populations—particularly populations with greater income constraints—would shed additional, important light on the extent to which price uncertainty leads to sub-optimal allocations of care. Second, while our results suggest that households would make different spending decisions without price uncertainty—in particular, consuming less care on average—we are unable to say anything about the welfare effects of these decisions given our current data. Future work might attempt to disentangle over-consumption of wasteful services from the perceived relaxation of liquidity constraints, which may lead households to consume needed medical care and actually improve household welfare.

The analysis we present could be extended in several meaningful ways. First, future work could incorporate observed payment interactions between patients and physicians, rather than relying on claims data alone. Data on physician practices—including how quickly physicians submit claims to payers for medical claims and send bills to patients—may provide insights into both the source of variation in processing times as well as the potential policy benefits of reducing the length of provider billing cycles. Future work may also consider the spillover effects of bill shock from healthcare consumption on other, non-health household consumption decisions.

In addition, doing so would prevent any measurement error in the exposure variable arising from our imperfect proxy. In general, however, the measurement error associated with our proxy for bill arrival is likely to attenuate our estimates. This is not because the measurement error is classic, but instead based on the fact that any measurement error in the actual transmission of price information would result in contamination bias from the interim period, when households still do not know their OOP spending. If consumers over-estimate OOP prices before the bill arrives, any regressions misclassifying $\mathbb{1}\{\text{Post\_Bill}\} = 1$ when it should be 0 will attenuate the correction parameter $\beta_{\text{Post\_Bill}}$ to zero.

More generally, future research could build on the learning model presented here. This could include a more thorough treatment of heterogeneous learning across service types, or allowing the learning parameters to be covariate-dependent in other ways. In particular, exploring the health equity concerns associated with learning about prices could provide valuable insight in the persistence of health disparities in accessing even high-value preventive services (Teutsch et al., 2020; McMorrow et al., 2014). Finally, future work could explore the impact of real-times claim adjudication on consumer spending responses. This could be especially policy-relevant when exploring how heterogeneity across payers and providers (e.g., integrated care practices) could be used to leverage improved price transparency.

Increasing understanding of how consumers form expectations about their health needs and utilization is a vital component of designing optimal insurance contracts and health policies. Economic modeling and health policy alike are well-served from incorporating delayed learning as we assess how consumers make health decisions in real time.

References

Agency for Healthcare Research and Quality (2007). Guide to Prevention Quality Indicators: Hospital Admission for Ambulatory Care Sensitive Conditions. Technical Report AHRQ Pub. No. 02-R0203, Department of Health and Human Services, Washington, D.C.

Aron-Dine, A., Einav, L., Finkelstein, A., and Cullen, M. (2015). Moral hazard in health insurance: Do dynamic incentives matter? *Review of Economics and Statistics*, 97(4):725–741.

Arrow, K. J. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*.

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

Bloschichak, A., Milewski, A., and Martin, K. (2020). CMS-specified shoppable services accounted for 12% of 2017 health care spending among individuals with employer-sponsored insurance. HCCI Original Report, Health Care Cost Institute.

Boyd, C. M. and Bellemare, M. F. (2020). The microeconomics of agricultural price risk. *Annual Review of Resource Economics*, 12(1):149–169.

Brot-Goldberg, Z. C., Chandra, A., Handel, B. R., and Kolstad, J. T. (2017). What does a Deductible Do? The Impact of Cost-Sharing on Health Care Prices, Quantities, and Spending Dynamics*. *The Quarterly Journal of Economics*, 132(3):1261–1318.

Brown, Z. Y. (2017). An empirical model of price transparency and markups in health care. *Working Paper*.

Cabral, M. (2017). Claim Timing and Ex Post Adverse Selection. *The Review of Economic Studies*, 84(1):1–44.

Card, D., Dobkin, C., and Maestas, N. (2009). Does Medicare save lives? *The quarterly journal of economics*, 124(2):597–636.

Caskey, R., Zaman, J., Nam, H., Chae, S.-R., Williams, L., Mathew, G., Burton, M., Li, J. J., Lussier, Y. A., and Boyd, A. D. (2014). The Transition to ICD-10-CM: Challenges for Pediatric Practice. *Pediatrics*, 134(1):31–36.

Chandra, A., Gruber, J., and McKnight, R. (2010). Patient Cost-Sharing and Hospitalization Offsets in the Elderly. *The American economic review*, 100(1):193–213.

Chernew, M. E., Rosen, A. B., and Fendrick, A. M. (2007). Value-Based Insurance Design. *Health Affairs*, 26(Supplement 2):w195–w203.

CMS (2019). Medicare and Medicaid Programs: CY 2020 Hospital Outpatient PPS Policy Changes and Payment Rates and Ambulatory Surgical Center Payment System Policy Changes and Payment Rates. Price Transparency Requirements for Hospitals to Make Standard Charges Public.

Colla, C. H. and Mainor, A. J. (2017). Choosing Wisely Campaign: Valuable For Providers Who Knew About It, But Awareness Remained Constant, 2014–17. *Health Affairs*, 36(11):2005–2011.

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.

Cooper, Z., Craig, S. V., Gaynor, M., and Van Reenen, J. (2019). The price ain't right? Hospital prices and health spending on the privately insured. *The Quarterly Journal of Economics*, 134(1):51–107.

Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal*, 20(1):95–115.

Crawford, G. S. and Shum, M. (2005). Uncertainty and Learning in Pharmaceutical Demand. *Econometrica*, 73(4):1137–1173.

Cutler, D. M. and Zeckhauser, R. J. (2000). The anatomy of health insurance. In *Handbook of Health Economics*, volume 1, pages 563–643. Elsevier.

Dave, D. and Kaestner, R. (2009). Health insurance and ex ante moral hazard: Evidence from Medicare. *International journal of health care finance and economics*, 9(4):367–390.

Diaz-Campo, . (2022). Dynamic moral hazard in nonlinear health insurance contracts. *Working Paper*.

Duarte, F. (2012). Price elasticity of expenditure across health care services. *Journal of Health Economics*, 31(6):824–841.

Dunn, A. (2016). Health insurance and the demand for medical care: Instrumental variable estimates using health insurer claims data. *Journal of Health Economics*, 48:74–88.

Einav, L. and Finkelstein, A. (2018). Moral hazard in health insurance: What we know and how we know it. *Journal of the European Economic Association*, 16(4):957–982.

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.

Elshaug, A. G., Rosenthal, M. B., Lavis, J. N., Brownlee, S., Schmidt, H., Nagpal, S., Littlejohns, P., Srivastava, D., Tunis, S., and Saini, V. (2017). Levers for addressing medical underuse and overuse: Achieving high-value health care. *The Lancet*, 390(10090):191–202.

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

Fu, S. J., Rose, L., Dawes, A. J., Knowlton, L. M., Ruddy, K. J., and Morris, A. M. (2021). Out-of-pocket costs among patients with a new cancer diagnosis enrolled in high-deductible health plans vs traditional insurance. *JAMA network open*, 4(12):e2134282–e2134282.

Gaffney, A., White, A., Hawks, L., Himmelstein, D., Woolhandler, S., Christiani, D. C., and McCormick, D. (2020). High-deductible health plans and healthcare access, use, and financial strain in those with chronic obstructive pulmonary disease. *Annals of the American Thoracic Society*, 17(1):49–56.

Ganguli, I., Lupo, C., Mainor, A. J., Wang, Q., Orav, E. J., Rosenthal, M. B., Sequist, T. D., and Colla, C. H. (2020). Assessment of prevalence and cost of care cascades after routine testing during the Medicare annual wellness visit. *JAMA network open*, 3(12):e2029891–e2029891.

Geyman, J. P. (2012). Cost-Sharing under Consumer-Driven Health Care Will Not Reform U.S. Health Care. *The Journal of Law, Medicine & Ethics*, 40(3):574–581.

Globerman, S. (2016). Select cost-sharing in universal health countries. Technical report, Fraser Institute.

Goldman, D. and Philipson, T. (2007). Integrated Insurance Design in the Presence of Multiple Medical Technologies. *American Economic Review*, 97(2):427–432.

Gondi, S., Beckman, A. L., Ofoje, A. A., Hinkes, P., and McWilliams, J. M. (2021). Early Hospital Compliance With Federal Requirements for Price Transparency. *JAMA Internal Medicine*, 181(10):1396–1397.

Gourieroux, C., Monfort, A., and Trognon, A. (1984). Pseudo maximum likelihood theory. *Econometrica*, 52(3):681.

Grubb, M. D. (2015). Consumer Inattention and Bill-Shock Regulation. *The Review of Economic Studies*, 82(1):219–257.

Grubb, M. D. and Osborne, M. (2015). Cellular Service Demand: Biased Beliefs, Learning, and Bill Shock. *American Economic Review*, 105(1):234–271.

Gruber, J. (2022). Financing Health Care Delivery. *Annual Review of Financial Economics*, 14(1).

Hartzema, A. G., Racoosin, J. A., MaCurdy, T. E., Gibbs, J. M., and Kelman, J. A. (2011). Utilizing Medicare claims data for real-time drug safety evaluations: Is it feasible? *Pharmacoepidemiology and Drug Safety*, 20(7):684–688.

Henrikson, N. B., Chang, E., Ulrich, K., King, D., and Anderson, M. L. (2017). Communication with Physicians about Health Care Costs: Survey of an Insured Population. *The Permanente Journal*, 21.

Hettinger, K. (2022). Intertemporal substitution in response to non-linear health insurance contracts. *Working Paper.*

Hoagland, A. (2022). An ounce of prevention or a pound of cure? The value of health risk information. *Working Paper.*

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.

Hong, L. and Mommaerts, C. (2022). Time aggregation in health insurance deductibles. NBER Working Paper 28430, National Bureau of Economic Research, Cambridge, MA.

Hwee, J., Sung, L., Kwong, J. C., Sutradhar, R., Tu, K., and Pole, J. D. (2018). Use of physician billing claims to identify infections in children. *PLOS ONE*, 13(11):e0207468.

Johansson, N., Sonja, C., Kunz, J. S., Petrie, D., and Svensson, M. (2023). Reductions in out-of-pocket prices and forward-looking moral hazard in health care demand. *Journal of health economics*, 87:102710.

Karlsson, N., Loewenstein, G., and Seppi, D. (2009). The ostrich effect: Selective attention to information. *Journal of Risk and uncertainty*, 38(2):95–115.

Klein, T. J., Salm, M., and Upadhyay, S. (2022). The response to dynamic incentives in insurance contracts with a deductible: Evidence from a differences-in-regression-discontinuities design. *Journal of Public Economics*, 210:104660.

Knowles, J. C. (1995). Research note: Price uncertainty and the demand for health care. *Health Policy and Planning*, 10(3):301–303.

Korenstein, D., Falk, R., Howell, E. A., Bishop, T., and Keyhani, S. (2012). Overuse of health care services in the United States: An understudied problem. *Archives of Internal Medicine*, 172(2):171–178.

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

Kowalski, A. (2016b). Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care. *Journal of Business & Economic Statistics*, 34(1):107–117.

Kullgren, J. T., Galbraith, A. A., Hinrichsen, V. L., Miroshnik, I., Penfold, R. B., Rosenthal, M. B., Landon, B. E., and Lieu, T. A. (2010). Health care use and decision making among lower-income families in high-deductible health plans. *Archives of internal medicine*, 170(21):1918–1925.

Lieber, E. M. (2017). Does it pay to know prices in health care? *American Economic Journal: Economic Policy*, 9(1):154–79.

Manning, W. G. and Mullahy, J. (2001). Estimating log models: To transform or not to transform? *Journal of Health Economics*, 20(4):461–494.

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

McMorrow, S., Kenney, G. M., and Goin, D. (2014). Determinants of receipt of recommended preventive services: Implications for the Affordable Care Act. *American Journal of Public Health*, 104(12):2392–2399.

Mehrotra, A., Dean, K. M., Sinaiko, A. D., and Sood, N. (2017). Americans Support Price Shopping For Health Care, But Few Actually Seek Out Price Information. *Health Affairs*, 36(8):1392–1400.

Muir, M. A., Alessi, S. A., and King, J. S. (2012). Clarifying costs: Can increased price transparency reduce healthcare spending. *William & Mary Policy Review*, 4:319.

Mullahy, J. and Norton, E. C. (2022). Why Transform Y? A Critical Assessment of Dependent-Variable Transformations in Regression Models for Skewed and Sometimes-Zero Outcomes. *NBER Working Paper*.

Naqvi, S. A. A. (2021). SCHEMEPACK: Stata module providing ready-to-use graph schemes. Boston College Department of Economics.

Ngangoue, M. K. (2021). Learning From Unrealized versus Realized Prices. *American Economic Journal: Microeconomics*, 13(2):174–201.

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

Peng, L. (2005). Learning with information capacity constraints. *Journal of Financial and Quantitative Analysis*, 40(2):307–329.

Reed, M., Fung, V., Brand, R., Fireman, B., Newhouse, J. P., Selby, J. V., and Hsu, J. (2005). Care-Seeking Behavior in Response to Emergency Department Copayments. *Medical Care*, 43(8):810–816.

Santos Silva, J. M. C. and Tenreyro, S. (2006). The log of gravity. *The Review of Economics and Statistics*, 88(4):641–658.

Shafer, P. R., Hoagland, A., and Hsu, H. E. (2021). Trends in Well-Child Visits With Out-of-Pocket Costs in the US Before and After the Affordable Care Act. *JAMA Network Open*, 4(3):e211248.

Shafer, P. R., Horný, M., and Dusetzina, S. B. (2022). Monthly Cost-Sharing Limits and Out-of-pocket Costs for Commercially Insured Patients in the US. *JAMA Network Open*, 5(9):e2233006.

Siegfried, I., Jacobs, J., and Olympia, R. P. (2019). Adult emergency department referrals from urgent care centers. *The American Journal of Emergency Medicine*, 37(10):1949–1954.

Snowbeck, C. (2022). Risk adjustment, claims processing woes still troubling Bright Health Group. *Star Tribune*.

Stockley, K. (2016). Evaluating Rationality in Responses to Health Insurance Cost-Sharing: Comparing Deductibles and Copayments. *Working Paper*.

Teutsch, S., Carey, T. S., and Pignone, M. P. (2020). Health equity in preventive services: The role of primary care. *American Family Physician*, 102(5):264–265.

Webb Hooper, M., Mitchell, C., Marshall, V. J., Cheatham, C., Austin, K., Sanders, K., Krishnamurthi, S., and Grafton, L. L. (2019). Understanding multilevel factors related to urban community trust in healthcare and research. *International Journal of Environmental Research and Public Health*, 16(18):3280.

White, C. and Eguchi, M. (2014). Reference pricing: A small piece of the health care price and quality puzzle. *National Institute for Health Care Reform*, (18).

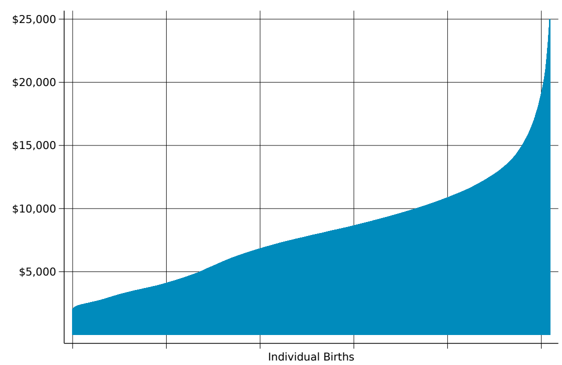
Zhang, A., Prang, K.-H., Devlin, N., Scott, A., and Kelaher, M. (2020). The impact of price transparency on consumers and providers: A scoping review. *Health Policy*, 124(8):819–825.

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.

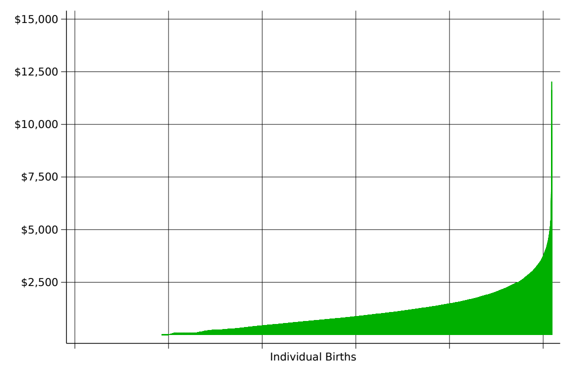
A Appendix

Figure A.7. Variation in Prices for CPT 59400: Routine Vaginal Delivery

(a) Total Costs



(b) OOP Costs

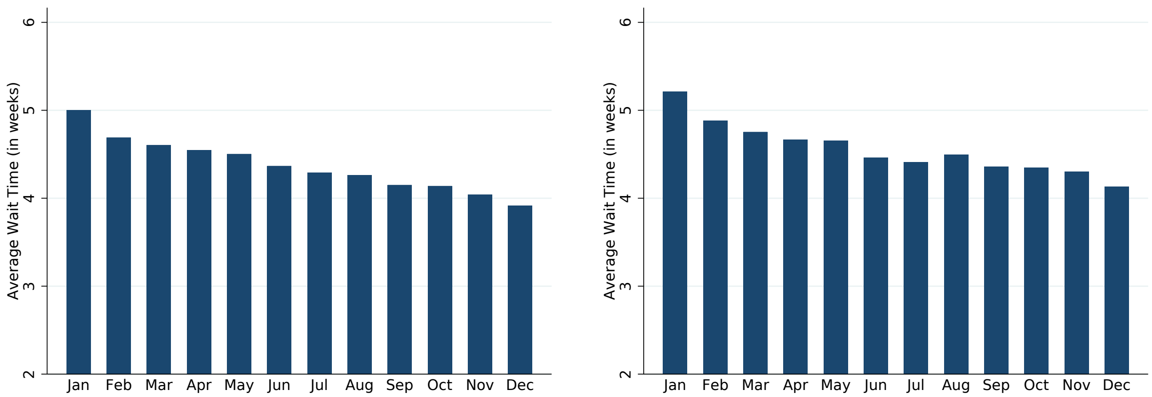


Notes: Figures show variation in total and OOP costs associated with CPT code 59400: "Routine obstetric care including antepartum care, vaginal delivery (with or without episiotomy, and/or forceps) and postpartum care." Each vertical bar represents a unique encounter in our analytical data set, with the height of the bar corresponding to the price (all measured in 2022 USD).

Figure A.8. Variation in Wait Times for Bills

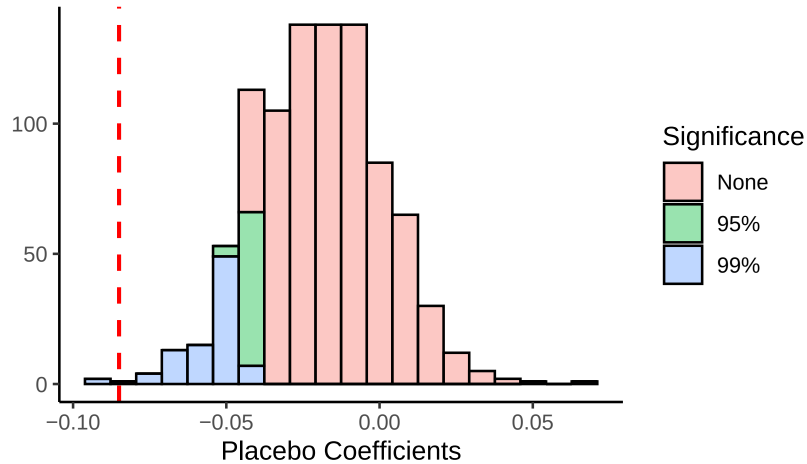
(a) All Services

(b) CMS Shoppable Services



Notes: Indicates average wait time (in weeks) between date of service and date the insurer paid their portion of the claim (the earliest date at which definitive OOP information is known). Panel (a) illustrates variation in average wait times across months of the year (pooled across all years) for all claims in the analytical data; panel (b) limits the sample to only the shoppable services used as major health events in the text.

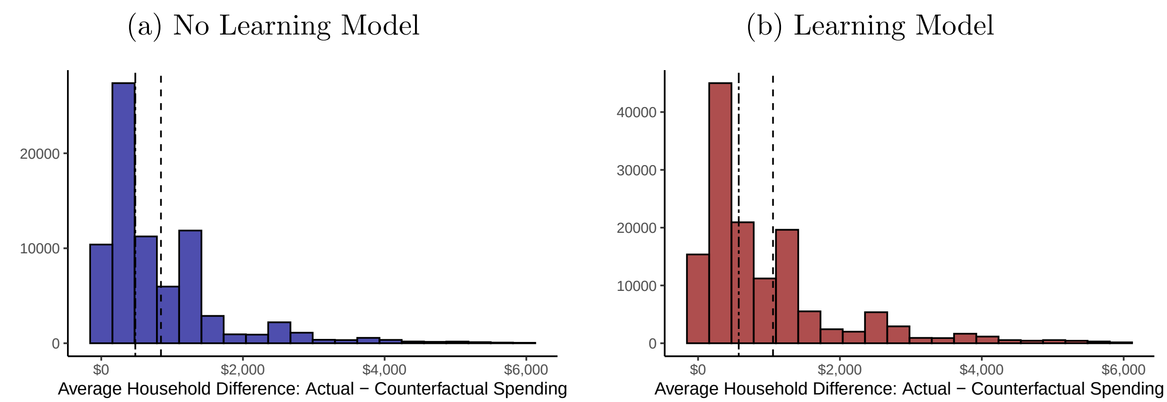
Figure A.9. Distribution of Placebo Regression Coefficients for $\beta_{\text{post\_bill}}$



NOTE: Dashed line indicates true regression coefficient.

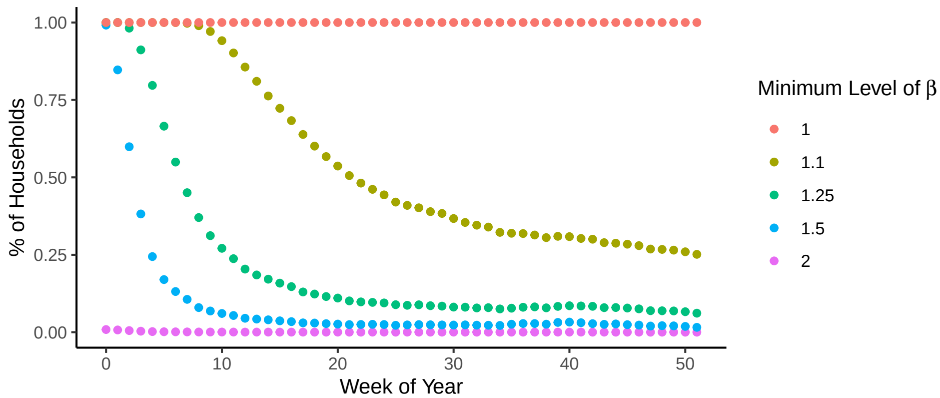
Notes: Figure shows the distribution of placebo regression coefficients for the dummy variable Post_Bill$_{it}$ in Equation 1 ($n = 1,000$). Each placebo is constructed by artificially varying wait times for bills based on the empirical distribution of wait times in the analytical sample. Standard errors are clustered at the household level. Coefficients are color-coded based on statistical significance. The vertical dashed red line indicates the estimated coefficient of the preferred specification (Table 3).

Figure A.10. Counterfactual Analysis: Change in Predicted Spending from Correcting $\beta = 1$



Notes: Figure shows estimated the differences in predicted (total) per-person health spending that arise from requiring that $\beta = 1$ in Equation 6, rather than the parameters estimated in the models (see Figure 4). Panel (a) shows results for the model without learning while panel (b) shows results for the generalized learning model. Histogram displays distribution of household-year average differences per person, conditional on a difference greater than 0. Note that for 87.5% of households in panel (a) and 77.6% of households in panel (b), no differences in spending are predicted. The dashed line indicates the average conditional difference in per-person spending while the dot-dashed line indicates the median in both groups.

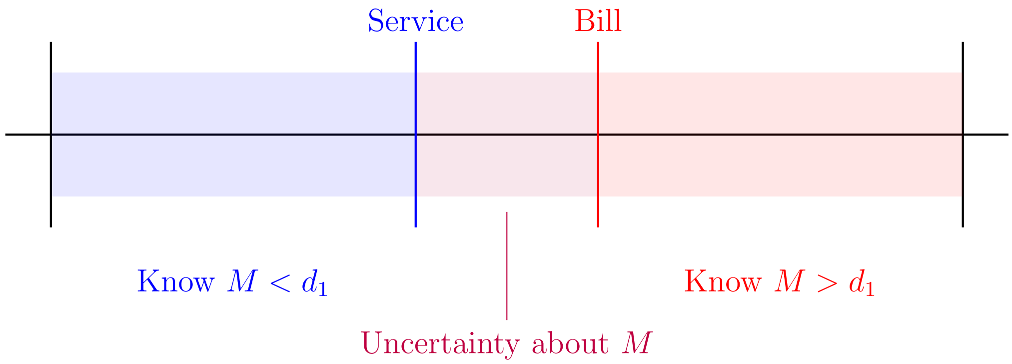
Figure A.11. Evolution of Beliefs about β Across Plan Year



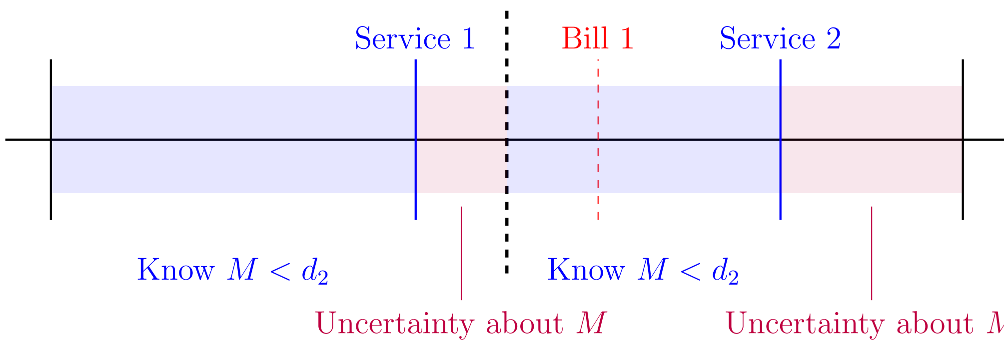
Notes: Figure depicts the fraction of households in the sample with simulated β greater than or equal to some threshold $\beta_{\min}$ for various thresholds. Simulations are performed based on the median equilibrium parameters of the model discussed in Section 5.2.

Figure A.12. Effect of Resetting Deductible More Frequently on Demand Uncertainty

Panel A: No Deductible Resetting



Panel B: Deductible Resetting



Notes: Figure illustrates the intuition behind the tradeoffs associated with deductibles of varying lengths. Larger deductibles covering a longer period of time may induce greater levels of uncertainty at specific, high-cost medical events (panel A); on the other hand, deductibles which reset more often (panel B) limit the over-consumption associated with a single period of uncertainty, but potentially induce multiple points across a plan year at which individuals are uncertain about whether or not they have met a deductible. Blue vertical lines indicate the point at which services are received which households may expect to change their marginal costs of future care (e.g., as deductibles are met), while red vertical lines indicate the bill arrival date. The dashed vertical line in panel B indicates the point at which the deductible resets within the plan-year.

Type	Code	Service Description
DRG	216	Cardiac valve and other major cardiothoracic procedures w/ cardiac catheterization
DRG	460	Spinal fusion, except cervical
DRG	470	Major joint replacement or reattachment of lower extremity
DRG	473	Cervical spinal fusion
DRG	743	Uterine and adnexa procedures for non-malignancy
CPT	19120	Removal of 1 or more breast growth, open procedure
CPT	29826	Shaving of shoulder bone using an endoscope
CPT	29881	Removal of one knee cartilage using an endoscope
CPT	42820	Removal of tonsils and adenoid glands (patient younger than age 12)
CPT	43235	Diagnostic examination of esophagus, stomach, and/or upper small bowel
CPT	43239	Biopsy of the esophagus, stomach, and/or upper small bowel using an endoscope
CPT	45378	Diagnostic examination of large bowel using an endoscope
CPT	45380	Biopsy of large bowel using an endoscope
CPT	45385	Removal of polyps or growths of large bowel using an endoscope
CPT	45391	Ultrasound examination of lower large bowel using an endoscope
CPT	47562	Removal of gallbladder using an endoscope
CPT	49505	Repair of groin hernia (patient age 5 years or older)
CPT	55700	Biopsy of prostate gland
CPT	55866	Surgical removal of prostate and surrounding lymph nodes using an endoscope
CPT	59400	Routine obstetric care for vaginal delivery
CPT	59510	Routine obstetric care for cesarean delivery
CPT	59610	Routine obstetric care for vaginal delivery after prior cesarean delivery
CPT	64483	Injections of anesthetic and/or steroid drug into lower or sacral spine nerve root
CPT	66821	Removal of recurring cataract in lens capsule using laser
CPT	66984	Removal of cataract with insertion of lens
CPT	93000	Electrocardiogram, routine, with interpretation and report
CPT	93452	Insertion of catheter into left heart for diagnosis
CPT	62322	Injection of substance into spinal canal of lower back or sacrum
CPT	62323	Injection of substance into spinal canal of lower back or sacrum

Notes: Table shows list of procedures used to identify non-urgent "shoppable services," which are the exposure of interest in the primary reduced-form specifications. Services are identified based on lists provided by the Center for Medicare and Medicaid Services (CMS), using the relevant Diagnostic Related Groups (DRGs) or Current Procedural Terminology (CPT) codes to identify procedures.

Table A.7. Shoppable Services Used in Analytical Sample

Service Description	**Code**
Panel A: Diagnosis Codes for Infections (ICD-9-CM)	
Acute Respiratory Infections	460-466
Pneumonia and Influenza	480-488
Nonsuppurative otitis media and eustachian tube disorders	381
Suppurative and unspecified otitis media	382
Streptococcal sore throat and scarlet fever	034
Whooping cough	033
Infectious mononucleosis	075
Chickenpox	052
Urinary Tract Infections	590, 595, 599
Panel B: Place of Service Codes (POS)	
Physician Office	11, 72, 95
Urgent Care Center	17,20
Emergency Department	23
Hospital (including on-campus outpatient)	21, 22, 28

Notes: Table shows list of diagnoses used to identify acute respiratory infections (Hwee et al., 2018) procedures used to identify non-urgent "shoppable services," which are the exposure of interest in the primary reduced-form specifications. Services are identified based on lists provided by the Center for Medicare and Medicaid Services (CMS), using the relevant Diagnostic Related Groups (DRGs) or Current Procedural Terminology (CPT) codes to identify procedures.

Table A.8. Identifying Respiratory Infections and Places of Service

Outpatient Category	CPT Codes	
	Code Range	Code Values
Behavioral Health	90000-99999	90791-90792, 90801-90802, 90805-90807, 90832-90834, 90836-90840, 90845-90847, 90849, 90853, 96105, 96112-96113, 96116, 96121, 96125, 96130-96133, 96136-96139, 96156, 96158-96159, 96164-96165, 96167-96168, 96170-96171, 99483-99494
Chiropractic Care	90000-99999	97001, 97010-97014, 97018, 97022, 97026, 97032-97035, 97039, 97110-97113, 97116, 97124, 97140, 97161-97162, 97530, 97535, 97750, 98940-98943, 99211
Evaluation & Management	10000-19999	11976, 11981-11983
	30000-39999	36415-36416
	40000-49999	44388-44389, 44392-44394, 45300, 45303-45309, 45315-45317, 45320, 45330-45335, 45338-45340, 45378-45386
	50000-59999	57170, 58300-58301, 58340, 58565, 58600, 58605, 58611, 58615, 58670-58671
	70000-79999	71250, 74263, 74740, 76070-76071, 76075-76078, 76497, 76977, 77078-77083, 78350
	80000-89999	80061, 82270, 82274, 82465, 82947-82952, 83036, 83718-83721, 84478, 86580, 86592-86593, 86631-86632, 86689, 86701-86703, 86803-86804, 87110, 87270, 87320, 87340-87341, 87390-87391, 87490-87492, 87590-87592, 87620-87622, 87801, 87810, 87850, 88141-88143, 88147-88155, 88164-88167, 88174-88175, 88304-88305
	90000-99999	92015, 92507, 92551-92553, 92558, 92567, 92585-92588, 96040, 96110, 96127, 96160-96161, 96372, 97802-97804, 99173-99174, 99201-99205, 99211-99215, 99381, 99385-99387, 99395-99397, 99401-99404, 99411-99412, 99420
Imaging	10000-19999	10005-10006, 19081-19084
	20000-29999	29881
	70000-79999	70030, 70110, 70130, 70150, 70160, 70200, 70210, 70220, 70260, 70330, 70336, 70360, 70450, 70460, 70470, 70480-70482, 70486-70491, 70496-70498, 70540, 70543-70553, 71010, 71020, 71045-71048, 71100-71101, 71110, 71120, 71130, 71250,

		71260, 71275, 71550-71552, 71555, 72040, 72050-72052, 72070, 72082, 72100, 72110, 72114, 72125-72132, 72141-72142, 72146-72149, 72156-72159, 72170, 72191-72202, 72220, 73000, 73010, 73030, 73050, 73060, 73070, 73090, 73100, 73110, 73120, 73130, 73140, 73200-73202, 73206, 73218-73225, 73501-73503, 73521-73523, 73552, 73560-73564, 73590, 73600, 73610, 73620, 73630, 73650, 73660, 73700-73702, 73706, 73718-73725, 74000, 74018-74021, 74150, 74160, 74170, 74174-74178, 74181-74185, 74210, 74220, 74241, 74245-74250, 74261-74263, 74270, 74280, 74400, 75635, 76010, 76390-76391, 76536, 76641-76642, 76645, 76700, 76705-76706,76770, 76775-76776, 76801, 76812, 76817, 76830, 76856-76857, 76870, 76881-76882, 76981, 77021, 77046-77049, 77052, 77057, 77063-77067, 77072-77077, 77080, 77085, 78012-78014, 78070-78071, 78206, 78215, 78226-78227, 78290, 78306, 78315, 78452, 78472, 78607-78608, 78707-78708, 78800, 78804, 78814-78816
	90000-99999	91200, 93000, 93005, 93010-93018, 93024-93025, 93040-93042, 93050, 93201-93205, 93208-93210, 93220-93237, 93241-93248, 93260-93261, 93264, 93268-93272, 93278-93299, 93303-93308, 93312-93321, 93325, 93350-93352, 93355-93356, 93451-93464, 93501-93505, 93508-93511, 93514, 93524-93533, 93536, 93539-93545, 93555-93556, 93561-93568, 93571-93572, 93580-93583, 93590-93603, 93607-93624, 93631, 93640-93644, 93650-93657, 93660-93662, 93668, 93701-93702, 93720-93724, 93727, 93731-93745, 93750, 93760-93762, 93770, 93784-93793, 93797-93799, 93880, 93926, 93970-93971, 93975
Lab Services	20000-29999	20610
	30000-39999	36415-36416
	80000-89999	80048, 80050, 80053, 80061, 80076, 81000-81003, 81025, 82000, 82003, 82009-82010, 82013, 82016-82017, 82024, 82030, 82040, 82042-82045, 82055, 82075, 82077, 82085, 82088, 82101, 82103-82108, 82120, 82127-82128, 82130-82131, 82135-82136, 82139-82140, 82143, 82145, 82150, 82154, 82157, 82160, 82163-82164, 82172, 82175, 82180, 82190, 82205,

	82232, 82239-82240, 82247-82248, 82250-82252, 82261, 82270-82274, 82286, 82300, 82306-82308, 82310, 82330-82331, 82340, 82355, 82360, 82365, 82370, 82373-82376, 82378-82380, 82382-82384, 82387, 82390, 82397, 82415, 82435-82436, 82438, 82441, 82465, 82480, 82482, 82485-82489, 82491-82492, 82495, 82507, 82520, 82523, 82525, 82528, 82530, 82533, 82540-82544, 82550, 82552-82554, 82565, 82570, 82575, 82585, 82595, 82600, 82607-82608, 82610, 82615, 82626-82627, 82633-82634, 82638, 82642, 82646, 82649, 82651-82654, 82656-82658, 82664, 82666, 82668, 82670-82672, 82677, 82679, 82681, 82690, 82693, 82696, 82705, 82710, 82715, 82725-82726, 82728, 82731, 82735, 82742, 82746-82747, 82757, 82759-82760, 82775-82777, 82784-82785, 82787, 82800, 82803, 82805, 82810, 82820, 82926, 82928, 82930, 82938, 82941, 82943, 82945-82948, 82950-82953, 82955, 82960, 82962-82963, 82965, 82975, 82977-82980, 82985, 83001-83003, 83006, 83008-83010, 83012-83015, 83018-83021, 83026, 83030, 83033, 83036-83037, 83045, 83050-83051, 83055, 83060, 83065, 83068-83071, 83080, 83088, 83090, 83150, 83491, 83497-83500, 83505, 83516, 83518-83521, 83525, 83527-83529, 83540, 83550, 83570, 83582, 83586, 83593, 83605, 83615, 83625, 83630-83634, 83655, 83661-83664, 83670, 83690, 83695, 83698, 83700-83701, 83704, 83715-83719, 83721-83722, 83727, 83735, 83775, 83785, 83788-83789, 83805, 83825, 83835, 83840, 83857-83858, 83861, 83864, 83866, 83872-83874, 83876, 83880, 83883, 83885, 83887, 83890-83894, 83896-83898, 83900-83909, 83912-83916, 83918-83919, 83921, 83925, 83930, 83935, 83937, 83945, 83950-83951, 83970, 83986-83987, 83992-83993, 84022, 84030, 84035, 84060-84061, 84066, 84075, 84078, 84080-84081, 84085, 84087, 84100, 84105-84106, 84110, 84112, 84119-84120, 84126-84127, 84132-84135, 84138, 84140, 84143-84146, 84150, 84152-84157, 84160, 84163, 84165-84166, 84181-84182, 84202-84203, 84206-84207, 84210, 84220, 84228, 84233-84235, 84238, 84244, 84252, 84255, 84260, 84270, 84275, 84285, 84295, 84300, 84302,

		84305, 84307, 84311, 84315, 84375-84379, 84392, 84402-84403, 84410, 84425, 84430-84432, 84436-84437, 84439, 84442-84443, 84445-84446, 84449-84450, 84460, 84466, 84478-84482, 84484-84485, 84488, 84490, 84510, 84512, 84520, 84525, 84540, 84545, 84550, 84560, 84577-84578, 84580, 84583, 84585-84586, 84588, 84590-84591, 84597, 84600, 84620, 84630, 84681, 84702-84704, 84830, 84999, 85007, 85014, 85018, 85025, 85027, 85610, 85651-85652, 85730, 86003, 86038, 86140, 86580, 86592, 86880, 86900-86901, 87040, 87070, 87077, 87081, 87086, 87088, 87186, 87491, 87591, 87621, 87804, 87880, 88142, 88175, 88304-88305, 88312-88313, 88342, 88720
	90000-99999	94760, 99000-99001
Low-Value Services	20000-29999	29877-29879
	30000-39999	36222-36224
	70000-79999	70450, 70460, 70470, 70498, 70547-70553, 71010, 71015, 71020-71023, 71030, 71034-71035, 72010, 72020, 72052, 72100, 72110, 72114, 72120, 72131-72133, 72141-72142, 72146-72149, 72156-72158, 72200-72202, 72220, 78451-78454, 78460-78461, 78464-78465, 78472-78473, 78481-78483, 78491-78492
	80000-89999	82306, 82652, 87620-87625, 88141-88143, 88147-88155, 88164-88167, 88174-88175
	90000-99999	93000, 93005, 93010, 93015-93018, 93303-93308, 93312, 93315, 93318, 93350-93351, 93880-93882, 94010
Preventive Care	10000-19999	11976, 11981-11983
	30000-39999	36415-36416
	40000-49999	44388-44389, 44392-44394, 45300, 45303-45309, 45315-45317, 45320, 45330-45335, 45338-45340, 45378-45386
	50000-59999	57170, 58300-58301, 58340, 58565, 58600, 58605, 58611, 58615, 58670-58671
	70000-79999	71250, 74263, 74740, 76070-76071, 76075-76078, 76497, 76977, 77078-77083, 78350

	80000-89999	80061, 82270, 82274, 82465, 82947-82952, 83036, 83718-83721, 84478, 86580, 86592-86593, 86631-86632, 86689, 86701-86703, 86803-86804, 87110, 87270, 87320, 87340-87341, 87390-87391, 87490-87492, 87590-87592, 87620-87622, 87801, 87810, 87850, 88141-88143, 88147-88155, 88164-88167, 88174-88175, 88304-88305
	90000-99999	92015, 92507, 92551-92553, 92558, 92567, 92585-92588, 96040, 96110, 96127, 96160-96161, 96372, 97802-97804, 99173-99174, 99201-99205, 99211-99215, 99381, 99385-99387, 99395-99397, 99401-99404, 99411-99412, 99420
Specialist Care	10000-19999	11100, 17000, 17003-17004, 17110-17111, 17250
	40000-49999	43239, 47562
	80000-89999	82962
	90000-99999	92012-92014, 92587, 93010, 94010

Table A.9. Identifying Types of Outpatient Services

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

Table A.10. Identifying Low-Value Health Services

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

Table Notes: Pediatric low-value services are defined based on Chua et al. (2016). Adult low-value services are based on definitions given in Bhatia et al. (2015), Chandra et al. (2021), and Colla et al. (2014).