Insurance-Induced Moral Hazard: A Dynamic Model of Within-Year Medical Care Decision Making Under Uncertainty*

Christopher J. Cronin†
University of Notre Dame

August 4, 2016

Abstract

Existing studies estimate health insurance-induced increases in medical care expenditure by examining medical care decisions that are aggregated to the annual level. Using employer-employee matched data from the Medical Expenditure Panel Survey, I quantify the moral hazard effect of insurance on medical care expenditure by estimating a dynamic model of within-year medical care consumption that allows for insurance selection, endogenous health transitions, and individual uncertainty about medical care prices in an environment where insurance has non-linear cost-sharing features. The results suggest that the additional consumption induced by moral hazard amounts to 53.1 percent, on average, of total annual medical care expenditure when insured. In order to understand the relationship between the dynamic features of the model and the estimated moral hazard effects, I estimate a second model that is representative of the annual decision-making models found in the literature. The within-year decision-making model produces a moral hazard effect that is significantly different, and generally larger, than the alternative model. To illustrate the importance of the within-year decision-making model, I quantify the welfare and spending implications of health insurance alternative sets with more and less generous options.

JEL Classification: C61, D81, G22, I12, I13
Keywords: Dynamic Discrete Choice, Choice Under Uncertainty, Risk Aversion, Health Insurance, Within-Year Decision Making, Health Production

*This research was conducted at the Triangle and Chicago Census Research Data Centers, and support from the Agency for Healthcare Research and Quality (AHRQ) is acknowledged. The results and conclusions in this paper are those of the author and do not indicate concurrence by AHRQ or the Department of Health and Human Services.

†I would like to thank my dissertation advisor Donna Gilleskie; committee members David Guilkey, Clement Joubert, Brian McManus, and Helen Tauchen; as well as Chuck Cortemanche, Michael Darden, Randy Ellis, Michael Grossman, Matt Harris, Tiago Pires, Dan Rees, Forrest Spence, Steve Stern, and Jessica Vistnes for their helpful comments and guidance. I also benefited greatly from conversations with my colleagues at the University of Notre Dame, as well as seminar participants at Census CES, ASHEcon-Minneapolis, AHEW-Chapel Hill, SEA-Tampa, American Enterprise Institute, IHEA-Milan, Econometric Society World Congress - Montreal, and the following universities: North Carolina-Chapel Hill, Kentucky, South Florida, McGill, WashU-Olin, Northeastern, Purdue, Clemson, Tulane, and DePaul. Email: ccronin1@nd.edu. Website: christophercronin.weebly.com.

1 Introduction

Economic theory suggests that health insurance may increase medical care consumption above the socially optimal level (Arrow, 1963; Pauly, 1968). The incentives that elicit this increase in consumption are often referred to as moral hazard (Cutler and Zeckhauser, 2000).[1] Empirical studies tend to estimate moral hazard effects using models that aggregate medical care decisions up to the annual level (Cardon and Hendel, 2001; Khwaja, 2001, 2010; Einav et al., 2013; Bajari et al., 2014; Kowalski, 2015). In this paper, I study insurance-induced moral hazard using a dynamic, stochastic model of within-year medical care consumption decisions. The within-year decision-making (WYDM) model more accurately captures the data generating process by relaxing several assumptions made frequently in the literature. Specifically, the model allows for endogenous health transitions, variation in medical care prices, and individual uncertainty within a health insurance year.[2] The primary objective of this research is to determine whether the WYDM model produces different, and more accurate, moral hazard estimates than an annual decision-making model.

Allowing an individual's medical care optimization problem to evolve endogenously over the course of a health insurance year has implications for the analysis of moral hazard. For example, if medical care consumption affects future health transitions, then a generous insurance plan, relative to no insurance or a less generous plan, may lead an individual to consume more medical care early in the year, experience health improvement, and spend less later in the year. Additionally, if an individual's medical care decisions are determined, in part, by his forecast of future health and medical care expenditure, then health insurance can affect spending patterns through its impact on the forecast horizon, in addition to its impact on out-of-pocket prices. As discussed in detail below, the cumulative effect that these mechanisms have on moral hazard estimates are ex-ante ambiguous, which motivates this empirical investigation.

The principle innovation of this research is the development and estimation of a within-year medical care decision-making model, which is motivated by theoretical (Grossman, 1972; Keeler et al., 1977) and empirical (Gilleskie, 1998; Khwaja, 2001, 2010; Blau and

[1] The term *moral hazard* is used rather loosely in the health economics literature. In this paper, I refer to moral hazard as a change in the incentives for purchasing medical care due to insurance acquisition. I focus on several *effects* of moral hazard, such as the change in total annual medical care expenditure that results from insurance acquisition.

[2] For an individual who carries health insurance in the United States, the *price* of medical care could refer to any of three different values. The *list* price can be thought of as the theoretical market price for care. The *transaction* price, which is the sum of the insurer's and insured's payments, is typically lower than the list price as insurance companies negotiate for reduced rates from some medical care providers. The *out-of-pocket* price is the portion of the transaction price that an insured individual pays for services. Throughout this manuscript, any reference to *price* or *total price* refers to the transaction price, as list prices are rarely relevant to an individual's decision making, which is the focus of this paper. When out-of-pocket prices are discussed, I reference them accordingly.

Gilleskie, 2008) models of health production and medical care demand. An individual's optimization problem consists of an annual health insurance decision, followed by a sequence of monthly medical care consumption decisions made over the course of a health insurance year. I model monthly medical care decisions to allow the unique benefits and costs associated with the timing of unexpected illness and dynamic insurance cost-sharing features to impact behavior within the model. Within each month, a forward-looking individual responds to an endogenous, stochastic health event by consuming medical care. The (anticipated) primary benefit of medical care consumption is improved future health. The (anticipated) primary cost is financial (i.e., a decrease in the current consumption of non-medical goods).[3] When health insurance has dynamic cost-sharing features (i.e., deductible and stop loss), an additional benefit of current medical care consumption is lower future out-of-pocket prices once accumulated expenditure crosses a threshold.[4] The model also allows for a direct contemporaneous utility benefit or cost of medical care consumption that is independent of the productive and financial effects.

The WYDM model is expressed as a dynamic programming problem, which is solved via backwards recursion and estimated via maximum likelihood (MLE) using employer-employee matched data from the 1996-1999 Medical Expenditure Panel Survey (MEPS). The use of nationally representative survey data is unique in this literature, as most recent research has been conducted using firm-level claims data. The MEPS data include dynamic medical care consumption and health information for individuals who work for different firms that offer a wide variety of (observable) health insurance alternative sets. This variation in alternative sets is important for the identification of moral hazard effects as it insures that observationally equivalent individuals hold different plans and, thus, face a different sequence of out-of-pocket medical care prices over the insurance year. Another advantage of the MEPS data is that illness episodes are reported even when medical care is not consumed, which allows endogenous illness transitions to be modeled and is not characteristic of firm-level claims data.

After estimating the structural parameters of the WYDM model, I quantify the effects of moral hazard by simulating an individual's health insurance selection, monthly medical care consumption, and health transitions over the course of the year. I then compare the behavior and outcomes of individuals when employer-sponsored health insurance is

[3] I qualify *anticipated* benefits and costs because the estimated model parameters determine both the sign and magnitude of medical care productivity/efficacy and the disutility from reduced consumption of non-medical goods. I expect to find that medical care is productive, that a reduction in non-medical consumption reduces utility, and that each plays a principal role in the medical care decision-making process.

[4] A deductible is a fixed amount of accumulated medical care expenditure that must be reached within an insurance year before the insurer covers any part of the total price of medical care. A stop loss is an accumulated out-of-pocket expenditure threshold at which an individual's share of the total price of any additional medical care consumed during that insurance year is zero.

required, to the behavior and outcomes of individuals who face the full transaction price of medical care (i.e., no cost sharing, no insurance). I find that moral hazard explains 53.1 percent of (insured) mean annual medical care expenditure. The mean effect, however, is heavily influenced by only a few individuals who are very sensitive to insurance coverage. If the moral hazard distribution is winsorized at the 99th percentile, the percentage of mean annual expenditure explained by moral hazard falls to 46 percent.[5] Furthermore, 54.4 percent of individuals do not alter their behavior when they become insured. The additional medical care consumption brought on by insurance acquisition decreases the average number of illnesses acquired over the course of the year by a small but statistically significant amount.

To isolate the effect that within-year dynamics have on the moral hazard estimates, I also estimate a representative annual expenditure (RAE) model. To form this model, I impose assumptions on the WYDM model so that the incentives faced by an individual and information available to him at the time of his health insurance and medical care decisions reflect those of an annual expenditure model. That is, both models allow for disaggregated (i.e., monthly) decision making, but those decisions reflect either the dynamic updating of expectations of the WYDM or retain the annual expectations of the RAE model. There are two benefits to this approach. First, by estimating a RAE model, as opposed to simply comparing my findings to those in the literature, I ensure that comparisons are made using the same sample of individuals, so that differences cannot be attributed to sample differences. Second, the RAE model is defined by the same functional form restrictions as the WYDM model, ensuring that differences in findings cannot be attributed to functional form differences.

The RAE model produces a moral hazard distribution that is similar in shape to the WYDM model, though the tails of the RAE distribution are heavier. Specifically, the RAE model predicts that 51.9 percent of (insured) mean annual medical care expenditure is explained by moral hazard, 42.6 percent of mean annual expenditure is explained by moral hazard upon winsorizing the moral hazard distribution at the 99th percentile, and 64.2 percent of individuals do not change their behavior when covered by insurance.[6] I describe below how these distributional differences can be explained by differences in information available to individuals at the time of a medical care decision. Ultimately, I find that the annual expenditure model, which assumes that individuals make medical care consumption decisions under perfect information, understates the effect that moral hazard has on medical

[5] Expenditure is simulated for each individual in each insurance state: insured and uninsured. The difference between these expenditures creates a distribution of insurance-induced spending, which I refer to as the *moral hazard distribution*.

[6] These three statistics are each (statistically) significantly different from those produced by the WYDM model.

care expenditures by failing to account for an important form of risk protection provided by health insurance.

The WYDM and RAE models produce mean moral hazard estimates that are statistically different, but similar economically (53.1 percent vs. 51.9 percent, respectively). That said, the magnitude of the differences in the models' predictions are greater in the tails of the moral hazard distribution, which is important for a number of reasons. For example, it is often cited that 65 percent of the medical care expenditure in the US is consumed by those in the top decile of the expenditure distribution (Kaiser Family Foundation). Both the WYDM and RAE models suggest that the popularity of health insurance significantly increases the number of extreme spenders; however, the role of insurance is more pronounced for the RAE model. As a result, my findings suggest that the effectiveness of insurance based policies meant to limit high cost elective procedures, such as capitation payments, will be overstated by annual expenditure models. Another reason to be interested in the full distribution of moral hazard effects is because the mean effect can be quite sensitive to behavior in the right tail, which is shown above to be true for both the WYDM and RAE models. As such, a more conservative estimate of the mean moral hazard effect may limit the relative importance of right tail spenders. Using the (99th percentile) winsorized moral hazard distribution produced by the models and a series of back-of-the-envelope calculations described in Sections 6, I predict the increase in total US medical care expenditure caused by forcing the roughly 45 million uninsured individuals in the US (in 2013) into health insurance coverage. I find the WYDM model prediction to be 12.2 billion dollars larger than that of the RAE model, which represents an economically meaningful difference in the models' predictions.

The paper concludes with a series of policy experiments that analyze the expenditure and welfare implications of more/less generous health insurance alternative sets. The first experiment adds a high deductible health insurance plan (HDHP) to every individual's alternative set. These plans have become increasingly popular in the US as employers look to expand their insurance offerings while minimizing premiums and medical care spending (Towers Watson, 2014). I find the inclusion of a HDHP to be most beneficial for (1) uninsured individuals, who tend to receive less generous insurance offers from their employers, and (2) relatively healthy insured individuals, who are willing to sacrifice coverage for a lower premium. Importantly, the HDHP addition has little impact on aggregate medical care expenditures as the spending changes from these two groups offset. The second experiment adds a HDHP and removes the most generous current offering from individuals' alternative sets. In recent years, many employers have adopted similar changes in their insurance offerings both to curb medical care costs and to avoid the "Cadillac Tax" imposed by the 2010 Patient Protection and Affordable Care Act (42 U.S.C. §18001). Anecdotally,

the singleton firms studied by Einav et al. (2013) and Handel and Kolstad (2015) made similar changes to their insurance offerings during and after the study periods, respectively. My findings suggest that the new alternative set leads to a decrease in medical care spending (of roughly 13 percent) because relatively sick individuals select into less generous coverage under the constrained set, and therefore face higher out-of-pocket prices. Abstracting from the multi-year dynamic implications of the experiment, I find that the welfare losses caused by removing the most generous plan are 2.7 times larger than the welfare gains associated with adding the HDHP.

This research is related to several distinct literatures. First, I estimate the effects of insurance-induced moral hazard using a structural model of individual-level health insurance and medical care consumption decisions. There is a robust literature on the estimation of moral hazard effects in health insurance markets using both reduced-form (e.g., Manning et al., 1987; Finkelstein et al., 2012) and structural (e.g., Cardon and Hendel, 2001; Einav et al., 2013) techniques, which I describe in the following section. Importantly, this paper is the first to quantify insurance-induced moral hazard by solving and estimating a dynamic model of insurance selection, health transitions, and medical care decisions of forward-looking individuals over the course of the health insurance year.

Second, in modeling within-year forecast horizons, the paper assumes implicitly that individuals respond to the *effective* (or future) price of medical care, rather than the *spot* price.[7] This price discrepancy, which results from the non-linear budget constraint created by the dynamic cost-sharing features of health insurance plans, was characterized in the theoretical work of Keeler et al. (1977) and later Ellis (1986). While early work by Keeler and Rolph (1988) did not find evidence of effective price (vs. spot price) responsiveness in the RAND Health Insurance Experiment (HIE) data, recent research by Aron-Dine et al. (2015) provides empirical support for the assumption. The annual expenditure models of Cardon and Hendel (2001) Einav et al. (2013), and Kowalski (2015) allow for this behavior by estimating price elasticities in the presence of non-linear price scheduling; however, the medical care optimization problem is static. The WYDM model, on the other hand, allows for uncertainty over the effective price of care as well as dynamic updating over the course of the insurance year. Aron-Dine et al. (2015) address the issue of price uncertainty by estimating price elasticities based on expected end-of-the year prices; however, their strategy assumes risk neutrality and that individuals have no private information about health shocks, both of which are relaxed by the WYDM model.

Third, much like Gilleskie (1998) and Khwaja (2010), this paper can be viewed as an

[7] Dynamic cost-sharing features of insurance contracts imply that the out-of-pocket price of medical care declines as total expenditure accumulates. Thus, at any point in time, the *effective* price of a unit of medical care is the marginal or *spot* price minus the bonus for moving closer to the deductible or stop loss threshold, past which cost-sharing by the individual falls or is eliminated.

empirical application of the influential Grossman (1972) health capital model. Zweifel and Manning (2000) identify an empirical test for Grossman's model, stating that acquisition of a generous insurance plan should result in a sudden increase in expenditure, which ultimately declines after health stock accumulates. Manning et al. (1985) find support for this hypothesis in the RAND HIE data, but only for dental care. My results show that the consumption of doctor's office visits also responds to insurance acquisition in a way that is consistent with Grossman's health capital model.

The paper is organized as follows. Section 2 provides motivation for research on moral hazard effects and the within-year decision-making model. Section 3 details the theoretical model of insurance and within-year medical care demand. Section 4 describes the data and the sample used in estimation. Section 5 details the estimation procedure and discusses identification. Sections 6 contains all empirical results. Section 7 concludes.

2 Motivation and Background

Health insurance generates welfare by protecting a risk averse individual from medical expenses associated with unforeseen health shocks (Arrow, 1963). These welfare gains from risk protection are potentially mitigated by changes in individual behavior after becoming insured. For example, insurance lowers the out-of-pocket price of medical care, which can lead to excess consumption when sick, known as ex-post moral hazard (Pauly, 1968). Also, a reduction in the expected cost of curative medical care can reduce participation in healthy behaviors (e.g., preventative medical care, diet, exercise, etc.) leading to worse health outcomes and potentially greater medical care consumption in the future, known as ex-ante moral hazard (Cutler and Zeckhauser, 2000).[8] Each of these behavioral responses affects how much, if any, welfare is generated by insurance coverage.[9] Therefore, efficient health insurance plan design requires an understanding of how health insurance leads to changes in individual medical care consumption behavior.

[8] The empirical health economics literature has focused primarily on ex-post moral hazard; some exceptions are Dave and Kaestner (2009) and Kelly and Markowitz (2009). Ex-ante moral hazard is difficult to study because it generally involves changes in many non-medical behaviors (e.g., exercise, diet, smoking, etc.), which lead to worse health outcomes, meaning endogenous health transitions must be modeled. The model presented in this paper allows for ex-post moral hazard and limits the effect of ex-ante moral hazard to changes in medical care consumption. That is, an individual may respond to insurance acquisition by consuming medical care less frequently, which leads to poor health outcomes and greater medical care consumption in the future.

[9] Cutler and Zeckhauser (2000) provide a comprehensive review of the welfare implications of insurance-induced moral hazard.

2.1 Estimating Moral Hazard Effects

The primary challenge in estimating the increase in medical care expenditure that is caused by health insurance possession is the endogenous selection of health insurance. If those who expect to consume more medical care during a health insurance year select generous health insurance coverage, which is known as adverse selection (Akerlof, 1970), then the positive correlation observed between medical care expenditure and insurance possession/generosity could be due to either moral hazard or selection effects.[10]

One method that has been used to identify moral hazard effects is a randomized experiment, which eliminates the endogenous selection problem by randomizing insurance possession. A well known example is the 1971 RAND Health Insurance Experiment (HIE), which randomly distributed health insurance plans to participants in six U.S. cities and collected health and medical care consumption data in the years following (Newhouse, 1974, 1993). A more recent example is the 2008 Oregon HIE, which expanded the state's Medicaid program to 10,000 additional low-income adults via lottery (Finkelstein et al., 2012; Baicker et al., 2013). There are also numerous quasi-experimental studies that have leveraged plausibly exogenous changes in public insurance programs, such as Medicaid expansion (Currie and Gruber, 1996; Dafny and Gruber, 2005) or the Massachusetts market reforms (Miller, 2012; Kolstad and Kowalski, 2012), to separately identify moral hazard effects from selection effects.

This experimental and quasi-experimental research is primarily focused on measuring the change in outcome variables caused by the introduction of a specific public policy, which makes the results from this literature difficult to generalize.[11] Moreover, a consequence of utilizing random insurance assignment to identify moral hazard effects is that an individual's health insurance decision-making process cannot be modeled, as the individual does not make an insurance decision. Therefore, this empirical strategy is not optimally suited to study a number of important economic topics related to insurance-induced moral

[10] Whether this correlation is driven by moral hazard or adverse selection has important policy implications. If moral hazard drives this correlation, then policy makers should encourage less risk protection through greater cost-sharing, reducing an individual's incentive to overconsume care. If adverse selection is the dominant force, then policy makers should encourage greater risk pooling.

[11] Interestingly, the results ofManning et al. (1987) and Keeler and Rolph (1988), who estimate a single price/co-insurance elasticity of medical care demand, are frequently used by researchers and policy makers to predict changes in medical care expenditure levels that would result from insurance plans and policies not observed in the marketplace. However, as is pointed out by Aron-Dine et al. (2013), application of the single price elasticity measure requires a researcher to characterize a health insurance plan by a single price. Because modern health insurance plans are characterized by many cost-sharing features that change the out-of-pocket price of medical care over the course of a health insurance year, there is no obvious way to summarize a plan by a single price. Aron-Dine et al. (2013) conduct an empirical exercise where they predict medical care expenditure using the single price elasticity and implement several common strategies for determining a single price. Their results show wide variation in predicted results depending on the strategy used.

hazard, such as estimating the welfare implications of moral hazard as well as the optimal design of health insurance plans and alternative sets. In this pursuit, some researchers have turned to structural modeling.[12] Importantly, structural models have allowed researchers to both control for adverse selection in order to quantify moral hazard effects and to calculate the welfare implications of these effects. Furthermore, because insurance decisions are typically modeled and insurance cost-sharing characteristics are allowed to impact optimal medical care decision making through the budget constraint, the models are well suited to study behavior and welfare responses to counterfactual insurance plans, insurance alternative sets, and regulatory policies.

The related structural models that have been designed and estimated (Cardon and Hendel, 2001; Khwaja, 2001, 2010; Einav et al., 2013; Handel, 2013; Kowalski, 2015; Bajari et al., 2014; Handel and Kolstad, 2015) have all aggregated medical care expenditures and health outcomes up to the annual or biennial level. Annual expenditure models have been popular here and elsewhere in the health economics literature, partly due to data limitations. Annual medical care expenditure data are accessible. Large public data sets, which contain total annual expenditure variables that have been cleaned and are ready for immediate use, are used by many empirical researchers and allow for nationally representative findings.[13] Also, estimation of annual expenditure models can be achieved without high frequency explanatory data, such as illness state, which is both difficult to find and necessary when estimating a model of within-year medical care decisions.[14] This research builds on these structural annual expenditure models by allowing for monthly medical care consumption decisions to be made over a health insurance year and by relaxing several assumptions commonly made in annual decision-making models.[15]

[12] See Chiappori and Salanie (2002) and Einav et al. (2010) for a review of this technique.

[13] Examples are: the MEPS; the Health and Retirement Survey (HRS), which has been cleaned by RAND; and the Medicare Current Beneficiary Survey (MCBS).

[14] Note that while insurer claims data (or claims data from a large self-insuring company, which are used in Einav et al. (2013), Kowalski (2015), Bajari et al. (2014), Handel (2013), and Handel and Kolstad (2015)) allow for the observation of high-frequency medical care consumption decisions, illness state is only observed when an individual chooses to consume care. Therefore, endogenous health transitions cannot be modeled well using claims data.

[15] Several researchers have studied health insurance and/or medical care demand using within-year behavior as an outcome. Keeler and Rolph (1988), Keeler et al. (1988), and Lin and Sachs (2016) examine medical care treatment episodes and/or expenditure in the RAND HIE data in order to study the role of insurance deductibles and stop losses. Aron-Dine et al. (2015) and Ellis (1986) study forward-looking price behavior by analyzing medical care consumption at different points in the health insurance year. Gilleskie (1998) estimates a structural model of daily medical care consumption and absenteeism during acute illness episodes. In all of these studies, insurance possession is either randomized or assumed to be exogenously determined.

2.2 Moral Hazard and Within-Year Decision Making

There are a number of reasons why we may expect the WYDM and an annual expenditure model to produce different moral hazard effects. One such reason is that the WYDM model allows for endogenous health transitions that create dynamic incentives for medical care consumption within the insurance year. As health insurance acquisition leads to greater medical care consumption early in the year due to lower out-of-pocket prices, an individual's health may improve, lessening the need for medical care consumption late in the year. Because this dynamic shift in the incentives for medical care consumption does not exist in an annual expenditure model, we may expect the annual expenditure model to estimate a larger moral hazard effect.

Another reason that the two models may produce different moral hazard estimates is that they make different assumptions about the information available to an individual when medical care decisions are made. For example, the WYDM model assumes that the effective price of medical care is unknown when an individual makes a medical care consumption decision.[16] The competing model assumes that an individual makes an annual expenditure decision while facing a non-linear budget constraint, implicitly assuming that individuals only respond to the effective price of care at the end of the year, which is known with certainty.[17]

The impact of this difference on moral hazard estimates is best understood with a simple example. Assume that the price of a hospital visit can take on two values: p_l, which occurs θ percent of the time, and p_h, which occurs $(1-\theta)$ percent of the time, where $p_h > p_l$.[18] Let Individual U be risk neutral and uncertain about the price of medical care when he makes a medical care decision; he represents an individual solving the WYDM model. Let Individual C be identical to Individual U, except that Individual C is certain about his price draw prior to making a medical care decision; he represents an individual solving an annual expenditure model. Consider the consumption decisions of Individuals U and C with and without insurance. If fully insured (i.e., the out-of-pocket price of a visit

[16] There are two reasons that these prices are unknown in the WYDM model. First, the model assumes that the contemporaneous transaction price (i.e., the sum of insurer and insured's payments) of a unit of medical care is unknown at the time of consumption. Support for this assumption is provided below, but is not a necessary condition of a general WYDM model. Second, because future health shocks and medical care preference shocks are unknown, the benefit of moving closer to one's deductible or stop loss threshold in any given month is unknown.

[17] Bajari et al. (2014) is somewhat of an exception, as the authors assume that an individual selects his total annual medical care expenditure while knowing only the distribution of the proportion of that expenditure he must pay out of pocket; thus, individuals face some uncertainty about their out-of-pocket payments.

[18] The following discussion assumes that an individual receives no benefit from the high price draw that he does not also receive from the low price draw, which is also an assumption of the WYDM model described in Section 3.

equals 0), then both individuals will visit the hospital if their willingness to pay for medical care, W, is greater than zero. There are two scenarios to consider when the individuals are uninsured: First, suppose the individuals' willingness to pay for medical care, W, is less than the expected price of medical care, $E[P] = \theta p_l + (1-\theta)p_h$, but greater than p_l. In this scenario, Individual U never visits the hospital when uninsured, as the expected price of medical care is greater than his willingness to pay, while Individual C visits the hospital any time he gets a price draw of p_l. As a result, the moral hazard effect for Individual U would be larger than that for Individual C. On the other hand, suppose $p_h > W > E[P]$.[19] In this scenario, Individual U always visits the hospital when uninsured, while Individual C again only visits the hospital when he gets a price draw of p_l; thus, the moral hazard effect for Individual C would be larger than that for Individual U.

The level of risk aversion in the sample should also impact the moral hazard effects predicted by the models, which can be seen in Figure 1. The figure features a strictly concave utility function for a non-medical good X (with price normalized to one) and the same lottery that is described above. The utility received from the expected value of the price lottery, $U(E[X])$, which is re-interpreted here as a lottery over two levels of non-medical good consumption (i.e., $E[X] = \theta(I - p_l) + (1-\theta)(I - p_h)$, where I is income), is greater than the expected utility of the gamble, $E[U(X)]$, suggesting that the individuals are risk averse. With this utility function, nothing about Individual C's medical care decision-making problem changes because he faces no risk - i.e., in both insured and uninsured states, he makes the same decisions that he made when risk neutral. Risk aversion also has no impact on Individual U's decisions when he is insured. However, consider the behavior of Individual U when he is uninsured. Under risk neutrality, Individual U did not visit the hospital when $W < E[P]$. In Figure 1, the group of risk neutral individuals not visiting the hospital are represented by area A, where $I - W > E[X]$. Under risk aversion, some individuals who value medical care slightly more than the expected price will still choose not to visit due to their risk aversion (i.e., their fear of a high price draw). The new group of individuals not purchasing care would include those both in area A and area B of Figure 1, as these individuals have income and medical care preferences such that $I - W > CE$, which is the certainty equivalent value of the non-medical care good. In other words, for a fixed W distribution in the population, the moral hazard effect predicted by the WYDM model should be increasing in the level of risk aversion in the population, while risk aversion should have no effect on the annual expenditure model.

[19] The two omitted scenarios are trivial. If $W > p_h$ then both individuals always consume when uninsured and if $W < p_l$ then both individuals never consume when uninsured. In both scenarios, the moral hazard effects for both individuals are identical.

3 Model

This section describes the optimization problem of an unmarried, childless, employed individual who makes an annual health insurance decision followed by a sequence of medical care consumption decisions to maximize the value of his expected discounted future utility.[20] The timing of the model can be observed in Figure 2. At the beginning of each year, y, a forward-looking individual observes the set of health insurance alternatives offered by his employer and the presence of any illnesses. Before the start of the first month, $t = 1$, he selects the health insurance alternative that maximizes his expected discounted future utility. Among other things, this expected utility is a function of anticipated medical care behavior within the year conditional on insurance coverage. In this paper the within-year medical care behavior is modeled explicitly.

At the beginning of each month, an individual learns his illness state, which evolves stochastically over the course of the year and is influenced by illness history and previous medical care consumption. After learning his current illness state, the individual decides how much and what types of medical care to consume. The amount he pays for a unit of medical care depends on the unit price, the cost-sharing characteristics of his health insurance plan, and his accumulated medical care expenditure within the coverage year. Much like the price uncertainty an individual faces in the US medical care market, the total price of care is stochastic over time and unknown prior to consumption. Conditional on consuming medical care, the individual learns the price he paid and updates his stock of accumulated out-of-pocket expenditure before transitioning into the next month. The remainder of this section explains the model and solution in greater detail.

3.1 Annual and Monthly Decisions

At the beginning of each year, an individual observes the set of health insurance plans offered to him by his employer. Each plan is defined by its premium, network type, and a set of cost-sharing characteristics. The cost-sharing features enter an individual's budget constraint throughout the year, determining how much is paid out of pocket for medical care. The following plan characteristics enter the model: out-of-pocket premium, composite

[20] The model focuses on employed individuals who receive an employer-sponsored health insurance offer (ESHI) because health insurance information is only available for these individuals in the data. ESHI is the most popular mechanism by which an individual obtains health insurance in the United States. Of the non-elderly population in 2014: 55% held ESHI, 12% were uninsured, 26% were insured by state or federal governments, and 7% were privately insured (Kaiser Family Foundation). The model focuses on a single, childless individual in order to explicitly capture dynamic health, non-linear out-of-pocket prices, and medical care demand throughout the insurance coverage period, which would be computationally prohibitive if it involved more than one individual. The estimation sample includes men and non-pregnant women.

annual deductible, doctor's office deductible, hospital deductible, stop loss, hospital co-insurance rate, hospital co-pay level, doctor's office co-insurance rate, doctor's office co-pay level, prescription drug co-insurance, and the extent to which the plan restricts coverage to a network of physicians (e.g., HMO, PPO, or FFS).[21] I define an indicator function, I_{iy}^{j}, that equals one if individual i selects insurance plan j in year y and zero otherwise.[22] Only one plan can be held at a time, so that

$$\sum_{j \in J^i} I_{iy}^{j} = 1 \quad \forall\, i \,\forall\, y \tag{1}$$

where J^i is the set of exogenously determined ESHI plans that includes the option to decline all plans.[23]

In each month, an individual learns his illness state (defined below) before making a medical care consumption decision. He chooses the number of doctor visits, d_{it}; hospital days, h_{it}; and whether or not to consume prescription drugs, r_{it}.[24] The monthly medical care decision is represented by an indicator function, m_{it}^{dhr}, that equals one if an individual chooses the bundle $\{d, h, r\}$ and zero otherwise. Bundles are mutually exclusive with a maximum of D doctor visits and H hospital days in each month, such that

$$\sum_{d=0}^{D} \sum_{h=0}^{H} \sum_{r=0}^{1} m_{it}^{dhr} = 1 \quad \forall\, i \,\forall\, t. \tag{2}$$

[21] A deductible and a stop loss are described in footnote 4. A co-insurance rate is the share of the medical care price that an individual must pay out of pocket; the remainder is paid by the insurer. A co-pay level is a fixed dollar amount that an individual must pay out of pocket for a unit of medical care; again, the remainder is paid by the insurer. A health maintenance organization (HMO) here refers to an insurance plan that limits its enrollees to receiving medical care from a specified group of providers. A preferred provider organization (PPO) is a plan that defines a preferred network of providers from which care can be purchased less expensively. If an enrollee chooses to seek care outside of this network, coverage is still provided but at a higher out-of-pocket price. A fee-for-service (FFS) plan covers an enrollee equally at all medical care providers.

[22] For notational simplicity and consistency, I include the subscript i to describe individual-level variables only when defining the variable. The subscript i is suppressed thereafter.

[23] Some individuals select a job based, at least partially, on the health insurance offered. However, modeling an individual's decision to accept a particular job, with health insurance options as a job characteristic, requires modeling the employment decision as a function of health insurance characteristics. Thus, this exogeneity assumption is made frequently in the literature.

[24] According to the Centers for Disease Control and Prevention (CDC) these three types of care account for over 80 percent of personal medical care expenditures in the United States. For the population of study, the percentage is even higher because individuals are non-elderly and unlikely to consume nursing home care or home health care. Other relevant medical care products, such as dental and optical, are unlikely to be covered by standard ESHI plans and are thus excluded from the study.

3.2 Illness Transitions and Probabilities

An individual's illness state evolves stochastically over the course of the insurance year. The illness state in month t is defined by the number of acute illnesses, A_{it}, and chronic illnesses, C_{it}.[25]

I define an acute illness as any medical condition that eventually subsides and, under normal conditions, has no permanent effect on an individual's health or medical care consumption. This characterization describes both short-natured ailments, such as a common cold or influenza, as well as persistent but non-permanent conditions, such as a pneumonia or a broken bone. In estimation, the number of acute illnesses that an individual has in month $t+1$ is determined by the latent variable A^*_{t+1}, which can be written as

$$A^*_{t+1} = G_A(\mathbf{W_{t+1}}, A_t, C_t, m_t^{dhr}; \alpha) + \mu_1^k + \psi_{t+1} \tag{3}$$

where $\mathbf{W_{t+1}}$ is a vector of individual characteristics such as sex, race, income, education, MSA indicator, age, initial self-reported health status, and calendar month indicators; α is a vector of linear parameters; μ_1^k captures permanent unobserved heterogeneity for an individual of type k, where $k = 1, ..., K$; and ψ_{t+1} is an i.i.d error.[26] The function $G_A(\cdot)$ is linear in parameters and includes both main and interaction effects, which allows the productivity/efficacy of medical care to vary by illness state entering the month.

I assume that A_t follows an ordered structure, so that $A_t = q$ if $\kappa_q^a < A_t^* \leq \kappa_{q+1}^a, \forall q \in \{0, \ldots, 4\}$. Assuming ψ_t follows a logistic distribution, the (ordered logit) probability of having q acute illnesses is

$$P(A_t = q) = \pi_t^q = \Lambda(\kappa_{q+1}^a - \widetilde{A}_t^*) - \Lambda(\kappa_q^a - \widetilde{A}_t^*) \tag{4}$$

where $\widetilde{A}_t^* = A_t^* - \psi_t$ and $\Lambda(\cdot)$ is the logistic function.[27]

I define a chronic illness to be any medical condition that never subsides (e.g., diabetes, asthma, AIDS) or, under normal conditions, has a permanent effect on an individual's health or medical care consumption (e.g., cancer, stroke, hypertension). Given the long-lasting effect of these ailments on health and/or medical care purchasing behavior, the occurrence of a chronic illness is modeled as a permanent, absorbing state. Medical care is then used to prevent or control a chronic illness, lessening the impact that one chronic illness has on the development of another chronic or acute illness.

[25] Death is only observed once in the data because the estimation sample includes ages 19-64, so it is not modeled as a possible health outcome. A *death state* would be a simple addition with alternative data.

[26] See Section 5.2 for a discussion of estimation and interpretation of unobserved permanent individual heterogeneity in this model.

[27] Note that $\{\alpha, \mu_1^k, \kappa_2^a, \kappa_3^a, \kappa_4^a\}$ are parameters to be estimated, while κ_0^a, κ_1^a and κ_5^a are normalized to $-\infty$, 0, and ∞, respectively.

Because chronic illnesses are permanent, the model must determine how many *additional* chronic illnesses an individual acquires in each month, which I define as C_t^+. In estimation, additional chronic illnesses in month $t+1$ are determined by the latent variable C_{t+1}^{+*}, which can be written as

$$C_{t+1}^{+*} = G_C(\mathbf{W_{t+1}}, A_t, C_t, m_t^{dhr}; \delta) + \mu_2^k + \zeta_{t+1} \tag{5}$$

As with acute illnesses, I assume C_t^+ follows an ordered (logit) structure. The probability of acquiring q additional chronic illnesses in month t is written as $\gamma_t^q, \forall q \in \{0, 1, 2\}$.

3.3 Utility Function and Budget Constraint

Preferences for medical care consumption bundle $\{d, h, r\}$ in month t are described by the following contemporaneous utility function[28]

$$U(m_t^{dhr}, \mathbf{S_t}|\mathbf{p_t}) = \frac{X_t^{\omega_0 \mathbf{R}}}{\omega_0 \mathbf{R}} + G_U(\mathbf{W_t}, A_t, C_t, m_t^{dhr}; \omega) + \mu(m_t^{dhr}, k) + \epsilon_t^{dhr} \tag{6}$$

where $\mathbf{S_t} = (\mathbf{W_t}, A_t, C_t, ADE_t, AHE_t, I_y^j, \mu^k, \epsilon_t^{dhr})$ denotes the set of information known by an individual at the time of his medical care consumption decision (i.e., his *state*); $\mathbf{p_t} = (p_t^d, p_t^h, p_t^r)$ is a vector of (known) doctor, hospital, and prescription drug prices, respectively;[29] X_{it} represents consumption of non-medical goods, which is determined by the budget constraint defined in Equation 7; $\mathbf{R} = (1, sex, race, age)$; ω is a vector of linear parameters; $\mu(m_t^{dhr}, k)$ captures permanent unobserved heterogeneity in medical care preferences for an individual of type k; and ϵ_{it}^{dhr} is the unobserved utility received from d doctor visits, h hospital days, and consuming prescription drugs ($r = 1$) or not ($r = 0$).[30] For simplicity, Equation 6 can be rewritten as the sum of deterministic and random components, $\overline{U}(m_t^{dhr}) + \epsilon_t^{dhr}$.

The monthly budget constraint is

$$X_t = Y_t - P_{jt} - O(m_t^{dhr}, \mathbf{p_t}, ADE_t, AHE_t, I_y^j) \tag{7}$$

where Y_{it} is monthly income; P_{ijt} is the month t premium paid out of pocket for plan j; $O(\cdot)$ is the out-of-pocket expenditure on medical care in month t; and ADE_{it} and AHE_{it}

[28] This utility function is representative when X_t is greater than or equal to zero. If medical care expenditure becomes so great that X_t is negative, then the first term $\frac{X_t^{\omega_0 R}}{\omega_0 R}$ is replaced by $\omega_{70} * X_t$ to capture the (dis)utility of negative non-medical good consumption.

[29] Because an individual faces a binary decision on whether or not to consume any prescription drugs, p_{it}^r measures total monthly expenditure on prescription drugs rather than the unit price per prescription.

[30] I allow for individual permanent unobserved heterogeneity in direct utility from medical care consumption by parameterizing $\mu(\cdot)$ such that $\mu(m_t^{dhr}, k) = d\mu_{3a}^k + d^2\mu_{3b}^k + h\mu_{4a}^k + h^2\mu_{4b}^k + r\mu_5^k + (1 - m_{it}^{000})\mu_6^k$.

represent accumulated out-of-pocket medical care expenditure for doctor visits and hospital days entering month t, respectively. This structure assumes that an individual consumes all income by the end of each month, as monthly saving decisions are not observed in the data.[31]

Having specified the contemporaneous utility function, budget constraint, and all transitions between uncertain illness states, it remains to describe what the model assumes about an individual's knowledge of medical care prices and out-of-pocket expenditure.

3.4 Medical Care Prices and Expenditure

Two characteristics of the medical care marketplace make the within-year medical care decision-making problem an important economic construct. First, most individuals do not pay the total price of medical care because of a cost-sharing arrangement with their health insurance provider. Rather, an individual pays a dollar amount out of pocket that is determined by the total price of medical care, insurance plan characteristics, and accumulated medical care expenditure during the coverage year. For example: an individual with a \$300 deductible, 10 percent co-insurance rate, and \$0 of accumulated expenditure who is charged \$100 for a doctor visit pays the full \$100 out of pocket. However, if the same individual were to have accumulated \$250 in medical care expenditure prior to the visit, then he would pay only \$55 out of pocket for the visit (\$50 pre-deductible + \$5 $[= 0.1 * (\$100 - \$50)]$ post-deductible). An individual with health insurance characterized by this cost-sharing structure, a deductible with a co-insurance rate, faces a non-linear budget constraint. The out-of-pocket expenditure function, $O(\cdot)$, is constructed so that the budget constraint in Equation 7 contains these non-linearities. Precise calculations of contemporaneous and accumulated out-of-pocket expenditure are detailed in Section 3 of the web appendix.

A second unique characteristic of the US medical care marketplace is that individuals are typically uncertain of the total price of medical care prior to consumption (United States Government Accountability Office, 2011; Robert Wood Johnson Foundation, 2012; Rosenthal et al., 2013). The lack of menu prices, uncertainty of diagnosis prior to a visit, and wide price variation even in local medical care markets contribute to price uncertainty.[32] Despite the evidence, surprisingly few models of medical care demand allow for

[31] A monthly savings decision cannot be observed in the data. French and Jones (2011) examine the effects of health insurance and self-insurance (i.e., savings) on retirement behavior. The authors explain that omitting savings from an individual's dynamic problem ignores the ability to smooth consumption through savings, which can potentially overstate the value of insurance. In simulation, they find that omitting savings from the model does increase the value of insurance, but retirement decision making is unchanged in the no-savings model. Applied to this model, the result suggests that one should expect the insurance decision to change little if savings were to be observed and controlled for.

[32] In May 2013, the Centers for Medicare and Medicaid Services (CMS) released data showing wide

this uncertainty. To address the reality of uncertain prices, I assume that an individual does not observe total medical care prices prior to making a medical care decision in each month. Rather, an individual knows the conditional distributions from which doctor visit prices, hospital day prices, and prescription drug prices are drawn. An individual makes medical care decisions by integrating over the three conditional price distributions, which are estimated from the data.[33]

The total price distributions are defined as $F^d(p_t^d|\mathbf{S_t};\lambda^d)$, $F^h(p_t^h|\mathbf{S_t};\lambda^h)$, and $F^r(p_t^r|\mathbf{S_t};\lambda^r)$, where $\{\lambda^d,\lambda^h,\lambda^r\}$ are parameters to be estimated. An individual's state, $\mathbf{S_t}$, contains the variables HMO_j, PPO_j, and FFS_j, which are indicators of the plan's coverage type. I control for coverage type to capture the lower rates negotiated by insurance providers who contract with a network of physicians.[34] An MSA indicator is included in $\mathbf{S_t}$ as well to capture urban area variation in prices. These distributions also depend on individual observed illness states and initial self-reported health status. Finally, these medical care price shocks are likely to be correlated with unobserved illness shocks. An individual who receives an exceptionally bad illness shock is also likely to experience a price distribution that is shifted rightward or has a fatter right tail. For this reason, the three medical care price shocks are likely to be correlated with one another as well. In order to control for these potential sources of unobserved correlation, I allow the permanent unobservables that influence preferences, illness states, and initial conditions (i.e., μ^k) to also influence the price distributions.

3.5 The Optimization Problem

An individual's objective is to maximize his expected discounted future utility by selecting the optimal sequence of medical care bundles, m_t^{dhr}, for $t=1,...,T$ and insurance plans, I_y^j, for $y=1,...,Y$ conditional on his state variables in $\mathbf{S_t}$. I describe an individual's dynamic optimization problem in two stages, as insurance decisions are made at the beginning of a year and medical care decisions are made repeatedly over the course of a year.

variation in medical care prices in local medical care markets. Such variation makes it difficult for an individual to know medical care prices prior to consumption. Recent articles in Time Magazine and The New York Times have also highlighted the issue of price uncertainty in medical care markets.

[33] An equilibrium model of the medical care market could conceivably allow for price determination in solution. Such a model, to be realistic, would have to include as players individuals, providers, hospitals, insurance companies, employers, and the government since interactions between all of these entities determine prices in the market.

[34] The model does not differentiate between in-network and out-of-network medical care consumption. All medical care is assumed to be in-network. Insurance cost-sharing characteristics are specific to in-network consumption.

3.5.1 The Optimal Monthly Decision Rule

Let $V^{ac}_{dhr}(\cdot_t)$ be the month t value of expected discounted future utility for medical care decision m^{dhr}_t and illness state $(A_t = a, C_t = c)$. Using Bellman's (1957) Equation, this value is constructed as the sum of contemporaneous utility and the expected discounted future utility yielded by the alternative. Conditional on unobserved heterogeneity type k (where $\mu^k = (\mu^k_1, ..., \mu^k_{12})$), insurance plan j, and medical care price vector $\mathbf{p_t}$, the alternative-specific value function can be written, for $t < T$

$$V^{ac}_{dhr}(\mathbf{S_t}, \epsilon^{dhr}_t | \mu^k, I^j_y, \mathbf{p_t}) = \overline{U}(m^{dhr}_t) + \epsilon^{dhr}_t + \beta \left[\sum_{a'=0}^{4} \pi^{a'}_{t+1}(\mathbf{S_t}, m^{dhr}_t) \sum_{c'=0}^{2} \gamma^{c'}_{t+1}(\mathbf{S_t}, m^{dhr}_t) \left[V^{a', c+c'}(\mathbf{S_{t+1}} | \mu^k, I^j_y) \right] \right], \quad (8)$$

and for $t = T$

$$V^{ac}_{dhr}(\mathbf{S_t}, \epsilon^{dhr}_t | \mu^k, I^j_y, \mathbf{p_t}) = \overline{U}(m^{dhr}_t) + \epsilon^{dhr}_t + \beta Q_{y+1}(\mathbf{S_0}) \quad (9)$$

where β is the monthly discount factor and $Q_{y+1}(\mathbf{S_0})$ is the value of expected discounted future utility in month $t = 0$ of year $y+1$, (i.e., prior to the year $y+1$ insurance decision). Maximal expected utility, in illness state $(A_{t+1} = a, C_{t+1} = c)$, in month $t+1$ is

$$V^{ac}(\mathbf{S_{t+1}} | \mu^k, I^j_y) = E_t \left[\max_{dhr} V^{ac}_{dhr}(\mathbf{S_{t+1}}, \epsilon^{dhr}_{t+1} | \mu^k, I^j_y) \right] . \quad (10)$$

The expectation operator is subscripted by t because an individual must form this expectation prior to learning month $t+1$ medical care preference shocks, ϵ^{dhr}_{t+1}.

The value function in Equation 8 is written conditional on realized medical care prices; however, it is assumed that an individual knows only the conditional distributions from which these prices are drawn. Solution to the optimization problem requires integration over these price distributions. The value function written unconditional on prices is

$$V^{ac}_{dhr}(\mathbf{S_t}, \epsilon^{dhr}_t | \mu^k, I^j_y) = \int_{\mathbb{R}^3_+} f^*(\mathbf{p_t}) V^{ac}_{dhr}(\mathbf{S_t}, \epsilon^{dhr}_t | \mu^k, I^j_y, \mathbf{p_t}) d\mathbf{p_t} \quad (11)$$

where $f^*(\mathbf{p_t}) = f^d(p^d_t) * f^h(p^h_t) * f^r(p^r_t)$ and $f^d(\cdot)$, $f^h(\cdot)$, and $f^r(\cdot)$ are the conditional density functions from which p^d_t, p^h_t, and p^r_t are drawn.[35]

Conditional on the prior insurance decision and unobserved heterogeneity, a utility

[35] Conditional on μ^k (i.e., unobserved permanent individual heterogeneity) these distributions are independent; however, their dependence on μ^k allows some correlation.

maximizing individual selects medical care consumption bundle $\{d, h, r\}$ with probability

$$P(m_t^{dhr} = 1) = P\left[V_{dhr}^{ac}(\mathbf{S_t}, \epsilon_t^{dhr}|\mu^k, I_y^j) \geq V_{d'h'r'}^{ac}(\mathbf{S_t}, \epsilon_t^{d'h'r'}|\mu^k, I_y^j) \quad \forall \quad d'h'r'\right] . \tag{12}$$

3.5.2 The Optimal Annual Decision Rule

The dynamic programming problem can be solved backwards to recover the time $t = 0$, year y value function conditional on any chosen health insurance alternative $j \in J_y^i$. That is,

$$V(\mathbf{S_0}|\mu^k, I_y^j) = \sum_{a'=0}^{4} \pi_1^{a'}(\mathbf{S_0}, m_{T,y-1}^{dhr}) \sum_{c'=0}^{2} \gamma_1^{c'}(\mathbf{S_0}, m_{T,y-1}^{dhr}) \left[V^{a',c+c'}(\mathbf{S_1}|\mu^k, I_y^j)\right] \tag{13}$$

where $m_{T,y-1}^{dhr}$ is the medical care consumption decision in the last month of year $y-1$ and c is the number of chronic illnesses possessed entering year y. Stated explicitly, Equation 13 represents the discounted value of optimal future behavior calculated at the beginning of year y unconditional on the first month acute and chronic illness state but conditional on insurance plan j (i.e., the expected discounted future value of plan j). I assume that this value, plus an additively separable i.i.d random error component, determines the expected future discounted value of plan j.

$$Q_y^j(\mathbf{S_0}, \phi_y^j|\mu^k) = V(\mathbf{S_0}|\mu^k, I_y^j) + \phi_y^j . \tag{14}$$

Thus, a utility maximizing individual selects insurance plan j with the probability[36]

$$P(I_y^j = 1) = P\left[Q_y^j(\mathbf{S_0}, \phi_y^j|\mu^k) \geq Q_y^{j'}(\mathbf{S_0}, \phi_y^{j'}|\mu^k) \quad \forall\, j'\right] . \tag{15}$$

Several papers in the health insurance choice literature have argued against the use of an i.i.d random error component (Einav et al., 2013; Handel, 2013; Kowalski, 2015). Importantly, each of these papers uses data from a single firm where plans are financially rankable and identical on non-financial characteristics. In such a circumstance, it is not clear what an error component would capture that is not in the model but, given the rankability of plans, the i.i.d assumption is likely inappropriate. Unlike these papers, I use data from a national sample of employees who work for hundreds of different firms

[36] In the optimization problem, an individual has knowledge of $\mathbf{S_0}$ at the time of an insurance decision, where $\mathbf{S_0}$ contains the chronic illness state learned in the last month of the previous insurance year. In the data, it cannot always be determined whether a chronic illness present in first month of the insurance year, which is often the first month of the survey period, began in the first month or a previous month. Therefore, in order to allow an individual to make health insurance decisions with knowledge of existing chronic illnesses, I assume in estimation that he learns his first month chronic illness state prior to his health insurance decision in the first year of optimization.

offering many different insurance alternative sets. Therefore, I cannot verify that plans within a firm are vertically ranked and there is little evidence to suggest that this is a good assumption generally. Moreover, recent findings by Handel (2013) and Handel and Kolstad (2015) suggest that inertia, mis-information about stated plan features, and perceptions of unobserved features (e.g., time and hassle costs associated with a plan) are important determinants of health insurance choice. Since the MEPS data reveal only a single insurance choice and no information about individual perceptions of plan features, I cannot control for these factors explicitly in estimation, further rationalizing the use of an unobserved preference component.

An i.i.d. random error component is not without presidence in the literature (Cardon and Hendel, 2001; Khwaja, 2001, 2010). Inclusion of a plan-specific error allows the model to rationalize observed insurance choices using the future expected value of the plan, which is produced by the model, as well as these unobserved determinants. Absent this error, the model alone must rationalize all observed choices, potentially tasking structural parameters with explaining behavior outside of the model. The assumption that errors are independent also has the benefit that it smooths out the likelihood function, which is computationally attractive for numerical optimization.

4 Data

4.1 Description of MEPS

My empirical analysis uses data from the 1996-1999 cohorts of the Medical Expenditure Panel Survey (MEPS).[37] MEPS contains detailed health, medical care expenditure, health insurance, and demographic information for a nationally representative sample of families and individuals in the United States. New participants are added annually, drawn randomly from the previous year's National Health Interview Survey sample. Individuals in each cohort are interviewed five times over the two years that follow January 1st of their cohort year.

The MEPS has two features that make it particularly well suited for the purposes of this research. First, detailed employer-level insurance information exists for the individuals comprising the 1996-1999 cohorts. Data collectors used information gathered in the first interview to contact current main employers, from which they obtained premium and cost-sharing characteristics for all plans offered to the employee. This data feature, which is

[37]Data collection began in 1996 and continues today. The data are collected and maintained by the Agency for Healthcare Research and Quality (AHRQ). All data used in estimation are publicly available, with the exception of an individual's insurance plan information. These restricted files may only be accessed through a Census Bureau Research Data Center (RDC). As a result, estimation of the structural model and all simulations were conducted in an RDC lab.

unique in national survey data, enables me to model a health insurance decision from the full set of available alternatives for an individual with a participating employer. However, roughly 50 percent of individuals participating in MEPS lack insurance information in this link file due to employee and/or employer refusal to disclose the information.[38] Also, while individuals are interviewed over the course of two years, few employers agree to provide health insurance plan information at the beginning of each year. Therefore, analysis concentrates on one health insurance decision and the medical care decisions in the year that follows for each individual. Second, illness information in the MEPS file is initially self-reported. An individual is asked to describe any, "health problems that may have bothered (person)," over the survey period, independent of whether they sought medical treatment.[39] Professional coders convert these descriptions into three-digit ICD-9-CM codes, some of which are later verified by medical professionals if treatment is sought. The ability to observe illness episodes even when medical care is not consumed is an important characteristic of these data, which is not characteristic of firm-level claims data, as it allows endogenous illness transitions to be modeled. Moreover, note that roughly 33 percent of the individual illnesses reported by the estimation sample do not result in any type of medical care consumption, suggesting that these illnesses represent a significant source of unobserved heterogeneity in studies that utilize firm-level claims data.

A number of important assumptions are required to prepare the data for estimation. For example, each illness, which is defined in the data by an ICD-9-CM medical code, must be interpreted as an acute or chronic illness. Also, medical care consumption dates and partially observable illness dates must be used to determine the starting and ending month of illnesses reported five times over the course of two years. I also face the challenge that at least one insurance cost-sharing features is missing in 47 percent of the 5,284 insurance plans observed in the data; imputations were made in these situations using a number of different techniques. The magnitude of these complications, and others, and the assumptions required to overcome them are discussed at length in Section 1 of the web appendix.

[38] There was one significant change to the collection process that took place after 1996. The 1996 MEPS asked each employer questions regarding their specific participating employee. This method caused many employees to refuse to provide their employers information as employees wished to remain anonymous. The method was also inconvenient for employers because it was much more difficult to provide information about a particular employee than employees in general. Therefore, in 1997 the collection process was altered such that employers were asked about their general insurance offerings, but not employee specific offerings. AHRQ then used a matching procedure to identify which offered plan was reportedly chosen by employees.

[39] Health problems are described as, "physical conditions, accidents, or injuries that affect any part of the body as well as mental or emotional health conditions, such as feeling sad, blue, or anxious about something." Individuals are instructed to, "please include all ... conditions ... regardless of whether (person) saw a medical provider, received treatment, or took medications...".

4.2 Determination of the Sample

The sample used in estimation is taken from the nationally representative sample of single and childless individuals included in the 1996-1999 cohorts of the MEPS survey. (See Table 1 for sample size by inclusion criteria.) I focus on employed individuals between the ages of 19 and 64 whose employers sponsor health insurance coverage.[40] I exclude the unemployed and those employed without an insurance offer because only general insurance information was gathered for these individuals (e.g., coverage status, coverage source, etc.). Employed individuals who receive an insurance offer but choose to be uninsured *are* included in the estimation sample. These omissions are representative of the sample restrictions found in similar work.[41]

Sample inclusion also requires that the ESHI plans offered to an individual are observed in the link file described above. Individuals must also participate in all interviews during the insurance year. The final restriction limits individuals in the sample to one of two types: (1) individuals taking up ESHI, holding it for an entire year, and holding no outside coverage; or (2) individuals remaining completely uninsured all year. I do not model insurance switching during an insurance year and cannot observe privately purchased or government provided (e.g., Medicare, Medicaid, etc.) plan characteristics.

The final estimation sample contains 1,232 individuals or 14,784 person-month observations. Table 2 compares the 4,859 individuals remaining at line three of Table 1, which is a nationally representative sample of single, childless, 19-64 year olds, who are employed and offered health insurance from their employer, to those in the estimation sample. The table reveals few differences between the samples. The estimation sample is slightly older, a little wealthier, and is comprised of a larger proportion of females. These differences contribute to higher medical care expenditure in the estimation sample. The estimation sample is also comprised of more federal employees, which is expected, as no federal employees are excluded due to employer non-response.

[40] I study individuals over 18 years old to avoid the unique decision-making process of an adolescent with possible access to his parents health insurance plans. I also exclude full time students under the age of 24 because 1996-1999 federal law allowed these individuals to stay on their parent's insurance plan. Individuals under 64 are targeted because they do not yet have access to coverage through Medicare.

[41] Cardon and Hendel (2001) limit their sample to single, childless, employed individuals who are between the ages of 18 and 64. However, these authors include individuals who are not offered health insurance by their employer. Einav et al. (2013), Kowalski (2015), and Bajari et al. (2014) estimate their models using a sample of individuals employed by one firm. None of the papers include uninsured individuals in the analysis. Einav et al. (2013) model the decisions of families. Kowalski (2015) models the behavior of individual employees, but allows individuals to be in a family of three or fewer people. Bajari et al. (2014) limits analysis to those holding single health insurance coverage.

4.3 Sample Statistics

This section summarizes the key variables used in estimation. Table 3 compares insured and uninsured individuals in the estimation sample.[42] The insured are older, more educated, wealthier, and are more likely to be white and female. The insured are also more likely to enter the year with a chronic illness, are more likely to get an acute illness at some point during the insurance year, and have more months where some acute illness is experienced. Each of these facts provides evidence of adverse selection. The insured also consume more units and greater values of doctor and prescription drug care.[43] The percentage of the population that consumes at least one hospital day during the year seems large (25% of insured and 17% of uninsured individuals), but includes emergency room visits as well as outpatient and inpatient visits. To reduce estimation time, I limit the maximum number of doctor visits and hospital days in a month (i.e., D and H from Equation 2) to nine and five, respectively.[44] The insured face lower transaction prices for doctor visits and hospital days and higher prices for prescription drugs.[45] The large variance in prices is due to the broad classification of medical care consumption types. High priced procedures (e.g., outpatient hospital surgery) and low priced procedures (e.g., emergency room visit for a sprained ankle) contribute to the same price distribution.

Chosen and rejected plans are compared in Table 4. The table suggests that individuals have a general preference for lower premium and therefore less generous plans. Compared to the average rejected plan, held plans are more likely to have a deductible, less likely

[42] Most variables are self-explanatory. Income is calculated as the sum of post-tax income, sale earnings, and tax refund. Medical care prices are only observed when medical care is consumed. An individual is considered to have *consumed any preventative care* if during the sample period they consume any form of medical care in a month in which they have no acute or chronic illnesses. For more detail on medical care prices, medical care consumption, and illness occurrence see Section 1 of the web appendix.

[43] While the uninsured are less likely to have at least one hospital day, the average number of hospital days for the uninsured is greater than that of the insured. This finding is likely due to emergency room usage among the uninsured.

[44] Of the 14,784 person-month observations in the data, the number of doctor visits exceeds the maximum of nine only 33 times and the number of hospital days exceeds the maximum of five only 30 times. In these instances, the number of visits/days is set to the maximum and the average price paid for a unit of medical care is adjusted accordingly. For example, if an individual visits the doctor 12 times in a month with an average price of $100 then the data are adjusted so that he visits the doctor nine times with an average price of $133.

[45] The (somewhat counterintuitive) fact that insured individuals pay higher monthly transaction prices for prescription drugs than uninsured individuals is an expected by-product of simplifying assumptions imposed on the model. The model does not allow for selection on prescription quality and, within a month, only allows for a prescription drug decision on the extensive margin (i.e., an individual decides whether to consume prescription drugs, not how much to consume). If the average uninsured individual is more likely to purchase generic medications and/or purchases fewer prescriptions in a month conditional on purchasing any, then (monthly) transaction prices should be lower than that of the average insured individual. The consequence of this simplifying assumption is that the model may over-state the benefit of prescription drug consumption for the uninsured; however, I find little evidence that this has a significant effect on moral hazard estimates.

to have a stop loss, and set higher thresholds when the plan has a deductible or stop loss. Held plans also feature higher co-insurance rates and co-pay levels for both doctor and hospital care, with the exception of hospital per day co-pay. Note that the survey does not elicit cost-sharing information for prescription drugs. Therefore, I assume that HMO, PPO, and FFS plans have a fixed prescription drug co-insurance rate of 13%, 17%, and 19%, respectively, and that prescription drug expenditure is completely unrelated to a plan's deductible and stop loss. These assumptions are justified in Section 1.3.2 of the web appendix.

While a formal discussion of identification is presented in Section 5.3, identification of many of the WYDM model's parameters requires individual-level month-to-month variation in state variables. For example, estimates of the productivity/efficacy of medical care consumption are identified by the month-to-month covariance between observed medical care consumption and illness transitions. As such, I provide transition matricies that show substantial month-to-month variation in illness and medical care consumption within an individual in Section 1.5 of the web appendix. Also in this section, I highlight the month-to-month variation in expected effective medical care prices by showing how the average proportion of total medical care prices paid out of pocket changes over the insurance year. To summarize this analysis, I find that 33% of the sample faces a budget constraint where the proportion of total medical care prices paid out of pocket is fixed over the course of the year.[46] The rest of the sample faces the possibility of a change in the proportion, of which 26% actually experience a change. Among these individuals, the average proportion of prices paid out of pocket falls from 97% in the first month of the insurance year to 13% by the final month.

5 Empirical Implementation

For each individual surveyed in the 1996-1999 MEPS, I observe one health insurance decision followed by medical care consumption, prices, and illness states over the following year. Therefore, I estimate the structural parameters of the model described in Section 3 using one year of data for each individual. In what follows, I describe the challenges in estimation, discuss identification and unobserved heterogeneity, and construct the estimated likelihood function.

[46] For this group, the expected effective price of a particular type of medical care consumption varies over the course of the year only due to changes in illness status. This group is comprised of 113 uninsured individuals, who always pay 100%; 28 fully insured individuals, who always pay 0%; and 269 insured individuals with no dynamic cost-sharing features. On average, this last group of individuals pays 7% (s.d. 6.6) of total medical care prices out of pocket.

5.1 Approximating the Future Value of a Medical Care Alternative

Solving the optimization problem requires calculating an individual's value function for each medical care bundle in each month. According to Equation 8, the value of bundle $\{d, h, r\}$ in month t, $V_{dhr}(\mathbf{S_t})$, is a function of the maximal expected utility in the next month, $V(\mathbf{S_{t+1}})$, where the future state vector, $\mathbf{S_{t+1}}$, is unknown. Thus, in order to calculate $V_{dhr}(\mathbf{S_t})$ in practice, a value $V(\mathbf{S_{t+1}})$ is needed for every potential outcome of $\mathbf{S_{t+1}}$ following every potential history of outcomes $(\mathbf{S_0}, \ldots, \mathbf{S_t})$. Given the number of months in a year, the number of variables in the state vector, and the fact that several of the state variables are continuous, the number of required future values grows exponentially.

To avoid what Bellman (1957) refers to as the "curse of dimensionality," I use an interpolation technique developed by Keane and Wolpin (1994) to approximate an individual's maximal expected future value in each month. The method works as follows: beginning in the last month of a year, $t = T$, I draw 3,500 random outcomes of the state vector, $\mathbf{S_T}$. I then calculate maximal expected future utility for each draw according to Equation 10. By estimating a linear regression of these values on the state variables in $\mathbf{S_T}$, I generate a mapping from any possible state to expected future values.[47] This mapping can be used in month $T-1$ to approximate the maximal expected future value of month T, because each alternative in month $T-1$ generates a probability distribution over $\mathbf{S_T}$. By repeating this process backwards, I can solve the model back to month $t = 0$.

A related challenge in solving any finite horizon dynamic problem is determining the maximal expected future value in the final period; $Q_{y+1}(\mathbf{S_0})$ from Equation 9. I take a popular approach, which is to formulate a closing function to approximate this value. I assume the value is determined by a non-stochastic linear function of the state variables entering the first month of the following year, medical care consumed in month T, and a vector of parameters. I estimate these parameters as part of the MLE procedure.[48]

[47] This regression includes as controls (1) state variables, which includes insurance plan characteristics; (2) interactions between state variables; and (3) the deterministic components of the current month utility function, which is advised by Keane and Wolpin (1994). I do not report parameter estimates, as the regressions include 37 parameters, which are estimated for all four unobserved types in each of the 12 months. Parameter signs and magnitudes are intuitive and the average R^2 of from these regressions at the maximum of the likelihood function is .99. These results are available upon request.

[48] As seen in Equations 8 and 9 an individual seems to integrate over future illness states when calculating his expected future value of medical care consumption when $t < T$ and not when $t = T$, but this is not the case. The insurance decision for year $y+1$ takes place prior to the realization of month $t = 1$ illnesses, meaning the integration over possible future illness states simply takes place as part of the insurance decision (see Equation 13). That said, because I only model one insurance decision, I assume that an individual *does* in fact integrate over possible future illness in month $t = T$ in estimation.

5.2 Unobserved Heterogeneity

In this model, illness outcomes, medical care prices, medical care decisions, and initial conditions are each partially determined by unobservables, which can be decomposed into two, additively separable components. The first component is a non-linear discrete factor, μ^k, that represents individual-specific persistent unobserved heterogeneity. The second component is the remaining i.i.d. serially-uncorrelated random error. I allow the distribution of persistent unobserved heterogeneity to be approximated by a discrete step-wise function, which is sometimes referred to as the discrete factor random effects (DFRE) method (Heckman and Singer, 1984; Mroz and Guilkey, 1992).[49] In practice, the population is assumed to have K unobserved types, which are drawn from discrete distribution Θ. The estimation technique determines the value of each discrete factor, $\{\mu_1^k, ..., \mu_{12}^k\}$, for K-1 types (identification requires that one type has all discrete factors set to zero) and the probability of each type, θ^k; where $\sum_{k=1}^{K} \theta^k = 1$.

Intuitively, the DFRE technique allows individuals to vary along a permanent unobserved dimension that is correlated with the endogenous variables of the model. While it is not possible to determine exactly what these unobservables are, the technique identifies unobserved *types*, which vary in their preferences for medical care, illness transition probabilities, and medical care price draws. Importantly, the expected value of an insurance plan is allowed to vary by these types; therefore, the modeling of permanent unobserved heterogeneity plays an important role in controlling for adverse selection in the model, as heterogeneous preferences for medical and unobserved health expectations are thought to be important determinants of insurance selection.

5.3 Identification

The challenge of separately identifying moral hazard effects from adverse selection is straight-forward. An individual selects a health insurance plan based on his expected future health and medical care spending. The health insurance election alters medical care demand through its effect on out-of-pocket prices (i.e., the budget constraint). Because expected expenditure at the beginning of the year, which is unobserved, correlates with both the insurance decision and medical care demand, any naive estimate of the impact of insurance on medical care demand is biased.

There are two *ideal* solutions to this problem. The first solution is random assignment of health insurance plans (Manning et al., 1987), which is discussed in Section 2. Though

[49] Mroz (1999), and more recently Guilkey and Lance (2013), use Monte Carlo simulation in a two-equation setting to show that when the true error distribution is joint normal, DFRE estimates are comparable to those derived using the correct distribution. However, when the true error distribution is not normal, the DFRE outperforms all other tested estimation methods.

health insurance plans are endogenously selected in the MEPS data, substantial variation across alternative sets aids in identification in a way that is similar to random assignment.[50] Assuming firm-specific insurance alternative sets are exogenously determined, variation across firms leads observationally equivalent individuals to possess different plans and, therefore, face a different sequence of out-of-pocket medical care prices over the insurance year. In a more reduced-form model of medical care spending, this strategy would be similar to instrumenting for plan generosity with the generosity of the insurance alternative set. Cardon and Hendel (2001) identify moral hazard effects utilizing similar variation in the National Medical Expenditure Survey (NMES). In a more recent study, Einav et al. (2013) utilize within-individual variation in the insurance alternative set, which occurs over time within a firm, but must also deal with limited variation in cost-sharing characteristics between plans, which is needed to identify price sensitivity. Ultimately, while plausibly exogenous variation in insurance alternative sets aids in identification, I do not expect this strategy to fully solve the selection problem as some individuals select their place of employment and, therefore, their insurance alternative set as well.

A second ideal solution is to jointly estimate insurance and medical care demand while conditioning the insurance decision on an observed report of expected medical care expenditure, thereby controlling for the source of endogeneity. Unfortunately, health surveys rarely inquire about an individual's expected annual medical care expenditure at the time of an insurance purchase. Therefore, the solution employed here and in other related structural work is to use the model to calculate an individual-specific expected value of each plan (see Equation 13) and condition the insurance decision on *this* expectation. With regard to this research, a key identifying assumption is then that the determinants of insurance choice which are not modeled, ϕ^j, are orthogonal to the unobserved determinants of illness, medical care prices, and medical care demand in the future.

When utilizing this strategy, constructing a realistic model of individual behavior is critical to identification. Backwards solution of the model must produce expected plan values that capture heterogeneity in individual preferences and health expectations. This paper is unique in its use of within-year medical care decisions and observed health transitions to construct the expected value of a plan. By modeling medical care decisions over the insurance year, I am able to allow for heterogeneity in individual preferences for medical care types, illness outcomes, and risk. Preferences over risk are particularly important in this setting because of the medical care price and health uncertainty that individual's face over the course of the year. By modeling observed health transitions, I ensure that

[50] The MEPS data contain a national sample of individuals who work for different employers, meaning each individual is offered a different set of health insurance alternatives from which to choose. Furthermore, employers subsidize these plans at different rates, creating additional variation in premiums across firms.

individuals who enter the year in a poor health state, or who possess observed and/or permanent unobserved characteristics that are associated with future poor health states, place a higher value on generous insurance. In comparison, Cardon and Hendel (2001) and Einav et al. (2013), who each assume that medical expenditure and health are perfect substitutes, calculated the expected value of a plan by integrating over a distribution of future annual health outcomes. The parameters of this distribution are then estimated to rationalize the insurance and expenditure decisions observed in the data but, unlike in this analysis, the distribution is not informed by actual health outcomes. Handel (2013) and Handel and Kolstad (2015) calculate the expected value of a plan by integrating over the distribution of expected medical care utilization, which is informed by prior year claims data and ex-post expenditure realizations. While this technique does utilize individual-level health data, it also assumes that an individual's medical care expenditure does not respond to the characteristics of his health insurance plan (i.e., no moral hazard) and that the individual has no private information about his health or medical preferences. Both of these assumptions are relaxed in this work.[51]

5.4 Estimation Procedure

I estimate the model's parameters, Ω, using a nested fixed point solution algorithm (Rust, 1987). The inner algorithm solves the dynamic programming problem for a given set of parameters and for each mass point in the unobserved heterogeneity distribution. The outer algorithm uses the resulting probabilities and densities to calculate the likelihood function, $L(\Omega)$, and attempts to improve the likelihood value using a BHHH gradient method (Berndt, Hall, Hall, and Hausman, 1974).

An individual's contribution to the likelihood function is the product of his observed illness state, medical care price, medical care choice, and insurance choice probabilities. Illness state probabilities take on closed forms due to the logit assumptions in Section 3. I assume that doctors prices are drawn from a Singh Maddala distribution and that hospital

[51] In addition to modeling medical care decision making over the insurance year, this paper is unique in its explicit modeling of individual illness transitions. As such, the separate identification of medical care productivity, medical care preferences, and illness preferences is worth discussing. The effect of medical care on illness probabilities is identified by individuals in the same illness state making different medical care consumption decisions and getting different illness outcomes the following month. The marginal (dis)utility of medical care is identified by variation in the *types* of medical care consumed, holding state variables fixed. Insurance cost-sharing features help identify these parameters by often making the expected out-of-pocket price of medical care equal across care types. The marginal (dis)utility of illness is identified by the optimization framework and by variation in *total* monthly medical care consumption for individuals in different wellness states. An acute illness yields disutility if an individual with an acute illness consumes more medical care (and therefore takes on the costs of consuming care) than an identical individual without an acute illness, in an attempt to cure the illness.

and prescription drug prices are distributed log-normal.[52] Each of these distributions is characterized by both scale and shape parameters; I allow the scale parameter to vary by state variables.

I assume ϵ_t^{dhr} and ϕ_y^j each follow a Type 1 Extreme Value (T1EV) distribution. This assumption simplifies estimation in two ways. First, when ϵ_t^{dhr} is T1EV the expectation in Equation 10 is equal to

$$V(\mathbf{S_t}|\mu^k, I_y^j) = EC + ln\left(\sum_{d=0}^{D}\sum_{h=0}^{H}\sum_{r=0}^{1} exp\left(\overline{V}_{dhr}(\mathbf{S_t}|\mu^k, I_y^j)\right)\right) \quad \forall t \tag{16}$$

where EC is Euler's Constant. The assumption simplifies solution to the optimization problem, as calculation/simulation of a $[(D+1)*(H+1)*2]-1$ dimensional integral would be required if another popular continuous distribution without a closed form (e.g., normal) were chosen (see Keane and Wolpin, 1994). Second, the additive T1EV distribution assumptions yield choice probabilities that have the following closed form structures.

$$P(m_t^{dhr}=1|\mathbf{S_t}, \mu^k, I_y^j) = \frac{exp\left(\overline{V}_{dhr}(\mathbf{S_t}|\mu^k, I_y^j)\right)}{\sum_{d'=0}^{D}\sum_{h'=0}^{H}\sum_{r'=0}^{1} exp\left(\overline{V}_{d'h'r'}(\mathbf{S_t}|\mu^k, I_y^j)\right)} \quad \forall t, \forall dhr \tag{17}$$

$$P(I_y^j=1|\mathbf{S_0}, \mu^k) = \frac{exp\left(\overline{Q}_j(\mathbf{S_0}, \mu^k)\right)}{\sum_{j'=0}^{J^i} exp\left(\overline{Q}_{j'}(\mathbf{S_0}, \mu^k)\right)} \quad \forall y, \forall j. \tag{18}$$

The likelihood contribution for individual i in month t conditional on μ^k and I_y^j is

$$\begin{aligned} L_{it}(\Omega|\mu^k, I_y^j) = &\prod_{a=0}^{4}\left[\pi_t^a(\cdot|\mu^k)\right]^{\mathbb{1}_{A_t=a}} \prod_{c=0}^{2}\left[\gamma_t^c(\cdot|\mu^k)\right]^{\mathbb{1}_{C_t^+=c}} \\ &\prod_{d=0}^{D}\prod_{h=0}^{H}\prod_{r=0}^{1}\Big[f^d(p_t^d|\mu^k)^{[1-m_t^{0hr}]} f^h(p_t^h|\mu^k)^{[1-m_t^{d0r}]} f^r(p_t^r|\mu^k)^{[1-m_t^{dh0}]} \\ &P(m_t^{dhr}=1|\mathbf{S_t}, I_y^j, \mu^k)\Big]^{m_t^{dhr}}. \end{aligned} \tag{19}$$

The first row contains the illness state contribution for month t. The price densities are in the second row. The $[1-m_t^{0hr}]$ exponent ensures that the price of a doctor visit in month t contributes to the likelihood function only if an individual actually visits the doctor, which is the only time that I observe this price. (The same is true for each type of care.) The third row contains month t medical care choice probabilities.

[52] For each medical care price, I tested a total of eleven parametric distributions for goodness-of-fit following the methodology of Jones et al. (2014). Details are provided in Section 2 of the web appendix.

The model is estimated using one year of data from individuals at various stages of life. An individual's health and illness states entering the year are a function of prior medical care consumption and related behaviors, as well as his personal preferences for good health and medical care, making these initial conditions endogenous. To control for this endogeneity, $L_{i1}(\Omega|\mu^k, I_y^j)$ appears as above with the first row replaced by

$$\prod_{s=0}^{4}\left[\tilde{\eta}^s(\cdot|\mu^k)^{\mathbb{1}_{HS=s}}\right]\prod_{a=0}^{4}\left[\tilde{\pi}_1^a(\cdot,s|\mu^k)\right]^{\mathbb{1}_{A_1=a}}\prod_{c=0}^{4}\left[\tilde{\gamma}_1^c(\cdot,s|\mu^k)\right]^{\mathbb{1}_{C_1=c}}. \tag{20}$$

These initial probabilities $(\tilde{\eta}, \tilde{\pi}, \tilde{\gamma})$ are separately estimated from transition probabilities, with exclusion restrictions, and are allowed to vary by modeled permanent unobserved heterogeneity.[53] Notice, first, that I model the probability that an individual has between zero and four chronic illnesses, rather than *additional* chronic illnesses, as the total sum of and individual's chronic illnesses entering the year is likely endogenous. Second, though I do not model transitions in overall health status over the course of an insurance year, I do allow initial self-reported health to impact illness transition probabilities. That said, initial self-reported health status HS_0 must be modeled explicitly. I use ordered logit probabilities, as self-reports are of the form: excellent, very good, good, fair, poor.[54]

The total likelihood contribution for individual i conditional on μ^k is then written

$$L_i(\Omega|\mu^k) = \prod_{j=1}^{J^i}\left[P(I_{iy}^j = 1|\mathbf{S_0}, \mu^k)\prod_{t=1}^{T}L_{it}(\Omega|\mu^k, I_y^j)\right]^{I_{iy}^j}. \tag{21}$$

The contribution of individual i unconditional on the unobserved heterogeneity is

$$L_i(\Omega) = \sum_{k=1}^{K}\theta^k L_i(\Omega|\mu^k). \tag{22}$$

6 Results

This section begins with a discussion of the estimated model parameters and model fit. The objective of this discussion is to provide evidence that (1) the estimated parameters are

[53] Income last year, an indicator for missing last years income, veteran status, and foreign birth status are included in the estimation of initial condition probabilities but not in transition probabilities.

[54] A model that is more consistent with Grossman (1972) might allow for self-reported health status, or *health stock*, to update in each month after medical care is consumed. Unfortunately, health status is reported only once per interview period, meaning I observe the measure a maximum of three times over the course of the health insurance year. As such, I allow initial health status to impact illness transition probabilities and rely on the sum of acute and chronic illnesses to capture changes in an individual's health stock over the course of the insurance year.

consistent with theoretical priors on individual behavior relating to medical care consumption and illness and (2) the model can explain unique features of the data; each of which suggests that the model adequately represents the true unknown data generating process. I then provide estimates of the impact that moral hazard has on medical care expenditure and health and compare these estimates to the literature. The section closes with a simple test of the Grossman (1972) health capital model and a series of policy experiments.

6.1 Parameter Estimates

Table 5 reports estimated preference parameters. A constant relative risk aversion (CRRA) parameter, RA, can be calculated for each individual using $\{\omega_{00}, \omega_{01}, \omega_{02}, \omega_{03}\}$.[55] Non-whites and males are found to be less risk averse than whites and females. Risk aversion is decreasing in age.[56] At the sample mean, $RA = 0.909$. This estimate is between the Blau and Gilleskie (2008) estimate of 0.96 and the Imai and Keane (2004) estimate of 0.74, though these estimates do not allow for heterogeneity in risk preferences.[57]

Parameters $(\omega_{10}, \ldots, \omega_{22})$ capture *contemporaneous* disutility from acute and chronic illness. The parameters suggest that for a 40-year-old individual, a single acute illness causes roughly three times as much contemporaneous disutility as a single chronic illness;[58] however, the expected lifetime disutility due to a single chronic illness is larger than that of an acute illness by several orders of magnitude.[59] This finding is consistent with the prior that most acute illnesses cause intense short-term discomfort (e.g., influenza, bronchitis, etc.) but have little impact on future health or utility, while many chronic illnesses cause

[55] The CRRA risk aversion parameter is calculated as $RA = [1 - \omega_{00} - \omega_{01} * age - \omega_{02} * nonwhite - \omega_{03} * male]$. Age is scaled in estimation so that the youngest individual included (19) has an age of 0. A 40 year-old, then, has an age of 21.

[56] There is a large literature on the relationship between risk aversion and demographics. Croson and Gneezy (2009) summarize the literature on gender and risk and cite many papers that find women to be more risk averse. Eckel (2008) also finds that women are more risk averse than men, but find no significant race effect and mixed age effects. Rosen et al. (2003) find gender and race effects that are similar to those that I find, but do not study age.

[57] Cohen and Einav (2007) and Handel (2013) also estimate risk preferences in an insurance contract setting, but assume that individual preferences are constant in the level of absolute risk aversion (CARA). To compare our estimates, I calculate the x that makes the mean individual in the estimation sample indifferent to a 50/50 gamble where he either wins \$100 or looses \$$x$ (i.e., the x that solves $\{u(w) = \frac{1}{2}u(w+100) + \frac{1}{2}u(w-x)\}$). I estimate $x = \$95.4$ at the sample mean. Cohen and Einav (2007) and Handel (2013) estimate $x = \$76.5$ and $x = \$91.0$, respectfully, at the sample mean, but allow for substantially more heterogeneity in risk preferences that I do, which explains some of this difference. These papers estimate $x = \$99.7$ and $x = \$96.3$, respectfully, at the sample median, while my estimate is $x = \$96.3$.

[58] Parameter ω_{10} is large relative to ω_{20}, which can be misleading. Note that the disutility of acute illness is decreasing in age, while the disutility of chronic illness is increasing in age; thus ω_{10} (ω_{20}) is significantly smaller (larger) without the age interaction.

[59] Using parameters $(\omega_{10}, \ldots, \omega_{22})$ and the closing function parameters presented in Table 11, one can determine that the expected lifetime disutility from a single chronic illness is almost nine times greater than that of a single acute illness in the last month of the insurance year.

less discomfort on a daily basis (e.g., hypertension, heart disease, etc.) but shorten one's expected lifespan, which dramatically reduces expected lifetime utility. The estimates also suggest that the disutility of acute illness is decreasing in age and in the number of acute illnesses, while the disutility of chronic illness is increasing in age and the number of chronic illnesses.

The direct utility effect of medical care consumption is captured by $(\omega_{30}, \ldots, \omega_{61})$. These parameters can be interpreted as the *net* direct effect of medical care consumption on utility - the physical, psychological, and time cost of medical care consumption may have negative effects on these parameters, while some individuals may enjoy consuming medical care, independent of its productive health effects, which has a positive effect on these parameters. Interpreting the linear and quadratic consumption terms is not useful without also considering the discrete factor terms discussed in the next paragraph. However, these parameters do reveal that, relative to women, men have significantly lower preferences for prescription drugs and any medical care consumption, yet stronger preferences for doctor and hospital care. The parameters also suggest that preferences for hospital visits and prescription drugs are increasing in age.

The parameters in Table 6 describe the discrete step-wise function used to approximate the joint distribution of unobservables in the model. Recall that the technique identifies unobserved *types* and estimates the relative effect that these types have on the model's probabilities, along with the probability of being a particular type.[60] For identification, I fix the mass points and probability parameter of Type 1 individuals to zero. Estimation reveals that 51 percent of the population is Type 1. The medical care preference parameters in Table 5 are fully representitive for these individuals; therefore, for the most representative unobserved type, preferences for doctor, hospital, and prescription drug consumption are negative. Individuals represented by the other three unobserved groups are sicker, have less distaste for doctor and hospital visits, and receive higher price draws for doctor and hospital visits (on average) than Type 1 individuals. Type 4 individuals are unique in their strong distaste for prescription drugs and abnormally low doctor's office price draws.[61]

Estimating the productive effects of medical care on health and/or illness is challeng-

[60] The number of points of support (i.e., types) is chosen by the econometrician. It is suggested by Mroz (1999) that this number should be chosen using an "upwards-testing approach based on the increase in the quasi-likelihood function value when one adds an additional point of support." This technique is cost prohibitive in this work because additional mass points increase estimation time substantially. Instead, I use an upwards-testing approach that requires a significant improvement in the likelihood function *and* model fit to add additional points of support. I arrive at four mass points using this technique.

[61] The parameters of the model estimated without any permanent unobserved heterogeneity (i.e., one mass point) are similar to those of the preferred model; parameter signs are mostly identical, though magnitudes vary. That said, allowing for permanent unobserved heterogeneity significantly improves model fit. Both parameter estimates and model fit analysis without permanent unobserved heterogeneity are available upon request.

ing for a number of reasons. The unconditional correlation between medical consumption and wellness is usually negative because an individual consumes more medical care when sick. The negative correlation likely reflects bias associated with selection into consumption and omitted health or medical care heterogeneity. In this research, I address the first issue by modeling medical care consumption and allowing for common unobserved individual heterogeneity that affects both medical care decisions and health outcomes. I address the second issue by simultaneously controlling for the number of acute illnesses and chronic illnesses entering the month, as well as general health status at the beginning of the insurance year, and by allowing the productivity of each type of medical care to vary by illness state. Table 7 reports acute and chronic illness probability parameter estimates. I find evidence of medical care productivity across all types of care and illness. Both acute and chronic illness probabilities decrease with additional doctor and hospital visits ($\alpha_{40}, \alpha_{42}, \delta_{40}, \delta_{42}$); however, each additional visit is less productive than the previous ($\alpha_{41}, \alpha_{43}, \delta_{41}, \delta_{43}$). Prescription drugs seem to be the most productive form of medical care in curing and preventing illness (α_{44}, δ_{44}), but these parameters are difficult to interpret given that prescription drugs are only consumed on the extensive margin.[62]

Price, initial condition, and closing function parameters can be found in Tables 9, 10, and 11, but are not discussed for brevity.

In summary, the estimated parameters support a set of theoretical priors used to develop this economic model of dynamic decision making. Specifically, the model suggests that the (explained) contemporaneous effect of medical care consumption on utility is negative - individuals have negative preferences for medical care and the consumption of medical care reduces one's ability to consume non-medical goods. Individuals are found to consume medical care because it promotes wellness in future months, from which they derive utility, and because of heterogeneous unobserved preference shocks.

6.2 Model Fit

To assess the model's ability to explain unique features of the data, I use the model to simulate a new dataset and compare key moments of the observed and simulated data. The simulated data are constructed by sampling from the joint error distribution and permanent unobserved heterogeneity distribution 100 times for each individual. I then forward simulate annual insurance decisions and monthly medical care decisions, illness transitions, and medical care prices for each of the 1,232 observed individuals, using observed insurance offer sets, demographics, and initial conditions. The resulting dataset contains 1,478,400 $[= 1,232 * 12 * 100]$ person-month observations.

[62] Medical care consumption types are not interacted with illness states in the chronic illness probability because there is not enough variation in the data to identify the parameters.

Table 12 summarizes the observed and simulated data. The model is able match illness and price moments well, which is unsurprising, given that these moments are fairly stable over the insurance year. Medical care consumption patterns are more difficult to explain, as consumption varies substantially by illness state and the annual distribution is non-normal. The first row of the third section in Table 12 shows that the likelihood of medical care consumption in any given month in the simulated data matches that of the observed data well. More importantly, the next three rows show that the model is capturing the relationship between illness and consumption, suggesting that much of the variation in simulated medical care consumption is being created by illness and not just the logit structure. Furthermore, mean and median annual medical care consumption counts in the simulated data are close to those of the observed data, though the model under-predicts prescription drug consumption by a small amount.

While the model matches mean annual medical care consumption well for all three types of care, the variances of the consumption distributions are larger in the observed data. The source of this discrepancy is better understood by looking at the entire distribution of annual medical care consumption in Table 13. Notice that for each type of medical care consumption, at least 28% of the observed population consumes zero units in a year; however, 14.12% and 27.52% of the observed population consume more than 10 doctor visits and prescription drug months in a year, respectively. As the table reveals, the model does a fairly good job of capturing the fact that many individuals consume no medical care in a calendar year, while many others consume large amounts.[63] That said, the model tends to under-predict both the share of the population with zero medical care consumption and the maximum amount of consumption for each type of care.[64]

Figure 3 shows both observed and simulated annual expenditure distributions on a log scale. Total annual medical care expenditure is a good indicator of overall model fit, as it encompasses the fit of all three consumption and price distributions. The figure reveals an under-prediction of zero medical care expenditure (i.e., roughly 17% of the simulated sample vs. 22% of the observed sample), an over-prediction of the median level of expenditure, and an under-prediction of the mean level of expenditure (see Table 12). These differences in the expenditure distribution can be explained by the differences in medical care consumption discussed above. Furthermore, Table 12 reveals that the model

[63]The DFRE method improves the fit of all three consumption distributions. Namely, the model that allows for four unobserved types generates significantly more zero consumers and right-tail consumers than the model with only one unobserved type. Adding a fifth mass point did not further improve the fit of these distributions, though it increased estimation time substantially.

[64]The maximum number of doctor and hospital visits in a year observed in the data 103 and 37, respectively. The corresponding maximums in the simulated data are 67 and 22. Furthermore, 22.7% of the observed population consumes prescription drugs in every month while only 13.2% of the simulated population does.

under-predicts insurance possession by 2.3 percentage points, which makes sense given that the simulated expenditure distribution has a mean and standard deviation that are smaller than what is observed in the data. In other words, the model slightly under-predicts insurance possession because individuals are not worried enough about being in the right tail of the expenditure distribution.

6.3 Moral Hazard

In concert with the literature, the moral hazard effect that I focus on is the percentage of mean annual medical care expenditure that would not occur in the absence of insurance. To calculate this statistic, I use the estimated model to forward simulate the behavior of individuals when they are uninsured (i.e., are responsible for the full costs of medical care consumed). Next, I simulate behavior when individuals are required to select a health insurance plan from the alternative set offered by their employer.[65] I consider two scenarios when insured: either the individual selects his most preferred plan or he is forced into *full* insurance coverage (i.e., the individual pays zero dollars out-of-pocket at all times during the insurance year.) Mean total expenditure is \$663 when uninsured and \$1,435 when insured in the preferred plan. Column 1 of Table 14 reports the preferred plan moral hazard distribution (i.e., the distribution of the difference in total annual medical care expenditure when insured versus uninsured). The additional consumption induced by moral hazard amounts to 53.1 percent (s.d. = 4.0), on average, of total annual medical care expenditure when insured. Put differently, individuals consume 2.16 times more medical care (in total value terms) when insured than when they do not participate in cost-sharing with an insurance company. These moral hazard effects vary widely across the population and are driven by those with very large increases in expenditure. Upon winsorizing the moral hazard distribution at the 99th percentile (i.e., spending increases are capped at \$10,906), the percentage of mean expenditure explained by moral hazard falls to 46 percent (s.d. = 3.1). Furthermore, 54.4 percent (s.d. = 10.3) of the sample does not increase their spending at all in response to coverage. Column 2 reports the predicted increase in expenditure when individuals are moved from no coverage to full coverage.

Columns 1-3 of Table 15 report changes in medical care consumption, prices, and illness due to the change in insurance coverage.[66] Insurance acquisition increases the consump-

[65] In both stages, for each individual I sample 100 times from the joint error distribution, unobserved heterogeneity distribution, and multivariate normal parameter distribution. (Under regularity conditions, $\sqrt{n}(\widehat{\theta}_{MLE} - \theta_0) \rightarrow N(0, \{I(\theta_0)\}^{-1})$, where θ_0 is the true parameter vector and $I(\theta_0)$ is Fisher's information matrix). A standard error for the mean moral hazard estimate can then be calculated from the distribution of means collected from each simulation draw.

[66] Results are shown for the counterfactual that forces individuals into their preferred plan. Results for the full insurance counterfactual are available upon request.

tion of all three types of medical care. The most dramatic percentage increase is in total annual hospital consumption, though the level increase is small (i.e., an additional 0.67 hospital visits a year on average). The increase in prescription drug consumption is small because drugs are both inexpensive and productive, making them the most attractive form of medical care consumption for the uninsured. Insurance acquisition also increases the average price paid for both doctor and hospital visits, despite the fact that an individual cannot observe prices. This finding suggests that an individual who expects higher price draws consumes care when insured that he would not when uninsured, ceteris paribus. I also find that the increase in medical care consumption that results from insurance acquisition has only a small impact on illness. The average sum of acute illness months (i.e., $\frac{1}{N}\sum_{i=1}^{N}\sum_{t=1}^{12} A_{it}$) and average number of chronic illnesses at the end of the year fall by only 1.16% and 0.23%, respectively.

Several alternative measures of moral hazard effects are reported in the literature. By simulating behavior under various insurance conditions, I am able to calculate these measures using the estimated WYDM model and MEPS data. Comparisons are reported in Table 16. Row 1 contains the estimates of Manning et al. (1987) who use the experimental RAND HIE data and "two- and four-part" models to estimate a co-insurance (arc) elasticity of medical care demand of 0.17 for the 0-25% co-insurance range and 0.22 for the 25-95% co-insurance range. To calculate comparable measures, I simulate the within-year decision-making model under three cost-sharing arrangements. All three arrangements feature no deductible, no stop loss, and no premium; differing only by a universal co-insurance rate, which is set to 0%, 25%, or 95%.[67] I estimate a co-insurance elasticity of medical care demand of 0.20 for the 0-25% co-insurance range and 0.35 for the 25-95% co-insurance range. The remaining three comparisons are straightforward. The measures used by Keeler and Rolph (1988) and Einav et al. (2013) compare expenditure under full insurance and no insurance.[68] The measure used by Bajari et al. (2014) compares expenditure, at the individual level, under no insurance and observed/preferred insurance status.[69]

[67] The RAND HIE plans had no premium and those used for price elasticity estimation had no deductible. The plans did feature a $1000 stop loss, but efforts are made by the researchers to avoid the distortions in price elasticity estimates caused by this dynamic incentive for reasons explained in Keeler et al. (1977). Specifically, Manning et al. (1987), "examine demand for episodes of treatment by individuals who are more than $400 from their (stop loss) limit. This strategy gives an approximation of the true price effect if such people treat the true probability of exceeding their limit as nearly zero." Rather than recreating their approximation technique, I eliminate the stop loss in simulation.

[68] In Row 3, I estimate the percentage increase in mean annual expenditure caused by switching from no insurance to full coverage to be 135%, which seems large relative to Einav et al. (2013). However, note that the same simulation results can be used to calculate the percentage of full insurance expenditure not explained by moral hazard (i.e., 663/[663+899] = 42%), which is near the preferred measure of Keeler and Rolph (1988), who study the RAND HIE data. In that sense, my estimate is in line with the rest of the literature.

[69] Kowalski (2015) also estimates the average increase in expenditure when moved from no insurance and

While the above exercise highlights the flexibility of the WYDM model to produce a variety of moral hazard estimates, the measures reported in Table 16 are not truly comparable, as each study uses a different sample population in its empirical analysis. These populations differ by observable characteristics, likely differ by unobservable characteristics, and allow health insurance to be selected from different sets of alternatives; all of which are likely to impact the estimated effect of moral hazard on medical care consumption.[70]

6.4 A Representative Annual Expenditure Model

An important objective of this research is to determine whether a model that features an annual medical care expenditure decision produces a biased estimate of moral hazard effects. I use the following strategy to answer this question. First, I impose a set of assumptions on the WYDM model that are representative of a general annual expenditure model (described below). I refer to this model as the representative annual expenditure (RAE) model. Second, using the same data and imposing the same functional form assumptions as the WYDM model, I estimate the structural parameters of the RAE model. Third, I calculate the moral hazard effects via simulation, in the same manner as the WYDM model (see Section 6.3). Fourth, I compare the various moral hazard effects produced by the two models.[71]

full coverage. She finds an increase of only $16 (in 2003 dollars), meaning a comparable moral hazard estimate to Einav et al. (2013) would be less than 1%. Also, note that Bajari et al. (2014) do not allow for income effects in their estimate, by adjusting the budget constraint in the no-insurance case so that the observed consumption bundle is guaranteed to be affordable. Their estimate would increase without this adjustment.

[70] E.g., the RAND HIE data used by Manning et al. (1987) and Keeler and Rolph (1988) include both children and married individuals, while my data do not. A number of studies have show children to be less sensitive to medical care prices, which could explain why my estimates are larger. Einav et al. (2013) and Bajari et al. (2014) use claims data from a single employer, though not the same employer, which contain a more homogenous population and limited set of health insurance alternatives than the MEPS data. Analysis is also limited to insured individuals, as claims data are not collected for those declining coverage. Individuals in each of these studies also face vastly different healthcare environments. The MEPS subsample used in this study were collected from 1996-1999, the RAND HIE data used by Manning et al. (1987) and Keeler and Rolph (1988) were collected from 1974-1980, and the claims data used by Einav et al. (2013) and Bajari et al. (2014) were collected from 2003-2006 and 2002-2004, respectively.

[71] There are two seemingly viable alternative strategies. First, I could specify and estimate a true annual expenditure model. However, specifying such a model would require functional form assumptions that differ from the within-year decision-making model, which could lead to divergent results independent of the economic constraints/incentives imposed on the models. Second, I could impose the annual expenditure assumptions on the WYDM model and simulate a new moral hazard effect (i.e., not re-estimate the model parameters). However, this procedure ignores the fact that the estimated parameters discussed in Section 6.1 are partially identified by the structure imposed on the model (i.e., the incentives an individual is assumed to internalize and the information he has at his disposal). Therefore, when the structure of the model changes, the previously estimated parameters do not describe the behavior of an individual according to the new model, making the resulting simulations nonsensical. To highlight the consequences of such an error, I used this alternative approach to see how the resulting moral hazard effect would differ. The model predicts that 46.1 percent (s.d. = 3.6) of insured expenditure is due to moral hazard, which is significantly

6.4.1 Model Construction

In order to transform the WYDM model into a representative annual expenditure model, I impose three assumptions. First, I assume that illness transitions are exogenously determined. Each of the annual expenditure models discussed in this paper lack health production, leaving preferences as the only motivation for medical care consumption.[72] Second, I assume that contemporaneous medical care prices are known to an individual. Third, I assume that in each month after the insurance decision is made, an individual knows all of his future price and medical care preference shocks.[73]

After imposing these three assumptions, the RAE model represents an annual expenditure model in the sense that an individual solves his optimization problem under the same incentives and information as a true annual expenditure model. In a true annual expenditure model, an individual makes a health insurance decision without knowing the expenditure preference shock he will receive in the future, which is sometimes interpreted as uncertain future health. In the RAE model, an individual makes a health insurance decision without knowing the medical care consumption and price shocks he will receive in the future. In a true annual expenditure model, an individual makes an annual medical care expenditure decision where (1) the only benefit to medical care is a contemporaneous utility gain and (2) there is no uncertainty about the current or future costs or benefits of his decision. In the RAE model, the same is true for decisions made over the course of the year.

6.4.2 Empirical Findings

Parameter estimates and a discussion of model fit for the RAE model can be found in Section 4 of the web appendix. The moral hazard effect produced by the RAE model is slightly lower than that of the WYDM model. For the RAE model, mean predicted expenditure is \$683 when individuals are uninsured and \$1443 when individuals are covered by their preferred plan; thus, the percentage of mean preferred plan expenditures explained by moral hazard is 51.9 percent (s.d. = 5.9), which is smaller than that produced by the

lower than the estimate produced by the RAE model, discussed below.

[72] An exception is Khwaja (2010), who allows for endogenous health; however, insurance and medical care decisions are made biennially in this model, so endogenous health transitions do not create dynamic medical care consumption incentives within the health insurance year.

[73] Assuming that an individual knows all future price and preference shocks significantly complicates the solution to the dynamic programming problem as the state space in any given month contains these future shocks. To simplify this solution, I assume that each individual receives only one price draw for each type of medical care and one preference draw for each medical care consumption bundle, rather than a new set of draws in each month of the year. These draws then become regressors in the interpolation procedure described in Section 5.1, so that in any month, t, the maximal expected future value of consumption in month, $t+1$ (see Equation 10), is a function of *known* future preference and price shocks.

WYDM model (53.1 percent, s.d. = 4.0) - the difference is statistically different from zero at the 10 percent level, but not the 5 percent level. Column 3 of Table 14 reports the moral hazard distribution. Like the WYDM model, the effect varies widely across the population. For the RAE model, upon winsorizing the moral hazard distribution at the 99th percentile (i.e., capping the spending increase at $12,295), the percentage of mean expenditures explained by moral hazard falls to 42.6 percent (s.d. = 5.2), which is significantly smaller than the estimate produced by the WYDM model (46 percent, s.d. = 3.1). Furthermore, 64 percent (s.d. = 8.1) of the sample does not increase their spending at all in response to coverage, which is significantly larger than the group predicted by the WYDM model (54.4 percent, s.d.=10.3).

In summary, there are two notable differences in the moral hazard distributions produced by the RAE and WYDM models. First, compared to the WYDM model, the RAE model predicts that more individuals do not increase their medical care consumption at all when they acquire insurance. Provided the discussion in Section 2.2, this finding is consistent with a sample population that is risk averse and has a low willingness to pay for medical care. The medical care consumption patterns presented in Table 15 are consistent with this hypothesis. The WYDM model predicts a larger percentage increase in average medical care consumption than the RAE model for all three types of care, with the largest differences for the most expensive types of care (i.e., hospital and doctor visits).

The second notable difference between the two models is that the moral hazard distribution produced by the RAE model has a longer right tail. This finding is consistent with Section 2.2 if a small group of individuals in the sample population face the possibility of very large price draws, p_h, yet also place a lot of value on medical care (i.e., are in the right tail of the willingness to pay distribution). According to this theory, when uninsured and facing price uncertainty these individuals are likely to purchase care despite the risk of a high price draw because of their strong preferences for medical care; however, when uninsured and facing known prices these individuals only purchase care when prices are low. Such individuals should have a higher moral hazard effect when prices are certain. Evidence of this behavior can also be see in Table 15. Notice that the WYDM model predicts an increase in the average price paid for doctor's visits and hospital visits, but the increase is much larger in the RAE. In other words, RAE model individuals never purchase very expensive care when uninsured, while those in the WYDM model with a large W may purchase care when uninsured and receive a high price draw.

While the differences between the models' predictions are statistically significant, it is not clear whether these differences are relevant economically. To determine this, I use each model to predict the total increase in US medical care expenditure that would result if every

uninsured individual in 2013 were forced into health insurance coverage.[74] To do this, I take the average medical care spending among all uninsured individuals in 2013, which is $2,443 (Kaiser Family Foundation), and determine the average predicted increase in spending for each individual under each model - $2,834 (WYDM) and $2,712 (RAE).[75] To calculate the difference in the predicted increase in total US annual medical care expenditure, I multiply each of these by the size of the uninsured population in 2013, which is 45,327,936,[76] and take the difference. The WYDM model predicts an increase in expenditure that is 5.3 billion dollars larger than the RAE model, which is roughly 4 percent of the total increase in expenditure predicted by the WYDM model.

This back-of-the-envelope calculation suggests that the difference in the mean moral hazard estimates produced by the two models may not be economically meaningful; thus, if a researcher's only interest is an estimate of this mean moral hazard effect, one may conclude that an annual expenditure model is adequate. That said, Table 14 highlights the important role that the full distribution of the moral hazard effect plays in determining the sample mean. Given the sensitivity of the mean estimate to small changes in the right tail, one might prefer a more conservative estimate of the mean that limits the important of the right tail. As discussed above, the mean more hazard estimates for the WYDM and RAE models with each distribution winsorized at the 99th percentile are 46.0 and 42.6 percent, respectfully. Applied to the calculation above, these moral hazard estimates cause the WYDM model to predict an increase in total US medical care expenditure (resulting from forced insurance take-up) that is roughly 12.2 billion dollars larger than the increase predicted by the RAE model. This figure represents 13 percent of the total increase in expenditure predicted by the WYDM model, which reflects a more economically meaningful difference in the model's predictions.

6.5 An Empirical Test of Grossman 1972

The seminal Grossman (1972) model assumes that the primary motivation for medical care consumption is the production of a durable health capital stock, which produces "healthy days", from which individuals derive utility; thus, individuals have a derived demand for medical care. While Grossman's model does not address health insurance, Zweifel and Manning (2000) describe an empirical test of the health capital model that

[74] This calculation is particularly relevant given that most major provisions of the Affordable Care Act, which has a principle objective of reducing the number of uninsured in the US, took effect in January of 2014. That said, this calculation should be interpreted with caution, given that the sample used in estimation is not representative of the 2013 US population.

[75] For the WYDM model, I multiply $2,443 by a factor of 1.16 [=(1,435/663)-1]. For the RAE model, I multiply $2,443 by a factor of 1.11 [=(1,443/683)-1].

[76] 16.7 percent (Kaiser Family Foundation) of the 2,711,424,765 (Census) non-elderly individuals living in the US.

involves insurance acquisition:

> In the logic of a Grossman 1972 model, an unexpected reduction in price [of medical care] will induce individuals to want to increase their health stock above the preceding optimal level. To bring desired and current health into alignment requires an increase in medical care demand ... Once the new desired stock equals the current, demand would fall back to a lower level.

Using this logic and the WYDM model, I am able to test whether the Grossman (1972) health capital model is supported by the data. Specifically, I simulate insurance acquisition for each individual and examine the change in illness and spending patterns over the insurance year. As Zweifel and Manning (2000) explain, Grossman's health capital model predicts that this increase in coverage (i.e., decrease in the out-of-pocket price of medical care) should lead to an initial increase in medical care expenditure, which improves the stock of health; as health stock improves, medical care consumption should then fall to a lower level.[77]

My findings are presented in Figures 4 and 5. Both figures display the month of the insurance year on the horizontal axis and the average difference (between fully insured and uninsured individuals) in an outcome variable of interest on the vertical axis. Figure 4 shows that in every month of the year, average medical care consumption is higher for insured individuals than uninsured individuals. Figure 5 shows that this medical care consumption leads the average individual to possess fewer illnesses when insured than when uninsured. Moreover, over the course of the year, the health benefit of this additional medical care consumption accumulates; in each consecutive month, the average number of acute and chronic illnesses possessed by insured individuals decreases relative to uninsured individuals. Returning to Figure 4, one can see that for doctor's office visits, the additional consumption due to insurance possession falls over the course of the year as health improves. This pattern is not observed for hospital or prescription drug care.

To summarize, the WYDM model predicts the exact patterns of illness and doctor's office consumption that would be predicted by the Grossman (1972) health capital model. There are several explanations as to why hospital and prescription drug consumption do not follow suit. With regard to prescription drugs, while I do find that they are effective in reducing illness (see parameters α_{44} and δ_{44}), I also find that consumption is not very sensitive to changes in prices (see Table 15); given this, it is not surprising that consumption is not sensitive to very small changes in illness. With regard to hospital care, the WYDM

[77] Note that these patterns should be observed only for an individual who experiences an increase in insurance generosity. Because I do not observed insurance switching behavior in the data, it is not possible to show these patterns in the raw data.

model only predicts one year of consumption following insurance acquisition; thus, the eventual decline in hospital usage could potentially require a longer time horizon. Moreover, my coarse measure of health capital (i.e., number of acute and chronic illnesses) may miss important heterogeneity in illness severity and accidents, which may be less responsive to early-year increases in medical care consumption. If this is true, we might expect hospital care to be less responsive to changes in illness counts than other types of care.

It is also important to note that while the WYDM model is motivated by Grossman's model, it is sufficiently flexible to reject the health capital hypothesis. This is evident from Figure 4, as hospital and prescription drug consumption are not found to behave according to a health capital model. To understand this, note that as medical care consumption enters the utility function directly, the model does not assume implicitly that medical care has a productive effect on health. In that sense, whether the primary motivation for medical care consumption is a desire to improve health or is preference related is an empirical question.

6.6 Counterfactual Analysis: Insurance Alternative Set Generosity

As medical care expenditure per capita has grown over time in the United States, many employers have sought to limit their exposure to this growth by offering a less generous basket of health insurance alternatives. According to a 2014 report by Towers Watson, nearly three quarters of employers offered high deductible health insurance plans (HDHP) in 2014, and another nine percent expected to add a HDHP alternative the following year.[78] Moreover from 2012 - 2014 the percentage of firms replacing all of their health insurance offerings with HDHPs grew from 7% to 16%, with predicted growth to 30% by 2015. These transitions should have different welfare implications across income and health distributions, which can be examined using the WYDM model.

To be more specific, I analyze consumer behavior under two counterfactual insurance settings. The first counterfactual adds a HDHP to every individual's alternative set. Abstracting from supply-side effects, we should expect this inclusion to benefit uninsured individuals who have been priced out of the market and overinsured individuals who are currently constrained by their employers' limited offerings. The medical care spending effects of this experiment are ex-ante ambiguous, as one would expect the first group to spend more and the second group to spend less. The second counterfactual replaces the most generous current health insurance offering with a HDHP. Compared to the first counterfactual, we should expect a reduction in both spending and consumer welfare, as those previously holding generous coverage are forced into less-generous plans; however, the net

[78] In 2014, a HDHP was defined by the IRS as a plan with a minimum deductible of $1,250 and a maximum out-of-pocket expenditure level of $6,350. With regard to the health insurance exchanges created under the Affordable Care Act, Coe et al. (2014) finds that 76% of the plans offered in 2013 were HDHPs.

spending and welfare effects are again left as an empirical question.

6.6.1 Counterfactual Analysis: Results

I first examine behavior after adding a HDHP to every individual's insurance alternative set. The plan's total premium and cost-sharing characteristics are equal to those of the average bronze plan offered in 2016 in the 38 states with federally facilitated or partnership marketplaces.[79] To determine the out-of-pocket premium that each individual pays for the HDHP plan, I calculate the maximum subsidy that their employer offers among it's existing plans, and assume that their employer would be willing to subsidize the HDHP plan by this amount.

The first two columns of Table 17 describe the switching behavior of the previously uninsured population. In general, uninsured individuals are relatively low earning and receive a health insurance offer that is not very generous. However, by comparing switchers and non-switchers, we can further decompose the uninsured into two types. Non-switchers tend to be quite unhealthy, meaning that, compared to switchers, their rationalization for choosing to be uninsured is driven by their distaste for medical care. This idea is reflected in Column 1 of Table 17, which shows that non-switchers are more likely than switchers to be of unobserved types 2-4. Type 2-4 individuals have a strong distaste for large quantities of medical care (compared to type 1 individuals); thus, a HDHP is especially unattractive. Switchers, on the other hand, are relatively healthy and very likely to be of unobserved type 1. Type 1 individuals have a strong distaste for *any* medical care consumption, but conditional on consuming any, they consume more than other types. For these individuals, a HDHP is particularly attractive, as the plan provides coverage only when lots of medical care is consumed.

The next two columns of Table 17 describe the switching behavior of the previously insured population. Compared to non-switchers, switchers appear to be overinsured - they have better initial health, lower income, are offered a smaller set of health insurance alternatives with a higher minimum out-of-pocket premium, and are very likely to be of unobserved type 1. The HDHP offering is attractive to these individuals because they currently are not offered a low-cost high-risk plan, which suits their preferences.

Note that the decline in spending among the previously insured ($8.87 on average)

[79] These data are collected by Healthcare.gov. Cost-sharing information is summarized by the Kaiser Family Foundation. In 2016, the average bronze plan was characterized by a $5,765 deductible, a 33% hospital coinsurance rate, a 31% doctor's office coinsurance rate, a 33% prescription drug coinsurance rate, and a $6,646 maximum out-of-pocket expenditure level. All dollar amounts are discounted to 1996 levels using the annual medical-cost inflation rate. Premium information is summarized by HealthPocket.com. In 2016 dollars, the average monthly premium for a bronze plan for a 30, 40, 50, and 60 year old was $258, $290, $405, and $615, respectively. Premiums are discounted to 1996 levels using the annual health insurance inflation rate, which is generally higher than the annual rate of medical care inflation.

is mostly offset by the increase in spending among the previously uninsured ($11.59 on average), meaning average annual medical care expenditure in the population remains mostly unchanged - it decreases from $1,264 without the HDHP to $1,262 with the HDHP. That said, the policy is welfare improving for 22.4 percent of the population - on average, these individuals would need to be paid $234 to be just as well off prior to the addition of the plan as they are after the HDHP is offered.

The second counterfactual examines behavior after replacing every individual's most generous health insurance offering with the HDHP described above. I identify the most generous current offering as the plan requiring the lowest out-of-pocket expenditure given $2,000 in total medical care expenditure.[80] Results are summarized in Columns 5 and 6 of Table 17, which compares individuals selecting the most generous health insurance offering in the first counterfactual (i.e., after the HDHP has been added) with those selecting some other plan. Compared to other insured individuals, those who are forced out of generous coverage enter the year with a greater number of illnesses, have less income, and are offered a worse health insurance alternative set by their employer. Moreover, these individuals are most likely to be unobserved Types 2-4, which have a higher likelihood of illness and stronger preferences for any medical care consumption. When forced out of these plans, 28% move to another offered plan, 27% move to the HDHP, and 45% choose to be uninsured. As expected, this move to less generous coverage results in a decrease in mean annual medical expenditure in the population - from $1,264 without the policy to $1,115 with the policy, or 12.9%. However, the policy is welfare reducing for those wishing to purchase generous coverage, which is 39.7 percent of the population. In total, the welfare decline associated with removing the most generous plan is roughly 2.7 times larger than the welfare gain associated with adding the HDHP.

In summary, the above counterfactuals highlight the trade-off that employers face in transitioning to less generous health insurance alternative sets. The policy change has the potential to reduce total annual medical care expenditure and improve welfare for some, while reducing welfare for the sickest individuals in the population. In this particular setting, the (consumer) welfare losses significantly outweigh the gains. That said, it is important to interpreted these results with the limitations of the model in mind. The model does not internalize the effects that changes in insurance and medical care demand have on prices; as such, in all counterfactuals I assume that insurance premiums are fixed at their observed levels. This assumption eliminates the possibility that the inclusion of the HDHP plan has a significant impact on the risk pool, and therefore premiums, of more

[80] I assume that half of the expenditure is on hospital care and half is on doctors office care. Results are robust to expenditure levels of $1,500 and $5,000, as well as identifying the most generous plan as the plan with the highest total premium.

generous plans, as is the topic of Cutler and Reber (1998).[81] The model also predicts behavior for only a calendar year; thus, long-run health implications of the less-generous insurance alternative set are unknown.

7 Conclusion

In response to the 2010 passage of the Affordable Care Act, the economics literature has reflected significant interest in optimal health insurance plan (and alternative set) design. Building on early empirical studies of adverse selection and moral hazard that primarily relied on experimental and quasi-experimental techniques (e.g., Manning et al. (1987)), the recent literature has utilized structural models of insurance choices and medical care expenditure, which are more suitable for counterfactual experimentation and welfare analysis (e.g., Einav et al. (2013), Handel (2013), Bajari et al. (2014), Handel and Kolstad (2015), Kowalski (2015)). Thus far, these models have focused almost exclusively on annual medical care expenditure decisions, abstracting from several important features of an individual's optimization problem, including endogenous health evolution and uncertainty over the effective price of medical care at the time of consumption.

This paper contributes to the literature an estimate of insurance-induced moral hazard derived from a dynamic model of within-year medical care consumption decisions under uncertainty. An individual's optimization problem is defined by an annual health insurance decision, followed by a sequence of monthly medical care consumption decisions made over the course of a health insurance year. The model allows medical care consumption to alter future health outcomes, assumes medical care prices are unknown prior to consumption, and allows for uncertainty about future medical care price, consumption, and health outcomes. I find significant differences in the distribution of moral hazard effects produced by the preferred within-year decision-making model and a representative annual expenditure model. Most importantly, the distribution produced by the within-year decision-making model has thinner tails and a slightly larger mean. The key feature of the within-year decision-making model that leads to a larger moral hazard effect is the presence of uncertainty at the time of a medical care decision. In a typical annual expenditure model, there is no uncertainty about the costs or benefits of medical care at the time that a medical care decision is made, which lessens the value of insurance, leading an individual to respond less strongly to insurance acquisition.

[81] These general equilibrium effects would likely increase the estimated welfare loss. In the first counterfactual, insurance premiums of existing plans are likely to rise as relatively healthy individuals move into the HDHP, generating a welfare loss for individuals remaining in the old plans. In the second counterfactual, insurance premiums of all plans are likely to rise as the risk pool receives an influx of unhealthy individuals who previously purchased generous coverage.

The paper also offers empirical support for the Grossman (1972) health capital model. Medical care is found to be productive in reducing the probability of future illness; thus, insurance acquisition is followed by an initial surge in medical care consumption, which gradually falls as health improves. This analysis was made possible by unique data that include self-reported illnesses, which is needed to model endogenous health/illness transitions and medical care productivity. That said, this analysis could be improved with better data. Ideally, illnesses would be measured with higher frequency to limit the potential for measurement error due to imperfect recall. A model more representative of Grossman's may also include an overall measure of health stock, which could be proxied for by a high frequency self-reported health status. The estimation of medical care productivity could also benefit from better instruments that alter medical care decision making, but not health/illness outcomes.

Within a broader literature on structural estimation, this research highlights a potential consequence of utilizing data on aggregate decision variables (e.g., annual expenditure) that necessarily prohibit one from designing a behavioral model that accurately describes the incentives faced by and information available to an individual who, in reality, makes decisions with higher frequency (e.g., daily medical care decisions).[82] In this research, the studied consequence is a biased predicted response to an environmental change in simulation (i.e., biased moral hazard effects), though other consequences could exist as well.

References

Akerlof, G. A. (1970): "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism," *The Quarterly Journal of Economics*, 84, 488–500.

Aron-Dine, A., L. Einav, and A. Finkelstein (2013): "The RAND Health Insurance Experiment, Three Decades Later," *Journal of Economic Perspectives*, 27, 197–222.

Aron-Dine, A., L. Einav, A. Finkelstein, and M. Cullen (2015): "Moral Hazard in Health Insurance: Do Dynamic Incentives Matter?" *The Review of Economics and Statistics*, 97, 725–741.

Arrow, K. J. (1963): "Uncertainty and the Welfare Economics of Medical Care," *The American Economic Review*, 53, 941–973.

[82] An example of a frequently modeled aggregate decision variable from the literature is an annual savings decision (e.g., Van der Klaauw and Wolpin (2008), French and Jones (2011)). In reality, an individual makes multiple savings decisions over the course of the year and must respond to changing interest rates, forecast future changes in interest rates, and respond to expenditure shocks (e.g., unexpected medical expenditure) that impact the savings decision.

Baicker, K., S. Taubman, H. Allen, M. Bernstein, J. Gruber, J. P. Newhouse, E. Schneider, B. Wright, A. Zaslavsky, A. Finkelstein, and the Oregon Health Study Group (2013): "The Oregon Experiment - Medicaid's Effects on Clinical Outcomes," *New England Journal of Medicine*, 368.

Bajari, P., H. Hong, A. Khwaja, and C. Marsh (2014): "Moral Hazard, Adverse Selection, and Health Expenditures: A Semiparametric Analysis," *RAND Journal of Economics*, 45, 747–763.

Bellman, R. (1957): *Dynamic Programming*, Princeton University Press.

Berndt, E., B. Hall, R. Hall, and J. Hausman (1974): "Estimation and Inherence in Nonlinear Structural Models," *Annals of Economic and Social Measurement*, 3, 103–116.

Blau, D. and D. Gilleskie (2008): "The Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review*, 49, 475–514.

Brill, Steven (2013): "Bitter Pill: Why Medical Bills are Killing Us," http://time.com/198/bitter-pill-why-medical-bills-are-killing-us/, [Online; accessed 2016-05-11].

Cardon, J. H. and I. Hendel (2001): "Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey," *The RAND Journal of Economics*, 32, 408–427.

Centers for Disease Control and Prevention (2011): "Health, United States, 2011: with Special Feature on Socioeconomic Status and Health," http://www.cdc.gov/nchs/data/hus/hus11.pdf#129, [Online; accessed 2016-05-11].

Chiappori, P. and B. Salanie (2002): "Testing Contract Theory: A Survey of some Recent Work," in *Advances in Economics and Econometrics – Theory and Applications, Eighth World Congress*, ed. by M. Dewatripoint, L. Hansen, and P. Turnovsky, Cambridge: Cambridge University Press, 115–149.

Coe, E. H., J. Oatman, M. Rayasam, and T. B. Wright (2014): "Exchange Product Benefit Design: Consumer Responsibility and Value Consciousness," *McKinsey on Healthcare*, 1–13.

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

Croson, R. and U. Gneezy (2009): "Gender Differences in Preferences," *Journal of Economic Literature*, 47, 448–474.

Currie, J. and J. Gruber (1996): "Health Insurance Eligibiity, Utilization of Medical Care, and Child Health," *The Quarterly Journal of Economics*, 111, 432–466.

Cutler, D. and S. Reber (1998): "Paying for Health Insurance: The Trade-Off between Competition and Adverse Selection," *The Quarterly Journal of Economics*, 113, 433–466.

Cutler, D. and R. Zeckhauser (2000): "The Anatomy of Health Insurance," *Handbook of Health Economics*, 1, 563–643.

Dafny, L. and J. Gruber (2005): "Public insurance and child hospitalizations: access and efficiency effects," *Journal of Public Economics*, 89, 109–129.

Dave, D. and R. Kaestner (2009): "Health Insurance and Ex-Ante Moral Hazard," *International Journal of Heathcare Finance and Economics*, 9, 367–390.

Eckel, C. C. (2008): "Men, Women and Risk Aversion: Experimental Evidence," *Handbook of Experimental Economics Results*, 1, 1062–1073.

Einav, L., A. Finkelstein, and J. Levin (2010): "Beyond Testin: Empirical Models of Insurance Markets," *The Annual Review of Economics*, 2, 311–336.

Einav, L., A. Finkelstein, S. Ryan, P. Schrimpf, and M. Cullen (2013): "Selection on Moral Hazard in Health Insurance," *American Economic Review*, 103, 178 – 219.

Ellis, R. P. (1986): "Rational Behavior in the Presence of Coverage Ceilings and Deductibles," *The RAND Journal of Economics*, 17, 158–175.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker, , and the Oregon Health Study Group (2012): "The Oregon Health Insurance Experiment: Evidence from the First Year," *Quarterly Journal of Economics*, 127.

French, E. and J. B. Jones (2011): "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," *Econometrica*, 79, 693–732.

Gilleskie, D. B. (1998): "A Dynamic Stochastic Model of Medical Care Use and Work Absence," *Econometrica*, 66, 1–45.

Grossman, M. (1972): "On the Concept of Health Capital and the Demand for Health," *Journal of Political Economy*, 80, 223–255.

Guilkey, D. K. and P. M. Lance (2013): "Program Impact Estimation with Binary Outcome Variables: Monte Carlo Results for Alternative Estimators and Empirical Examples," *Studies in Applied and Empirical Econometrics*, A Festschrift in Honor of Peter Schmidt (New York: Springer).

Handel, B. (2013): "Adverse Selection and Inertia in Health Insurance Markets: When Nudging Hurts," *The American Economic Review*, 103, 2643–2682.

Handel, B. R. and J. T. Kolstad (2015): "Health insurance for ?humans?: Information frictions, plan choice, and consumer welfare," *The American Economic Review*, 105, 2449–2500.

Heckman, J. and B. Singer (1984): "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 52, 271–320.

Imai, S. and M. P. Keane (2004): "Intertemporal Labor Supply and Human Capital Accumulation," *International Economic Review*, 45, 601 – 641.

Jones, A. M., J. Lomas, and N. Rice (2014): "Applying Beta-Type Size Distributions to Healthcare Cost Regressions," *Journal of Applied Econometrics*, 649–670.

Kaiser Family Foundation (2014): "Health Insurance Coverage of Nonelderly 0-64," http://kff.org/other/state-indicator/nonelderly-0-64/, [Online; accessed 2016-05-11].

Keane, M. P. and K. I. Wolpin (1994): "The Solution and Estimation of Discrete Choice Dynamic Programming Models by Simulation and Interpolation: Monte Carlo Evidence," *The Review of Economics and Statistics*, 76, 648–672.

Keeler, E., J. Buchanan, J. Rolph, J. Hanley, and D. Reboussin (1988): *The Demand for Episodes of Medical Treatment in the Health Insurance Experiment*, R-3454-HHS, Santa Monica, CA: RAND Publishing.

Keeler, E. B., J. P. Newhouse, and C. E. Phelps (1977): "Deductibles and the Demand for Medical Care Services: The Theory of a Consumer Facing a Variable Price Schedule under Uncertainty," *Econometrica*, 45, 641–656.

Keeler, E. B. and J. E. Rolph (1988): "The Demand for Episodes of Treatment in the Health Insurance Experiment," *Journal of Health Economics*, 7, 337 – 367.

Kelly, I. and S. Markowitz (2009): "Incentives in Obesity and Health Insurance," *Inquiry*, 46, 418–432.

Khwaja, A. (2001): "Health Insurance, Habits, and Health Outcomes: Moral Hazard in a Dynamic Stochastic Model of Investment in Health," *Working Paper*.

——— (2010): "Estimating Willingness to Pay for Medicare using a Dynamic Life-Cycle Model of Demand for Health Insurance," *Journal of Econometrics*, 156, 130 – 147.

Kolstad, J. T. and A. E. Kowalski (2012): "The Impact of Health Care Reform on Hospital and Preventive Care: Evidence from Massachusetts," *Journal of Public Economics*, 96, 909–929.

Kowalski, A. (2015): "Estimating the Tradeoff Between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance," *International Journal of Industrial Organization*, 43, 122–135.

Lin, H. and D. Sachs (2016): "Dynamics of Demand in the RAND Health Insurance Experiment," *Working Paper*.

Manning, W. G., H. L. Bailit, B. Benjamin, and J. P. Newhouse (1985): "The demand for dental care: evidence from a randomized trial in health insurance," *The Journal of the American Dental Association*, 110, 895–902.

Manning, W. G., J. P. Newhouse, N. Duan, E. B. Keeler, and A. Leibowitz (1987): "Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment," *The American Economic Review*, 77, 251–277.

Meier, B., J. C. McGinty, and J. Creswell (2013): "Hospital Billing Varies Wildly, Government Data Shows," http://www.nytimes.com/2013/05/08/business/hospital-billing-varies-wildly-us-data-shows.html?pagewanted=all&_r=1, [Online; accessed 2016-05-11].

Miller, S. (2012): "The Impact of the Massachusetts Health Care Reform on Health Care Use Among Children," *American Economic Review: Papers and Proceedings*, 102, 502–507.

Mroz, T. A. (1999): "Discrete Factor Approximations in Simultaneous Equation Models: Estimating the Impact of a Dummy Endogenous Variable on a Continuous Outcome," *Journal of Econometrics*, 92, 233 – 274.

Mroz, T. A. and D. Guilkey (1992): "Discrete Factor Approximations for Use in Simultaneous Equation Models with Both Continuous and Discrete Endogenous Variables," *University of North Carolina Working Paper*.

Newhouse, J. P. (1974): "A Design for a Health Insurance Experiment," *Economic Inquiry*, 11.

——— (1993): *Free For All?: Lessons from the Rand Health Insurance Experiment*, Harvard University Press.

Pauly, M. V. (1968): "The Economics of Moral Hazard: Comment," *The American Economic Review*, 58, 531–537.

Rae, M., L. Levitt, G. Claxton, C. Cox, M. Long, and A. Damico (2016): "Patient Cost-Sharingin Marketplace Plans," http://kff.org/health-costs/issue-brief/patient-cost-sharing-in-marketplace-plans-2016/, [Online; accessed 2016-05-11].

Robert Wood Johnson Foundation (2012): "Counting Change - Measuring Health Care Prices, Costs, and Spending," .

Rosen, A. B., J. S. Tsai, and S. M. Downs (2003): "Variations in Risk Attitude across Race, Gender, and Education," *Medical Decision Making*, 23, 511–517.

Rosenthal, J. A., X. Lu, and P. Cram (2013): "Availability of consumer prices from US hospitals for a common surgical procedure," *JAMA Internal Medicine*, 173, 427–432.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica*, 55, pp. 999–1033.

Towers Watson (2014): "The New Health Care Imperative Driving Performance, Connecting to Value," https://www.towerswatson.com/en-US/Insights/IC-Types/Survey-Research-Results/2014/05/full-report-towers-watson-nbgh-2013-2014-employer-survey-on-purchasing-value-in-health-c [Online; accessed 2016-05-10].

United States Government Accountability Office (2011): "Health Care Price Transparency - Meaningful Price Information is Difficult for Consumers to Obtain Prior to Receiving Care," .

Van der Klaauw, W. and K. I. Wolpin (2008): "Social security and the retirement and savings behavior of low-income households," *Journal of Econometrics*, 145, 21–42.

Zweifel, P. and W. G. Manning (2000): "Moral hazard and consumer incentives in health care," *Handbook of Health Economics*, 1, 409–459.

A Tables and Figures

Figure 1: Risk Aversion and Moral Hazard

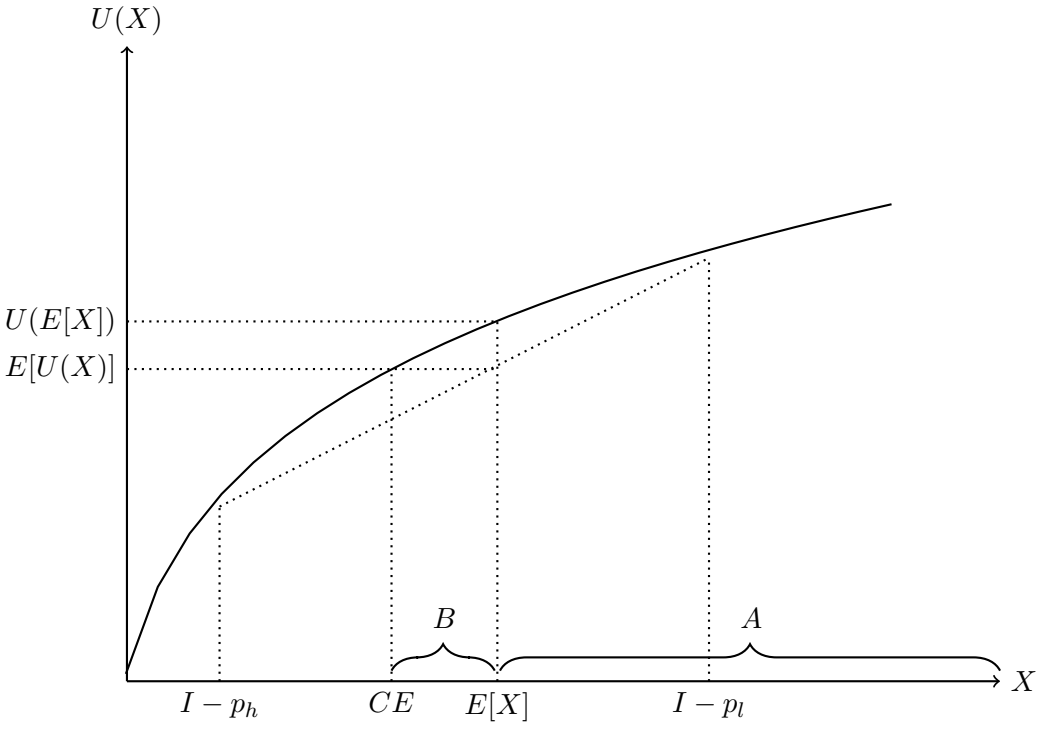


Figure 2: Timing of the Model

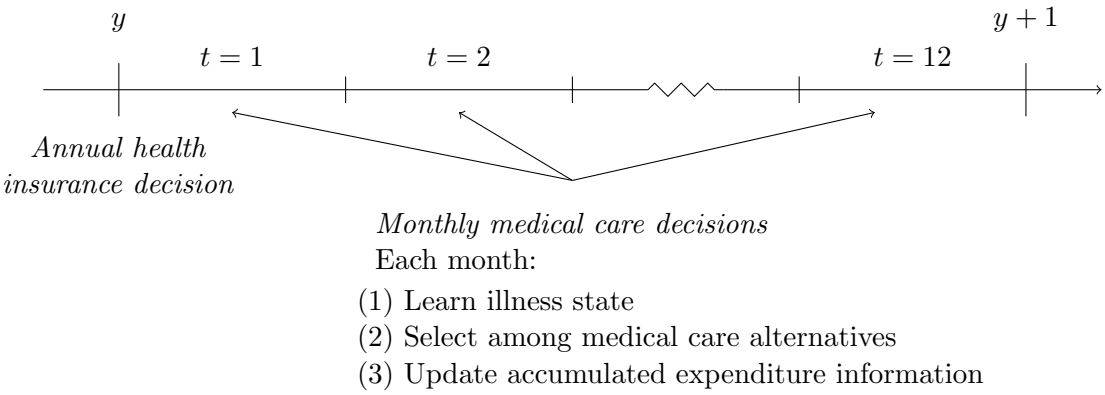


Table 1: Sample Inclusion Criteria

	1996	1997	1998	1999	Total
1996-1999 MEPS Household Component	22601	13683	11137	14178	61599
and single, childless, 19-64 yrs old	4406	2534	2169	2589	11698
and employed in first interview period with offer	1821	923	987	1128	4859
and matches to link file†	749	516	156	688	2109
and no missing interviews	693	472	139	636	1940
and stable insurance status	455	290	98	389	1232

† AHRQ only attempted to collect insurance information for a 25% random sample of individuals in 1998, which explains why so many individuals are lost because of no linkage in that year.

Table 2: Representativeness of the Sample

	Whole Sample		Estimation Sample	
	Mean	S.D.	Mean	S.D.
age	36.88	12.21	39.56	11.70
education (highest grade completed)	13.40	2.53	13.62	2.46
income (in 1996 dollars)	30985.62	21432.55	33713.84	21953.46
male	0.53	0.50	0.48	0.50
lives in a MSA	0.84	0.36	0.83	0.38
Hispanic	0.14	0.35	0.12	0.33
black	0.16	0.37	0.14	0.35
federal employee	0.03	0.18	0.09	0.28
northeast	0.19	0.40	0.19	0.40
midwest	0.24	0.43	0.25	0.43
south	0.35	0.48	0.34	0.47
west	0.22	0.41	0.22	0.41
excellent health status	0.34		0.32	
good health status	0.35		0.35	
poor health status	0.31		0.33	
total annual expenditure †	1490.37	4033.53	1636.53	3063.95
insured all year	0.81	0.39	0.91	0.29
Sample Size	4859		1232	

† This expenditure level was take directly from the MEPS data and includes types of medical care spending not included in this analysis (e.g., dental and eye care, home healthcare, medical equipment, etc.).

Table 3: Sample Statistics by Insurance Status

	Insured		Uninsured	
	Mean	S.D.	Mean	S.D.
Demographics (time invariant)				
age	40.04	11.62	34.73	11.44
education (highest grade completed)	13.73	2.43	12.46	2.49
income (in 1996 dollars)	35198.52	22120.24	19011.55	13144.27
male	0.48		0.52	
lives in a MSA	0.82		0.83	
Hispanic	0.11		0.21	
black	0.14		0.19	
federal employee	0.09		0.09	
Initial Self-Reported Health (time invariant)				
excellent	0.35		0.36	
good	0.34		0.32	
fair	0.23		0.22	
poor	0.07		0.09	
very poor	0.01		0.01	
Illness (time varying)				
entered year with chronic illness	0.40		0.24	
chronic illness by years end	0.50		0.32	
average number of chronic illnesses†	1.84	1.24	1.79	1.18
at least one acute illness during sample year	0.82		0.74	
total months with acute illness	5.60	4.52	3.04	3.21
average number of acute illnesses†	1.56	0.90	1.31	0.67
Medical Care Prices (time varying)				
transaction price for a doctor visit	89.57	159.09	112.43	160.52
transaction price for a hospital day	823.72	1357.51	878.72	2127.58
transaction price for a Rx month‡	75.47	115.57	51.00	54.60
Medical Care Consumption (time varying)				
at least one doctor visit in sample year	0.74		0.50	
total doctor visits in sample year	5.13	8.95	2.51	4.89
at least one hospital day in sample year	0.25		0.17	
total hospital days in sample year	0.63	1.73	0.81	3.76
at least one Rx month in sample year	0.66		0.45	
total Rx months in sample year	4.69	5.00	2.19	3.87
consumed any preventative care	0.17		0.14	
probability of consumption in ill month	0.59		0.41	
probability of consumption in well month	0.12		0.03	
total annual expenditure on doctor visits	455.72	951.61	289.67	759.21
total annual expenditure on hospital days*	660.46	2706.30	1108.39	6320.84
total annual expenditure on Rx consumption	353.28	819.20	111.48	275.49
Other				
number of offered plans	4.41	5.99	3.05	4.46
Sample				
individuals	1119		113	
person-month observations	13428		1356	

† Conditional on having any acute/chronic illness.

‡ In accordance with the model, the prescription drug transaction price is the "price" of one month of drug consumption. In practice, it is calculated at an individuals total expenditure on prescription drugs in a month.

* One uninsured individual had hospital expenditures totaling $52,032.16. Removing this outlier lowers the mean to $647.93 for the uninsured.

Table 4: Insurance Plan Summary

	Held Plans			Rejected Plans		
	Plans	Mean	S.D.	Plans	Mean	S.D.
Premium						
total premium	1119	2057.19	819.09	4165	2207.13	715.25
out-of-pocket premium	1119	343.83	540.31	4165	519.46	610.28
Deductible†						
defined by total expenditure	397	283.59	272.99	772	293.25	348.26
defined by doctor expenditure only	59	191.38	97.04	655	215.72	50.65
defined by hospital expenditure only	28	252.63	218.10	62	150.58	41.24
plan has no deductible	648			2684		
Stop loss						
stop loss	729	1512.82	1077.97	2775	1689.06	1077.63
plan has no stop loss	390			1390		
Hospital‡						
co-insurance rate	417	17.02	9.91	797	15.44	8.01
co-pay level (per stay)	199	258.46	346.70	768	159.41	194.45
co-pay level (per day)	85	52.61	60.25	142	58.70	85.24
free care past the deductible	435			2577		
Doctor						
co-insurance rate	208	18.91	8.69	755	12.43	7.83
co-pay level	858	10.19	4.53	3110	8.41	4.25
free care past the deductible	53			300		
Network Type						
HMO	1119	0.42		4165	0.51	
FFS	1119	0.47		4165	0.44	
PPO	1119	0.11		4165	0.06	

† These categories are not mutually exclusive. Some of the plans that feature a doctor specific deductible also feature a hospital specific deductible.

‡ These categories are not mutually exclusive. Some of the plans are structured so that an individual pays a daily co-pay, plus a percentage (i.e., co-insurance rate) of the remainder of the bill.

Table 5: Preference Parameter Estimates

	Parameter	Estimate	SE
Utility Function			
RA constant	ω_{00}	0.0594	0.0066
RA age	ω_{01}	0.0010	0.0004
RA non-white (black or Hispanic)	ω_{02}	0.0040	0.0114
RA male	ω_{03}	0.0213	0.0139
acute illnesses	ω_{10}	-28.5934	4.4581
acute illnesses (squared)	ω_{11}	2.7975	0.7919
acute illnesses*age	ω_{12}	0.1894	0.0370
chronic illnesses	ω_{20}	-3.5427	5.208
chronic illnesses (squared)	ω_{21}	-0.7278	0.6208
chronic illnesses*age	ω_{22}	-0.1531	0.1761
doctor visits	ω_{30}	-1.5748	0.1832
doctor visits (squared)	ω_{31}	-0.1772	0.0276
doctor visits*age	ω_{32}	-0.0008	0.0012
doctor visits*male	ω_{33}	0.0408	0.0203
hospital days	ω_{40}	-5.4660	0.4032
hospital days (squared)	ω_{41}	0.8503	0.0986
hospital days*age	ω_{42}	0.0082	0.0034
hospital days*male	ω_{43}	0.1998	0.0586
any Rx consumption	ω_{50}	-5.8783	0.5979
any Rx consumption*age	ω_{51}	0.0258	0.0068
any Rx consumption*male	ω_{52}	-0.3121	0.1361
any consumption	ω_{60}	-2.1247	0.1151
any consumption*male	ω_{61}	-0.4811	0.0753
negative consumption†	ω_{70}	0.1000	
Other			
discount factor†	β	0.9960	
log-likelihood value‡	$L(\Omega)$	-56,885.2	

† These parameters are not estimated. Note that β is set to 0.996, instead of the traditional 0.95, because decisions are made monthly in the model. $0.996^{12} \approx 0.95$. The parameter ω_{70} measures the disutility of each dollar of negative consumption (i.e., outspending monthly income, requiring an individual to use savings or to borrow) and is not estimated due to weak identification. Other model parameters do not seem to be sensitive to different fixed values for this parameter.

‡ The log-likelihood value with only one unobserved mass point is -60,525.3.

Table 6: Permanent Unobserved Heterogeneity Parameter Estimates

	Param.	Type 1	Type 2		Type 3		Type 4	
			Est.	S.E.	Est.	S.E.	Est.	S.E.
Mass Point Location								
acute illness prob.	μ_1	0.0	0.569	0.059	0.556	0.066	0.573	0.064
chronic illness prob.	μ_2	0.0	0.985	0.099	0.978	0.101	1.109	0.106
doctor visit pref. (linear)	μ_{3a}	0.0	-1.599	0.144	-1.636	0.150	-1.336	0.169
doctor visit pref. (squared)	μ_{3b}	0.0	0.423	0.025	0.397	0.028	0.412	0.025
hospital day pref. (linear)	μ_{4a}	0.0	-1.536	0.274	-1.543	0.297	-1.352	0.323
hospital day pref. (squared)	μ_{4b}	0.0	0.243	0.082	0.230	0.097	0.282	0.086
any Rx pref.	μ_5	0.0	0.788	0.549	0.759	0.563	-3.788	0.652
any consumption pref.	μ_6	0.0	1.213	0.227	0.832	0.275	1.631	0.146
doctor visit price dist.	μ_7	0.0	-3.161	2.402	-2.816	2.729	-13.005	2.330
hospital day price dist.	μ_8	0.0	-0.120	0.198	-0.110	0.213	-0.445	0.180
monthly Rx price dist.	μ_9	0.0	0.770	0.032	-0.747	0.033	-0.090	0.040
initial acute illness prob.	μ_{10}	0.0	1.360	0.195	1.155	0.234	0.899	0.232
initial chronic illness prob.	μ_{11}	0.0	2.300	0.222	2.102	0.254	1.460	0.251
initial health status prob.	μ_{12}	0.0	0.387	0.173	0.385	0.208	0.192	0.212
Type Probabilities†								
parameter estimate	θ	0.0	-0.961	0.094	-1.363	0.106	-1.129	0.122
type probability		51.0	19.5		13.0		16.5	

† Probabilities are derived from estimating parameter θ^k such that $Pr(\mu^k) = \frac{exp(\theta^k)}{\sum_{k'=1}^{4} exp(\theta^{k'})}$.

Table 7: Acute Illness Probability Parameter Estimates

	Parameter	Estimate	SE
constant	α_{00}	-2.1088	0.0940
male	α_{10}	-0.3296	0.0374
non-white (black or Hispanic)	α_{11}	-0.0824	0.0373
education (highest grade completed)	α_{12}	0.0016	0.0053
age	α_{13}	-0.0030	0.0018
lives in a MSA	α_{14}	-0.0016	0.0465
income (in 1996 dollars)	α_{15}	0.0045	0.0011
initial health	α_{16}	0.0081	0.0156
acute illness	α_{20}	3.2414	0.0560
acute illness (squared)	α_{21}	-0.1965	0.0160
consecutive mons. with acute illness†	α_{22}	0.0179	0.0131
consecutive mons.* initial month	α_{23}	0.0115	0.0122
chronic illness	α_{30}	0.1515	0.0374
chronic illness (squared)	α_{31}	-0.0201	0.0088
doctor visits	α_{40}	-0.0481	0.0071
doctor visits (squared)	α_{41}	0.0053	0.0009
hospital days	α_{42}	-0.0946	0.0181
hospital days (squared)	α_{43}	0.0193	0.0040
Rx consumption	α_{44}	-0.1475	0.0212
acute illness*doctor visits	α_{50}	-0.0058	0.0010
acute illness*hospital days	α_{51}	-0.0016	0.0021
acute illness*Rx consumption	α_{52}	-0.0041	0.0040
chronic illness*doctor visits	α_{53}	0.0013	0.0006
chronic illness*hospital days	α_{54}	-0.0017	0.0014
chronic illness*Rx consumption	α_{55}	0.0063	0.0036
threshold 2	κ_2^a	3.2557	0.0415
threshold 3	κ_3^a	6.0154	0.0727
threshold 4	κ_4^a	8.2134	0.1261

Month indicators are included in regression but are not reported here.

† I control for the number of consecutive months with *any* acute illness leading up to the current month to capture the impact of a more severe acute illnesses on illness transition probabilities. Due to the nature of the data, this count begins in the first month of the insurance year, meaning an individual entering the third month of the year with two consecutive acute illness months may actually have had an acute illness even longer. To control for this measurement error, I interact consecutive months with an acute illness with an indicator that equals one if an individual has had an acute illness in every month since the first month of the year.

Table 8: Additional Chronic Illness Probability Parameter Estimates

	Parameter	Estimate	SE
constant	δ_{00}	-5.0130	0.1723
male	δ_{10}	-0.0182	0.0281
non-white (black or Hispanic)	δ_{11}	-0.0173	0.0130
education (highest grade completed)	δ_{12}	0.0188	0.0044
age	δ_{13}	0.0047	0.0044
lives in a MSA	δ_{14}	-0.0326	0.0177
income (in 1996 dollars)	δ_{15}	0.0002	0.0004
initial health	δ_{16}	0.0546	0.0114
acute illness	δ_{20}	-0.2161	0.0445
acute illness (squared)	δ_{21}	0.0561	0.0112
consecutive mons. with acute illness	δ_{22}	-0.0143	0.0067
consecutive mons.*initial period	δ_{23}	-0.0002	0.0044
chronic illness	δ_{30}	0.3153	0.0444
chronic illness (squared)	δ_{31}	-0.0137	0.0062
doctor visits	δ_{40}	-0.0346	0.0076
doctor visits (squared)	δ_{41}	0.0028	0.0007
hospital days	δ_{42}	-0.0392	0.0129
hospital days (squared)	δ_{43}	0.0072	0.0031
Rx consumption	δ_{44}	-0.1869	0.0378
threshold 2	κ_2^c	2.8027	0.2648

Month indicators are included in regression but are not reported here.

Table 9: Structural Price Parameter Estimates

	Doctor Price		Hospital Price		Prescription Price	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
constant	46.0687	6.1075	4.8073	0.4075	2.7601	0.0669
male	0.6465	1.3201	0.2742	0.1164	0.1557	0.0148
non-white (black or Hispanic)	0.9701	1.7763	0.1114	0.1387	-0.0900	0.0161
education (highest grade completed)	-0.0634	0.2940	0.0368	0.0230	0.0001	0.0032
age	0.1573	0.0687	0.0116	0.0054	0.0033	0.0007
lives in a MSA	12.3417	1.5972	-0.2760	0.1441	0.1272	0.0200
income (in 1996 dollars)	0.0934	0.0379	0.0096	0.0040	0.0027	0.0003
initial health	-0.1577	0.7375	0.0174	0.0582	0.0899	0.0077
HMO	-3.2647	1.4617	0.2219	0.1433	-0.1424	0.0140
PPO	-3.4918	2.4293	-0.2162	0.1868	0.0852	0.0296
no insurance	-15.3438	2.2738	-2.4222	0.2072	-0.4426	0.0307
acute illnesses	-4.6980	1.4736	0.0810	0.1398	0.1036	0.0180
acute illnesses (squared)	1.2520	0.3657	0.0057	0.0347	-0.0071	0.0049
chronic illnesses	-1.8965	1.1936	-0.0358	0.1096	0.3636	0.0142
chronic illnesses (squared)	0.4138	0.2183	0.0006	0.0151	-0.0177	0.0024
shape 1†	2.1544	0.0391	1.1566	0.0385	0.8033	0.0041
shape 2†	0.9031	0.0359	*	*	*	*

Month indicators are included in all regressions but are not reported here.
† The hospital and prescription drug price data are fit to log-normal distributions. The estimated *shape* parameters of these distributions correspond to the standard error of the distribution. The doctor price data is fit to a Singh-Maddala distribution, which has two shape parameters.

Table 10: Initial Condition Probability Parameter Estimates

	Health Status		Acute Illness		Chronic Illness	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
Initial Condition						
constant	2.1265	0.4270	-1.5873	0.6021	-5.2351	0.7405
male	-0.1440	0.1339	-0.5475	0.1530	-0.3923	0.1797
non-white (black or Hispanic)	0.2043	0.1356	-0.4217	0.1754	-0.3111	0.2061
education (highest grade completed)	-0.1005	0.0269	0.0390	0.0315	0.1512	0.0410
age	0.0204	0.0056	0.0051	0.0147	0.0568	0.0169
lives in a MSA	-0.1345	0.1629	0.0895	0.2026	-0.3269	0.2185
income (in 1996 dollars)	-0.0056	0.0041	-0.0005	0.0049	-0.0042	0.0050
March/April (indicator)	*	*	-0.0777	0.2817	*	*
May/June (indicator)	*	*	0.3950	0.2912	*	*
July/August (indicator)	*	*	0.3756	0.2082	*	*
September/October (indicator)	*	*	0.1512	0.2692	*	*
November/December (indicator)	*	*	0.0232	0.3829	*	*
initial health	*	*	0.2107	0.1605	0.5382	0.1918
initial health*age	*	*	0.0010	0.0064	0.0032	0.0072
last year income	-0.0748	0.0404	-0.0557	0.0494	-0.0275	0.0567
last year income missing	-0.2294	0.2686	-0.4274	0.3344	-0.5290	0.3826
veteran (indicator)	0.1353	0.2168	0.0150	0.2621	-0.1838	0.3026
foreign born (indicator)	-0.2105	0.2212	0.2807	0.2486	-0.1373	0.2836
threshold 1	1.5161	0.0736	1.6354	0.1066	1.7092	0.1230
threshold 2	3.3160	0.1437	3.0146	0.1997	3.0549	0.1970
threshold 3	6.4692	0.6487	4.1915	0.3699	4.3233	0.3311

Table 11: Closing Function Structural Parameter Estimates

	Estimate	SE
Closing Function		
doctor visits	-0.0741	0.0970
doctor visits (squared)	0.0102	0.0153
hospital days	-0.0723	0.3118
hospital days (squared)	-0.0594	0.0922
Rx consumption	-0.3432	0.1669
acute illness	-105.9790	15.0042
acute illness (squared)	3.3967	1.7980
acute illness*age	0.6927	0.1733
chronic illness	-884.2178	348.2993
chronic illness (squared)	-26.8637	14.8387
chronic illness*age	-0.8261	2.7786

Table 12: Observed and Simulated Outcomes

	Observed			Simulated		
	Mean	Median	S.D.	Mean	Median	S.D.
Illness						
% entered year with chronic illness	38.23			37.42		
% with chronic illness by year's end	48.54			47.15		
ave. number of chronic illnesses†	1.78	1.00	1.05	1.71	1.00	1.05
% with ≥ 1 acute illness in year	81.41			85.29		
ave. total months with acute illness	5.37	4.00	4.69	5.27	4.00	4.19
ave. number of acute illnesses†	1.52	1.00	0.81	1.45	1.00	0.74
Medical Care Prices‡						
doctor visit price	90.61	55.00	159.18	88.17	55.60	191.66
hospital day price	828.48	329.60	1435.91	807.03	355.30	1637.26
prescription drug price	74.37	41.93	113.63	75.25	40.36	109.03
Medical Care Consumption						
% consuming (any month)	43.53			42.62		
% consuming when $A_{it} \neq 0$	68.06			64.82		
% consuming when $C_{it} \neq 0$	72.80			66.27		
% consuming when $A_{it} = C_{it} = 0$	8.69			13.94		
annual doctor visits	4.89	2.00	8.69	4.86	2.00	6.62
annual hospital days	0.65	0.00	2.00	0.65	0.00	1.51
months with Rx cons.	4.45	2.00	4.96	4.28	2.00	4.66
Medical Care Expenditure						
total annual medical care exp.	1473.13	358.34	3656.88	1267.70	406.25	3062.88
Held Insurance Types						
% No Insurance	9.17			11.34		
% HMO	37.99			37.46		
% PPO	10.14			12.46		
% FFS	42.69			38.74		

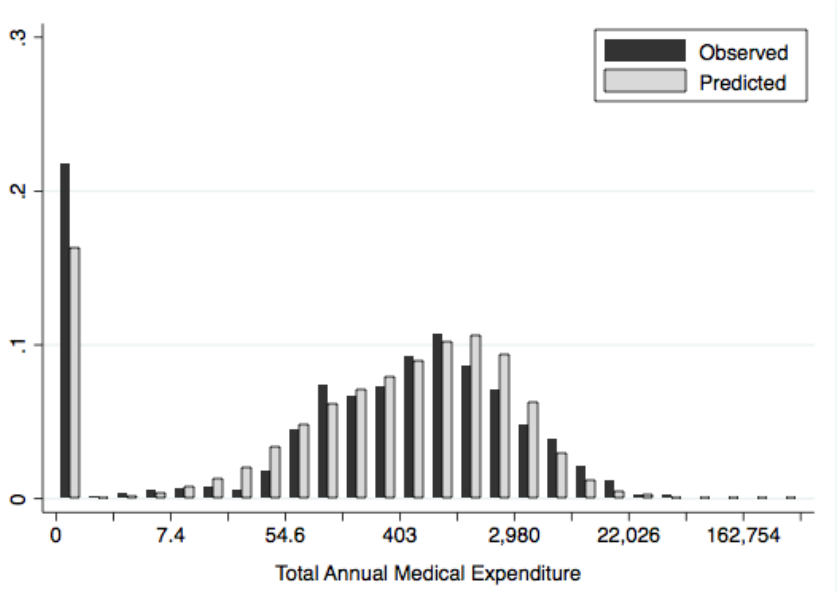
† Conditional on any chronic/acute illness.
‡ Prices are only observed when an individual consumes medical care; thus, the simulated mean and standard deviation are calculated only from individuals consuming care in simulation.

Table 13: Observed and Simulated Annual Consumption

	Doctor visits		Hospital days		Prescriptions	
	Obs.	Sim.	Obs.	Sim.	Obs.	Sim.
0	28.17	24.22	75.81	71.93	35.47	31.14
1	16.15	16.04	12.09	15.43	13.31	16.22
2	12.66	11.47	4.87	4.75	5.60	7.95
3-5	18.43	19.34	4.79	5.88	10.47	9.67
6-9	10.46	12.29	1.46	1.60	7.63	10.47
10+	14.12	16.67	0.97	0.41	27.52	24.55

Notes: The table reports annual consumption levels for the observed and simulated samples. The values are the percentage of the samples consuming at each level. For example, the top left entry states that 28.17% of the observed sample visited a doctor's office zero times during the insurance year. The prescription drug levels are measured in consumption months.

Figure 3: Annual Medical Care Expenditure



Notes: This figure compares observed and simulated annual expenditure distributions. The horizontal axis is log-scaled, so that markers are spaced $exp(1.0)$ apart.

Table 14: Predicted Increase in Annual Expenditure

	WYDM Model		RAE Model	
Percentile	Force	Full	Force	Full
50	0	0	0	0
55	17	69	0	0
60	83	144	0	84
65	164	250	48	203
70	285	403	176	380
75	462	616	346	630
80	722	927	602	987
85	1,143	1,414	1,004	1,529
90	1,904	2,265	1,720	2,467
95	3,651	4,172	3,437	4,791
99	10,906	12,044	12,295	16,266
Mean	772	899	760	1,115

Notes: This table displays various percentiles of the distribution of the predicted increase in mean annual medical care that results from insurance acquisition. The first and third columns, labeled *Force*, correspond to a counterfactual where an individual is forced into his most preferred plan among those offered to him by his employer. The second and fourth columns, labeled *Full*, correspond to a counterfactual where an individual is forced into full coverage. Note that the out-of-pocket insurance premium is assumed to be zero with full coverage. These expenditure values are then compared to those calculated in a third counterfactual, where individuals are forced to be uninsured.

Table 15: Change in Endogenous Variables with Insurance Acquisition

	WYDM Model			RAE Model		
Variable	Uninsured	Insured	% Δ	Uninsured	Insured	% Δ
Medical Care Consumption						
average annual doctor visits	3.85	5.41	+40.52	4.17	4.97	+19.18
average annual hospital visits	0.05	0.72	+1,340.00	0.33	0.72	+118.18
average annual Rx months	4.15	4.27	+2.89	4.34	4.46	+2.76
Medical Care Prices						
average doctor price	88.56	93.95	+6.07	75.85	90.57	+19.41
average hospital price	610.94	869.03	+42.24	267.54	869.69	+225.07
average Rx price	61.01	60.52	-0.80	65.65	68.92	+4.98
Illness†						
total acute illness months	7.75	7.66	-1.16	*	*	*
average number of chronic illnesses by years end	0.860	0.858	-0.23	*	*	*

† Illness transitions are treated as exogenous in the RAE model, so there is no change due to insurance acquisition.

Table 16: Other Measures of the Effects of Moral Hazard

	Measure	Reported Estimate	Within-Year Model Estimate
Manning et al. (1987)[a]	Co-insurance arc elasticity of medical care demand 0%→25%, 25%→95%[b]	0.17, 0.22	0.20, 0.35
Keeler and Rolph (1988)	Percentage of full insurance expenditure not explained by moral hazard[c]	55%	44%
Einav et al. (2013)	Percentage increase in mean annual expenditure, no insurance to full coverage[d]	30%	135%
Bajari et al. (2014)	Percentage of preferred insurance expenditure explained by moral hazard[e]	45%	24%

a These results are also reported in Keeler and Rolph (1988).

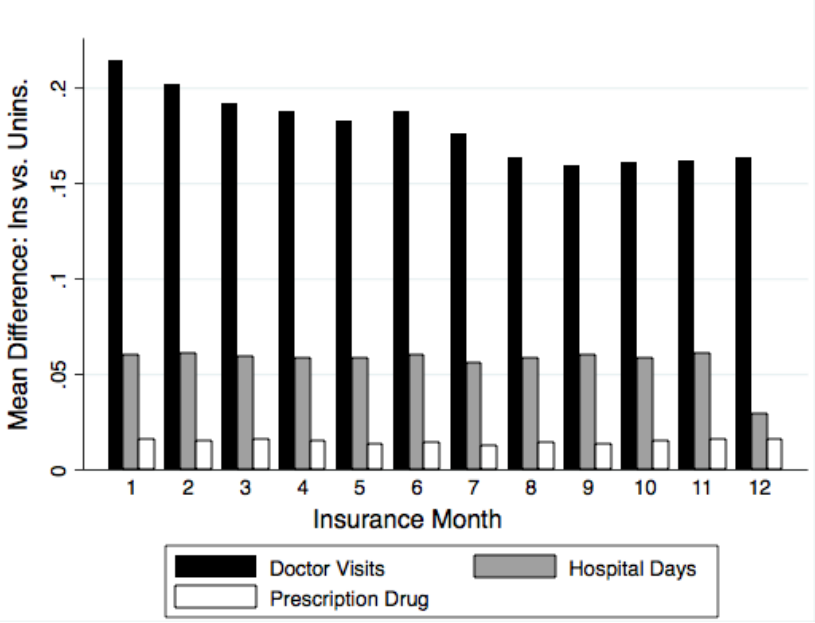
b The arc elasticity is calculated as $E_{arc} = ((q_2 - q_1)/(p_l - p_1)) \times ((p_2 + p_1)/2)/((q_2 + q_1)/2)$, where q is mean annual medical care expenditure and p is the co-insurance rate. Manning et al. (1987) make this calculation for each type of care and then weight elasticities for various types of care by share of spending.

c q_2/q_1 is calculated, where q_2 is mean annual medical care expenditure with a 95% co-insurance rate (i.e., near no insurance) and q_1 is mean annual medical care expenditure with full insurance.

d $(q_2 - q_1)/q_1$ is calculated, where q_2 is mean total annual medical care expenditure for the population under full coverage and q_1 is mean total annual medical care expenditure for the population under no coverage.

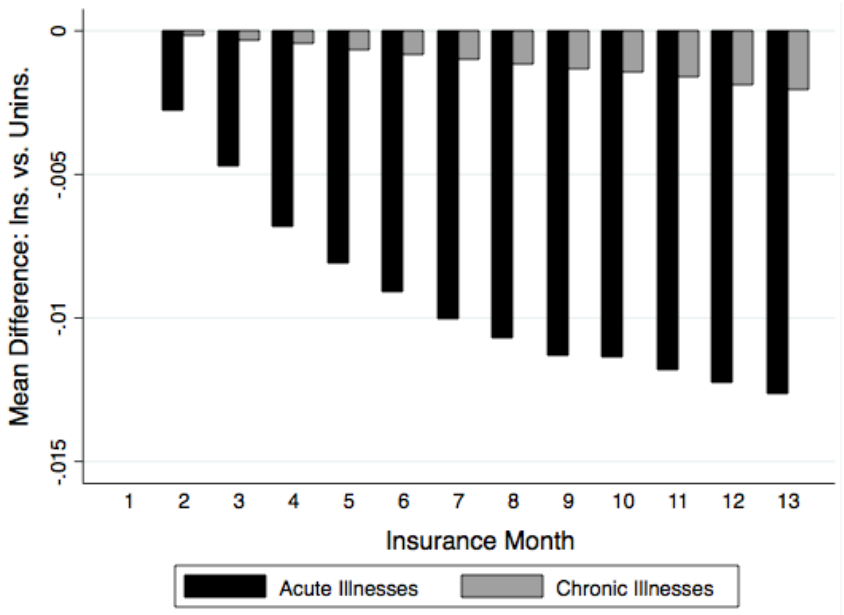
e $(q_{2i} - q_{1i})/q_{2i}$ is calculated for each individual, where q_{2i} is total annual medical care expenditure with preferred/chosen coverage and q_{1i} is total annual medical care expenditure with no coverage for individual i. Note that this measure differs from my preferred estimate, which compares mean expenditure levels, rather than the average of individual spending ratios.

Figure 4: Medical Care Consumption Difference by Insurance Month



Notes: This figure compares the difference in mean medical care consumption between fully insured and uninsured individuals in each month of the insurance year. The horizontal axis measures the number of months since the beginning of the the health insurance year. The vertical axis measures, for example, the average number of doctor's office visits for a fully insured individual minus the average number of doctor's office visits for an uninsured individual in each month. For prescription drugs, the vertical axis measures the difference in the proportion of insured and uninsured individuals consuming any drugs. Note that this image is produced from simulated data, so that the insured vs. uninsured comparison is done for the same set of individuals. Most importantly, the figure illustrates that the increase in doctor's office visits that results from insurance acquisition decreases over the course of the year, presumably due to an improvement in health.

Figure 5: Illness Difference by Insurance Month



Notes: This figure compares the difference in the average number of illnesses between fully insured and uninsured individuals in each month of the insurance year. The horizontal axis measures the number of months since the beginning of the the health insurance year. The vertical axis measures, for example, the average number of acute illnesses for a fully insured individual minus the average number of acute illnesses for an uninsured individual in each month. Note that this image is produced from simulated data, so that the insured vs. uninsured comparison is done for the same set of individuals. Most importantly, the figure illustrates that the full health improvement caused by insurance acquisition is not immediate, but accumulates over time.

Table 17: Characteristics of Insurance Switchers vs. Non-Switchers

	Counterfactual 1				Counterfactual 2	
	Uninsured		Insured		Insured	
	No Switch	Switch to HDHP	No Switch	Switch to HDHP	Non-Generous	Generous
Δ in expenditure	*	11.594	*	-8.867	*	-419.14
Δ in welfare (CV)	*	215.68	*	242.33	*	-361.64
Δ in doctor visits	*	0.070	*	-0.183	*	-1.27
Δ in hospital visits	*	0.024	*	0.002	*	-0.41
Δ in Rx months	*	0.082	*	-0.023	*	-0.08
Acute illnesses entering year	59.4	35.7	65.0	35.6	51.5	66.8
Chronic illnesses entering year	74.5	9.0	73.9	6.0	43.6	76.8
Income	2308.0	2220.2	2905.4	2729.6	2924.0	2735.3
Number of plans†	2.0	1.4	5.14	2.42	6.2	2.1
Minimum monthly OOP premium	37.58	39.10	20.99	24.76	23.84	22.82
Unobserved type						
Type 2:	19.9	1.3	25.0	1.1	13.3	27.4
Type 3:	16.3	1.5	16.5	1.2	9.1	17.6
Type 4:	16.6	0.7	21.7	0.6	11.3	23.9
Observations	7,993	4,403	87,564	23,240	66,241	48,966
Percentage of Population	64.5	35.5	79.0	21.0	57.5	42.5

† In all columns, number of plans refers to the number of plans actually offered by the employer.

* Notes: In the first two columns, I compare the characteristics of uninsured individuals who remain uninsured to uninsured individuals who decide to purchase the HDHP when offered. In the second two columns, I compare insured individuals who remain in their plan to insured individuals who switch to the HDHP when offered. In the final two columns, I limit the sample to all individuals selecting any plan in Counterfactual 1 (i.e., when the HDHP is available). I then compare those selecting the most generous offered coverage to those selecting a plan other than the most generous one they are offered.