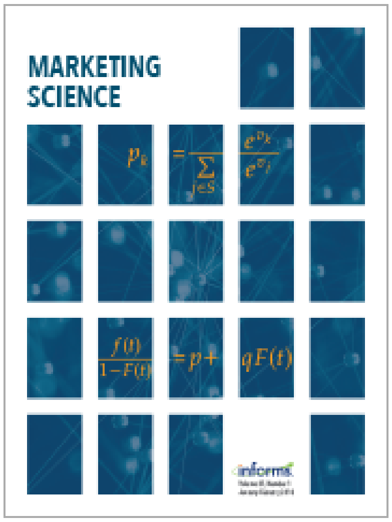


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

A Dynamic Model of Health Insurance Choices and Healthcare Consumption Decisions

Nitin Mehta, Jian Ni, Kannan Srinivasan, Baohong Sun

informs
http://pubsonline.informs.org/journal/mksc/
MARKETING SCIENCE
Vol. 36, No. 3, May–June 2017, pp. 338–360
ISSN 0732-2399 (print), ISSN 1526-548X (online)

A Dynamic Model of Health Insurance Choices and Healthcare Consumption Decisions

Nitin Mehta,[a] Jian Ni,[b] Kannan Srinivasan,[c] Baohong Sun[d]

[a] Rotman School of Management, University of Toronto, Toronto, Ontario M5S 3E6, Canada; [b] Carey Business School, Johns Hopkins University, Baltimore, Maryland 21202; [c] Tepper School of Business, Carnegie Mellon University, Pittsburgh, Pennsylvania 15213; [d] Cheung Kong Graduate School of Business, New York, New York 10022

Contact: nmehta@rotman.utoronto.ca (NM); jni@jhu.edu (JN); kannans@cmu.edu (KS); bhsun@ckgsb.edu.cn (BS)

Received: January 26, 2014
Revised: December 11, 2015; April 9, 2016
Accepted: April 14, 2016
Published Online in Articles in Advance: March 6, 2017

https://doi.org/10.1287/mksc.2016.1021

Abstract. Chronic diseases, which account for 75% of healthcare expenditure, are of particular importance in trying to understand the rapid growth of healthcare costs over the last few decades. Individuals suffering from chronic diseases can consume three types of services: secondary preventive care, which includes diagnostic tests; primary preventive care, which consists of drugs that help prevent the illness from getting worse; and curative care, which includes surgeries and expensive drugs that provide a quantum boost to the patient's health. Although the majority of cases can be managed by preventive care, most consumers opt for more expensive curative care that leads to a substantial increase in overall costs. To examine these inefficiencies, we build a model of consumers' annual medical insurance plan decisions and periodic consumption decisions and apply it to a panel data set. Our results indicate that there exists a sizable segment of consumers who purchase more comprehensive plans than needed because of high uncertainty vis-à-vis their health status, and that once in the plan, they opt for curative care even when their illness could be managed through preventive care. We examine how changing cost-sharing characteristics of insurance plans and providing more accurate information to consumers via secondary preventive care can reduce these inefficiencies.

History: John Hauser served as the guest editor-in-chief and Peter Fader served as associate editor for this article.
Funding: This paper was funded by the Social Sciences and Humanities Research Council [Grant 410-2011-0189] to N. Mehta and an Andrew Mellon doctoral fellowship to J. Ni.
Supplemental Material: Data and the web appendix are available at https://doi.org/10.1287/mksc.2016.1021.

Keywords: insurance choice • healthcare service • nested dynamic decisions • primary and secondary preventive care • curative care • informative effect • investment effect • dynamic programming

1. Introduction

Healthcare is one of the most important personal services that consumers buy, and it has a pervasive impact on quality of life and the economy. In the United States, insurance firms are the major providers of health insurance and cover 55.4% of the population through employer-sponsored health insurance plans (Smith and Medalia 2015). During the past decade, healthcare costs incurred by the providers have more than doubled, resulting in a price tag of $3 trillion in 2014, which constitutes 17.5% of U.S. gross domestic product (Centers for Medicare and Medicaid Services 2015). The increase in healthcare costs is particularly relevant in the context of chronic diseases such as heart disease, cancer, diabetes, hypertension, etc., which afflict more than 133 million people in the United States. Unlike other ailments and acute diseases, they require healthcare over a long time horizon and account for 86% of the overall healthcare costs in the United States (Gerteis et al. 2014).

The focus of this paper is an investigation of the increase in healthcare costs associated with chronic disease in the context of consumers enrolled in employer-sponsored insurance plans. This increase is relevant from both the consumer's and the insurance firm's standpoint. It is relevant for consumers because the increase in healthcare costs has resulted in insurance firms passing on the costs to consumers in terms of increased premiums. This imposes a huge burden on the consumer's wallet and has led to a dramatic increase in the number of people who cannot pay the high premiums. It is relevant for insurance firms because a part of the increase could be due to the inefficient design of insurance plans, which results in consumers choosing plans and treatment options that are not cost effective. Decreasing this inefficiency would help insurance firms increase profits by either increasing their profit margin or improving their competitive position vis-à-vis other insurance firms by passing on the lowered costs to the consumers.

In this paper, we posit that in the context of chronic disease, a part of the increase in healthcare costs is due to demand-side inefficiencies that stem from the moral hazard problem faced by employer-based insurance firms. Our goal is to quantify the extent of this inefficiency and explore policies that managers in employer-based insurance firms can employ in the design of health plans to reduce it. In what follows, we explain the moral hazard problem faced by insurance firms in the context of chronic diseases. We then segue into a discussion of our specific objectives.

1.1. Moral Hazard in the Context of Chronic Diseases

Unlike individuals with other ailments, individuals suffering from chronic diseases can consume two types of services to manage their illnesses: preventive and curative care (Kenkel 2000). Preventive care, as defined by the Centers for Disease Control and Prevention, consists of two types: primary and secondary prevention. Primary prevention refers to services that include certain types of drugs that prevent diseases from getting worse. Secondary prevention refers to services that help detect illness and detect any adverse changes that are symptomatic of worsening of the disease; these services include diagnostic tests and screening exams. Curative care refers to treatments such as surgery and certain types of drugs that help improve symptoms or cure the condition. Table 1 provides examples of curative, primary preventive, and secondary preventive care. Although primary preventive and curative care serve a similar purpose, they differ in two respects. First, curative care is more expensive and provides a greater boost to a consumer's health compared to primary prevention. Second, primary prevention is useful when the illness is detected at an early stage and its severity ranges from low to moderate. On the other hand, curative care is useful when the severity of the illness ranges from moderate to high.

The distinction between preventive and curative care is crucial in the context of chronic diseases, because a vast majority of chronic disease cases can be effectively managed by preventive care, especially those with moderate levels of severity (Grossman and Rand 1974, Thorpe 2008). A study by the Milken Institute (DeVol et al. 2007) claims that by investing in prevention of the most common chronic diseases, the United States could decrease treatment costs by $218 billion per year and reduce the economic impact of disease by $1.1 trillion annually. However, the reality is that more than 96% of overall healthcare expenditure for chronic diseases goes to curative care (Thorpe 2008). For instance, consider prostate cancer, which affects around 200,000 men every year in the United States (Beck 2010). An estimated 85% of prostate tumors grow so slowly that they will never cause problems if they are effectively managed through preventive care. This includes taking primary preventive treatments such as statins that control the spread of cancer (along with prostate-specific antigen (PSA) tests on a periodic basis). However, less than 10% of such consumers opt for primary preventive care; instead, most undergo curative care such as surgery and radiation therapy—care that substantially increases overall healthcare costs.

It follows that there could be inefficiency in managing chronic diseases: instead of seeking less expensive primary prevention, consumers suffering from chronic diseases with moderate levels of severity opt for more expensive curative care. In this paper, we investigate an explanation for this inefficiency: moral hazard (Rothbaum 2006, Brennen and Reisman 2008, Aaron and Ginsburg 2009, Korobkin 2014). Recall that consumers suffering from chronic diseases with moderate levels of severity can be treated by either curative care or primary prevention, where curative care provides a greater boost to the consumer's health compared to primary prevention, but is substantially more expensive. Thus, if cost of treatment to consumers (i.e., their out-of-pocket expenses) is not an issue, such consumers will always opt for curative care. Since most consumers in the United States obtain healthcare services through insurance plans, it follows that the more

Table 1. Examples of Preventive and Curative Care

Secondary preventive	Primary preventive	Curative
Blood pressure test, cholesterol test, Pap smear, preconception care programs for women with diabetes, mammogram, prostate-specific antigen test, annual checkup, electrocardiogram (ECG, EKG), cardiovascular stress test, electrocardiographic monitoring for 24 hours, echocardiography, coronary angiography, thallium stress test, cardiac computerized tomography angiography, breast ultrasound, breast magnetic resonance imaging, bone scan	Pneumococcal vaccine, tetanus vaccine, hepatitis A vaccine, simvastatin, Lipitor, Crestor, Zocor, pravastatin, lovastatin, gemfibrozil, Pravachol, Simcor, Lopid, atorvastatin, Caduet, rosuvastatin, Mevacor, Advicor, Lescol XL, Lescol, amlodipine-atorvastatin, niacin-simvastatin, Altoprev, fluvastatin, niacin-lovastatin, chlorthalidone, methyclothiazide, bromocriptine, Kombiglyze XR, Prandin	Coronary angioplasty, coronary stent placement, coronary artery bypass, mastectomy simple complete, excision of breast lesion, partial mastectomy (lumpectomy), radical mastectomy, modified radical mastectomy, radiotherapy, prostatectomy, external beam radiation, brachytherapy, orchiectomy, cryoablation, colonoscopy with polyp removal, radical surgical resection, hemicolectomy, colectomy

comprehensive the plan, the lower the share of expense incurred by consumers, and the greater the incentive for consumers with moderate health status to opt for curative care.

In summary, moral hazard arises when consumers suffering from chronic diseases with moderate levels of severity select comprehensive insurance plans and opt for the more expensive curative care to manage their illnesses because they do not bear most of the cost. Our goal is to quantify the extent of this moral hazard and investigate the policies that managers can employ to reduce it. To achieve this goal, we have three objectives.

1.2. Research Objectives and Main Findings

The *first* objective is to model consumers' annual insurance plan decisions and periodic healthcare consumption decisions conditional on their chosen insurance plans. The insurance plan decision entails choosing between various types of health plans that differ in terms of annual premiums and cost-sharing characteristics. The periodic consumption decision entails choosing between primary preventive, secondary preventive, curative care, and the no-consumption option. We model the consumption decision as a trade-off between the out-of-pocket expenses an individual would incur when consuming each of these services and the extent to which different healthcare services impact a consumer's health status. We model the impact of curative, primary, and secondary preventive care services on a consumer's health status via two mechanisms: investment and informative effects. The investment effect refers to an increase in the consumer's health after consuming the healthcare service. The informative effect refers to the information that the healthcare service provides to the consumer about her health status that helps her make future decisions about treatment. As for the annual insurance plan decision, we model it as a trade-off between the annual premium charged in the health plan and the consumer's expected future healthcare consumption.

The *second* objective is to estimate the model on panel data that consist of consumers' insurance plan and consumption choices over time. The panel data contain detailed information on insurance plans available to each consumer each year, the characteristics of each plan, the plan chosen by each consumer each year, healthcare consumption decisions on a weekly level, and the cost incurred by the consumer and insurance company for each consumption decision. Our key empirical results are as follows: (a) The informative effect of secondary preventive care is significant, and whereas both primary preventive and curative care have significant investment effects, the investment effect is higher for curative care. (b) Consumers' health depreciates over time in the absence of consumption. (c) There is considerable heterogeneity in consumers' health status, their uncertainty over health status, price sensitivity, and degradation rate. This heterogeneity results in consumers self-selecting different insurance plans. (d) Moral hazard plays a crucial role in creating inefficiency. Our results indicate that there exists a sizable segment of "problem" consumers (14.3% of consumers) who have moderate health status, high uncertainty about their health status, and low price sensitivity. Because of high uncertainty, these consumers choose comprehensive plans and not medium plans that would be a better fit for their moderate health status; once in the comprehensive plan, they opt for the more expensive curative care instead of primary preventive care.

The *third* objective is to explore policies that managers can employ to reduce the inefficiency caused by moral hazard. We examine two policy routes. The first is the "immediate route," in which we change the cost-sharing characteristics of insurance plans at the beginning of the policy year, which incentivizes problem consumers to choose the medium plan in that policy year itself. The second is the "delayed route," in which we incentivize the problem consumers to consume more secondary preventive care in the first year (and onward), which decreases their uncertainty and induces them to choose medium plans in the future. Our results indicate that the immediate route is not feasible, because these consumers have low price elasticities in insurance plan decisions. We conduct two counterfactuals in the context of the delayed route, in which we examine the impact of the following on healthcare costs: (a) changing cost-sharing characteristics that encourage the use of secondary preventive care and (b) providing more accurate information through secondary preventive care. In both counterfactuals, we find that inefficiency reduces substantially as a result of the problem consumers choosing medium plans in the future.

1.3. Related Literature

We start with prior empirical papers that have modeled consumers' healthcare consumption and/or insurance plan decisions based on a joint utility maximization framework. Most papers that have modeled healthcare consumption decisions have treated insurance plan decisions as exogenous. The few empirical papers that have endogenized insurance plan decisions are Cameron et al. (1988), Cardon and Handel (2001), Carlin and Town (2008), Khwaja (2010), Bajari et al. (2014), Einav et al. (2013), and Handel (2013). The main distinction between our paper and these papers is that we have access to a richer panel data set that informs us about consumers' annual insurance plan decisions and periodic healthcare consumption decisions within the policy year. This distinction, as we explain next, allows us to estimate a more comprehensive model of consumers' decisions, which in turn allows us to address the research questions we posed in Section 1.1: To what extent does the increase in healthcare costs stem from moral hazard? What policies can managers in insurance firms employ to reduce this inefficiency?

Our data set differs from those used in the prior papers on two accounts. *First*, regarding healthcare consumption decisions, the prior papers had information only about the annual healthcare expenditure incurred by each consumer. They did not have information about whether or not the consumer purchased primary preventive, secondary preventive, or curative care on a per-period basis in the given policy year. Consequently, prior papers modeled the healthcare consumption decision as a one shot decision in terms of how much money to spend on healthcare services in a given policy year. Since it is not possible to identify investment effects of healthcare services from their informative effects if the researcher has access to only the information on consumers' annual healthcare expenditures, prior papers modeled only the investment effects of healthcare services and not their informative effects. As a result of these differences, the emphasis of prior papers was primarily on demonstrating the existence of moral hazard. On the other hand, since our data set informs us whether a consumer chooses primary preventive care or curative care, we can quantify the moral hazard that stems from consumers with moderate health status choosing curative care over primary preventive care. Along the same lines, since our data set informs us whether a consumer chooses secondary preventive care that provides informative effects, we are able to examine how moral hazard can be reduced by encouraging consumers to use healthcare services that provide informative effects.

Second, all prior papers with the exception of Khwaja (2010) used cross-sectional data that covered consumers' purchase decisions for one policy year only. On the other hand, we have panel data that cover consumers' purchase decisions over multiple years. This allows us to control for unobserved heterogeneity along multiple dimensions (such as health status, uncertainty, price sensitivity, etc.), which, as we will explain in Section 2.1, is crucial for accurately estimating the extent of moral hazard.

Finally, a related stream of literature has investigated whether consumers behave in a bounded-rational fashion when making insurance plan decisions (e.g., Johnson et al. 1993, Strombom et al. 2002, Shapira and Venezia 2008, Schram and Sonnemans 2011, Johnson et al. 2013). The findings of this literature are that consumers can succumb to psychological biases when making insurance plan decisions; the reasons include that they exhibit status quo bias, overweigh out-of-pocket expenses, are unfamiliar with how deductibles contribute to the overall cost, etc. Although these biases are important, they are of less concern in our context for two reasons. *First*, Rabin and Thaler (2001) have shown that people are most likely to depart from a rational utility maximization model when stakes are small—for instance, when choosing between insurance products that cover against small losses. Since the stakes are much higher when it comes to one's own health, especially in the case of chronic diseases, it follows that consumers would be less likely to depart from the utility maximization model in our context. *Second*, Cunningham et al. (2001) and Harris (2002) have shown that the propensity to succumb to psychological biases in insurance plan decisions is higher for individuals with low cognitive ability and low financial literacy. Note that all consumers in our data have college degrees, which implies that their propensity to succumb to biases will be small.

2. Data

We use a proprietary data set provided by an anonymous health insurer. This insurer offers employer-sponsored preferred provider organization (PPO) plans to consumers through their employers, typical of the American healthcare system. Around 87% of the insurer's group business is in a PPO product category. In our data, consumers come from different employers, but their health plans are offered by the same single insurer from 2005 to 2007.[1] Across these three years, the insurer offered consumers three plans labeled as basic, medium, and comprehensive. All consumers in the data were enrolled in one of the three plans. All three plans have the same hospital/physician network, the same covered medical services, and the same contractual agreement with healthcare providers (such as discounts for services provided). The three plans differ in terms of cost-sharing characteristics and annual premiums, which determine the mapping from total medical expenditure from a given treatment to out-of-pocket expense incurred by the consumer. The characteristics of the three plans vary annually and also across consumers who choose the same type of plan, since the insurer customizes these three plans for different employers.

The cost-sharing characteristics of a plan include the copayment, deductible, coinsurance rate, and out-of-pocket maximum. The copayment is the specified dollar amount that a consumer must pay out of her own pocket for a doctor visit at the time the service is rendered. The deductible is the flat amount that the consumer must pay before the insurer will make any benefit payments. The coinsurance rate is the percentage of all remaining eligible healthcare expenses that the consumer has to pay after the deductible amount has been paid. The out-of-pocket maximum is the dollar amount set by the insurer that limits the amount the consumer has to pay out of her own pocket for particular healthcare services during a certain time period. As we move from the basic to the comprehensive plan, the annual premium increases, but the coinsurance rate, deductible, and out-of-pocket maximum decrease. See Figure 1 for an illustration of the pricing structure.

We observe every consumer's annual health insurance plan choice from 2005 to 2007 and her detailed claim history over the three years. For each consumer, we have information on demographics, annual premiums, and cost-sharing characteristics of each insurance plan in each consumer's choice set. For each claim, we have information on when the specific healthcare service was consumed; the diagnostic, procedure, and therapeutic codes for that specific service; and the payment information such as the copayment, deductible paid, coinsurance paid, insurer paid, and total bill charged. The typical chronic diseases are heart disease, cancer, hypertension, respiratory diseases, diabetes, Alzheimer's disease, and kidney disease. We classify healthcare consumption services into three types: (i) secondary preventive care, which includes tests such as PSA tests and mammograms; (ii) primary preventive care, such as statins (Lipitor) used for lowering blood cholesterol; (iii) curative care, which includes surgeries such as coronary stent placement and coronary angioplasty, and curative drugs. The classification of different treatments into primary preventive, secondary preventive, and curative care is based on the definitions by the American Medical Association (International Classification, Ninth Revision, Clinical Modification 2008 and Current Procedural Terminology 2008) and the internal manual provided by the insurance company.

Table 2 reports the number of different treatments that fall under a given consumption option for three chronic diseases. Note that for each disease, there are multiple treatments within a given consumption option. For instance, there are five different treatments that fall under secondary preventive care for type 2 diabetes. Thus, to keep the model tractable, for each disease, we aggregate all treatments that fall under a given consumption option into a single consumption option with a single aggregate cost. The procedure is explained in Section 1 of the web appendix and is similar to what prior papers have done to operationalize the aggregate price of a category/brand composite when using scanner data sets. Essentially, the aggregate cost of a consumption option for a given disease is calculated as a weighted average cost of all treatments that fall under that consumption option and disease, where the weights are based on frequencies of usage of the different treatments. To see whether or not this would lead to large aggregation biases, we compute the weighted average cost (and standard deviation of costs) of different treatments that fall under each consumption option for the same three diseases. The weighted average cost (and standard deviation) of a given consumption option and disease is calculated based on costs recorded in our data for all observations in which the consumers consumed any of the treatments that fall under that given consumption option and disease. The weighted average costs and standard deviations of costs are reported in Table 2. Observe in Table 2 that standard deviations of all costs are an order of magnitude lower than the average costs. This implies that using a weighted average measure to aggregate the costs across treatments within a given consumption option and disease, although not perfect, is reasonable.

Moving on to other aspects, around 95% of consumers in our data consumed at most one type of service in a given week; in our estimation sample, this number is 100%. Hence, in our model, we assume that consumers make discrete choices about healthcare consumption (whether to choose primary preventive care,

Figure 1. (Color online) Cost-Sharing for Individuals During the Consumption Stage

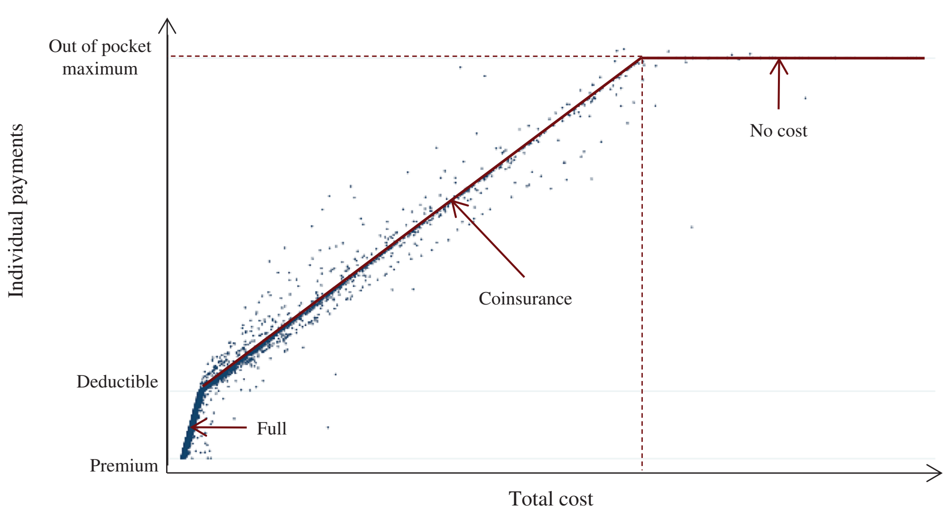


Table 2. Summary Statistics of Consumption Options for Some Chronic Diseases

	Type 2 diabetes	Heart disease	Prostate cancer
Secondary preventive care			
Number of different treatments used by consumers in the data	5	6	4
Weighted average cost of treatments (std. dev. of costs)	159.2 (21.3)	410.5 (29.9)	408.6 (24.9)
Primary preventive care			
Number of different treatments used by consumers in the data	8	10	6
Weighted average cost of treatments (std. dev. of costs)	458.9 (22.4)	862.7 (37.8)	752.5 (37.1)
Curative care			
Number of different treatments used by consumers in the data	4	8	4
Weighted average cost of treatments (std. dev. of costs)	3,423.7 (112.9)	5,382.9 (153.6)	3,682.4 (129.4)

secondary preventive care, curative care, or no care) on a weekly basis. Furthermore, we model only the consumption decisions that are related to chronic diseases. We do not model consumers' visits that are not related to chronic diseases, for two reasons. First, the fraction of such visits (5.3%) is small, and the expenditure on these visits is even smaller. Second, as we discussed in Section 1, the consumer's consumption choice set for nonchronic diseases does not consist of primary preventive, secondary preventive, and curative care. Therefore, the modeling framework that we propose in this paper does not apply to nonchronic diseases.

Our data consist of individual insurance policy holders who were continuously enrolled for the entire period. We consider only those consumers who have purchased individual plans and not family plans. This is because in our modeling framework, we model the insurance plan and consumption decisions based on the health status of a single individual. This leaves us with 2,833 insurance contract holders for estimation. In our estimation sample, 92% of consumers have taken preventive or curative care prescribed for chronic diseases, and 34.2% of consumers have taken preventive or curative care for more than one chronic disease. The dependent variables are consumers' annual insurance plan choices (i.e., whether to choose basic, medium, or comprehensive plan) and weekly consumption choices (i.e., whether to choose curative care, primary preventive care, secondary preventive care, or no care). The independent variables are annual premiums and cost-sharing characteristics for each insurance plan in a consumer's choice set and the actual price paid by each consumer for each option. From the total cost for each option and cost-sharing characteristics of the chosen plan, we construct the actual price paid by each consumer for each option. The procedure for doing so is explained in Section 3.

In our sample, 2.9% of consumers did not purchase any of the three consumption options, 52.8% purchased all three options at least once, and 79.6% purchased secondary preventive care (i.e., diagnostic tests) at least once. Between primary preventive and curative care, 14.3% of consumers did not use curative care but used primary preventive care, 20.1% did not use primary preventive care but used curative care, and 58.9% used both primary preventive and curative care at least once during their entire purchase history. Tables 3 and 4 present consumers' insurance plan choices and plan-switching behavior over three years, respectively, and Table 5 provides the summary statistics of health insurance and claim information across the three types of plans. We next provide the model-free and reduced-form insights that follow from the data.

2.1. Model-Free and Reduced-Form Insights

Observe in Table 3 that, on average, across the three years, 22% of consumers chose basic plans, 35% chose medium plans, and 43% chose comprehensive plans. Also, observe in Table 5 that as we move from the basic plan to the comprehensive plan, the number of visits per consumer per year and healthcare cost per consumer per year increase for all three healthcare services; importantly, as we move from the basic plan to the comprehensive plan, the increase in number of visits is the largest for curative care and the smallest

Table 3. Percentage of Plan Choices

	Basic (%)	Medium (%)	Comprehensive (%)
2005	22.2	33.9	43.9
2006	22.5	34.1	43.4
2007	21.7	35.7	42.6

Table 4. Switching Patterns of Insurance Purchases

	2006 (%)		
2005	Basic	Medium	Comprehensive
Basic	91.5	3.2	5.3
Medium	2.9	94.3	2.8
Comprehensive	3.0	3.1	93.9
	2007 (%)		
2006	Basic	Medium	Comprehensive
Basic	91.1	3.3	5.6
Medium	1.9	93.2	4.9
Comprehensive	1.1	7.3	91.6

Table 5. Summary Statistics Across Insurance Plans

	Basic	Medium	Comprehensive
Premium ($)	873.2	1,683.8	2,797.5
Copayment ($)	19.2	29.8	15.8
Coinsurance rate (%)	27	14	2
Deductible ($)	1,008.8	356.8	112.6
Out-of-pocket maximum ($)	3,107.6	1,020.5	275.3
Avg. no. of primary preventive care visits/year	8.7	12.7	15.3
Avg. no. of secondary preventive care visits/year	1.4	1.5	2.2
Avg. no. of curative care visits/year	0.9	1.3	4.5
Total cost paid by consumer and insurer for primary preventive care/year ($)	150.2	340.1	1,683.7
Total cost paid by consumer and insurer for secondary preventive care/year ($)	28.1	28.1	54.9
Total cost paid by consumer and insurer for curative care/year ($)	1,427.5	3,928.5	40,327.6

for primary preventive care. These observations shed some light on high healthcare costs—a majority of consumers choose comprehensive plans, and among these consumers, there is high usage of curative care.

There are several reasons that consumers choose comprehensive plans. For instance, they could have low health status, or they could have high risk aversion toward uncertainty in their health status, or they could have low price sensitivity, etc. Furthermore, there are two reasons that consumers in comprehensive plans consume curative care. First, they may have low health status, which necessitates their consumption of curative care; second, they may have moderate health status (along with high uncertainty in their health status and/or low price sensitivity), which makes them consume curative care because they do not bear most of its cost. Recall that the inefficiencies due to moral hazard stem from the second reason. This implies that to accurately identify the extent of moral hazard, we need a structural approach that can separately identify heterogeneity in health status from other sources of heterogeneity (such as heterogeneity in price sensitivity, uncertainty in health status, etc.), which is what we will try to accomplish in our proposed model in Section 3.

We next provide reduced-form insights for healthcare consumption decisions. We estimate a hazard model with multiple destinations (Lancaster 1990), where the destinations are purchases of the informative service (i.e., secondary preventive care) or either of the two investment services (i.e., curative care or primary preventive care). Consider the following conditional hazards that capture consumer i's propensity to purchase curative, primary preventive, or secondary preventive services in week t after her last purchase of any service

$$\theta_{i,c}(t) = \exp(\alpha_{i,c} + \gamma_{1,c} t)\exp\big(\gamma_{2,c}\tau + \beta_{c,c}N_{\tau,c} + \beta_{pp,c}N_{\tau,pp} + \beta_{sp,c}N_{\tau,sp} + \beta_{p,c}p_{c,\tau}\big), \quad (1)$$

$$\theta_{i,pp}(t) = \exp(\alpha_{i,pp} + \gamma_{1,pp} t)\exp\big(\gamma_{2,pp}\tau + \beta_{c,pp}N_{\tau,c} + \beta_{pp,pp}N_{\tau,pp} + \beta_{sp,pp}N_{\tau,sp} + \beta_{p,pp}p_{pp,\tau}\big), \quad (2)$$

$$\theta_{i,sp}(t) = \exp(\alpha_{i,sp} + \gamma_{1,sp} t)\exp\big(\gamma_{2,sp}\tau + \beta_{c,sp}N_{\tau,c} + \beta_{pp,sp}N_{\tau,pp} + \beta_{sp,sp}N_{\tau,sp} + \beta_{p,sp}p_{sp,\tau}\big), \quad (3)$$

where $q_{i,c}(t)$, $q_{i,pp}(t)$, and $q_{i,sp}(t)$ are consumer i's conditional hazards for curative, primary preventive, and secondary preventive care, respectively. The first term on the right-hand side (RHS) of Equations (1)–(3) constitutes the baseline conditional Gompertz hazards, which are functions of weeks lapsed since the last purchase of any of the three healthcare services, t. The second term on the RHS of Equations (1)–(3) captures the impact of the following covariates: (i) weeks elapsed from the first period in the data to the period in which any of the three services was last purchased, τ; (ii) the number of curative, primary preventive, and secondary preventive care services purchased by the consumer in the first τ weeks of data, $\{N_{t,c}, N_{t,pp}, N_{t,sp}\}$; and (iii) the out-of-pocket expenses incurred by the consumer if she were to choose any of the three services in period $\tau + t$, $\{p_{c,\tau}, p_{pp,\tau}, p_{sp,\tau}\}$.

The estimates of all parameters are reported in Table 6. In what follows, we discuss only the key parameters of interest. The parameters $\{\alpha_{i,c}, \alpha_{i,pp}, \alpha_{i,sp}\}$ capture the consumer's propensity to buy curative, primary preventive, and secondary preventive care in the first period of the data. These are assumed to be distributed across the population as $\{\alpha_{0,i,c}, \alpha_{0,i,pp}, \alpha_{0,i,sp}\} \sim N(\{\bar{\alpha}_c, \bar{\alpha}_{pp}, \bar{\alpha}_{sp}\}, \Omega)$. From the covariance matrix, Ω, we get significant estimates (at $p < 0.05$) of correlations between curative and primary preventive care use (0.191), curative and secondary preventive care use (0.352), and secondary and primary preventive care use (0.372). The larger positive correlations between the informative service (i.e., secondary preventive care) and each of the two investment services (i.e., curative and primary preventive care) implies that consumers who have a high preference for curative or primary preventive care also have a high preference for secondary preventive care.

The parameters $\{\gamma_{1,c}, \gamma_{1,pp}, \gamma_{1,sp}\}$ capture the impact of t (time elapsed since the last purchase of any service) on the consumer's propensity to purchase curative, primary preventive, and secondary preventive care, respectively, in the focal period $\tau + t$. Since the consumer's propensity to choose an investment service is inversely related to her health status, positive values of $\gamma_{1,c}$ and $\gamma_{1,pp}$ will imply that the health status of consumers degrades over time in the absence

Table 6. Parameter Estimates of the Multiple-Destination Hazard Model

Parameter	Estimate (standard error)		
	Curative	Primary preventive	Secondary preventive
Baseline hazard			
Mean intercepts (γ_0)	1.247 (0.385)	1.528 (0.478)	1.517 (0.621)
Time t for any service purchased lapsed (γ_1)	0.317 (0.117)	0.214 (0.091)	0.142 (0.089)
Impact of the other covariates			
Time τ for the particular service purchased lapsed (γ_2)	0.354 (0.125)	0.185 (0.069)	0.164 (0.101)
Number of past purchases of curative care ($\beta_{c,.}$)	−0.349 (0.112)	−0.185 (0.068)	−0.102 (0.062)
Number of past purchases of primary preventive care ($\beta_{pp,.}$)	−0.226 (0.106)	−0.121 (0.050)	−0.084 (0.049)
Number of past purchases of secondary preventive care ($\beta_{sp,.}$)	−0.064 (0.028)	−0.067 (0.044)	−0.075 (0.030)
Out-of-pocket expense of the specific service (β_p)	−0.085 (0.035)	−0.142 (0.044)	−0.151 (0.068)
Unobserved heterogeneity (Cholesky)			
Diagonal	C_{11} = 0.527 (0.231)	C_{22} = 0.546 (0.252)	C_{33} = 0.532 (0.241)
Off-diagonal	C_{21} = 0.106 (0.051)	C_{31} = 0.215 (0.095)	C_{32} = 0.187 (0.069)

of consumption of the investment services. Similarly, since the propensity to choose an informative service is positively related to the extent of the consumer's uncertainty over her health status, a positive value of $\gamma_{1,sp}$ would imply that consumers' uncertainty increases over time in the absence of consumption. We get significant estimates (at $p < 0.05$) of $\{\gamma_{1,c}, \gamma_{1,pp}\}$ of $\gamma_{1,c} = 0.317$, and $\gamma_{1,pp} = 0.214$. However, the estimate of $\gamma_{1,sp} = 0.142$) is not significant at $p < 0.05$. This indicates that whereas the health status of consumers degrades significantly over time in the absence of consumption, their uncertainty over health status does not.

The parameters $\{\gamma_{2,c}, \gamma_{2,pp}, \gamma_{2,sp}\}$ capture the impact of τ (weeks elapsed from the first period in the data to the period in which any of the three services was last purchased) on the consumer's propensity to purchase curative, primary preventive, and secondary preventive care, respectively, in focal period $t+\tau$. They have a similar interpretation as parameters $\{\gamma_{1,c}, \gamma_{1,pp}, \gamma_{1,sp}\}$. Similar to $\{\gamma_{1,c}, \gamma_{1,pp}, \gamma_{1,sp}\}$, we get significant positive estimates (at $p < 0.05$) of $\gamma_{2,c}$ and $\gamma_{2,pp}$, but not $\gamma_{2,sp}$.

The parameters $\beta_{c,c}$ and $\beta_{pp,c}$ capture the impact of the number of past purchases of curative and primary preventive care, respectively, on the consumer's current propensity to purchase curative care, and the parameters $\beta_{c,pp}$ and $\beta_{pp,pp}$ capture the impact of number of past purchases of curative and primary preventive care, respectively, on the consumer's current propensity to purchase primary preventive care. We get significant negative estimates (at $p < 0.05$) of these parameters of $\beta_{c,c} = -0.349$, $\beta_{pp,c} = -0.226$, $\beta_{c,pp} = -0.185$, and $\beta_{pp,pp} = -0.121$. This shows that, all else being the same, the consumer's propensity to repurchase an investment service decreases with every additional (past) purchase of an investment service. This in turn indicates that each successive use of an investment service has a diminishing marginal impact on the consumer's utility. Because $|\beta_{c,c}| > |\beta_{pp,c}|$ and $|\beta_{c,pp}| > |\beta_{pp,pp}|$, it follows that a prior curative care service decreases the propensity of repurchasing an investment service more than a prior primary preventive care service. Since the propensity to choose an investment service is inversely related to the consumer's health, this indicates that curative care provides a much greater boost to the consumer's health compared to primary preventive care.

The parameter $b_{sp,sp}$ captures the impact of the number of past purchases of secondary preventive care on the consumer's current propensity to purchase secondary preventive care. We get a significant negative estimate (at $p < 0.05$) of this parameter of $b_{sp,sp} = -0.075$. This implies that, all else being the same, the consumer's propensity to repurchase secondary preventive care decreases with every additional (past) purchase of secondary preventive care, which in turn indicates that each successive use of secondary preventive care has a diminishing marginal informational value.

In summary, the key insights are that the usage of curative care increases substantially as we move from the basic plan to the comprehensive plan, consumers' health degrades over time in the absence of treatments, their uncertainty over health does not change significantly over time in the absence of informative services, each successive use of an investment service has a diminishing marginal impact on a consumer's utility, and each successive use of secondary preventive care has a diminishing marginal informational value.

3. Model

We propose an economic model of consumers' healthcare decisions in which their periodic (weekly) consumption decisions are nested within their annual insurance plan decisions. Consumers decide the type of health insurance plans on an annual basis, based on the premium charged and their expected healthcare consumption in the future. Conditional on the health insurance plan, they then decide their periodic consumption decisions, which include taking primary preventive care, secondary preventive care, curative care, or the no-consumption option. We assume that the physician is a perfect agent and acts in the best interest of the patient. In other words, the healthcare consumption decisions are made by the consumer based on her beliefs about her health status and the out-of-pocket expenses she would incur while consuming the different options. This is a standard assumption in the healthcare literature (Felder and Mayrhofer 2011) and a reasonable one in our context, because the level of informational asymmetry between doctor and patient will be small in the context of chronic diseases (Vera-Hernandez 2003).

3.1. Insurance Plan and Healthcare Consumption Decisions

Consider $i = 1, \ldots, I$ consumers who make annual insurance choices from a set of $j = 1, \ldots, J$ health insurance plans at year $a = 1, \ldots, A$. In our data, the number of available insurance plans for each consumer is $J = 3$, where $j = 1$ indicates the basic plan, $j = 2$ the medium plan, and $j = 3$ the comprehensive plan. The characteristics of the three plans differ in terms of annual premiums, deductibles, coinsurance rates, and out-of-pocket maximums. Since all consumers are enrolled in one of the three plans over the entire period of the data, we model only the insurance plan decisions in terms of which of the three plans to purchase (i.e., we do not allow for an outside option in insurance plan decisions). We use d_{ija} to denote the consumer's choice of insurance plan j at year a as

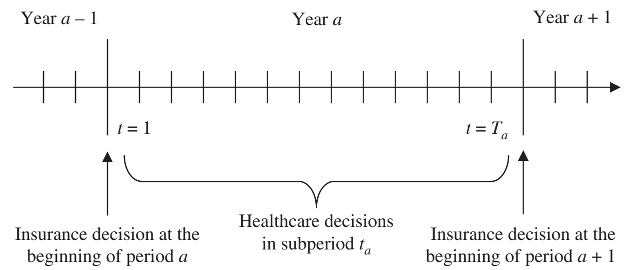
$$d_{ij,a} = \begin{cases} 1 & \text{if consumer } i \text{ chooses plan } j \text{ at} \\ & \quad \text{the beginning of year } a, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Conditional on the insurance plan chosen in year a, the consumer makes her consumption decision k in each period (i.e., week) $t = 1, \ldots, T$. The consumption decision entails choosing either primary preventive care (*pp*), secondary preventive care (*sp*), curative care (*c*), or the no-consumption option, (*no*). Let another variable c_{ikt} represent the healthcare consumption decision, where $k \in \{pp, sp, c, no\}$

$$c_{ikt} = \begin{cases} 1 & \text{if consumer } i \text{ chooses healthcare} \\ & \quad \text{service } k \text{ at week } t, \\ 0 & \text{otherwise.} \end{cases} \tag{5}$$

In a given period, these alternatives are mutually exclusive such that $\sum_k^{c_{ikt}} = 1$. See Figure 2 for an illustration of the timeline of the two decisions. In what follows, we discuss the factors that influence the two decisions: the consumer's health status, the investment and informative effects of the consumption options, and the insurance plan's pricing components. For notational ease, we omit the subscript i for the consumer in the remainder of this section. We will introduce the subscript for the consumer when we discuss heterogeneity in Section 4.

Figure 2. Timeline of Sequence of Health Insurance/Healthcare Decision



3.2. True Health Status, Degradation, and Investment Effects

The consumer's health status is a key factor that influences her consumption and health insurance decisions. Following Grossman's (1972) health production framework, we treat health status as a (latent) human capital stock. Let H_t be the consumer's true health status at the end of period t. Unlike quality learning literature in which the true quality is time invariant, consumers' true health status can evolve over time due to degradation and investment effects. Starting with degradation, we allow for a consumer's health to degrade over time in the absence of consumption of any healthcare service. The rationale for allowing for degradation in health status follows from the reduced-form evidence presented in Section 2.1, where we showed that a consumer's propensity to purchase an investment service increases with time elapsed since its last purchase, and also follows from the prior empirical literature, which has shown that consumers' health degrades in absence of any treatment (e.g., Khwaja 2010, Gupta and Li 2011, Einav et al. 2013). We model the degradation in a consumer's true health status over time in the absence of consumption of any healthcare service as

$$H_t = H_{t-1} - \delta_d, \tag{6}$$

where δ_d represents the degradation rate, which is the deterioration in the consumer's true health status over each period (i.e., each week). Moving on to investment effects, we allow for healthcare services to provide a quantum boost to the consumer's health status. We assume that the investment effect of any healthcare service impacts the consumer's health status in a linear fashion. We thus model the evolution of the consumer's true health status in the presence of degradation and investment effects as

$$H_t = H_{t-1} - \delta_d + \sum_k \delta_k c_{kt}, \tag{7}$$

where δ_k is the investment effect of option $k \in \{pp, sp, c, no\}$ consumed in period t. Recall that secondary preventive care comprises diagnostic tests and screening exams, which do not provide any boost to health status.

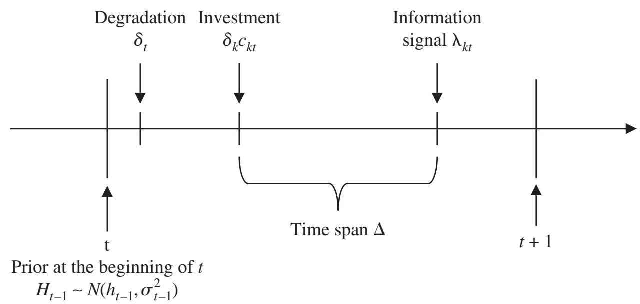
Accordingly, we set the investment effect of secondary preventive care to zero (i.e., $\delta_{sp} = 0$). Similarly, we set the investment effect of the no-consumption option to zero (i.e., $\delta_{no} = 0$) and estimate only the investment effects of primary preventive and curative care. Note that setting the investment effect of the no-consumption option to zero does not imply that the consumer's health status will necessarily degrade in the absence of consumption. This is because we do not impose any a priori restriction on the degradation parameter δ_d when estimating the model. If the estimate of δ_d is positive, it will imply that the consumer's health status will degrade in the absence of consumption; if it is negative, it will imply that the consumer's health status will increase in the absence of consumption, which could be because of exercise, proper diet, etc.

Before we proceed, it is worth pointing out two assumptions implicit in Equation (7). *First*, we assume that each successive use of a given healthcare service generates a constant boost to the health status. This may not be the case if the consumer chooses the most effective medical procedure to begin with and follows it up with less effective medical procedures, which would imply a diminishing impact of each subsequent service on the consumer's health status. The reason for this assumption is that we can identify the impact of a healthcare service only on the consumer's utility (which is a function of the health status) and not on her health status per se. Thus, in our model, while we assume that each successive use of a given healthcare service generates a constant boost to the health status, we allow for a diminishing impact of each subsequent service on the consumer's utility by allowing for the consumer's utility to be concave in her health status (more on this in Section 3.5). *Second*, we assume that the benefits of a treatment are obtained right after the purchase of the treatment, which implies that the consumer gets the quantum boost in her health status on the day when she purchases the treatment. The reason for this assumption is that our data do not indicate whether the consumer purchased a one-month supply or a two-month supply, etc., or whether the consumer was compliant with the medication or not. Therefore, for simplicity, we assume that the benefit of any treatment, including drugs, is obtained right after the purchase of the treatment.

3.3. Perceived Health Status and Informative Effects

Because consumers are not well informed about their true health status, they will have uncertainty about it. Therefore, their insurance plan and consumption decisions will be based on their beliefs about their true health status. Consumers can learn about their true health status via the informative effects of consumption options. To model informative effects, we assume

Figure 3. Illustration of Degradation and Investment/Information Effects in Period t



that a consumption option provides a noisy signal to the consumer about her true health status, and the consumer learns about her true health status from these signals in a Bayesian fashion. The rationale for this assumption follows from the reduced-form evidence presented in Section 2.1, where we showed the diminishing marginal informational value of each successive use of diagnostic tests/screening exams. To understand the evolution of consumers' beliefs about their health status from the end of period $t-1$ to the end of period t, see Figure 3, which illustrates the sequence of the consumer's periodic degradation and investment and informative effects. In this sequence, we start with the consumer's beliefs about her true health status at the beginning of period t (or end of period $t-1$) as

$$H_{t-1} \sim N(h_{t-1}, \sigma^2_{t-1}). \tag{8}$$

Following that, we update the priors in Equation (8) based on degradation of the health status in period t and the investment effects from consumption of option k in period t. By the end of this step, we get the consumer's beliefs using Equations (7) and (8) as

$$H_t^{\Delta} \sim N(h_{t-1} - \delta_d + \delta_k c_{kt}, \sigma^2_{t-1}). \tag{9}$$

In the next step, we update the priors in Equation (9) based on the informative effects from the consumption of option k in period t. Let λ_{kt} be the noisy consumption signal that the consumer receives about her true health status from the consumption of option k in period t. This is given as

$$\lambda_{kt} = H_t + \sigma_k \eta_{kt}. \tag{10}$$

The first term on the RHS of Equation (10), H_t, represents the consumer's true health status at the end of period t after she has experienced degradation in her health in period t and the investment effect of the option consumed in period t. The second term, $\sigma_k \eta_{kt}$, represents the signal noise of option k, where η_{kt} is a standard normal random variable, and σ_k is the standard deviation of the signal noise, which is a measure of the informational inaccuracy of service k. If $\sigma_k = \infty$, it implies that healthcare option k does not provide any information to the consumer about her health status.

We assume that consumers do not get to learn about their health status in the absence of consumption of a healthcare service. We thus set the informative effect of the no-consumption option to zero, i.e., $\sigma_{no} = \infty$. The rationale for this assumption follows from the reduced-form evidence presented in Section 2.1, which indicated that the consumer's uncertainty over her health status does not significantly change over time in the absence of consumption of any healthcare service. Since we would expect self-learning (i.e., being able to learn about one's true health in the absence of consumption of any healthcare service) to decrease the consumer's uncertainty over time in the absence of consumption, this reduced-form result implies that the consumer's self-learning about her health status is not significant. Moreover, our paper focuses on chronic diseases, where it is challenging for consumers to self-learn. For instance, it is difficult for consumers to know their cholesterol levels or sugar levels without blood tests. Among the three healthcare services, note that it is only the secondary preventive care that provides information to consumers about their current health status, and not curative care or primary preventive care. We thus set the informative effects of curative and primary preventive care to zero,[2] i.e., $\sigma_c = \sigma_{pp} = \infty$, and estimate only the informative effects of secondary preventive care, σ_{sp}. Substituting H_t given in Equation (7) into Equation (10), we get the following expression for λ_{kt} in terms of the consumer's true health status at the beginning of period t:

$$\lambda_{kt} = H_{t-1} - \delta_d + \delta_k c_{kt} + \sigma_k \eta_{kt}. \tag{11}$$

Given the prior in Equation (6) and the signal in Equation (8), the consumer's posterior beliefs at the end of period t will be distributed as $H_t \sim N(h_t, \sigma_t^2)$, where

$$h_t = \frac{(h_{t-1} - \delta_d + \sum_k^{\delta_k c_{kt}})/\sigma_{t-1}^2 + \sum_k (c_{kt}\lambda_{kt}/\sigma_k^2)}{1/\sigma_{t-1}^2 + \sum_k (c_{kt}/\sigma_k^2)}, \tag{12a}$$

$$\frac{1}{\sigma_t^2} = \frac{1}{\sigma_{t-1}^2} + \sum_k \frac{c_{kt}}{\sigma_k^2}. \tag{12b}$$

Equations (12a) and (12b) represent the evolution of the mean and variance of health status from the consumer's perspective. To complete the specification of evolution of health status, we represent the consumer's prior beliefs at the beginning of period $t = 1$ in year $a = 1$ by $H_0 \sim N(h_0, \sigma_0^2)$. To deal with the initial conditions problem with respect to these prior beliefs, we follow the prior literature that has modeled Bayesian learning of consumers' beliefs (Crawford and Shum 2005, Mehta et al. 2008, Narayanan and Manchanda 2009, Zhang 2010). We assume that at the beginning of $\{a = 1, t = 1\}$, the consumer in a given segment k has rational expectations about her health status, which is that the mean and variance of prior beliefs of a consumer in a given segment k vis-à-vis her health status (i.e., $h_{0,k}$ and $\sigma_{0,k}^2$) are also the mean and the variance of the distribution of the true health status of consumers in segment k at the beginning of $\{a = 1, t = 1\}$.

3.4. Out-of-Pocket Expenses

The pricing structure of health plans offered by the insurer affects consumers' health insurance and consumption decisions through out-of-pocket expenses incurred by the consumers when purchasing the insurance plans and when consuming the healthcare services. If a consumer chooses a plan j, her out-of-pocket expense will be the premium ($Prem_{j,a}$) that she pays at the beginning of year a. Once enrolled in plan j, her out-of-pocket expense when choosing a healthcare consumption option will depend on the actual price of the chosen option and cost-sharing features of the plan. The out-of-pocket expenses for consumption decisions are constructed as follows. Let p_{k,j,t_a} be the consumer's out-of-pocket expense if she were to consume healthcare option k in period t of year a if she were enrolled in insurance plan j. We represent p_{k,j,t_a} as $p_{k,j,t_a} = p'_{k,j,t_a} + cop_{k,j,a}$, where p'_{k,j,t_a} is consumer i's marginal effective price for option k in period t, and $cop_{k,j,a}$ is the copayment amount incurred by consumer i for using option k in year a if she were enrolled in plan j. The consumer's marginal effective price is given by

$$p'_{k,j,t_a} = \begin{cases} cost_{k,j,t_a}, \\ \qquad \text{if } cost_{k,j,t_a} \le r_{j,t_a} \text{ and } \sum_{\tau=1}^{t_a-1} p_{k,\tau} < opm_{j,a}, \\ r_{j,t_a} + (cost_{k,j,t_a} - r_{j,t_a}) ci_{k,j,a}, \\ \qquad \text{if } cost_{k,j,t_a} > r_{j,t_a} > 0 \text{ and } \sum_{\tau=1}^{t_a-1} p_{k,\tau} \ge opm_{j,a}, \\ 0, \quad \text{if } \sum_{\tau=1}^{t_a-1} p_{k,\tau} \ge opm_{j,a}, \end{cases} \tag{13}$$

where $opm_{j,a}$ denotes the annual out-of-pocket maximum of plan j chosen by consumer i in year a, and $ci_{k,j,a}$ is the coinsurance rate faced by consumer i in year a for option k. Furthermore, $cost_{k,j,t_a}$ is the overall cost (sum of healthcare costs incurred by the consumer and insurer) of healthcare option k in period t of year a, whose operationalization is discussed in Section 1 of the web appendix. Finally, r_{j,t_a} is the dollar amount remaining in the consumer's annual deductible in period t, which is defined as

$$r_{j,t_a} = \begin{cases} ded_{j,a} & \text{if } t = 1, \\ \max\{0, r_{j,t-1_a} - p_{k,j,t-1_a} c_{k,t-1_a}\} & \text{if } t = 2, \ldots, T_a, \end{cases} \tag{14}$$

where T_a denotes the last period of year a, and $ded_{j,a}$ denotes the total deductible amount paid by consumer i for the chosen plan j in year a before the insurer makes any benefit payments. Note that although we do not model doctor visits related to nonchronic diseases, we take into account the expenditures incurred by the consumer on such visits when calculating r_{j,t_a} in Equation (14).

3.5. Per-Period (Weekly) Consumption Utility

We define the consumer's period t utility for choosing healthcare option $k \in \{pp, sp, c, no\}$ as a function of her out-of-pocket expenses p_{k,j,t_a} and true health status H_t that would result if she were to consume that option k in period t. We assume that the utility is concave with respect to the health status (i.e., $u''(H_t) < 0$), which is synonymous with risk aversion in health. The rationale for assuming concavity follows from the reduced-form evidence presented in Section 2.1, where we showed that each successive use of an investment service has diminishing marginal impact on the consumer's utility. Our rationale also follows from the prior medical literature as well as the prior empirical literature in health economics, which has shown that consumers are risk averse in their health status (Llewellyn-Thomas et al. 1982, Hellinger 1989, Manning and Marquis 1996, Khwaja 2010, Einav et al. 2012, 2013). Moreover, in Section 3.6, we will discuss some commonplace observations in the healthcare market that can be explained only by risk aversion in health. We consider the Constant Absolute Risk Aversion (CARA) functional form for the consumption utility for choosing option k, which satisfies this assumption (Chan and Hamilton 2006). This is given as[3]

$$u_{kt} = -\exp(-rH_t) - \alpha p_{k,j,t_a} + \varepsilon_{kt}. \tag{15}$$

In Equation (15), r (>0) is the consumer's degree of absolute risk aversion, α is the consumer's price sensitivity, and ε_{kt} is the econometrician's error that is assumed to be independent and identically extreme value distributed and captures factors such as health shocks that are unobserved to the researcher but that can influence the consumption decisions. For each consumption option k, the consumer faces a different out-of-pocket expense and a different evolution of her health status in the future. The consumer's information set when making a decision at the beginning of period t consists of (i) her beliefs of her health status at the end of period $t-1$, which are distributed as $H_{t-1} \sim N(h_{t-1}, \sigma_{t-1}^2)$, (ii) her out-of-pocket expenses for all consumption options in period t and onward, and (iii) unobserved factors that influence her consumption decision in period $t, \{\varepsilon_{kt}\}$. Conditional on this information set, her period t expected utility for choosing option k follows from Equation (15) as

$$\begin{aligned} E_t u_{kt} = &-E\big(\exp(-rh_t + r^2\sigma_t^2/2) \,\big|\, h_{t-1}, \sigma_{t-1}, c_{kt} = 1\big) \\ &- \alpha p_{k,j,t_a} + \varepsilon_{k,t}. \end{aligned} \tag{16}$$

In Equation (16), the term $r^2\sigma_t^2/2$ denotes the consumer's risk premium in period t, in which the specification of σ_t^2 conditional on $\{h_{t-1}, \sigma_{t-1}, c_{kt}\}$ follows from Equation (12b) as

$$\sigma_t^2 = \frac{\sigma_{t-1}^2 \sigma_k^2}{c_{kt}\sigma_{t-1}^2 + \sigma_k^2}. \tag{17}$$

Equation (17) shows that informative effects, as captured by the inverse of σ_k^2, serve to decrease the risk premium in the expected utility by decreasing uncertainty in perceived health status. In Equation (16), the specification of h_t conditional on the information set at the beginning of time t, $\{h_{t-1}, \sigma_{t-1}, c_{kt} = 1\}$, can be derived using Equation (12a), which specifies h_t as a function of h_{t-1}, δ_d, δ_k, σ_{t-1}, σ_k, and the signal λ_{kt}. Note that among these variables, although the consumer knows the values of h_{t-1}, δ_d, δ_k, σ_{t-1}, and σ_k at the beginning of period t, she does not know the value of the signal λ_{kt}. This is because the signal from consuming option k is realized by the consumer only at the end of period t. Thus, from the consumer's perspective at the beginning of period t, λ_{kt} will be a random variable that will be distributed as per Equation (11). Substituting Equation (11) into (12a), we get the specification of h_t conditional on $\{h_{t-1}, \sigma_{t-1}, c_{kt} = 1\}$ from the consumer's perspective at the beginning of period t

$$h_t = h_{t-1} + (\delta_k c_{kt} - \delta_d) + \left(c_{kt} \frac{\sigma_{t-1}^2}{\sqrt{\sigma_k^2 + \sigma_{t-1}^2}} \upsilon_{kt} \right), \tag{18}$$

where the term υ_{kt} is a standard normal random variable. In Equation (18), the second term on the RHS, $\delta_k c_{kt} - \delta_d$, represents the change in the mean health status that stems from degradation and the investment effect of option k; the third term, $c_{kt}(\sigma_{t-1}^2/\sqrt{\sigma_k^2 + \sigma_{t-1}^2})\upsilon_{kt}$, represents the change that stems from the informative effect of option k (more on this in Section 3.6). This completes the discussion on the per-period utility. We next discuss the behavioral implications of the model.

3.6. Implications

Thus far, two implications on consumption decisions follow from the model discussed, and these pertain to the concavity of the utility and use of investment and informative services.

3.6.1. Concavity and the Use of Investment Services.

The concavity of the utility implies that the marginal utility of health decreases with the increase in health status. Since the marginal utility of health is directly related to the consumer's incentive for seeking an investment service, it implies the following: (a) The sicker the consumer, the greater her incentive to choose an investment service over the no-consumption option (and vice versa). This makes sense because consumers undergo treatments when they are sick and not when they are healthy. (b) The greater the degradation rate, the greater is the consumer's incentive to seek an investment service. For instance, if the degradation rate is very high and the consumer might die, then this will affect her decision, and she will be more likely to get a curative care operation to boost her health status. (c) The sicker the consumer, the greater her incentive to choose an investment service with high investment effects over an investment service with low investment effects. This also makes sense because consumers typically purchase curative care when their health status is low to moderate and choose primary preventive care when their health status is moderate to high.

3.6.2. Concavity and the Use of Informative Services. Recall from Section 3.5 that when a consumer consumes an informative service at the beginning of period t, she receives a learning draw about her health status at the end of period t, $(\sigma_{t-1}^2/\sqrt{\sigma_k^2+\sigma_{t-1}^2})\upsilon_{kt}$. This learning draw provides an option value to the consumer, because it enables her to decide her future course of treatments in periods $t+1$ and onward. For instance, if the consumer receives a very negative value of the learning draw at the end of period t, she will opt for a treatment with large investment effects in period $t+1$, which will increase her optimal utility in period $t+1$. Similarly, if she receives a positive learning draw, she will opt for the no-consumption option in period $t+1$. This showcases the importance of assuming concavity of the utility. If the utility were not concave and were instead linear in the health status, there would be no option value of the learning draw, because the probability of purchasing an investment service in period $t+1$ would be the same, regardless of whether the consumer received a positive or negative learning draw from the informative service.

3.7. Weekly Consumption and Annual Insurance Plan Purchase Decisions

We assume that consumers are forward-looking in their consumption and insurance plan decisions. The rationale for this assumption follows from Aron-Dine et al. (2015), who find statistically significant elasticities of healthcare consumption with respect to future prices, thereby rejecting the null model of myopic behavior in the healthcare market. We thus model the weekly consumption and annual insurance plan decisions in a nested dynamic framework. We start with the consumer's weekly consumption decision in period t given her insurance plan coverage and her perceived health status at the beginning of period t, h_{t-1}. Following that, we specify her annual insurance choice decisions.

3.7.1. Consumption Decisions in Period t. The consumer chooses the healthcare option k, which maximizes her discounted lifetime expected utility in period t in year a derived from both consumption and subsequent insurance purchases. Denote the state variables at the beginning of period t in year a as $S_t^a \equiv \{h_{t-1}, \sigma_{t-1}, \{p_{lt}\}, \{\varepsilon_{lt}\} l\}$. Given the period t expected utility for choosing option k in Equation (16), we get the value function at period $t<T$ (where T is the last period of the insurance year) conditional on insurance plan j, and the state variables S_t^a is given as

$$V(S_t^a \mid j) = \max_k E\big[\big(-\exp\big(-rh_t + r^2\sigma_t^2/2\big) \\ -\alpha p_{k,j,t_a} + \varepsilon_{kt} + \beta V(S_{t+1}^a)\big) \,\big|\, S_t^a, k, j\big]. \quad (19)$$

In this equation, β is the weekly discount factor, set at 0.999 (which corresponds to an annual discount factor of 0.95). This is a standard value assumed in the healthcare literature, which includes papers that have modeled consumption and/or insurance plan decisions (e.g., Gilleskie 1998, Khwaja 2010), drug choice decisions (Crawford and Shum 2005), observational learning in the transplant market (Zhang 2010), and decisions on clinical trial participation (Chan and Hamilton 2006).

The RHS of Equation (19) is a function of state variables at the end of period t (or beginning of period $t+1$), $S_{t+1}^a \equiv \{h_t, \sigma_t, \{p_{l,t+1}\} l, \{\varepsilon_{l,t+1}\} l\}$, in which $\{h_t, \sigma_t\}$ are related to $\{h_{t-1}, \sigma_{t-1}\}$ via Equations (17) and (18) as $h_t = h_{t-1} + c_{kt}\delta_k - \delta_d + c_{kt}(\sigma_{t-1}^2/\sqrt{\sigma_k^2+\sigma_{t-1}^2})\upsilon_{kt}$ and $\sigma_t^2 = \sigma_{t-1}^2\sigma_k^2/(c_{kt}\sigma_{t-1}^2+\sigma_k^2)$. We next substitute these into Equation (19) and take the expectation over the learning draw, $(\sigma_{t-1}^2/\sqrt{\sigma_k^2+\sigma_{t-1}^2})\upsilon_{kt}$, that stems from informative effects of the chosen option k in period t and the period $t+1$ econometrician's errors, $\{\varepsilon_{l,t+1}\}$. As shown in Section 2 of the web appendix, this yields the consumer's value function based on her information set at the beginning of period t, $S_t^a \equiv \{h_{t-1}, \sigma_{t-1}, \{p_{lt}\} l, \{\varepsilon_{lt}\} l\}$, as

$$V(S_t^a \mid j) = \max_k\big\{-\exp\big(-r(h_{t-1} - \delta_d + \delta_k) + r^2\sigma_{t-1}^2/2\big) \\ -\alpha p_{k,j,t_a} + \varepsilon_{kt} + \beta E_{\upsilon_{kt}, \varepsilon_{lt+1}} V(S_{t+1}^a \mid S_t^a, k, j)\big\}.[4] \quad (20)$$

In the last time period $t=T$ of insurance year a, when an insurance decision will be made in the next time period, the value function is given by

$$V(S_T^a \mid j) = \max_k\big\{-\exp\big(-r(h_{T-1} - \delta_d + \delta_k) + r^2\sigma_{T-1}^2/2\big) \\ -\alpha p_{k,j,T_a} + \varepsilon_{kT} + \beta E_T W(S_1^{a+1} \mid S_T^a, k)\big\}, \quad (21)$$

where $W(S_1^{a+1} \mid S_T^a, k)$ represents the value of the future utilities at the beginning of the next year when the consumer makes another insurance decision, which we define in a moment. To compute the value functions, we employ a variant of the Keane and Wolpin (1994) approximation method, which we explain in Section 3 of the web appendix.

3.7.2. Insurance Plan Choice at the Beginning of Year a.

At the beginning of each year, the consumer chooses one of the $J=3$ insurance plans in her choice set based on premium ($Prem_{j,a}$) and her expected utilization of healthcare for the coming year. Thus, the value of utility of a particular insurance plan j in period $t=1$ of year a will be

$$VI_j(S_t^a) = -\alpha\, Prem_{j,a} + V(S_1^a \mid j) + \varphi_{j,a}, \quad (22)$$

where the coefficient for the premium, α, measures how much the consumer i values the premium and is the same as the parameter that captures the out-of-pocket expenses in consumption decisions, and $\varphi_{j,a}$ is an additive idiosyncratic error that is independent across time and consumers and follows a type 1 extreme value distribution. The term $V(S_1^a \mid j)$ is the continuation value at the beginning of the year from the consumption. From Equation (22), the expected lifetime utility at the beginning of year a is

$$W(S_1^a) = E_{T_{a-1}+1}(\max VI_j(S_1^a)). \quad (23)$$

4. Econometric Specification

4.1. Choice Probabilities

The choice probabilities of insurance plan and consumption decisions follow from Equations (19)–(23). However, note that the period t value functions in Equations (19)–(23) are defined from the consumer's perspective, where the consumer's health status at the beginning of period t, h_{t-1}, is known to the consumer. On the other hand, h_{t-1} is not known to the researcher. This is because h_{t-1} is a function of the past signals received by the consumer from periods $s=1$ to $s=t-1$ (see Equation (12a)), and, unlike the consumer, the researcher does not observe the values of these past signals. Therefore, before we use Equations (19)–(23) to specify the choice probabilities in period t, we need to first specify h_{t-1} from the researcher's perspective. The evolution of the health status from the researcher's perspective is derived in Section 4 of the web appendix. This is given as

$$h_{t-1} = h_0 - \delta_d(t-1) + \delta_c N_{c,t-1} + \delta_{pp} N_{pp,t-1} + \frac{1}{\beta_0 + N_{sp,t-1}} \cdot \left(\frac{\sigma_{sp}^2}{2} + \sigma_0 N_{sp,t-1}\eta_0 + \sum_{s=1}^{t-1} c_{sp,s}\sigma_{sp}\eta_{sp,s}\right), \quad (24)$$

where $\{h_0, \sigma_0^2\}$ are the mean and variance of the consumer's prior beliefs at the beginning of $\{t=1, a=1\}$; c_{ks} is a 0-1 indicator variable that takes a value of 1 if the consumer consumes option k ($\forall k \in \{sp, pp, c, no\}$) in period s; the terms $\{N_{sp,t-1}, N_{c,t-1}, N_{pp,t-1}\}$, respectively, represent the number of times the consumer has consumed secondary preventive care, curative care, and primary preventive care from periods $s=1$ to $s=t-1$; and $\beta_0 = \sigma_{sp}^2/\sigma_o^2$. The terms $\omega_t \equiv \{\eta_0, \{\eta_{ks}\}_{s=1,\ldots,t-1,k}\}$ are standard normal random variables that reflect the researcher's uncertainty in knowing the true value of h_{t-1}. Substituting h_{t-1} as given in Equation (24) into the value functions in Equations (20) and (21) will yield the period t value functions from the researcher's perspective. Given the period t value functions, the relevant probabilities follow immediately. These are the probability of choosing an insurance plan and the probability of seeking healthcare consumption during the policy year. Conditional on the insurance plan chosen by an individual and the errors ω_t, the probability of choosing no consumption, primary preventive, secondary preventive, or curative care is given by

$$\Pr(c_{kt_a} = 1 \mid d_{j,a}, \omega_{t_a}) = \frac{\exp(V_k(S_t^a(\omega_{t_a}) \mid d_{j,a}))}{\sum_{k'=\{pp,sp,c,no\}} \exp(V_{k'}(S_t^a(\omega_{t_a}) \mid d_{j,a}))}, \quad (25)$$

where $V_k(S_t^a(\omega_{t_a}) \mid d_{j,a})$ denotes the choice-specific value function for primary preventive care, secondary preventive care, and the no-consumption option (conditional on the errors ω_t and the insurance plan choice) as defined in Equations (20) and (21) after substituting the expression for h_{t-1} given in Equation (24). The choice probabilities of alternative insurance plans at the beginning of year a conditional on the errors are given as

$$\Pr(d_{ja} = 1 \mid \omega_{1_a}) = \frac{\exp(W_j(S_1^a(\omega_{1_a})))}{\sum_{j=1,\ldots,J} \exp(W_{j'}(S_1^a(\omega_{1_a})))}. \quad (26)$$

4.2. Parameters and Heterogeneity

We have introduced the following parameters till now: (i) degree of risk aversion, r; (ii) price sensitivity, α; (iii) weekly degradation in the consumer's health, δ_d; (iv) investment effects of primary preventive and curative care, $\{\delta_{pp}, \delta_c\}$; (v) noise in the secondary preventive care signal, σ_{sp}; and (vi) the mean and variance of the consumer's prior beliefs in the beginning of the first year of the estimation sample, $\{h_0, \sigma_0\}$. Next note that the risk aversion parameter r and the other parameters, $\{\delta_d, \delta_{pp}, \delta_c, \sigma_{sp}, h_0, \sigma_0\}$, always enter the utility in a multiplicative fashion as $\{r\delta_d, r\delta_{pp}, r\delta_c, r\sigma_{sp}, rh_0, r\sigma_0\}$. This can be seen in Equation (27), where we derive the period t value function from the researcher's perspective using Equations (20) and (24) in terms of the parameters $\{r, \alpha, \delta_d, \delta_{pp}, \delta_c, \sigma_{sp}, h_0, \sigma_0\}$

$$V(S_t^a) = \max_k \left\{ -\exp\left[-(rh_0) + (r\delta_d)t - (r\delta_c)N_{c,t} - (r\delta_{pp})N_{pp,t} \frac{(r\sigma_{sp})^2/2 + (r\sigma_0)N_{sp,t}\eta_0 + \sum_{s=1}^{t} c_{sp,s}(r\sigma_{sp})\eta_{sp,s}}{\beta_0 + N_{sp,t}} \right] - \alpha p_{k,j,t_a} + \varepsilon_{k,t} + \beta E_t V(S_{t+1}^a \mid S_t^a, k) \right\}. \quad (27)$$

Equation (27) shows that r cannot be identified from the other parameters, $\{\delta_d, \delta_{pp}, \delta_c, \sigma_{sp}, h_0, \sigma_0\}$. We thus normalize r =1.[5] To capture unobserved heterogeneity, we use latent class segmentation in which we allow for the following parameters to differ across segments: (i) the prior mean and variance of the consumer's health status, $\{h_0, \sigma_0\}$; (ii) price sensitivity, α; (iii) weekly degradation rate, δ_d; (iv) investment effects of primary preventive and curative care, $\{\delta_{pp}, \delta_c\}$; and (v) noise in the secondary preventive care signal, σ_{sp}. Note that since we normalize $r = 1$, we capture heterogeneity in risk premiums by the heterogeneity in uncertainty about their health status.

The number of segments is chosen so that the Bayesian information criterion (BIC) is optimized. To estimate the parameters, let Θ denote the vector of parameters to be estimated. The likelihood for consumer i is

$$L_i(\Theta) = \int \cdots \int \prod_a \prod_{j=1}^{J} \Pr(d_{ijt_a} = 1)^{d_{ijt_a}} \cdot \left(\prod_t \prod_k \Pr(c_{ikt_a} = 1 \mid d_{ijt_a})^{c_{ikt_a}} \Phi(dh_{ik}; \Theta) \right) G(d\omega_i). \tag{28}$$

4.3. Discussion on Identification of Parameters

We start with the identification of the distribution of price sensitivities across the population. The price sensitivity parameter, α, enters as the coefficient of the out-of-pocket expenses in both the insurance plan and healthcare consumption decisions. Hence, its identification follows from the intertemporal and cross-sectional variation in the out-of-pocket expenses and the impact of this variation on both the insurance plan and consumption decisions. In insurance plan decisions, the out-of-pocket expenses vary over time for a given consumer because the insurance premiums vary from year to year, and they also vary across consumers in any given year because the premiums for the same plan are different across various employers. In consumption decisions, the out-of-pocket expenses vary over time for a given consumer because of nonlinear pricing induced by the deductibles and the out-of-pocket maximum limits in each plan and because these cost-sharing characteristics vary from year to year; they also vary across consumers with the same plan because the cost-sharing characteristics of the same plan are different across various employers. This cross-sectional and longitudinal variation of out-of-pocket expenses allows us to identify the distribution of price sensitivity.[6]

Given the price sensitivity, we next discuss the identification of parameters that determine the usage of an informative service (i.e., secondary preventive care). After controlling for the price effects, the usage of service k with only informative effects depends on (i) the informative effect of the service k, as captured by the inverse of the noise in option k's signal, σ_k, and (ii) the consumer's uncertainty about her health status at the beginning of $\{t = 1, a = 1\}$, σ_0. Recall from our discussion in Section 2.1 that the informative effect of a service k is identified by the impact of the number of past purchases of an informative service k on the consumer's current propensity to repurchase that service k, where the greater the informative effect of a service k, the greater the impact of number of prior purchases of that service k on the consumer's propensity to repurchase that service. This allows us to separately identify the informative effect of service k from the consumer's prior uncertainty in her health status.

We next discuss the identification of parameters that determine the usage of an option with only investment effects (i.e., curative and primary preventive care). After controlling for prices, the usage of option k with only investment effects depends on (i) the investment effect of that option, δ_k; (ii) the consumer's prior mean health status at the beginning of $\{t = 1, a = 1\}$, h_0; and (iii) her degradation rate, δ_d. We start with the identification of the degradation rate. Recall from Section 2.1 that the greater the degradation rate, the greater the increase in the consumer's propensity to repurchase an investment service with the time elapsed since its last purchase. Note that it is only the degradation rate, and not prior mean health status or investment effect, that results in this time dependence. Consequently, the degradation rate is identified by the impact of time elapsed since the last purchase of a treatment on the consumer's current propensity to purchase the investment service. Given the degradation rate, the investment effect of service k is identified by the impact of the number of past purchases of service k on the consumer's current propensity to repurchase that service k, where the greater the investment effect of a service k, the greater the impact of the number of prior purchases of that investment service k on the consumer's propensity to repurchase that service. This allows us to separately identify the investment effect from the prior mean health status.

5. Empirical Results

5.1. Model Fit and Comparison

Our data consist of purchase decisions of 2,833 consumers across a period of 3 years. We longitudinally split the data into two parts: the first two years are taken as the estimation sample and the third year is taken as the holdout sample. We compare the goodness of fit and predictive ability of our proposed model (Model I) with those of the three competing models.

Recall from Section 1 that prior literature has modeled only investment effects of healthcare consumption decisions and not their informative effects. Therefore, our first competing model (Model II) is in the spirit of what has been done in the prior literature—unlike the proposed model, it allows only for the investment effects of all three healthcare services and not their informative effects.

The second competing model (Model III) is a two-stage hidden Markov model that hews closely to the model proposed by Schweidel et al. (2011). The two-stage decision process in this model is also related to the category consideration model proposed by Ching et al. (2009). Its details are provided in Section 5 of the web appendix. It differs from our proposed model in two respects. *First*, in Model III, the consumer's health status does not evolve as per Equation (7). Instead, it evolves as a hidden Markov process where the consumers' propensities to consume any of the three services vary across the different hidden states. *Second*, it relaxes the assumption that we make in our proposed model that consumers make a multinomial healthcare consumption decision in each week (i.e., whether to choose primary preventive care, curative care, secondary preventive care, or the no-consumption option). It thereby allows for consumers to have periods of "no decision," in which they stay with the no-consumption option by default. This is done by modeling the consumer's consumption decision. Each week is modeled in two stages. The first stage determines whether or not the consumer is in a "no-decision phase." If the consumer is in a no-decision phase, then she will stay with the no-consumption option by default; if not, then she goes to the second stage, where she decides whether or not she should consume a healthcare service and, if so, then which service.[7] (In the second stage, the consumer can still choose the no-consumption option.)

The parameter estimates of the competing models are reported in Section 5 of the web appendix. Table 7 compares our proposed model with the alternative models for the estimation and the holdout samples on the basis of the log likelihood (LL), Akaike information criterion (AIC), BIC, and U^2 statistic (Hauser 1978). The U^2 statistic tells us the percentage of variability in behavior explained by the model, and the null model used for calculating it is the model with only intercepts in utilities of all purchase options. The comparison in Table 7 shows that our proposed model performs better than Models II and III in both samples. The value of U^2 for our proposed model in both samples is around 0.42, which implies that our proposed model does a reasonably good job in explaining the variation in the data.[8]

For further validation, we undertake two additional exercises. The first exercise pertains to how the proposed model explains the data summaries given in Tables 3 and 4, which consist of the average plan choices and the switching patterns of the plans over the years in the data. The results of this exercise are reported in Tables 8 and 9. The second exercise pertains to tracking plots for healthcare consumption decisions in the longitudinal holdout sample, where we compare the actual number of each of the treatments over each week with those predicted by the proposed model. The results are reported in Figure 4. Observe that the proposed model does a reasonably good job in predicting the actual shares of the insurance plans and the switching probabilities of the plans in the first exercise and in predicting the actual shares of the three consumption options over time.

Table 7. Goodness of Fit and Predictive Ability of the Proposed and Competing Models

	Model I (Proposed model)	Model II (Investment effect only)	Model III (2-stage hidden Markov model)
LL			
Estimation	18,592.3	19,562.2	19,109.1
Holdout	10,121.2	10,582.1	10,354.6
AIC			
Estimation	37,246.6	39,164.4	38,342.2
Holdout	20,304.2	21,204.2	20,833.2
BIC			
Estimation	37,541.1	39,354.4	38,931.2
Holdout	20,577.0	21,394.2	21,378.4
U^2			
Estimation	0.426	0.396	0.411
Holdout	0.420	0.394	0.407

Table 8. Percentage of Plan Choices Predicted by the Proposed Model

	Basic (%)	Medium (%)	Comprehensive (%)
2005	21.7	34.1	44.2
2006	22.4	34.7	42.9
2007	21.5	35.9	42.6

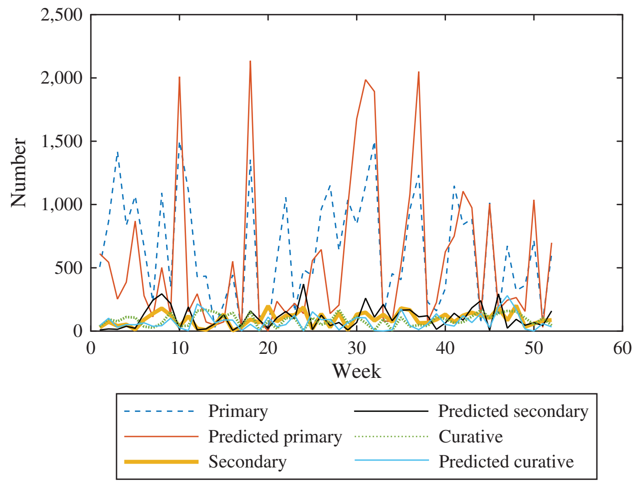
Table 9. Insurance Plan-Switching Patterns Predicted by the Proposed Model

	2006		
2005	Basic (%)	Medium (%)	Comprehensive (%)
Basic	90.8	3.9	5.3
Medium	3.0	94.1	2.9
Comprehensive	3.3	3.2	93.5
	2007		
2006	Basic (%)	Medium (%)	Comprehensive (%)
Basic	91.3	3.3	5.6
Medium	1.9	93.6	4.5
Comprehensive	2.7	7.0	90.3

5.2. Parameter Estimates of the Proposed Model

Table 10 reports the parameter estimates of our proposed model. There are four segments. Segment A accounts for 21.2%, segment B accounts for 29.4%, seg-

Figure 4. (Color online) Tracking Plot of Actual Weekly Consumption of Primary Preventive, Secondary Preventive, and Curative Care Services vs. Those Predicted by the Proposed Model in the Longitudinal Holdout Sample



ment C accounts for 14.3%, and segment D accounts for 35.1% of the consumer population. The estimates of the prior mean health status across the four segments are $h_{0,A} = 2.825$, $h_{0,B} = 1.601$, $h_{0,C} = 1.586$, and $h_{0,D} = 0.909$. This shows that segment A is the healthiest at $\{t = 1, a = 1\}$, followed by segments B and C, which are similar in their prior mean health status, and finally followed by segment D. The estimates of degradation rates are positive and significant (at $p < 0.05$) for segments B, C, and D ($\delta_{d,B} = 0.0011$, $\delta_{d,C} = 0.0012$, and $\delta_{d,D} = 0.0051$), while not significant for segment A. This implies that whereas the health of consumers in segments B, C, and D degrades in the absence of consumption, the health status of consumers in segment A does not. Note that across the four segments, the degradation rates are inversely correlated with prior mean health status. This implies that segment A is the healthiest segment, followed by segments B and C, and finally followed by segment D. We thus refer to segment A as the healthy segment, B and C as the moderately healthy segments, and D as the least healthy segment.

The estimates of price sensitivity (which is the sensitivity for out-of-pocket expenses and the premium) across the four segments are $\alpha_A = 0.151$, $\alpha_B = 0.179$, $\alpha_C = 0.139$, and $\alpha_D = 0.101$. This shows that the healthy segment A is more price sensitive compared to the least healthy segment D and that the moderately healthy segment B is more price sensitive compared to the moderately healthy segment C. The estimates of the prior uncertainty in the health status across the four segments are $\sigma_{0,A} = 2.854$, $\sigma_{0,B} = 2.926$, $\sigma_{0,C} = 4.401$, and $\sigma_{0,D} = 3.037$, which shows that the prior uncertainty is similar in segments A, B, and D and smaller compared to the prior uncertainty of segment C.

These results imply the following regarding insurance plan decisions. If we consider heterogeneity only in prior mean health status and degradation rate, we would predict that in the first year, consumers in healthy segment A would primarily prefer the basic plan, consumers in the moderately healthy segments B and C would primarily prefer the medium plan, and consumers in the least healthy segment D would prefer the comprehensive plan. However, in our data, 22.2%, 33.9%, and 43.9% of consumers choose basic, medium, and comprehensive plans, respectively, in the first year. Comparing these percentages with percentages implied by heterogeneity in prior mean health status and degradation rate, we can see a misalignment: only 35.1% of consumers are in the weakest segment D, but a much higher percentage (43.9%) of consumers in the data choose the comprehensive plan; a total of 43.7% of consumers are in the moderate health segments B and C, but a much lower percentage (33.9%) of consumers choose the medium plan. This issue resolves once we also consider heterogeneity in prior uncertainty. Since segments A, B, and D have low uncertainties, their insurance plan decisions are dictated more by their mean health status, and as a result, consumers in the healthy segment A primarily choose the basic plan, consumers in the moderately healthy segment B primarily choose the medium plan, and consumers in the least healthy segment D primarily choose the comprehensive plan. On the other hand, because

Table 10. Parameter Estimates of the Proposed Model

	Estimate (standard error)			
Parameter	Segment A	Segment B	Segment C	Segment D
Segment size	0.212 (0.101)	0.294 (0.126)	0.143 (0.071)	0.351 (0.164)
Mean of prior beliefs of health status at $\{t = 1, a = 1\}$, h_0	2.825 (1.276)	1.601 (0.742)	1.586 (0.763)	0.909 (0.413)
Std. dev. of prior beliefs of health status at $\{t = 1, a = 1\}$, σ_0	2.854 (1.152)	2.926 (1.451)	4.401 (1.805)	3.037 (1.297)
Degradation, δ_d	0.0004 (0.0003)	0.0011 (0.0005)	0.0012 (0.0006)	0.0051 (0.0024)
Price coefficient, α	0.151 (0.071)	0.179 (0.085)	0.139 (0.066)	0.101 (0.049)
Informative effect: Standard deviation of the consumption signal of secondary preventive care, σ_{sp}	6.601 (3.115)	6.914 (3.057)	6.956 (3.145)	7.056 (3.246)
Investment effect of primary preventive care, δ_{pp}	0.007 (0.005)	0.007 (0.003)	0.008 (0.004)	0.015 (0.007)
Investment effect of curative care, δ_c	0.043 (0.023)	0.046 (0.021)	0.049 (0.022)	0.077 (0.030)

consumers in the moderately healthy segment C have high uncertainty in their health status, they choose the comprehensive plan instead of the medium plan.

The estimates of investment effects of curative care across the four segments are $\delta_c = \{0.043, 0.046, 0.049, 0.077\}$, and the estimates of investment effects of primary preventive care across the four segments are $\delta_{pp} = \{0.007, 0.007, 0.008, 0.015\}$. Note that for each segment, the estimates of the investment effects of curative and primary preventive care are much greater than the degradation rate of any segment, which implies that consumers in all four segments will improve their health if they take any of the two treatments. Moreover, for any given segment, the estimate of investment effect of curative care is significantly greater than the investment effect of primary preventive care. This makes intuitive sense because curative care consists of surgeries and drugs that provide a quantum boost to the consumer's health, whereas primary preventive care consists of drugs that prevent the disease from getting worse. Finally, across the four segments, the investment effects of curative care are much higher for segment D compared to the other three segments (and, similarly, for primary preventive care). This implies that consumers in the least healthy segment D experience the greatest increase in their health status by consuming an investment service compared to the other three segments. Moving on to informative effects, the estimates of the noise in the signal of secondary preventive care across the four segments are $\sigma_{sp} = \{6.601, 6.914, 6.956, 7.056\}$, which yield noise to information ratio (i.e., the ratio of the noise in the signal, σ_{sp}, for segment j to the prior uncertainty in health status for segment j, $\sigma_{0,j}$) values for segments A, B, C, and D of 2.336, 2.362, 1.580, and 2.323, respectively. A small value of noise to information ratio means that there will be a large reduction in uncertainty after consuming secondary preventive care. The values of noise to information ratios imply that all four segments can reduce their uncertainty by consuming secondary preventive care, but segment C will reduce it the most.

Finally, we discuss the impact of pricing components on insurance choice and consumption decisions in each of the four segments. The pricing components include copayment, coinsurance, deductible, out-of-pocket maximum, and premium. Tables 11–13 report the elasticities of each of the pricing components of each of the three plans on the insurance plan decisions of consumers in the four segments. These elasticities represent the percentage change in the average probability of each choice following a 1% change in the pricing component. Furthermore, these are contemporaneous elasticities; i.e., they capture the impact of the change in the pricing component at the beginning of a

Table 11. Price Elasticities of Insurance Choice: Comprehensive Plan

	Insurance purchase			
	Segment A	Segment B	Segment C	Segment D
Copay of comprehensive plan	0.084	0.112	0.063	0.053
Deductible of comprehensive plan	0.113	0.141	0.085	0.078
Coinsurance of comprehensive plan	0.113	0.142	0.097	0.064
Premium of comprehensive plan	0.052	0.069	0.044	0.036

Table 12. Price Elasticities of Insurance Choice: Medium Plan

	Insurance purchase			
	Segment A	Segment B	Segment C	Segment D
Copay of medium plan	0.114	0.137	0.085	0.072
Deductible of medium plan	0.131	0.157	0.104	0.081
Coinsurance of medium plan	0.138	0.174	0.118	0.095
Premium of medium plan	0.072	0.089	0.056	0.051

Table 13. Price Elasticities of Insurance Choice: Basic Plan

	Insurance purchase			
	Segment A	Segment B	Segment C	Segment D
Copay of basic plan	0.128	0.155	0.083	0.076
Deductible of basic plan	0.149	0.196	0.111	0.095
Coinsurance of basic plan	0.168	0.194	0.131	0.104
Premium of basic plan	0.089	0.112	0.064	0.051

Table 14. Price Elasticities of Consumption Decisions

	Healthcare consumption											
	Secondary preventive				Primary preventive				Curative			
	Segment A	Segment B	Segment C	Segment D	Segment A	Segment B	Segment C	Segment D	Segment A	Segment B	Segment C	Segment D
Copay	0.542	0.559	0.349	0.321	0.491	0.521	0.377	0.306	0.362	0.384	0.231	0.218
Deductible	0.454	0.471	0.259	0.211	0.437	0.475	0.297	0.248	0.316	0.341	0.185	0.163
Coinsurance	0.369	0.412	0.215	0.204	0.401	0.425	0.237	0.226	0.391	0.417	0.235	0.206

given year on the insurance plan choices in that year only. Observe that the elasticities are smaller for segments C and D compared to those for segments A and B. Moreover, although all of the pricing component elasticities are statistically significant, they are small in magnitude. This implies that the consumer's health status, her uncertainty about her health status, and her degradation rate play a more important role in her insurance plan decision compared to the pricing components. Table 14 reports the contemporaneous elasticities of each of the cost-sharing characteristics on consumption choices of the four segments. Unlike those in Tables 11–13, these elasticities are larger in magnitude, which implies that cost-sharing characteristics have a greater impact on consumption choice decisions compared to insurance choice decisions. Across the four segments, the usage of both primary and secondary preventive care is influenced most by changes in the copayment, followed by changes in the deductible and coinsurance rate. The opposite is true for curative care, which makes intuitive sense—since curative care is expensive, a decrease in the coinsurance rate translates to greater savings for the consumer compared to a similar percentage decrease in the deductible or the copayment.

6. Policy Simulations

In this section, we examine the policies that managers in insurance companies can employ to reduce inefficiency that stems from moral hazard. As discussed earlier, such policies would help them increase profits by either increasing their profit margin or improving their competitive position vis-à-vis other insurance firms. An important point that follows from the discussion in Section 5.2 is that segment C is a potential source of moral hazard. To see why that is so, recall that segment C consumers have moderate health status, which implies that they can manage their well-being by either curative or primary preventive care. However, because they also have the largest uncertainty, they purchase the comprehensive plan and not the medium plan; consequently, they choose to consume the more expensive curative care, which increases overall healthcare costs. This overconsumption of curative care can be seen in the following exercise. Consider two cases: Case 1, in which segment C consumers are enrolled in the medium plan, and Case 2, in which they are enrolled in the comprehensive plan. For both cases, we simulate their purchases of the three consumption options over the insurance year. The results in terms of usage in dollars per person for each service over the policy year are reported in Table 15. Observe that as we go from Case 1 to Case 2, there is a significant increase in usage of curative care and a decrease in usage of primary preventive care.

Table 15. Consumption of Services Over the Policy Year for Consumers in Segment C

	Annual consumption of segment C ($)	
	Enrolled in comprehensive plan	Enrolled in medium plan
Primary preventive care	328.1	513.9
Secondary preventive care	28.2	26.3
Curative care	6,912.5	3,012.7

Two policy routes can be employed to incentivize segment C consumers to choose the medium plan and thereby reduce inefficiency. The first is the "immediate route," in which we change cost-sharing characteristics of insurance plans at the beginning of the policy year, which will incentivize segment C consumers to choose the medium plan in that policy year itself. The second is the "delayed route," in which we provide incentives for segment C consumers to consume more secondary preventive care in the first year (and onward), which will reduce their uncertainty and in turn induce them to choose the medium plan in the future. Regarding the immediate route, our results in Section 5.2 indicate that this may not be feasible, because the contemporaneous elasticities for insurance choice decisions are small. Thus, changing cost-sharing characteristics of any of the three plans in a given year will not be sufficient to induce segment C consumers to switch to the medium plan in that year itself. We discuss below two policy simulations in the context of the delayed route. In the first one, we investigate the impact of changing cost-sharing characteristics of insurance plans that encourage use of secondary preventive care. In the second one, we investigate the impact of providing more accurate information to consumers via secondary preventive care.

6.1. Changing Cost-Sharing Characteristics to Encourage Use of Secondary Preventive Care

Recently, many health service providers have advocated that costs incurred by consumers for secondary preventive care should be reduced. Recall from Section 5.2 that copayment has the strongest impact on the use of secondary preventive care. Thus, in this section, we examine the extent to which the overall healthcare costs (where the overall healthcare costs are the sum of the healthcare costs incurred by consumers and by the insurance company) would decrease in the long run if the insurance firm were to reduce the copayment for secondary preventive care. Decreasing the copayment for secondary preventive care can have two opposing effects on overall healthcare costs: On one hand, it can encourage consumers to consume more secondary preventive care, which will result in an increase in healthcare costs. On the other hand, it can decrease the inefficiency, because it will encourage consumers in segment C to consume more secondary preventive care, which will decrease their uncertainty and make them choose medium plans in the future.

To see how reducing the copayment impacts the overall costs, we run a counterfactual in which the insurance company reduces the copayment for secondary preventive care in 2005 and onward by 50%. Table 16 presents the results on the change in consumers' insurance plan choices and change in overall healthcare costs of each of the segments in the subsequent year, 2006, with respect to the baseline case (in which there is no reduction in the copayment).

Observe that plan choice probabilities are similar between the baseline and counterfactual cases in 2005. This result follows from our discussion in Section 5.2. Since the contemporaneous insurance choice elasticities are small, changing the cost-sharing characteristics will not have a strong immediate impact on insurance choices. On the other hand, there is a greater divergence in the plan choice probabilities across the two cases in 2006, which shows that the delayed route is a much more effective tool for decreasing moral hazard. Consequently, we see that the overall healthcare costs in 2006 are lower in the counterfactual case compared to the baseline case. To understand the source of this decrease, we next look at the changes in healthcare costs on a segment-level basis.

Observe in Table 16 that decreasing the copayment for secondary preventive care leads to a small increase in overall healthcare costs for segments A and B, a small decrease in overall healthcare costs for segment D, and a large decrease in overall healthcare costs for segment C. The reason for the increase in overall healthcare costs for healthy segment A and moderately healthy segment B is that most consumers in these segments choose medium or basic plans. Because consumers in medium and basic plans incur high out-of-pocket expenses when consuming secondary preventive care, a 50% decrease in copay amounts to greater savings, which encourages them to use more secondary preventive care. The reason for such a large decrease in overall healthcare costs for the moderately healthy segment C follows from our earlier discussion—a 50% decrease in copayment of secondary preventive care increases the probability of consumers in segment C to consume more preventive care in 2005, which in turn increases their probability of purchasing medium plans in 2006. This can be seen in Table 16—the overall probability in 2006 of choosing a comprehensive plan reduces from 43.3% to 41.4%, and the overall probability of choosing a medium plan increases from 34.1% to 36.1%.

6.2. Improving Health Status Learning (Personalized Medicine)

In the second simulation, we investigate the impact of providing more accurate information to the consumer via secondary preventive care on overall healthcare costs. This is in the spirit of personalized medicine, in which healthcare providers use precise molecular profiling technologies that can provide much more accurate information to consumers about their health compared to standard diagnostic tests. To do so, we run a counterfactual in which we consider an alternative scenario where the noise in the secondary preventive care signal decreases by half in 2005 and onward. Since we do not have information on the costs of improved secondary preventive care in the counterfactual case, we assume that its costs are the same as the costs of secondary preventive care in the data.[9]

Table 17 presents the results on the change in consumers' insurance plan choices and change in overall healthcare costs of each segment in the counterfactual case vis-à-vis the baseline case. The pattern of results

Table 16. Decreasing Copayment of Secondary Preventive Care by 50%

Plan	Plan choice probability in 2005: Baseline (%)	Plan choice probability in 2006: Baseline (%)	Plan choice probability in 2005: Counterfactual (%)	Plan choice probability in 2006: Counterfactual (%)	Segment (size)	Change in healthcare costs between baseline and counterfactual in 2006 (%)
Basic	22.2	22.6	22.1	22.5	A (21.2%)	+1.0
Medium	33.9	34.1	34.8	36.1	B (29.4%)	+0.6
					C (14.3%)	−6.9
Comprehensive	43.9	43.3	43.1	41.4	D (35.1%)	−1.4

Table 17. Personalized Medicine: Increasing Informational Value of Secondary Preventive Care

Plan	Plan choice probability in 2005: Baseline (%)	Plan choice probability in 2006: Baseline (%)	Plan choice probability in 2005: Counterfactual (%)	Plan choice probability in 2006: Counterfactual (%)	Segment (size)	Change in healthcare costs between baseline and counterfactual in 2006 (%)
Basic	22.2	22.6	22.3	22.7	A (21.2%)	−0.5
Medium	33.9	34.1	36.2	39.6	B (29.4%)	−1.5
					C (14.3%)	−11.1
Comprehensive	43.9	43.3	41.5	37.7	D (35.1%)	−4.4

is similar to that in Table 16. The plan choice probabilities are similar between the baseline and counterfactual cases in 2005, and there is a much greater divergence in the plan choice probabilities across the two cases in 2006. We next compare the change in the overall costs in 2006 across the two cases. On a segment-level basis, the overall healthcare costs for segments A, B, and D are marginally lower in the counterfactual case compared to the baseline case. On the other hand, overall healthcare costs for segment C are substantially lower in the counterfactual case compared to the baseline case. The cause for this variation follows from our earlier discussion. Providing more accurate information decreases consumers' uncertainty about their health status, which in turn decreases the probability that consumers in segment C will purchase more comprehensive plans than they actually need. This can be seen in Table 17—note that when more accurate preventive care is offered, the probability of choosing the comprehensive plan in 2006 reduces from 43.3% to 37.7%, and the probability of choosing the medium plan increases from 34.1% to 39.6%.

7. Conclusions

In this paper, we posit that in the context of chronic diseases, a part of the increase in healthcare costs stems from the moral hazard problem faced by employer-based insurance firms, whereby consumers suffering from chronic diseases with a moderate level of severity select comprehensive insurance plans and opt for the more expensive curative care. We use a data set on consumers' insurance choices and healthcare consumption decisions to investigate the extent of this inefficiency and explore the policies that insurance firms can employ to reduce it. Our key results are as follows. *First*, the variation in insurance plan decisions across consumers stems from heterogeneity in health status, uncertainty in health status, and degradation in health status over time. *Second*, the source of moral hazard is the presence of a sizable segment (14.3%) of consumers with moderate health status, low price sensitivity, and a high uncertainty. Instead of buying the medium plan that better matches their health status, they purchase a comprehensive plan; once in the plan, they opt for more expensive curative care even when the illness could be managed through primary preventive care. *Third*, changing cost-sharing characteristics to induce these consumers to switch to medium plans in the same policy year may not work. A more feasible route for decreasing moral hazard is the delayed route, in which we incentivize consumers to purchase more secondary preventive care, which reduces their uncertainty and makes them choose medium plans in the future.

In this paper, we attempt to model an extremely complex decision-making process of insurance plan and consumption choices. Consequently, we make some simplifying assumptions. For instance, we assume that consumers do not succumb to psychological biases when making insurance plan and consumption decisions. Similarly, we aggregate multiple treatments into a single healthcare option. Ideally, we would have considered only consumers suffering from one chronic disease in which there are a relatively homogenous set of treatments. However, we would have been left with a very small sample size if we had done so. Finally, for the reasons discussed in Section 3, we assume that the physician is a perfect agent and acts in the best interest of the patient. Relaxing these assumptions would provide fruitful avenues for future research.

Acknowledgments

This paper is based on Chapter 2 of J. Ni's Ph.D. dissertation at Carnegie Mellon University, and the authors are listed in alphabetical order. The authors are grateful to Sunder Kekre for helping them obtain the data and to Tat Chan, Hanming Fang, George-Levi Gayle, Avi Goldfarb, and David Soberman for helpful comments. The authors also appreciate feedback from seminar participants at Arizona State University, Emory University, Georgia Tech, Johns Hopkins University, Hong Kong University of Science and Technology, the University of Minnesota, the National University of Singapore, New York University, the University of Southern California, the University of Washington, and Yale University. All errors are the responsibility of the authors.

Endnotes

[1] Although we do not observe employees' health insurance options outside the firm, according to the Kaiser Annual Survey of Employer Health Benefits (https://kaiserfamilyfoundation.files.wordpress.com/2013/04/7790.pdf), 82% of eligible workers enroll in plans offered by their own employers. This was also verified by the insurer who provided us with the data.

[2] One could potentially argue that curative care services like surgery could be informative, because the services would most likely come with a series of tests. However, note that the tests that accompany curative care are recorded separately in our data as secondary preventive care. As a further validation of our assumption, we estimated an alternative model in which we did not assume $\sigma_c^2 = \infty$. The alternative model did not perform significantly better than the proposed model.

[3] Note that in the consumption utility, we assume risk neutrality in money (i.e., we assume that the price term enters the utility linearly) and not risk aversion in money, in which the price term would enter in the exponential in the utility. As a validation for our assumption, we estimated an alternative model in which we assumed risk aversion in money. The alternative model did not perform as well as the proposed model in terms of goodness of fit. Furthermore, the results in the counterfactuals were similar across the two specifications, which shows that the main insights of the model are robust to whether or not we assume risk aversion in money.

[4] In Equation (20), note that the informative effect of the chosen option k in period t impacts the value function $V(S_t^a \mid j)$ through the expected future value term, $\beta E_t V(S_{t+1}^a \mid S_t^a, k, j)$. This is because the state variable vector, S_{t+1}^a, in the future value term is a function of the health status at the beginning of period $t+1$, h_t, which in turn is a function of the learning draw, $(\sigma_{t-1}^2/\sqrt{\sigma_k^2 + \sigma_{t-1}^2})v_{kt}$, that stems from the informative effect of the chosen option k in period t. As discussed in Section 3.6, this learning draw impacts the expected future utilities in periods $t+1$ and onward because it enables the consumer to make judicious consumption decisions in periods $t+1$ and onward. For instance, if the consumer receives a very negative value of the learning draw at the end of period t, she will opt for a treatment with large investment effects in period $t+1$, which will increase her optimal utility in period $t+1$. Similarly, if she receives a positive learning draw, she will opt for the no-consumption option in period $t+1$.

[5] Some prior papers that have used the CARA utility have estimated r as a separate parameter. These papers include those by Chan and Hamilton (2006) and Ching and Ishihara (2010). The reason Chan and Hamilton (2006) were able to identify r is that in their data, they observed the values of the signals. The reason Ching and Ishihara (2010) were able to identify r is that they instead normalized the true quality of one of the alternatives to one.

[6] This explains how we can identify whether a consumer who is in a comprehensive plan and consumes costly curative care has low price sensitivity and moderate health or moderate price sensitivity and poor health—unlike a consumer with low price sensitivity, a consumer with high price sensitivity will show differences in consumption behavior before and after the deductible is reached and before and after the out-of-pocket maximum is reached.

[7] We also estimated a two-stage version of our proposed model where the first stage is similar to that of Model III. Similar to Model III, this extended model relaxes the assumption that we make in our proposed model that consumers make a multinomial healthcare consumption decision in each week. The extended model performs slightly better than our proposed model in both the estimation and the holdout sample. However, the results of the counterfactuals that follow from the extended model are similar to those in our proposed model, which indicates that the aforementioned assumption that we make in our model, although a simplification, does not influence the results.

[8] This has to be seen in the context of the values of U^2 reported in the discrete choice models in prior literature, which range from 0.20 to 0.55. The value of U^2 tends to be on the lower end of this range when one of the alternatives in the data has a predominantly larger share, such as the no-purchase option. For example, in Chintagunta (1993), $U^2 \approx 0.25$ because he includes the no-purchase option as one of the alternatives in the discrete choice model, which has a very high share. Similarly, in our data, the no-consumption option (among the four consumption options in each week) has a predominantly high share.

[9] While the costs would typically be higher for a more precise diagnostic test, the reason we make this assumption is that we do not know how costs will decrease as the tests become more precise. Thus, our results should be interpreted as the change in revenues when the tests become more precise. A manager or policy maker who wishes to use our methodology would have access to the cost information, and can thus use our methodology to see whether it is profitable to come up with a more precise diagnostic test.

References

Aaron HJ, Ginsburg PB (2009) Is health spending excessive? If so, what can we do about it? *Health Affairs* 28(5):1260–1275.

Aron-Dine A, Einav L, Finkelstein A, Cullen M (2015) Moral hazard in health insurance: Do dynamic incentives matter. *Rev. Econom. Statist.* 97(4):725–741.

Bajari P, Dalton C, Hong H, Khwaja A (2014) Moral hazard, adverse selection and health expenditures: A semi-parametric analysis. *RAND J. Econom.* 45(4):747–763.

Beck M (2010) The prostate cancer quandary. *Wall Street Journal* (June 29), http://www.wsj.com/articles/SB10001424052748703279704575334982806405218.

Brennan T, Reisman L (2007) Value-based insurance design and the next generation of consumer-driven health care. *Health Affairs* 26(2):204–207.

Cameron A, Trivedi PK, Milne F, Piggott J (1988) A microeconometric model of demand for health care and health insurance in Australia. *Rev. Econom. Stud.* 55(1):85–106.

Cardon J, Hendel I (2001) Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *RAND J. Econom.* 32(3):408–427.

Carlin C, Town R (2008) Adverse selection, welfare and optimal pricing of employer-sponsored health plans. Working paper, University of Minnesota, Minneapolis.

Centers for Medicare and Medicaid Services (2015) https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html.

Chan TY, Hamilton B (2006) Learning, private information, and the economic evaluation of randomized experiments. *J. Political Econom.* 114(6):997–1040.

Ching A, Ishihara M (2010) The effects of detailing on prescribing decisions under quality uncertainty. *Quant. Marketing Econom.* 8(2):123–165.

Ching A, Erdem T, Keane M (2009) The price consideration model of brand choice. *J. Appl. Econometrics* 24(3):393–420.

Chintagunta PK (1993) Investigating purchase incidence, brand choice and purchase quantity decisions of households. *Marketing Sci.* 12(2):184–208.

Crawford G, Shum M (2005) Uncertainty and learning in pharmaceutical demand. *Econometrica* 73(4):1137–1173.

Cunningham PJ, Charles D, Sinclair M (2001) Do consumers know how their health plan works? *Health Affairs* 20(2):159–166.

DeVol R, Bedroussian A, Charuworn A, Chatterjee A, Kim I, Kim S, Klowden K (2007) An unhealthy America: The economic burden of chronic disease. Milken Institute Report, Santa Monica, CA. http://www.milkeninstitute.org/publications/view/321.

Einav L, Finkelstein A, Pascu I, Cullen M (2012) How general are risk preferences? Choices under uncertainty in different domains. *Amer. Econom. Rev.* 102(6):2606–2638.

Einav L, Finkelstein A, Ryan S, Schrimpf P, Cullen M (2013) Selection on moral hazard in health insurance. *Amer. Econom. Rev.* 103(1):178–219.

Felder S, Mayrhofer T (2011) *Medical Decision Making—A Health Economic Primer* (Springer, Berlin).

Gerteis J, Izrael D, Deitz D, LeRoy L, Ricciardi R, Miller T, Basu J (2014) Multiple chronic conditions chartbook. Agency for Healthcare Research and Quality Publications, Rockville, MD. https://www.ahrq.gov/sites/default/files/wysiwyg/professionals/prevention-chronic-care/decision/mcc/mccchartbook.pdf.

Gilleskie D (1998) A dynamic stochastic model of medical care use and work absence. *Econometrica* 66(1):1–45.

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

Grossman M, Rand E (1974) Consumer incentives for health services in chronic illness. Mushkin SJ, ed. *Consumer Incentives for Health Care* (Prodist, New York), 114–151.

Gupta A, Li Z (2011) Calibration of a stochastic health evolution model using NHIS data. *Physica A* 390(20):3524–3540.

Handel B (2013) Adverse selection and inertia in health insurance markets: When nudging hurts. *Amer. Econom. Rev.* 103(7): 2643–2682.

Harris KM (2002) Can high quality overcome consumer resistance to restricted provider access? Evidence from a health plan choice experiment. *Health Services Res.* 37(3):551–571.

Hauser JR (1978) Testing the accuracy, usefulness, and significance of probabilistic choice models: An information theoretic approach. *Oper. Res.* 26(3):406–421.

Hellinger FJ (1989) Expected utility theory and risky choices with health outcomes. *Medical Care* 27(3):273–279.

Johnson EJ, Hershey J, Meszaros J, Kunreuther H (1993) Framing, probability distortions, and insurance decisions. *J. Risk Uncertainty* 7(1):35–51.

Johnson EJ, Hassin R, Baker T, Bajger AT, Treuer G (2013) Can consumers make affordable care affordable? The value of choice architecture. *PLoS One* 8(12):e81521.

Keane M, Wolpin K (1994) The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. *Rev. Econom. Statist.* 76(4): 648–672.

Kenkel DS (2000) Prevention. Culyer AJ, Newhouse JP, eds. *Handbook of Health Economics*, Vol. 1, Part B (Elsevier, Amsterdam), 1675–1720.

Khwaja A (2010) Estimating willingness to pay for Medicare using a dynamic life-cycle model of demand for health insurance. *J. Econometrics* 156(1):130–147.

Korobkin R (2014) Health-care costs and the "moral hazard" problem. *Washington Post* (March 10) https://www.washingtonpost.com/news/volokh-conspiracy/wp/2014/03/10/health-care-costs-and-the-moral-hazard-problem/?utm_term=.e1f61d5c0d3e.

Lancaster T (1990) *The Econometric Analysis of Transition Data* (Cambridge University Press, Cambridge, UK).

Llewellyn-Thomas H, Sutherland J, Tibshirani R, Ciampi A, Till JE, Boyd NF (1982) The measurement of patients' values in medicine. *Medical Decision Making* 2(4):449–462.

Manning W, Marquis MS (1996) Health insurance: The tradeoff between risk pooling and moral hazard. *J. Health Econom.* 15(5):609–639.

Mehta N, Chen X, Narasimhan O (2008) Informing, transforming, and persuading: Disentangling the multiple effects of advertising on brand choice decisions. *Marketing Sci.* 27(3):334–355.

Narayanan S, Manchanda P (2009) Heterogeneous learning and the targeting of marketing communications for new products. *Marketing Sci.* 28(3):424–441.

Rabin M, Thaler RH (2001) Anomalies: Risk aversion. *J. Econom. Perspect.* 15(1):219–232.

Rothbaum E (2006) Cost-consciousness in the patient-physician relationship. *Virtual Mentor* 8(3):127–129.

Schram A, Sonnemans J (2011) How individuals choose health insurance: An experimental analysis. *Eur. Econom. Rev.* 55(6):799–819.

Schweidel DA, Bradlow ET, Fader PS (2011) Portfolio dynamics for customers of a multi-service provider. *Management Sci.* 57(3):471–486.

Shapira Z, Venezia I (2008) On the preference for full-coverage policies. Why do people buy too much insurance? *J. Econom. Psych.* 29(5):747–761.

Smith JC, Medalia C (2015) Health insurance coverage in the United States: 2014. United States Census Bureau Report, Washington, DC. https://www.census.gov/content/dam/Census/library/publications/2015/demo/p60-253.pdf.

Strombom BA, Buchmueller TC, Feldstein PJ (2002) Switching costs, price sensitivity and health plan choice. *J. Health Econom.* 21(1): 89–116.

Thorpe KE (2008) Prevention and public health: The key to transforming our sickcare system. Written testimony to the U.S. Senate Committee on Health, Education, Labor, and Pensions Hearing (December 10). http://www.help.senate.gov/imo/media/doc/Thorpe1.pdf.

Vera-Hernandez M (2003) Structural estimation of a principal-agent model: Moral hazard in medical insurance. *RAND J. Econom.* 34(4):670–693.

Zhang J (2010) The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* 29(2):315–335.