Investment in Human Capital, Longevity and Moral Hazard in a Stochastic Life-cycle Model of Demand for Health

Ahmed Khwaja*
Fuqua School of Business
Duke University
Email: ahmed.khwaja@duke.edu

Work in Progress

Abstract

This paper develops and estimates a life cycle dynamic stochastic discrete choice model of individual decisions about health insurance, exercise, smoking, alcohol consumption and medical treatment in a Grossman (1972) type framework where health is human capital. The model accounts for the effects of moral hazard from health insurance and longevity on health related behaviors. Furthermore, longevity is endogenous to behaviors. The estimation of the model controls for dynamic selection through mortality. In particular it develops and implements a simulation based procedure for dealing with an initial conditions problem that is common in estimating dynamic models. The model is estimated using data from the Health and Retirement Study. Model simulations imply that access to health insurance has considerable impact on medical utilization but there is little moral hazard with respect to other health related behaviors, i.e., alcohol consumption, smoking and exercise. The results also suggest that increases in longevity provide incentives to lead a healthier life-style (the 'Mickey Mantle' effect). Hence, one reason that the level of moral hazard in health behaviors may be small is that the dynamic effects of longevity counteract the moral hazard generated by health insurance. The effects of health insurance on health outcomes are found to be small because utilization is largely for *palliative* rather than *preventive* or *curative* purposes.

Keywords: Life-cycle human capital, health production, health insurance, moral hazard, survival incentives, mortality selection, alcohol consumption, smoking, exercise, dynamic discrete choice

JEL Classification: C33, C35, C53, D12, D13, D91, I12, I18

*I am grateful to Mike Keane for his advice and suggestions. This paper has also benefitted from discussions with Abby Barker, Toshi Arimura, V. V. Chari, Tim Cheney, Andrew Ching, Roger Feldman, Tom Holmes, Leo Hurwicz, Eric Olson, and Reuven Shnaps. I acknowledge the help of the Minnesota Supercomputer Institute, the use of the HRS data and the help of its staff. I alone am responsible for any errors.

1 Introduction

This paper develops and estimates a life cycle dynamic stochastic discrete choice model of individual decisions about health insurance, exercise, smoking, alcohol consumption and medical treatment and the consequent effects of these choices on health outcomes. The model is in the Grossman (1972) type framework where health is human capital as well as a consumable commodity. The model accounts for the effects of moral hazard from health insurance and longevity on health related behaviors. Furthermore, longevity is endogenous to behaviors. The estimated model is used to examine the effects of potential changes in health insurance coverage on health outcomes and health related behaviors over the life cycle. In particular to determine the moral hazard associated with the provision of health insurance on medical treatment (Arrow 1963, Pauly 1968) and habits like smoking and alcohol consumption. The model is also used to analyze the effect of longevity on health behaviors.

An analysis of the individual as a (household) producer of his own health must recognize the significance of uncertainty in this process. An individual's decisions about health related choices (e.g. health insurance, medical care use, exercise) depend on his state of health, which is randomly determined. Conversely, his state of health is also affected by his health-related choices in a stochastic way. Thus the production of health and the health related decisions of individuals are inter-related. It is not just that the demand for medical care and health insurance is derived from the demand for health. More than that, the demand for health, demand for health insurance, medical utilization, and health related behaviors that also directly affect utility (e.g., smoking) are jointly determined. Consequently to analyze the production of health and the choices about health related behaviors one has to be cognizant of the endogeneity of these choices (see e.g., Mundlak 1963, Rosenzweig and Schultz 1983). In particular one important reason for formulating a dynamic life cycle model of investment in health is to understand the role of medical care on health outcomes. At any given point in time health status and medical care are negatively related because sicker people tend to seek more medical care, i.e. the medical treatment choice is endogenous. Further, since health itself is an endogenous commodity that can be produced or degraded by individual actions, there will be a systematic tendency for individuals with certain kinds of "unhealthy" choices to be sicker than others on average (Grossman, 1972). In other words there will be dynamic selection in to the sample that seeks medical treatment (Rosenzweig and Wolpin 1995). Given this correlation between individual behaviors, medical treatment and health status it is imperative to track the interaction of habits, medical care and health status over the life cycle in order to measure the benefits of medical care on health status.

A second reason for developing a dynamic model of life cycle investment in health is "survivorship bias" or dynamic selection (e.g., Olley and Pakes 1996, Cameron and Heckman 2001). In other words, a crucial determinant of individuals being included in a data sample is because they had not degraded their health to an extent that they died prior to the process of being sampled. Since in the model developed in this paper the life horizon of an individual is endogenously determined, on the basis of his life cycle health investment decisions, the issue of survivorship bias in empirical work can be resolved by accounting for the process that leads to censoring of individuals from the data.

A third reason for constructing a life cycle model of health production that incorporates health insurance decisions is to deal with the issue of adverse selection in the individual choices about health insurance (Phelps 1973, Rothschild and Stiglitz 1976) and the consequent endogeneity in estimating the effects of insurance on individual health related decisions and health outcomes. In addition if one is interested in analyzing the effects on health outcomes of individual choices like smoking that both provide utility through consumption as well as play a role in household production of health then one has to account for the endogeneity that arises through "joint production" (Pollak and Wachter 1975). The endogeneity makes the task of disentangling the individual preferences from household technology difficult in determining the marginal effects of these choices on health outcomes. The dual objectives of analyzing the effects of individual choices about health insurance and habits (that lead to joint production) on health outcomes can be tackled through the construction of a dynamic model of health outcomes that endogenizes the health insurance decision and clearly delineates the separate consumption and production roles of the habits. A fifth reason for developing and estimating a health as human capital model is to examine in a rigorous fashion whether Grossman type models, which are the work horse models in the health production literature, are able to explain observed individual behaviors well.

Grossman (1972), in a deterministic framework, analyzed the decisions of individuals to invest in their own health under both economic and bio-technological constraints. He noted that investment in health by individuals involved behavioral inputs (e.g. exercise) as much as the use of medical care. He also emphasized the dynamic and forward looking aspects of individual decisions in the production of health, i.e. past decisions affect the current health, which in turn affects current decisions and thus future health. This implies that in effect individuals can make choices about their desired health states through their decisions regarding behaviors that affect health. Some of these behaviors also provide direct utility through consumption, e.g. smoking. Thus the consumption choices and the production of desired health outcomes are made simultaneously by individuals. This "joint production" leads to endogeneity in the individual decisions, which is an issue of importance when considering empirical analysis (Barnett 1977, Pollak and Wachter 1977). In fact this means that a static (or cross-section) regression of health outcomes on medical care will often produce a negative relationship. This is because at any given point in time, it is more likely that the sicker individuals (who have made less investment in their health behaviorally, e.g. high smoking and low exercise) will be using more medical services compared to the healthier ones (to compensate for the other choices). The Grossman model also has implications for dynamic empirical analysis. It suggests that there will be attrition from a longitudinal sample of individuals on a systematic basis leading to selection in the data. This will happen because the individuals who invest less in their health (and are consequently sicker) or receive bad health shocks will die and drop out of the sample.

In the model developed in this paper I adopt a framework of health as human capital a la Grossman (1972). The model is dynamic and incorporates uncertainty with respect to future health outcomes. One novelty of the research is to explicitly include the choices over smoking, exercise and alcohol consumption. I allow for "rational addiction" (Becker and Murphy 1988) in these habits. Secondly, in estimating the benefits of medical care I differentiate between the mitigatory effects that alleviate short run discomfort and those that lead to long term investments in health. A third innovation is that in the estimation I allow for health to play the dual roles of a consumption good that provides utility and a human capital input in the income earning capacity of an individual. Fourthly, I model the decisions of individuals about health insurance. This provides me with a means of estimating the effects of health insurance on not just medical care but also health outcomes and habit decisions, which is an additional contribution of my research. Consequently one way of characterizing my model is to see it as an extension of Grossman (1972) and Phelps (1973) to a framework of uncertainty with the addition of an health insurance choice for the individual and an explicit incorporation of multiple inputs in the health production function.

To my knowledge there is almost no work on the joint estimation of the demand for health insurance and the demand for other services and behaviors that affect health outcomes. Cameron et. al. (1988) was an exception that developed a model of the joint decision by an individual to purchase health insurance and to seek medical care. However the demand for medical care was not estimated jointly with the demand for health insurance. A reduced form was used to estimate the demand for health insurance whereas the structural parameters determining the medical treatment decision were estimated.

The work in this paper also complements the methods developed by Manning et. al. (1987) and Duan et. al. (1983, 1984). These 'two/four part models' was developed in tandem with the RAND Health Insurance Experiment (Newhouse 1993). Since it is unlikely that such an enterprise will be undertaken in the near future, e.g., due to ethical and economic constraints, comparing the results of this research which relies on 'counter-factual simulations' to estimate the level of moral hazard to the results of the RAND study should be of considerable interest.

I use data from the Health and Retirement study (HRS) on 3671 males for the period 1991-1998. The model is estimated using a maximum likelihood procedure that nests the numerical solution of the mode (Rust 1987, 1994). The estimation controls for unobserved heterogeneity with respect to latent preferences for health, ability to produce health and income earning capacity. Failure to control for unobserved heterogeneity can result in misleading estimates of a health production function. This is because in the presence of unobserved heterogeneity the inputs in a health production function will be correlated with the econometrically specified residual. The consequence would be biased estimates of the production coefficients if the endogeneity is ignored (Rosenzweig and Schultz 1983). One reason unobserved heterogeneity may occur is because of differences in genetic endowments (that are known to the individuals but unknown to the econometrician and) that may make some individuals more resistant to the consequences of unhealthy behaviors and others more susceptible.

The methodology adopted in this paper may be traced in form to Wolpin (1984) which examined the relationship between fertility and child mortality. Rust and Phelan (1997) used similar estimation methodology to analyze the effects of social security policies and the Medicare system on the retirement behavior of the elderly. Similarly Lillard and Weiss (1997) have examined the financial behavior of the elderly vis a vis their health outcomes. Their model however did not allow for the health status of the individuals to change through their behavioral choices. Related work by Sickles and Yazbeck (1998) jointly estimated a health production function and the demand for consumption and leisure for elderly males. Their research however did not incorporate any decisions about habits and was limited to the input of medical care in the production of health. Recently Gilleskie (1998) analyzed medical care use and absenteeism and the effects of these on health care costs using closely related techniques. However her model was one of short run acute illnesses that did not change a person's fundamental state of health. My research extends this aforementioned work and in one sense may be seen as an extension of Gilleskie's model to a long run framework.

Estimates suggest that the model is able to replicate the data well. The model does well in matching the means, frequencies and transitions in the data for the health states, habits and medical care. Income and out of pocket expenditure are also matched well. The out of sample fit of the model is also examined using US mortality data from the CDC, Atlanta and data on males from the National Health Interview Survey 1998 for smoking, alcohol consumption and proportion of uninsured in the population. By these measures the model performs well too.

The estimates of the model are employed in evaluating the effects of two radically different counterfactual public health policies. Though it is not likely that these two policies per se are going to be introduced in to the real world in the near or far future they help provide bounds on what the greatest potential changes in individual behaviors and health outcomes would be in the face of the highly drastic changes in health insurance. The first policy is one in which every individual in the economy is mandated to be on a health insurance plan that charges an annual premium of $1000.0. This is approximately the average premium paid by the individuals in the HRS data. The insurance however covers all out of pocket costs. This experiment tries to simulate the effects of a policy that provides comprehensive health insurance coverage. The simulations suggest that there is not much change in average life cycle health outcomes. However there is a substantial increase in the demand for medical care. The proportion of individuals seeking medical treatment increases by as much as 49% while health outcomes improve negligibly (up to 0.06%). It is also found that there is a slight increase (up to 0.4%) in alcohol consumption compared to the baseline simulations. Smoking rates are slightly higher (as much as 0.5%) compared to the status quo. Importantly it is found that the provision of subsidized medical care does not lead to a large increase in the rates of smoking or alcohol consumption.

In the second policy experiment every individual in the economy is denied health insurance. In this policy regime the individuals do not incur any insurance costs but have to pay for any medical treatment they seek. This experiment tries to replicate the effects of the complete withdrawal of the provision of subsidized medical care. Not surprisingly it is found that the demand for medical care falls. The fall is most marked for the elderly. The proportion of individuals seeking medical care reduces as much as 95% compared to the status quo and there is a drop in life cycle health outcomes of as much as 0.9%. It is also found that the proportion of individuals consuming alcohol falls (up to 0.3%) compared to the baseline simulations. Smoking rates fall as much as 0.6% in comparison with the baseline simulations. This policy experiment indicates that withdrawal of medical coverage leads to a fall in the demand for medical care but not to a significant fall in the rates of smoking and alcohol consumption.

Taken together the two experiments suggest that there is moral hazard associated with the provision of subsidized medical care and the demand for medical treatment as found in the literature previously (Manning et al. 1987). However there is no moral hazard associated with the provision of subsidized medical care and the individual decisions about smoking and alcohol consumption.

The rest of the paper is organized as follows. Section 2 describes the model, while Section 3 provides a brief summary of the data. Section 4 discusses the estimation procedure. The estimation results and the consequences of two radically different policy experiments are analyzed in section 5. Finally section 6 concludes.

2 The Model

2.1 Structure and Specification

The stochastic life cycle model in which health is both a consumption good and human capital (Grossman 1972) employs a dynamic discrete choice framework used previously in estimating a durable goods model (e.g., Rust 1987, 1994a, 1994b). Individuals are assumed to be forward-looking with a finite lifetime, $t = 1, \ldots, T$. They maximize their lifetime discounted utility by making sequential choices about health insurance, I_t, alcohol consumption, a_t, smoking, c_t, exercise, e_t and medical care, m_t, in each time period, t. They derive utility from health, H_t, alcohol consumption, smoking, exercise and a composite consumption commodity, X_t. The model adopts a random utility specification (McFadden 1981, McFadden and Richter 1990) with a stochastic component, ζ_t, for the preferences associated with each of the choices. Defining β as the discount rate and $U(\cdot)$ as the single period utility function over decisions and states described below, the maximization problem is represented as,

$$\max_{\{I_t, a_t, c_t, , e_t, m_t\}} E\left[\sum_{t=1}^{T} \beta^t U(I_t, I_{t-1}, H_t, a_t, a_{t-1}, c_t, c_{t-1}, e_t, e_{t-1}, X_t, s_t, m_t, HHS_t, A_t, \zeta_t)\right] \quad (1)$$

subject to

$$H_t = h(H_{t-1}, a_{t-1}, c_{t-1}, e_{t-1}, s_{t-1}, m_{t-1}, \epsilon_t^H) \quad (2)$$

$$X_t = Y_t - P_I - OOP_t \quad (3)$$

$$Y_t = y(H_{t-1}, A_t, \epsilon_t^y) \quad (4)$$

$$OOP_t = o(I_t, H_t, m_t, A_t, HHS_t, \epsilon_t^{oop}) \quad (5)$$

$$s_t = s(H_t, A_t, \epsilon_t^s) \quad (6)$$

$$HHS_t = f(HHS_{t-1}, A_t, \epsilon_t^{HHS}) \quad (7)$$

$$H_0 = \overline{H}_o; \quad I_0 = \overline{I}_o; \quad a_0 = \overline{a}_o; \quad c_0 = \overline{c}_o; \quad e_0 = \overline{e}_o; \quad m_0 = \overline{m}_o \quad HHS_0 = \overline{HHS}_o \ .$$

The model may also be interpreted as a long run version of the model developed by Gilleskie (1998). The model accounts for the moral hazard because the endogenous alcohol consumption, smoking and exercise choices explicitly depend on health status and health insurance. The model also allows for addiction (Becker and Murphy 1988) in these behaviors. Thus the current choices depend on the past as well as the expected future choices. The alcohol consumption choices, a_t, take one of three values (1-"none," 2-"drinks per day ≤ 1," 3-"drinks per day > 1"); smoking choices, c_t, take one of three values (1-"none," 2-"packs per day < 1," 3-"packs per day ≥ 1"); exercise choices, e_t, may be one of two values (1-"no," 2-"yes").[1]

[1] There are only two choices for exercise due to data limitations in the HRS.

In the model, insurance protects against medical expenditure risk by reducing out of pocket costs. In turn, the demand for insurance depends on health outcomes and health behaviors (Phelps 1973), which also affect medical expenditures. Insurance does not provide any utility per se but there is a monetary equivalent of the cost of switching insurance plans. Hence I_t is an argument of the utility function (1). Insurance choices take one of *six* values if age ≤ 64 and *three* if age ≥ 65. The set of insurance choices is enumerated as:

(a) For age ≤ 64,

INS={1-"none," 2-"group," 3-"personal," 4-"VA/Champus," 5-"group/VA/Champus," 6-"group/personal."}

(b) For age ≥ 65,

INS={1-"Medicare," 2-"Medicare/Medigap/other-personal," 3-"Medicare/group."}

These alternatives are based on the empirical distribution of insurance choices in the HRS.[2] For those under 65, "1" is the choice to be uninsured, plan 2 represents group coverage, e.g., through employer, plan 3 is personal coverage and plan 4 is coverage for armed services personnel and veterans. Plan 5 represents a mix of group and VA/Champus coverage[3] and plan 6 represents a mix of group and personal insurance coverage. Plans 2 and 3 for those 65 and over are similarly defined. Individuals younger than 65 can be uninsured but those over 65 have at least Medicare coverage. An advantage of modelling insurance choices in a discrete choice framework is that it parsimoniously characterizes the non-linear expenditure contingent features of insurance plans (Keeler et al 1977). The supply of insurance is assumed to be exogenous because the HRS data lacks information about the insurance choices available to each individual, and the factors used by insurers (or employers) to determine the characteristics and premiums of plans they offer, and also for computational reasons.[4] Thus all individuals are assumed to have access to the same set of insurance choices conditional on age. The idiosyncratic random utility shocks can however account for the temporary unavailability of a particular plan, e.g., through the association between insurance and employment.

The model accounts for the "full price" of medical care by including a term for the monetary equivalent of the psychological cost of seeking medical care in the utility function (1), e.g., cost of scheduling and

[2]Medicaid is not included because of the small number of male HRS respondents on Medicaid and the difficulty of assessing Medicaid eligibility when asset formation and savings behavior are not modelled.

[3]E.g., veterans who have VA/Champus coverage and additional coverage through their current civilian employer.

[4]Dey and Flinn (2005) is the only example, to the best of my knowledge, of estimation of demand for health insurance in an equilibrium framework. They estimate an extremely sophisticated model of joint determination in equilibrium of employer provided health insurance and wages in a search, matching and bargaining framework. However they do not consider the life cycle aspect of health production nor the moral hazard and trade-offs inherent in the provision of health insurance. Further, "due to data limitations and also for reasons of tractability" even they assume that the "employer's direct cost of purchasing health insurance coverage is exogenously determined (p. 575)."

waiting for an appointment. The medical care choices, m_t, take one of three values: 1-"low," 2-"moderate," and 3-"high," depending on whether the total number of visits to a physician or medical facility are respectively fewer than approximately[5] 1/3, or between 1/3 and 2/3, or more than 2/3 of the empirical distribution. The model decomposes the effects of medical care in to *three* components. The first is *palliative* care. It is the pure consumption aspect that alleviates the short-run disutility of sickness (eq. 1). The second is *curative* care. It is the productive aspect that reduces the period specific depreciation in health stock (eq. 2). The third is *preventive* care. It is the productive net investment aspect that leads to improvements in future health (eq. 2). Disentangling these three components is insightful in understanding the effect of insurance on health outcomes. More details are provided below in the discussion on the health production technology, and in section 2.3 that discusses the specification of the utility function.

The model also distinguishes between the stock of health and a flow variable measuring sickness,[6] s_t. This helps to account for situations in which there may be consumption of medical care without any change in the underlying health stock, and also helps to disentangle the *mitigative*, *curative* and *preventive* components of medical care. It allows for disutility from sickness.[7] The sickness variable takes one of three values: 1-"none," 2-"moderate," 3-"high." The roles of household size, HHS_t, age, A_t, and lagged variables in the utility function (1) are explained later in section 2.3.

Health and longevity are endogenous in the model through a health production function (eq. 2). This allows for the current health to depend on one period lagged health, alcohol consumption, smoking, exercise, medical treatment, sickness and a random element, ϵ_t^H. Empirically health stock, H_t, is defined using two measures of health, i.e., (a) self-reported health status[8] (SRHS) and (b) mortality. Thus H_t takes one of six values: 1-"dead," 2-"poor," 3-"fair," 4-"good," 5-"very good," 6-"excellent." Death is an

[5]Given the discrete nature of the utilization variable the cut offs are not exact.

[6]For example, an Olympic skier would have a certain level of health stock while an AIDS patient would have a different level. The same sickness e.g., the common cold, could have very different effects on the health outcomes, medical utilization, OOP expenditures and health related behaviors of the two individuals.

[7]Conceptually the variable s_t may be considered an index of illnesses in the past two years. This is however difficult to implement empirically because the HRS data has limited information about illnesses in the past two years making it impossible to track every possible illness. Also explicit inclusion of multiple illnesses would expand the state space considerably creating a significant computational burden.

[8]SRHS is a rough but good measure of life cycle health that does well in predicting significant health events e.g., mortality. See e.g., Deaton and Paxson (1998a, 1998b), and Case and Deaton (2003), who discuss the pros and cons of the measure. The validity of the SRHS as an objective measure of health stock in the HRS data is strengthened because the survey question does not include a reference to "*people your age*." More complicated measures of health status could have been used but these would have made the model computationally intractable. Wallace and Herzog (1995) provide an overview of health measures in the HRS data.

absorbing state. The stochastic health production technology is specified to have a multinomial logit form with the index function for transition from health stock level H_{t-1} to health stock level $q = 1, ..., 6$, at each time t, $t = 1, \ldots, T$, given by,

$$Pr[H_t = q \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] = \frac{exp(\underline{\eta}'_q \cdot R_t)}{\sum_{q'=1}^{6} exp(\underline{\eta}'_{q'} \cdot R_t)} \tag{8}$$

where the index function $(\underline{\eta}'_q \cdot R_t)$ is parameterized for $q \neq 1$ as,

$$\begin{aligned} \underline{\eta}'_q \cdot R_t &= \eta_{1,q} \cdot (q - H_{t-1}) + \eta_{2,q} \cdot (q - H_{t-1})^2 \\ &\quad - [\eta_{3,q} - \eta_{4,q} \cdot m_{t-1}] \cdot [1\{s_{t-1} = 2\}] - \eta_{5,q} \cdot [\eta_{3,q} - \eta_{4,q} \cdot m_{t-1}] \cdot [1\{s_{t-1} = 3\}] \\ &\quad + \eta_{6,q} \cdot a_{t-1} + \eta_{7,q} \cdot c_{t-1} + \eta_{8,q} \cdot e_{t-1} + \eta_{9,q} \cdot m_{t-1} + \eta_{10,q} \quad . \end{aligned} \tag{9}$$

A multinomial specification is adopted (here and elsewhere below) rather than an ordered logit as the former allows for greater flexibility in replicating the inter-temporal transitions.[9] The current health outcome depends on the lagged health outcome via the terms $\eta_{1,q}$ and $\eta_{2,q}$. The quadratic specification allows for the persistence in health transitions. The effects of lagged choices on current health are represented by $\eta_{6,q}$ (alcohol), $\eta_{7,q}$(smoking), $\eta_{8,q}$ (exercise), and $\eta_{9,q}$ (medical treatment). The lagged sickness affects current health through $\eta_{3,q}$, in case of moderate sickness (i.e., $s_{t-1} = 2$), and $[\eta_{5,q} \cdot \eta_{3,q}]$ in case of high sickness (i.e., $s_{t-1} = 3$) respectively. The $\eta_{5,q}$ is a proportionality factor for high sickness. Utilization of medical care alleviates the effect of moderate sickness by the amount $\eta_{4,q}$. In case of high sickness, utilization alleviates the effect of sickness by the amount $[\eta_{5,q} \cdot \eta_{4,q}]$. Hence $\eta_{4,q}$ is the effect of medical care in alleviating the depreciation in health stock. It is referred to as the curative component of medical treatment. On the other hand $\eta_{9,q}$ is the net investment effect of medical care on future health. This is referred to as the preventive component of medical care.

The sickness variable (s_t) may be considered to be an index of illnesses in the past two years. It is assumed to be a latent variable that is unobserved to the econometrician.[10]. It is specified to evolve as a multinomial logit process that is a function of For all $t = 1, \ldots, T$ the probability of sickness level[11] $q = 2, 3$ is,

$$Pr[s_t = q \mid H_t, A_t] = \frac{exp(\phi_{1q} + \phi_{2q} \cdot H_t + \phi_{3q} \cdot A_t)}{1 + \sum_{q'=2}^{3} exp(\phi_{1q'} + \phi_{2q'} \cdot H_t + \phi_{3q'} \cdot A_t)} \quad . \tag{10}$$

[9] The benefit of using an ordered logit would likely be efficiency gains in estimation but preliminary work showed that this specification would do a poor job in replicating health transitions. Hence it was not adopted.

[10] The HRS data has limited information about illnesses in the past two years. It would be near impossible to track every possible illness that afflicted each individual. Explicit inclusion of multiple illnesses would create a significant computational burden.

[11] The sickness levels are 1-none, 2-moderate, 3-high.

with the standard normalization of the parameters, i.e., $\phi_{r,1} = \underline{0}$ for $r = 1, \ldots, 3$. the current health (ϕ_{2q}) and age (ϕ_{3q}). . There is conditional independence assumption between H_t and s_t. That is,

$$Pr[H_t, s_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] = Pr[s_t \mid H_t, A_t] \cdot Pr[H_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] \ .$$

$$\begin{aligned} Pr[H_t, s_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] \ &= \\ & Pr[s_t \mid H_t, A_t] \cdot Pr[H_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] \ . \end{aligned}$$

This is an exclusion restriction that implies that the current sickness does not affect the current health but, conversely, the current health does affect the current sickness. Current sickness does, however, affect the future health state. The parameters are identified using the variation in pre-determined health and age, and the exclusion restriction. The household size is specified to be a multinomial logit process that is a function of lagged household size and current age. For all $t = 1, \ldots, T$ the probability of moving from household size HHS_{t-1} at time $t-1$ to a household size level[12] $q = 1, \ldots, 4$ is,

$$Pr[HHS_t = q \mid HHS_{t-1}, A_t] = \frac{exp(\psi_{1q} + \psi_{2q} \cdot (q - HHS_{t-1}) + \psi_{3q} \cdot (q - HHS_{t-1})^2 + \psi_{4q} \cdot A_t + \psi_{5q} \cdot (A_t)^2)}{1 + \sum_{q'=2}^{4} exp(\psi_{1q'} + \psi_{2q'} \cdot (q - HHS_{t-1}) + \psi_{3q'} \cdot (q - HHS_{t-1})^2 + \psi_{4q'} \cdot A_t + \psi_{5q'} \cdot (A_t)^2)} \ . \quad (11)$$

As with multinomial logit specifications the parameters are normalized so that $\psi_{r,1} = 0$ for $r = 1, 4, 5$. The specification states that current household size depends in quadratic[13] fashion upon both the lagged household size (via ψ_{2q} and ψ_{3q}) and age (via ψ_{4q} and ψ_{5q}). Given an unique classification feature of the HRS data the household size process is endogenous to mortality. The HRS dictates that if a spouse in a two or more member household dies then in the next wave the dead spouse is classified as a *new single member* household. Identification comes from the the pre-determined lagged household size and current age variables.

The per-period budget constraint (eq. 3) is, $X_t + P_I + OOP_t = Y_t$, where X_t is the composite consumption good in monetary terms, P_I is the insurance premium in each period,[14] OOP_t is the current OOP medical expenditure and Y_t is the current income.[15]

The OOP cost function is used to model the affect of insurance coverage on individual health behaviors, and consequent health outcomes, i.e., through the budget constraint (eq. 3). The OOP expenditures (5)

[12] The household size levels are 1-one member, 2-two members, 3-three members, 4-four or more members.

[13] This specification allows for persistence in the household size transitions as well as the characteristic that the likelihood of a transition to household size closer to HHS_{t-1} is greater than to a size further away from HHS_{t-1}.

[14] It is assumed that the individual has perfect foresight about the insurance premia and these are constant over time. This restrictive assumption is adopted because information about plan characteristics in extremely limited in the HRS data.

[15] Information about geographical location of the HRS respondents was not available. Thus it was not possible to calculate expenditures on alcohol consumption and smoking. Exercise could not be priced e.g., as time costs, as information about the duration of exercise is not available. These costs are subsumed in the net indirect utility from these choices.

depend on the health and insurance status of the individual as well as the utilization of medical care, age, the household size,[16] HHS_t and a random element, ϵ_t^{oop}. The OOP cost function takes a flexible mixed-continuous form to allow for non-linearities in the insurance coverage and reimbursements conditional on the amount of medical treatment (Keeler et al. 1977). The model allows for OOP expenditures, OOP_t, to be zero conditional on medical utilization due to generous coverage. It also explicitly relates the non-use of medical care to zero OOP costs. The probability that the OOP costs are zero is a function of insurance status, medical utilization, health, household size and age. This probability is modelled as a logit.

If the OOP costs are positive at time t, the mean OOP expenditure, $\overline{OOP_t}$, is a function of insurance status, health status, age, household size[17] and medical care consumption. That is,

$$\begin{aligned}\overline{OOP_t} \;=\; & \left[\sum_{q=1,\dots,6} \mu_q \cdot 1\{I_t = q\} \cdot 1\{A_t < 65\}\right] \cdot \left[1\{m_t = 2\} + \left(\mu_7 \cdot 1\{m_t = 3\}\right)\right] \\ & + \left[\sum_{q=1,\dots,3} \mu_{7+q} \cdot 1\{I_t = q\} \cdot 1\{A_t \geq 65\}\right] \cdot \left[1\{m_t = 2\} + \left(\mu_{11} \cdot 1\{m_t = 3\}\right)\right] \\ & + \mu_{12} \cdot H_t + \mu_{13} \cdot HHS_t + \mu_{14} \cdot \log(A_t)\ .\end{aligned}$$

The terms $(\mu_1, \dots, \mu_6)$ represent the effect of insurance status[18] for individuals under the age of 65 who consume moderate (i.e., $m_t = 2$) level of medical care. For individuals under the age of 65 seeking high (i.e., $m_t = 3$) levels of medical care, the mean OOP expenditure is determined by the coefficients $(\mu_1, \dots, \mu_6)$ multiplied by a proportionality constant μ_7.[19] Similarly, the mean OOP expenditure of those over the age of 65 who seek a moderate amount of medical care is determined by the terms $(\mu_8, \dots, \mu_{10})$, while in the case of high medical care consumption these terms are multiplied by the proportionality constant μ_{11}. The terms μ_{12}, μ_{13} and μ_{14} represent the effects of health status, household size and age respectively. The law of motion for out of pocket expenditures states that in each time period, t, the individuals know their medical expenditures up to a random error. Thus if $OOP_t > 0$, $\log\ OOP_t = \overline{OOP_t} + \epsilon_t^{oop}$

$$\log\ OOP_t = \overline{OOP_t} + \epsilon_t^{oop} \tag{12}$$

[16]Household size is included because individuals purchasing insurance typically consider the future needs of their household members along with their own. In this they form expectations over not just their own but also the medical expenditures of the household. Income information exists at the household level and hence only a measure of household consumption is available in the HRS data. Thus a reason for including household size in the model is to normalize household consumption to create a measure of individual consumption. This is discussed further in section 2.3.

[17]The individuals may purchase insurance in anticipation of not just their own expected OOP costs but also for their household members.

[18]There are 6 insurance choices for those under the age of 65 and 3 for those older than 65.

[19]Including a more complicated non-linear form was not computationally feasible.

where $\epsilon_t^{oop} \sim N(0, \sigma_{oop}^2)$. The model allows for OOP expenditures, OOP_t, to be zero in spite of consumption of medical treatment. It also explicitly relates the non-use of medical care to zero OOP costs. The probability that the OOP costs are zero, $OOP_t = 0$, is a function of the insurance status, medical utilization, household size and age. This probability is modeled as a logit process. with the index function $(\underline{\gamma}' \cdot G_t)$ parameterized as,

$$\begin{aligned}\underline{\gamma}' \cdot G_t &= \gamma_1 + \gamma_2 \cdot H_t + \gamma_3 \cdot HHS_t \\ &+ \left[\sum_{q=2,\ldots,6} \gamma_{2+q} \cdot 1\{I_t = q\} \cdot 1\{A_t < 65\}\right] \cdot \left[1\{m_t = 2\} + \left(\gamma_9 \cdot 1\{m_t = 3\}\right)\right] \\ &+ \left[\sum_{q=1,\ldots,3} \gamma_{9+q} \cdot 1\{I_t = q\} \cdot 1\{A_t \geq 65\}\right] \cdot \left[1\{m_t = 2\} + \left(\gamma_{13} \cdot 1\{m_t = 3\}\right)\right] .\end{aligned}$$

The γ parameters have a similar interpretation as the μ parameters described above. The price effects of insurance (i.e., the μ and γ parameters) are identified off those individuals under the age of 65 who are uninsured and those over the age of 65 who have only Medicare compared to others who also have supplemental plans. The source of the variation is the pre-determined insurance status and medical treatment choice.

Income (eq. 4) is considered to be the sum of wage and non-wage income. It is determined endogenously with the law of motion of income is log $Y_t = \overline{Y_t} + \epsilon_t^y$,

$$\log\ Y_t = \overline{Y_t} + \epsilon_t^y \tag{13}$$

where $\epsilon_t^y \sim N(0, \sigma_y^2)$ and $\overline{Y}_t$ is the average income of an individual in each period. The mean income is specified to be a function of past health, age, and education. Thus,

$$\overline{Y_t} = \kappa_1 + \kappa_2 \cdot (H_{t-1}) + \kappa_3 \cdot A_t + \kappa_4 \cdot A_t^2\ \ . \tag{14}$$

The parameters are identified using the variation in pre-determined health and age variables. Further, it is assumed that the random components of out of pocket medical expenditures and income are uncorrelated.[20]

The current sickness (eq. 6) depends on the current health, age and a random term, ϵ_t^s. It is specified to evolve as a multinomial logit process. [21] The household size (eq. 7) depends on the lagged household size, HHS_{t-1}, current age and a random element, ϵ_t^{HHS}. It is specified to take one of four values: 1-"one

[20]This assumption is not unreasonable given that the current income depends on lagged health while current OOP expenditures depend on current health. The model's dynamics would capture a large part of the correlation between income and OOP expenditures through the effect of health on these variables. Similarly the correlation between income and health status, in particular the reverse effect of income on health would be captured by the model's dynamics.

[21]As mentioned previously, the palliative component is the effect of medical care in reducing the disutility of sickness. The

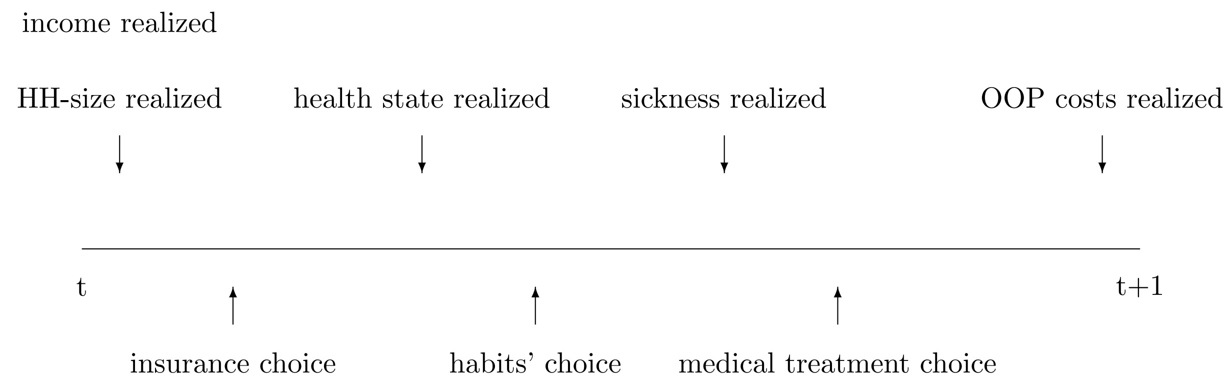


Figure 1: Timing of Per-period Decisions

member," 2-"two members," 3-"three members," 4-"four or more members," and to evolve as a multinomial logit process.

Since labor supply and savings are not the focus of the present study, these decisions are excluded from the model to keep it tractable.[22] Individuals can however save implicitly by investing in their health stock, which is the only asset in the model. Also the random utility component for each of the insurance choices can account for the dependence between insurance and employment e.g., the lack of access to a certain insurance choice because of limited employment opportunities.

2.2 The Timing Convention

The HRS surveys respondents at *two year* intervals so each time period t is assumed to be the same length. There are 324 potential choices for those under age 65 and 162 for those 65 and older. Hence each time period is divided into *three* sub-periods to reduce computational demands (fig. 1).This also provides a source of dynamic exclusion restrictions for identification (see section 5). At the start of each time period t, $t = 1, \ldots, T$, random draws associated with income and household size are realized concurrently but independently. In the *first* sub-period, the individual makes a decision about insurance.

curative component is the gross investment effect that counteracts the depreciation of the health stock through sickness in the production function. The preventive effect is the net investment aspect of medical care in the production function that leads to improvement in future health outcomes.

[22]Additionally, preliminary work indicated that inclusion of these decisions would make the model computationally intractable. The existing literature has also been unable to account for endogenous employment and savings decisions in examining the effects of Medicare. Palumbo (1999) is an exception, and examines the effects of medical expenditures on savings behavior of the elderly using a dynamic framework. However he does not analyze decisions regarding medical care or health insurance. In spite of excluding these decisions, the model is able to fit the data quite well, especially on OOP expenditures and income (see table **??** in section **??**).

Next the the health status draw is realized. If an individual dies no further decisions are made. In the *second* sub-period the individual makes choices about alcohol consumption, smoking and exercise and also derives utility from the accumulated stock of health. Next the random draw for sickness is realized. The assumption that insurance decisions precede the revelation of the current period health and sickness states prevents transitory adverse selection due to contemporaneous correlation between health shocks and insurance decisions.[23] In the *third* sub-period the individual makes the medical treatment choice *without* knowledge of the actual costs. At the end of the period the random draw for the OOP expenditures is realized and these are determined. The individual also derives utility from the composite consumption good.

2.3 Per Period Utility Functions

In the first sub-period an individual makes an insurance choice $i \in INS$.[24] The payoff from insurance choices is a switching (or transactions) cost if and only if the current choice is different from last period's choice. The utility function for insurance choice i at time t is, $U^*_{(i)}(I^i_t, A_t, \zeta) = U_{(i)}(I^i_t, A_t) + \zeta^i_t$. The deterministic component[25] of the utility function takes the form

$$U_{(i)}(I_t, A_t) = \alpha_1 \cdot 1\{I_{t-1} \neq I^i_t\} \cdot 1\{A_t < 65\} + \alpha_2 \cdot 1\{I_{t-1} \neq I^i_t\} \cdot 1\{A_t \geq 65\} \ . \qquad (15)$$

The term I^i_t denotes that the choice at time t is $i \in INS$ while $1\{x\}$ is an indicator function that takes the value 1 if the expression within the brackets is true and 0 otherwise. This specification allows the switching cost to to be different for those who are younger than 65 (i.e., α_1) from those 65 or older (i.e., α_2). These are the only utility parameters that vary by age. There is also the exclusion restriction that current income affects the current insurance choice but not vice versa. However lagged insurance may affect current income. The ζ^i_t is an additive period t and choice i specific stochastic component of the preferences for insurance (McFadden 1981, McFadden and Richter 1990). It is assumed to be IID and drawn from a multivariate extreme value distribution $(\rho_1\gamma, \frac{\pi^2\rho_1^2}{6})$, where $\gamma = 0.577$ is Euler's constant.

In the second sub-period the individual makes choices about alcohol consumption, smoking and exercise. He derives utility from each *combination*[26] $j \in J$ of these health related behaviors, and the health stock.

[23] As mentioned earlier, accounting for the endogenous supply of insurance would add considerable computational burden and also require information that is unavailable in the HRS.

[24] It should be noted that #INS is 6 if age $\leq$ 64 and #INS is 3 if age $\geq$ 65.

[25] Strictly the notation should be $U_{(i)}(I_t, I_{t-1}, A_t)$. The argument for the lagged insurance choice is omitted for brevity. A similar practice is adopted when representing the choices in the other sub-periods.

[26] There are 3 choices for alcohol consumption, 3 for smoking and 2 for exercise and so $\#J = 18$.

The utility function for each combination j depends on the current health status, the current health related choices, and the lagged health related choices, i.e., $U^*_{(j)}(H_t, a^j_t, c^j_t, e^j_t, A_t, \zeta^j_t) = U_{(j)}(H_t, a^j_t, c^j_t, e^j_t, A_t) + \zeta^j_t$. Inclusion of lagged choices allows for habit persistence and addiction in these behaviors (Becker, Grossman and Murphy 1994). The deterministic component of the utility function has the form,[27]

$$\begin{aligned} U_{(j)}(H_t, a^j_t, c^j_t, e^j_t, A_t) &= \alpha_3 \cdot H_t + \alpha_4 \cdot {H_t}^2 + \alpha_5 \cdot e^j_t + \alpha_6 \cdot 1\{e_{t-1} \neq e^j_t\} \\ &\quad + \alpha_7 \cdot c^j_t + \alpha_8 \cdot 1\{c_{t-1} \neq c^j_t\} + \alpha_9 \cdot 1\{c_{t-1} \neq c^j_t\} \cdot 1\{c_{t-1} = 1\} \cdot A_t \\ &\quad + \alpha_{10} \cdot a^j_t + \alpha_{11} \cdot 1\{a_{t-1} \neq a^j_t\} + \alpha_{12} \cdot 1\{a_{t-1} \neq a^j_t\} \cdot 1\{a_{t-1} = 1\} \cdot A_t \,. \end{aligned} \quad (16)$$

This function is quadratic in the health stock, with the parameters α_3 and α_4, to allow for risk aversion with respect to health. The net utility from the behaviors is given by α_{10} (alcohol), α_7 (smoking) and α_5 (exercise).[28] There is a switching cost associated with alcohol consumption (α_{11}), smoking (α_8) and exercise (α_6) decisions (note that a value of "1" indicates "no activity"). This represents the disutility of changing the level of consumption between periods. There is a start-up cost associated with the alcohol consumption (α_{12}) and smoking (α_9) decisions. This measures the disutility of initiating a behavior provided it did not occur in the previous period. Thus the specified start-up cost is more general than the cost of first time initiation, which is included as a special case.[29] The start-up cost for alcohol consumption and smoking is allowed to change proportionally with age. However the utility parameters *do not* vary by age. Analogous to addiction capital (Becker and Murphy 1988) this represents *non-addiction* capital that individuals may develop, i.e., as individuals age they are less likely to start smoking or consuming alcohol if they had not done so in the past.[30] The ζ^j_t is an additive period t and combination j specific stochastic component of the preferences for the three health related behaviors. It is IID with a multivariate extreme value distribution $(\rho_2\gamma, \frac{\pi^2\rho_2^2}{6})$. Though the taste shocks are independently drawn, the independence is across the *combinations* of the behaviors and not across each single behavior. This assumption restricts the level of correlation between behaviors but does allow for some level of "bundling."

In the third sub-period, $U^*_{(k)}(Y_t, P_I, OOP_t, s_t, m^k_t, HHS_t, \zeta^k_t) = U_{(k)}(Y_t, P_I, OOP_t, s_t, m^k_t, HHS_t) + \zeta^k_t$,

[27]The notation (a^j_t, c^j_t, e^j_t) represents $(a_t = a^j, c_t = c^j, e_t = e^j)$, where (a^j, c^j, e^j) are respectively the alcohol consumption, smoking and exercise choices associated with the j^{th} combination of the three behaviors.

[28]The prices for alcohol, smoking and exercise are not explicitly included in the budget constraint but are subsumed in the net utility parameters due to lack of such data in the HRS. Arcidiacono et al (2001) using data from other sources show that alcohol and cigarette prices are not significant determinants of the consumption decisions in the HRS sample.

[29]There are two choices for exercise (1-"no", 2-"yes"). Thus the start up and switching cost cannot be separately identified and are not distinguished.

[30]Including an explicit cumulative measure of past addiction (or non-addiction) would have expanded the state space considerably and added to the computational burden. Such information is also very limited in the HRS data.

is the indirect utility function for each medical care choice $k \in K$.[31] In making the medical treatment choice the individual knows the distribution but not the actual realization of the OOP cost for each choice. He also derives utility from the composite consumption commodity at this stage. The indirect utility function depends on income, insurance premium, OOP costs, current sickness, medical treatment choice and household size.[32] The deterministic component of the indirect utility function is,

$$\begin{aligned} U_{(k)}(Y_t, P_I, OOP_t, s_t, m_t^k, HHS_t) \quad = \quad & \frac{\alpha_{13} \cdot (Y_t - P_I - OOP_t)}{(HHS_t)^{\alpha_{19}}} + \frac{\alpha_{14} \cdot (Y_t - P_I - OOP_t)^2}{(HHS_t)^{2\alpha_{19}}} \\ & + \alpha_{15} \cdot 1\{m_t^k > 1\} \cdot (m_t^k)^2 + \left[\alpha_{16} \cdot (m_t^k) - \alpha_{17}\right] \cdot [1\{s_t = 2\}] \\ & + \alpha_{18} \cdot \left[\alpha_{16} \cdot (m_t^k) - \alpha_{17}\right] \cdot [1\{s_t = 3\}] \quad . \end{aligned} \tag{17}$$

The function is quadratic in consumption, with the coefficients α_{13} and α_{14}, to allow for risk aversion. The model extends previous empirical research by allowing for risk aversion in two dimensions, i.e., health and aggregate consumption.[33] The individual consumption is the household consumption normalized non-linearly, via α_{19}, by household size. Hence the model allows for the public good aspect of within household consumption. There is a monetary equivalent of a psychological cost for getting medical care, α_{15}, e.g., due to waiting time or scheduling costs, that increases convexly with medical utilization. The last set of terms on the second line and those on the third line of eq. 17 represent the net difference between the disutility of sickness and its mitigation through medical treatment. The monetary equivalent of the disutility of moderate sickness (i.e., $s_t = 2$) without medical treatment is $[-\alpha_{17}]$. Medical care sought by a moderately sick individual mitigates the disutility of sickness by the amount α_{16}. Similarly the disutility of high sickness (i.e., $s_t = 3$) is $[-\alpha_{18} \cdot \alpha_{17}]$, where α_{18} is a proportionality constant. This disutility is

[31] The choices for treatment are: 1-low, 2-moderate, 3-high. Thus $\#K = 3$.

[32] The household size variable is included because only a household measure of non-labor income is available in the HRS. To calculate individual consumption the household consumption is normalized by the household size. Alternatively, the household income could be directly normalized by the household size providing a measure of individual income. However that would disallow *self-insurance* within households.

[33] A quadratic utility specification may suggest limited risk aversion. However the Arrow-Pratt coefficient of absolute risk aversion for the aggregate consumption commodity is $\frac{-[\frac{2\alpha_{14}}{(HHS_t)^{2\alpha_{19}}}]}{[\frac{2\alpha_{14} X_t}{(HHS_t)^{2\alpha_{19}}} + \frac{\alpha_{13}}{(HHS_t)^{\alpha_{19}}}]}$. Since this depends on the parameters, the model can flexibly generate varying levels of risk aversion conditional on the estimates, in particular a level sufficient to make insurance valuable. In fact the model does perform well in replicating the life cycle income and OOP cost profiles (see section **??** below). The same is true for risk aversion regarding the health commodity. Experimentation with more general forms of risk preferences, e.g., a CRRA utility function, made simulating from the model computationally burdensome, which made the optimization of the likelihood function unfeasible. Details are provided for a related model in Khwaja (2001). The model does not allow for "prudence," that generates precautionary saving which is important in matching life cycle savings behavior (Hubbard, Skinner and Zeldes 1995). This implies the model may not account for any precautionary value of Medicare insurance to individuals. However savings are not the focus of the current research.

mitigated by the amount $[\alpha_{18} \cdot \alpha_{16}]$ in the case of high sickness. The α_{16} represents the pure consumption or mitigative component of medical care. The ζ_t^k is an additive period t and choice k specific stochastic component of the preferences for medical care. It is IID with a multivariate extreme value distribution $(\rho_3\gamma, \frac{\pi^2\rho_3^2}{6})$.

The stochastic components of the preferences ζ_t^i, ζ_t^j,and ζ_t^k represent information associated with a particular choice respectively of insurance, health related behaviors and medical treatment that is known to the individual but unknown to the econometrician. The distributional assumption implies that choice probabilities have the familiar multinomial logit closed form expression (Rust 1987, 1994a, 1994b). However the model does not suffer from the usual Independence of Irrelevant Alternatives (IIA) limitation because of its dynamic structure (Rust 1994a, p. 139, 1994b, pp. 3107-3108). It represents information about each medical care choice that is known to the individual but unknown to the econometrician. Hence the probability of the k^{th} choice takes the multinomial logit form.

2.4 The Dynamic Programming Problem

An individual's choice b in any sub-period from any set B in time period t is denoted by $d_t(b)$, where

$$d_t(b) = \begin{cases} 1 & \text{if the } b^{th} \text{ alternative is chosen from set } B \text{ in period } t \\ 0 & \text{otherwise.} \end{cases}$$

The choices are mutually exclusive, i.e., $\sum_{b=1}^{\#B} d_t(b) = 1$. The value function, $V(Z_t)$, is defined as the maximal expected present value of lifetime utility at age t when the individual is in state Z_t, where $Z_t = \{H_{t-1}, I_{t-1}, a_{t-1}, c_{t-1}, e_{t-1}, s_{t-1}, m_{t-1}, HHS_t\}$ is the predetermined component of the state space[34] for the insurance choices. It contains information known up to time t relevant for the insurance decision, including past endogenous transitions, random variable realizations, and decisions. Given the discount factor, β, the value function is,

$$\begin{aligned} V(Z_t) &= \max_{\{d_t(i), d_t(j), d_t(k)\}_{t=1}^T} E\left[\sum_{\tau=t}^{T} \beta^{\tau-t} \sum_{i=1}^{\#INS} \left\{N + U^*_{(i)}(I^i_\tau, \zeta^i_t) \cdot d_\tau(i) \mid Z_\tau\right\}\right] \qquad (18) \\ \text{where } N &= \sum_{j=1}^{18} \left[M + U^*_{(j)}(H_\tau, a^j_\tau, c^j_\tau, e^j_\tau, \zeta^j_t) \cdot d_\tau(j) \mid \widetilde{Z}_\tau\right] \\ \text{and } M &= \left(\sum_{k=1}^{3} \left(E_{\epsilon^{oop}}[U^*_{(k)}(Y_\tau, P_I, OOP_\tau, s_\tau, m^k_\tau, \zeta^k_t)] \cdot d_\tau(k)\right) \mid \overline{Z}_\tau\right). \end{aligned}$$

[34]Strictly speaking the state space is the Cartesian product of the discrete possibilities for each of the outcomes and choices. This notation is adopted to simplify the exposition.

The $\widetilde{Z}_t = \{H_t, I_t, a_{t-1}, c_{t-1}, e_{t-1}, HHS_t\}$ is the predetermined component of the state space relevant for the choice of health related behaviors and $\overline{Z}_t = \{H_t, I_t, a_t, c_t, e_t, s_t, HHS_t\}$ is the predetermined component of state space relevant for the medical treatment choices. At every age $t = 1, \ldots, T$ an individual makes decisions about health insurance ($d_t(i)$), health related behaviors ($d_t(j)$), and medical treatment ($d_t(k)$) that maximize the expected present value of his remaining lifetime rewards (18) subject to the constraints (2), (3), (4), (5) , (6) and (7) discussed earlier. The Bellman equations representing the maximization problem are formulated by defining the separate Bellman equations and the associated alternative specific value functions, i.e., $V^i(Z_t)$, $V^j(\widetilde{Z}_t)$, $V^k(\overline{Z}_t)$, for the choices related to health insurance, health related behaviors and medical treatment respectively.

The Value function for the insurance choices is,

$$V(Z_t) \quad = \quad E\left[\max_{i \in INS}\left\{V^i(Z_t)\right\}\right] \tag{19}$$

and the associated Bellman equation is,

$$\begin{aligned} V^i(Z_t) &= U_{(i)}(I_t^i, A_t) + \zeta_t^i \\ &\quad + \left[\Sigma_{q=1}^{6} \Pr(H_t = q) \cdot [\widetilde{V}(\widetilde{Z}_t) \mid d_t(i), H_t = q]\right], \text{ if } t \leq T \\ V^*(Z_{T+1}) &= U^*_{(I)}(I_{T+1}, \zeta^I_{T+1}) \equiv 0\ . \end{aligned} \tag{20}$$

The Bellman equation (20) represents the health insurance choice by an individual given information about his lagged states of health and sickness, the lagged choices about insurance, health related behaviors and medical care, and the current household size. In making this choice the individual considers the current payoff from each insurance option along with the the future expected maximum return from the next subperiod's optimal choice of health related behaviors conditional on his current insurance choice. Through this explicit dependence of the current insurance choice on the past as well as the future expected individual behaviors and outcomes the model accounts for the endogeneity of the demand for health insurance (Phelps 1973). A terminal value function represents the utility after the final time period T and is parameterized as a function of the health and sickness states in that period.

The Value function for the health related behavior choices, i.e., alcohol consumption, smoking and exercise is,

$$\widetilde{V}(\widetilde{Z}_t) = E\left[\max_{j \in J}\left\{V^j(\widetilde{Z}_t)\right\}\right] \tag{21}$$

and the associated Bellman equation is,

$$\begin{aligned} V^j(\widetilde{Z}_t) &= U_{(j)}(H_t, a_t^j, c_t^j, e_t^j, A_t) + \zeta_t^j \\ &\quad + \left[\Sigma_{q=1}^3 \Pr(s_t = q) \cdot [\overline{V}(\overline{Z}_t) \mid d_t(j), s_t = q]\right], \text{ if } t \leq T \\ V^j(\widetilde{Z}_{T+1}) &= U^*_{(j)}(H_{T+1}, a_{T+1}^j, c_{T+1}^j, e_{T+1}^j, A_{T+1}, \zeta_{T+1}^j) \equiv 0 \, . \end{aligned} \tag{22}$$

The Bellman equation (22) states that an individual with knowledge of his current health state, current insurance and lagged choice of health related behaviors, and current household size makes a decision about current health related behaviors. In this choice he considers the current payoff from the choice of health related behaviors and also the future expected maximum return from the next sub-period's optimal medical treatment choice conditional on his current health related behaviors. The explicit dependence of current choices on past and expected future behaviors and outcomes helps account for the endogeneity of health related behaviors.

The Value function for the medical treatment choice is,

$$\overline{V}(\overline{Z}_t) = E\left[\max_{k \in K}\left\{V^k(\overline{Z}_t)\right\}\right] \tag{23}$$

and the associated Bellman equation is,

$$\begin{aligned} V^k(\overline{Z}_t) &= \Pr(OOP_t = 0) \cdot U_{(k)}(Y_t, P_I, s_t, m_t^k, HHS_t) + \\ &\quad (1 - \Pr(OOP_t = 0)) \cdot E_{\epsilon^{oop}}[U_{(k)}(Y_t, P_I, OOP_t, s_t, m_t^k, HHS_t)] + \zeta_t^k \\ &\quad + \beta\left[\Sigma_{q=1}^4 \Pr(HHS_{t+1} = q) \cdot [V(Z_{t+1}) \mid d_t(k), HHS_{t+1} = q]\right], \text{ if } t \leq T \\ V^k(\overline{Z}_{T+1}) &= U^*_{(k)}(Y_{T+1}, P_I, OOP_{T+1}, s_{T+1}, m_{T+1}^k, HHS_{T+1}, \zeta_{T+1}^k) \equiv 0 \, . \end{aligned} \tag{24}$$

The Bellman equation (24) states that an individual with knowledge of his current health and sickness states, current choices of insurance and health related behaviors, and current household size makes a decision about medical care. In doing so he considers the current payoff from the medical treatment choice as well as the future discounted expected maximum return from the next period's optimal insurance choice conditional on his current treatment choice. This accounts for the endogeneity of the medical treatment choice. The Bellman equation involves expectations ($E_{\epsilon^{oop}}$) over the distribution of ϵ^{oop} that will be realized later in the *same* period, since the individual makes the medical treatment choice with knowledge of the distribution but not the exact OOP cost of treatment.

More generally, the expectations operator $E[\cdot]$ in equations (19), (21) and (23) signifies that expectations are taken over the distribution of the stochastic elements $\{\zeta_t^i, \epsilon_t^H\}$, $\{\zeta_t^j, \epsilon_t^s\}$, and $\{\zeta_t^k, \epsilon_t^{oop}, \epsilon_{t+1}^y, \epsilon_{t+1}^{HHS}\}$

respectively. In the empirical analysis the decisions of the individuals are modelled from ages 22 to 80 with a terminal value function representing utility after age 80.[35] The model has no closed form solution and is solved using numerical techniques of backward recursion.[36]

3 Data

The data comes from the Health and Retirement Study (HRS). The data and information about it is available from the following website, i.e. "http://www.umich.edu/ hrswww". This is a national panel study with an oversampling of Blacks, Hispanics and residents of Florida. In estimating the model I use data from the first four waves that were sampled in 1992, 1994, 1996 and 1998. Data has been collected from individuals who were between the ages of 51 and 61 in 1992, i.e. in the 1931-41 birth cohort. Each individual was surveyed every *two years*. Information has been collected on the following subjects among others, demographics, health, employment, disabilities, retirement plans, income and insurance. I use a longitudinal sample of 9627 observations on 3671 males from this data set for my research. This is a subset of individuals for whom there is complete information available regarding the variables of interest for at least two waves.

A summary of the data is provided in table 15. All values are in 1991 prices. An interesting fact to note is that according to the sample information even those who claim to be uninsured report an average two yearly insurance cost of $215.0. One possible explanation for this could be that some of these so called uninsured individuals were on a type of insurance plan that was not mentioned on the questionnaire. Thus the individuals whom I categorize as "uninsured" may be on some unknown insurance plan. Alternatively there could be some coding errors in the data. A third possibility is that some of the respondents where insured for part of the period the question addressed. For the purposes of estimating this model I ignore this information and assume that those who were "uninsured" paid no insurance premium. From tables 16 - 31 it is clear that there is a high degree of persistence in the health states and the choices of habits and insurance. Figure 2 provides information about the age distribution of my sample. Not unexpectedly the bulk of the distribution lies between the ages of 50 and 70 (inclusive). One implication of this fact is that I don't expect the model to do very well in matching the states and behaviors observed for the individuals in the tails of the age distribution in my sample (i.e. for those under 50 and those over 70). However the

[35] This start date is adopted because individuals typically leave college and start making independent decisions about health insurance and medical care at age 22. The terminal decision making age, T, is assumed to be 80 to reduce the computational burden.

[36] Khwaja (2001) provides details for a related model.

model should be able to match representative data for the ages in question (i.e. under 50 and over 70) obtained from some other source. In fact this would be an important test of the model. It is imperative for a structural model to be able to extrapolate beyond the data it is fit to if one has to have any faith in the model. Indeed I perform such a test and report the results in section 6.

I would also like to draw attention to one idiosyncracy of the HRS data. In the period that an individual is observed to die, the HRS data follows the convention of assigning the individual to a separate "single-member" household. Thus in the period of an individual's death I have two pieces of information about the individual, i.e. his health state and his household state. This information will be important later in the construction of the likelihood and consequently in the estimation.

The data used in the estimation is constructed in the following manner. The health stock of an individual was the self-reported health status[37] (for those who are alive). Thus the health states are enumerated to be excellent, very good, good, fair, poor or dead. There are six insurance states for those under 65 years of age, i.e. uninsured, personal insurance, group insurance, VA/Champus plans, Group/Personal insurance, Group/VA/Champus plans. These states were determined from the choices (and their combinations) observed in the data. Similarly there are 3 possible insurance states for those 65 and over, i.e. Medicare, Medicare/Group, Medicare/Medigap/Other-Personal insurance. I do not include Medicaid as an insurance state as less than 1% of the males in the sub-sample were observed to be on Medicaid and hence there is a paucity of information for such individuals.

The medical treatment variable was constructed by finding the total of all the visits to hospitals and nursing homes, the number of nights spent in hospitals and nursing homes and the number of other visits to doctors. The total was then discretized in to three categories representing "low", "medium" and "high" number of visits by splitting the distribution of the data on total visits in to three (almost equal) parts. The variable for smoking was constructed by determining the number of cigarettes smoked per day by an individual and then discretizing it to represent those who did not smoke at all, those who smoked less than a pack a day and those who smoked a pack or more a day.

The variable for alcohol consumption was constructed by calculating the number of drinks for an individual per day. The discretization was done to represent those individuals who don't drink at all. Those who reported having 1 or fewer drinks a day, and those who reported having more than a drink a day. The exercise variable was constructed using a "yes" or "no" response to question about whether individuals get "rigorous physical exercise" three or more times a week. Rigorous exercise included not

[37]This is a controversial assumption however there exists empirical research that has employed the same notion of health stock, e.g. Benitez-Silva et. al. (1999).

just sports but also heavy household work.

The income and out of pocket expenditure used in the estimation are the empirical household counterparts. The information on insurance premiums is very sketchy in the HRS data. The values of the insurance premiums were constructed by averaging over the household insurance payments after pooling the individuals by each type of insurance plan considered in the estimation. The average was taken over the number of individuals in each pool. This is admittedly a very rough estimate of the actual costs for health insurance and was truly a last resort given the nature of the data set.

As mentioned earlier the respondents were sampled every two years. Thus there are individuals in the data who are observed at a sequence of odd ages and others observed at even ages. To reduce the computational burden in estimating the model I modified the data so that every individual was coded to be observed at an even age. Care was however taken that the insurance states of the individuals were not affected, i.e. those who were observed to be 65 or older in the raw data and thus had access to Medicare were not coded to an age below 65 and vice versa for those under 64. One implication of this is that the estimates of the model using data in its current form would not be very precise with respect to any age effects. However this is not of prime concern as I have kept age effects to a bare minimum in constructing the model. It should also be noted that at a future date this restriction on the data could be removed and the model re-estimated using the original data. This will only entail a larger computational burden and pose no other methodological problems.

Figures 3 - 10 provide information about the general trends observed over the life-cycle for the health states and the choices. As seen in figure 3, average health status declines with age. It is also seen in figure 4 that mean household size follows an inverted U-shape. Average level of medical treatment rises with age while mean alcohol consumption is also an inverted U from figures 5 and 6 respectively. Figure 7 shows that mean smoking levels decline with age while mean exercise is relatively flat over the life cycle as seen in figure 8. Figure 9 shows that mean of log income declines with age though it is not clear what the trend is like in the initial years due to the lack of data. The mean of log of out of pocket expenditures is found to be increasing with age from figure 10.

4 Estimation

4.1 The Likelihood Function

The model's parameters are estimated by maximizing a likelihood function that nests the solution to the dynamic programming problem (see e.g., Rust 1987, 1994a, 1994b). The HRS provides data for a sample of $n = 1, \ldots, N$ individuals on the sequence of insurance, alcohol consumption, smoking, exercise and medical treatment choices, and health outcomes, household size, income and OOP costs, i.e., $\{I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, H_{n,t}, HHS_{n,t}, Y_{n,t}, OOP_{n,t}\}_{t=t_{n,0}}^{T_n}$, where $t_{n,0}$ is the first and $T_n \leq T$ the last observation for individual n. These variables are observed if an individual is alive, i.e., $H_t > 1$. However at the time of death only the health outcome and household size are observed. The sickness variable is not observed in any time period. Hence it is treated as a latent variable and integrated out in the likelihood function. The solution of the model provides the joint probability of observing the choices and outcomes for each individual n at time t conditional on the pre-determined state and the model's parameters, Θ. Suppressing the subscript n and parameters Θ, the likelihood function for each individual is,

$$\begin{aligned}\mathcal{L}_n(\Theta) = \Pr[H_{t_0}, I_{t_0}, a_{t_0}, c_{t_0}, e_{t_0}, HHS_{t_0} \mid (H_{t_0} > 1)] \cdot \prod_{t=t_0+1}^{T} \big(P[Y_t]\big)^{1\{H_t>1\}} \cdot \Pr[HHS_t] \\ \cdot \sum_{q=1}^{3} \Big\{ \Pr[s_{t-1} = q] \big(\Pr[I_t \mid Z_t^q] \big)^{1\{H_t>1\}} \Pr[H_t^q] \Big\} \cdot \big(\Pr[a_t, c_t, e_t \mid \widetilde{Z}_t] \big)^{1\{H_t>1\}} \\ \cdot \sum_{q=1}^{3} \Big\{ \Pr[s_t = q] \big(\Pr[m_t \mid \overline{Z}_t^q] \big)^{1\{H_t>1\}} \Big\} \cdot \big(P[OOP_t] \big)^{1\{H_t>1\}} \quad .\end{aligned} \tag{25}$$

The sample likelihood is $\mathcal{L}(\Theta) = \prod_{n=1}^{N} \mathcal{L}_n(\Theta)$.

The first term in (25) is the probability of observing the individual in the sample conditional on survival. The second term is the likelihood contribution of the individual's income. A term in (25) that has the indicator function, $1\{H_t > 1\}$, as its exponent is a likelihood contribution relevant only when an individual is alive. The third term is the likelihood contribution of the household size. The terms in the first set of brackets (on the second line of (25)) are the contributions of the individual's insurance choice and health outcome. The calculation of these terms involves integrating over the distribution of the latent sickness variable for the last period. The Z_t^q is state Z_t with the superscript denoting that the value of lagged sickness is q, i.e., $s_{t-1} = q$. The $\overline{Z}_t^q$ and $\Pr[H_t^q]$ (eq. 9) are similarly defined. The next term is the likelihood contribution of the choice of health related behaviors. The second set of brackets (on the third line of (25)) contains the likelihood contribution of the medical treatment choice. This involves integrating over the distribution of the latent sickness variable for the current period. The final term is the likelihood contribution of the OOP expenditure variable.

Maximum likelihood estimation does not impose the requirement that components associated with OOP costs, income and household size be included in the likelihood function (see e.g., Eckstein and Wolpin 1987, p. 588-589). In particular if OOP costs were observed regardless of treatment in all *four* waves then parameters of its distribution could be consistently estimated directly from the data. However the OOP cost data suffer from selection on medical utilization and insurance coverage.[38] Hence the OOP cost parameters are estimated jointly with the other parameters of the model. The model provides a means of correcting for selection in estimation (Heckman 1979).[39] Thus even in the absence of data on the OOP costs of all individuals[40] the parameters of the OOP cost distribution can be estimated consistently, in fact joint estimation provides efficiency. It is important to estimate the OOP cost parameters consistently because these affect estimation of other utility and belief parameters in the model. For the same reason the parameters of other outcome variables, e.g., health, are estimated jointly with the utility parameters.

4.2 Unobserved Heterogeneity and Measurement Error

Typically there is serially correlated unobserved (to the econometrician) heterogeneity between individuals which if ignored can bias estimates. Therefore the estimation procedure controls for unobserved heterogeneity through the well known procedure proposed by Heckman and Singer (1984).[41] Specifically, suppose there are $l = 1, \ldots, L$ types of individuals with the proportion of the l^{th} type in the population given by π_l. The model is modified to allow for unobservable differences in the health technology (i.e., $\{\eta_{10,q,l}\}_{q=2}^{6}$), preferences for health (i.e., $\alpha_{3,l}$) and income earning ability[42] (i.e., $\kappa_{1,l}$). It is assumed that preferences and technology are common to individuals of a given type, and they know their type. Hence conditional on an individual's observed characteristics and unobserved type, the unobserved stochastic components of preferences, health technology and income process are assumed to be *serially uncorrelated*.

Let the population probability of being a particular type, π_l, be defined as a multinomial logit in education and race with parameters λ. This allows the unobserved type of an individual to be correlated with education and race.[43] The likelihood is modified to integrate out the probability of an individual being a particular type and follows a mixture distribution given by $\mathcal{L}_n(\overline{\Theta}) = \sum_{l=1}^{L} \pi_l \cdot \mathcal{L}_{n,l}(\Theta_l, \lambda)$. The sample likelihood is $\mathcal{L}(\overline{\Theta}) = \prod_{n=1}^{N} \mathcal{L}_n(\overline{\Theta})$. This can be maximized with respect to the parameters $\overline{\Theta}$, where $\overline{\Theta} \equiv (\{\Theta_l\}_{l=1}^{L}, \lambda)$. The estimation also allows for additive measurement errors (see e.g., Wolpin 1987) in income $\epsilon_\kappa \overset{\text{IID}}{\sim} N(0, \sigma_\kappa^2)$ and in OOP expenditures $\epsilon_\mu \overset{\text{IID}}{\sim} N(0, \sigma_\mu^2)$.

[38]The OOP cost data is limited to waves 2 and 3. It could also be zero even in the case of treatment depending on the insurance coverage. This selection problem is similar to that of estimating a wage equation in a job search or labor participation model.

[39]The mixed-continuous distribution of OOP costs explicitly allows for these to be unobserved in the absence of treatment, and to potentially be zero even conditional on treatment due to generous insurance coverage. The estimation also allows for measurement error in the data.

[40]There are 4276 observations on OOP costs (table **??**) suggesting no cause for concern about statistical power.

[41]Applications of this procedure are, e.g., Keane and Wolpin 1997, Gilleskie 1998, and Cameron and Heckman 2001.

[42]The constant in the income function is allowed to differ by type.

[43]Lleras-Muney (2005) finds that there is a substantial effect of education on mortality in the U.S.

4.3 Initial Conditions and Dynamic Selection

Incorporating unobserved heterogeneity in the model implies that the entire sequence of choices and outcomes of each individual is correlated with those in the initial period and there is dynamic selection in behaviors. An initial conditions problem exists as data is unavailable for the individuals at younger ages.[44] Mortality is endogenous to behaviors so there is also selection on survival.

To correct for initial conditions and selection, the distribution of initial choices and outcomes conditional on unobserved heterogeneity is estimated jointly with the distribution of unobserved effects (Heckman 1981).[45] To illustrate, assume that all the individuals are first observed at the age of 50 or $t = 15$.[46] Suppressing the indices n for individuals and l for types, and the parameters $\overline{\Theta}$, the first term in the likelihood function (25) is modified to $\Pr[H_{15}, I_{15}, a_{15}, c_{15}, e_{15}, HHS_{15} \mid (H_{15} > 1)]$. This term corrects for the initial conditions and dynamic selection. It can be calculated by simulating choices and outcomes for a sample of individuals between the ages of 22 (i.e., $t = 1$) and 50 (i.e., $t = 15$) conditional on type from the model.[47] In doing so it is assumed that the state of health at $t = 1$ is determined by the health production technology (eq. 9) with the only non-zero regressor being the constant $\eta_{10,q}$, $q = 1, \ldots, 6$. It is also assumed that at $t = 0$, the values of the health, alcohol, smoking, exercise, sickness and medical care variables are zero. The initial household size is determined similarly through the household process. There is also an assumption that at $t = 1$ there is no start-up or switching cost for insurance or health related behaviors. The simulated sample can be used to obtain, $\Pr^*[H_{15}, I_{15}, a_{15}, c_{15}, e_{15}, HHS_{15} \mid (H_{15} > 1)]$, where "$*$" denotes that these simulated probabilities may not be "continuous," i.e., potentially a hyper-cell associated with a given element of the state space at $t = 15$ may contain no simulated data.

To obtain continuous initial probabilities a kernel smoothing procedure (Aitchison and Aitken 1976) is adopted.[48] Define the set of cell indices[49] as $C = \{\overline{c} = (c_1, c_2, c_3, c_4, c_5, c_6) : c_1, c_2 = 1, ..., 6; c_3, c_4 = 1, 2, 3; c_5 = 1, 2, c_5, c_6 = 1, ..., 4\}$. The smoothing of cell probabilities is done by non-parametrically estimating a linear combination of the simulated cell frequencies. The smoothed cell probabilities denoted by

[44] The HRS respondents were aged between 51 and 61 in 1991-92, however spouses of any age are included.

[45] Wooldridge (2000) solves the initial conditions problem based on the converse approach of specifying a distribution for unobserved heterogeneity conditional on initial observations. There is also a literature that relies on lagged regressors as instruments to solve the initial conditions problem, e.g., Blundell and Bond (1998).

[46] The problem can be generalized easily to include different ages at which an individual's history becomes observable.

[47] Olley and Pakes (1996) provide a semi-parametric procedure to correct for similar selection that does not rely on explicitly solving and simulating from a model. Their method however requires input (or investment) levels to be strictly positive in each period. Hence it is not applicable in this case.

[48] This procedure came to my attention through Gilleskie (1994).

[49] These dimensions follow from 6 possible health and insurance states, 3 possible alcohol consumption and smoking choices, 2 possible exercise choices and 4 possible states for the household size.

$\widetilde{p_{\overline{c}}} \equiv \widetilde{\Pr}[H_{15}, I_{15}, a_{15}, c_{15}, e_{15}, HHS_{15} \mid (H_{15} > 1)]$ are calculated as,

$$\widetilde{p_{\overline{c}}} = \sum_{\overline{l} \in C} a_{\overline{c}\overline{l}}(\omega) \cdot \widehat{p_{\overline{l}}} \ \ .$$

The $\widehat{p_{\overline{l}}} \equiv \Pr^*[H_{15}, I_{15}, a_{15}, c_{15}, e_{15}, HHS_{15} \mid (H_{15} > 1)]$ is the simulated frequency of cell $\overline{l}$. The $a_{\overline{c}\overline{l}}(\omega)$ is the weight attached to cell $\overline{l}$ in estimating the probability of cell $\overline{c}$. This depends on the parameters $\omega = (\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6)$ that represent the bandwidth in each direction of the "hyper-rows" (or the dimensions that define each hyper-cell). The smoothing technique assigns a smaller weight to the cells that are further away. The rate of decline is determined by the parameter ω. Thus if $\omega = 0$ then there is no smoothing while if $\omega = 1$ there is extreme smoothing and every cell gets equal probability. The following standard assumptions are made in the smoothing procedure: (i) $a_{\overline{c}\overline{l}} \geq 0 \ \forall \ \overline{c}, \overline{l}$ (ii) $a_{\overline{c}\overline{c}} \geq a_{\overline{c}\overline{l}} \ \forall \ \overline{c}, \overline{l}$ and (iii) $\sum_{\overline{c} \in C} a_{\overline{c}\overline{l}} = 1 \ \forall \ \overline{l} \in C$. In particular the weights according to the Aitchison and Aitken method are,

$$a_{\overline{c}\overline{l}}(\omega) = \omega_1^{|c_1 - l_1|} \cdot ... \cdot \omega_6^{|c_6 - l_6|} \cdot D(\omega)$$

where

$$D(\omega) = \left\{ \sum_{\overline{c} \in C} \omega_1^{|c_1 - l_1|} \cdot ... \cdot \omega_6^{|c_6 - l_6|} \right\}^{-1} \quad .$$

The $\widetilde{p_{\overline{c}}}$ is substituted in place of the first term in (25) to obtain a "smoothened" sample likelihood that is maximized using a Newton-Raphson hill climbing algorithm employing the BHHH method (Berndt et al 1974) to calculate the Hessian. The BHHH method is also used to compute the standard errors.

5 Identification

The model is identified[50] through a combination of exogenous variation in the data, economic assumptions about the individual's optimization problem, and parametric assumptions (see e.g., Rust 1994a, 1994b). Since estimation involves solving a dynamic programming problem it necessitates making parametric assumptions about the stochastic terms in the model. One important source of identification is the *exogenous* change in the set of insurance choices through eligibility for Medicare at age 65.[51] The identification assumption is that eligibility for Medicare at age 65 is independent of past health outcomes or health related behaviors of an individual.[52] Card, Dobkin and Maestas (2004) and Decker (2005) employ a similar identification strategy. This strategy is valid if the model's parameters *close* to age 65 are the *same* as at other

[50]The formal conditions for identification are the same as required for consistency in maximum likelihood estimation (see e.g., Newey and McFadden 1994). Unlike linear models there are no analytical conditions that can be used to establish identification (Eckstein and Wolpin 1989, p. 588). Hurwicz (1950) provides an excellent discussion of identification in structural non-linear models.

[51]Recall, there are 6 insurance choices under age 65 and 3 thereafter. This identification assumption also relies on *not all* individuals retiring at age 65, i.e., the retirement profile is smoother than eligibility for Medicare.

[52]The exceptions to this eligibility criterion are described in footnote **??**. These do not adversely affect the identification assumption.

ages.[53] Hence the parsimonious specification of the model with respect to age is a source of identification through the local discontinuity in insurance coverage at age 65 (Hahn, Todd and Van der Klaauw 2001).

Another identification assumption is that the random components of preferences, and the income, OOP cost, health, sickness and household size processes are IID *conditional* on the time invariant unobserved type (Heckman and Singer 1984). Identification also relies on exogenous variation through evolution of age and household size. These are assumed to be uncorrelated with the unobserved type. The evolution of age is deterministic in the model. Age directly affects income (eq. 4), OOP costs (eq. 5), sickness (eq. 6), and household size (eq. 7). The evolution of household size is stochastic but exogenous to the model. The household size directly affects OOP costs. The combined pre-determined variation in income, OOP costs and sickness provides variation in the respective decisions to purchase health insurance, the health related behaviors and medical utilization. The pre-determined variation in the individual behaviors in turn provides variation in the health outcomes. Identification also depends on the economic assumptions about the individual's optimization problem. These assumptions provide exclusion restrictions, e.g., the health stock directly affects income (eq. 4) and OOP cost (eq. 5) but income and OOP costs are excluded from the health production function (eq. 2). The timing convention is another economic assumption (section 2.2) providing a series of dynamic exclusion restrictions, e.g., the last period's medical treatment decision affects the current insurance choice, conversely the current insurance choice affects the current medical treatment decision.

The rest of this section describes the identification of the parameters in greater detail. In order to be able to interpret the utility parameters as monetary equivalents, and because it cannot be identified separately, the parameter, α_{13}, on the linear term for the composite consumption commodity (eq. 17) is fixed to be 1. All the utility parameters are normalized to zero in the dead state i.e., $H_t = 1$. The parameters, α_1 and α_2, for the switching cost of insurance in (15) are identified primarily through the persistence in the insurance choices (see data in tables **??** and **??** in section **??**) and the variation in the insurance choice set due to age. Similarly, the switching cost parameters for alcohol consumption (α_{11}), smoking (α_8) and exercise[54] (α_6) are identified off the persistence in the respective choices (see e.g., data on smoking in table **??** in section **??**).

The utility parameters associated with the health stock, α_3 and α_4, and alcohol consumption (α_{10}),

[53] The only parameters that change locally at 65 are the switching cost for insurance, and coefficients related to the dummy variables indicating insurance status in the OOP cost function. Insurance switching cost is α_1 under age 65 and α_2 at 65 or older (eq. 15). Insurance status parameters in the OOP cost processes are shown in tables **??** and **??**, which are self-explanatory. These changes reflect the exogenous change in the insurance choice set at age 65. Alternatively put, identification relies on the absence of any unrestricted age effects in the model, i.e., there are gradual changes over the life cycle in the behaviors and outcomes described by the model without any changes in the parameters with the exceptions described above.

[54] There are two choices for exercise (1-"no", 2-"yes"). Thus the start up and switching cost cannot be separately identified and are not distinguished in the specification.

Table 1: Demand for Medical Care Conditional on Change in Health: HRS Data

Change in Health Status $[H_t - H_{t-1}]$	Medical Care (row %)		
	Low	Moderate	High
-4	11.11	33.33	55.56
-3	18.37	14.29	67.35
-2	37.01	23.70	39.29
-1	39.46	32.71	27.83
0	41.47	33.82	24.72
1	40.95	35.66	23.39
2	43.62	35.64	20.74
3	50.00	20.00	30.00
4	0.00	40.00	60.00

Source: Sample used for estimation. See table **??** for more information.

smoking (α_7), and exercise (α_5) (eq. 16) are identified off the variation in pre-determined health status, and alcohol consumption, smoking and exercise choices. The entails the further assumption that conditional on unobserved type the random component of the choice of health related behaviors is IID. In addition the payoffs from alcohol consumption, smoking and exercise are normalized to zero when the choice of the respective behaviors is "1," i.e., "no activity." The start-up costs for alcohol consumption (α_{12}) and smoking (α_9) (eq. 16) are identified primarily through variation over age in the individual decisions to (re)start alcohol consumption or smoking conditional on not having done so in the last period.

The parameters α_{13} and α_{14} associated with utility from the composite consumption commodity (eq. 17) are identified primarily through variation in age, insurance premiums and household size, along with variation in pre-determined income, health and insurance status conditional on time invariant unobserved heterogeneity. The identification of the transactions cost for getting medical care (α_{15}), the mitigative component of medical care (α_{16}), the disutility from sickness (α_{17}), and the proportionality factor for 'high' sickness (α_{18}) relies primarily on the variation in the data conditional on the assumption about the underlying distribution of sickness. To illustrate, α_{15} is identified off the individuals who seek *low*[55] amount of treatment (i.e., are averse to seeking medical care), while α_{16} is identified off the individuals who seek *moderate* or *high* medical care (i.e., have a greater desire to seek medical care), conditional on their health getting *worse* between the last and current period (see table 1) and the underlying distribution of sickness.[56] Additionally, the transactions cost and mitigative component of medical care are normalized to zero when the choice of medical care is "1," i.e., "low." The variation in the household size helps to identify α_{19}. The discount factor is not estimated and is fixed to be 0.95 annually.

[55]The medical utilization choices are: 1-low, 2-moderate, and 3-high.

[56]The latter assumption is used in forming the terms in brackets on the second line of the likelihood function (25).

Table 2: Change in Health State Conditional on Medical Treatment: HRS Data

Lagged Medical Care(m_{t-1})	Change in Health Status $[H_t - H_{t-1}]$ (row %)									
	-5	-4	-3	-2	-1	0	1	2	3	4
Low	0.16	0.67	1.61	6.34	23.4	49.23	14.93	3.35	0.32	0.00
Moderate	0.32	0.69	1.38	5.32	21.16	51.09	16.48	3.35	0.16	0.05
High	0.15	0.51	1.98	7.27	21.51	45.89	18.8	2.94	0.66	0.29

Source: Sample used for estimation. See table **??** for more information.

The health production function (9) has a multinomial logit form so the coefficients differ by outcome and the parameters for the outcome $q = 1$, i.e., "dead," are normalized to be $\underline{\eta}_1' \equiv (\eta_{1,1}, \eta_{2,1}, 0, \ldots, 0)$ for identification. The identification assumption is that the current period health shock is IID conditional on the time invariant unobserved type. The parameters $\eta_{1,q}$ and $\eta_{2,q}$ are identified off the persistence and changes in health status over time (see data in table **??** in section **??**). The effects of lagged sickness on current health, $\eta_{3,q}$, the curative component of medical care, $\eta_{4,q}$, the proportionality factor for high sickness, $\eta_{5,q}$, and the preventive component of medical care, $\eta_{9,q}$, are primarily identified off variation in the data and the assumption about the underlying sickness distribution.[57] For example, $\eta_{4,q}$, is identified off the individuals whose health is *unchanged* or *worsens*, while $\eta_{9,q}$ is identified off those whose health *improves*, between the last and current period after seeking medical care (see table 2) conditional on the underlying distribution of sickness. The effects of alcohol consumption, $\eta_{6,q}$, smoking, $\eta_{7,q}$, and exercise, $\eta_{8,q}$, on current health are identified off the variation in health outcomes induced by the pre-determined lagged alcohol, smoking and exercise variables conditional on unobserved type. Furthermore, the effects of the inputs in the health production function are normalized to zero when the lowest level of a input is chosen, i.e., the choice is "1."

The income parameters are identified through variation in pre-determined (one period) lagged health and age. Specifically, the lagged health and current age variables are assumed to be uncorrelated with the current income draw, ϵ_t^y, conditional on the time invariant unobserved type.[58] There is also an exclusion restriction that current income depends on lagged health and not current health, but current health is permitted to depend on current income. The identification of the OOP cost parameters is aided by the exclusion restrictions and timing assumptions. The current insurance status affects the current OOP costs but not vice versa (eq. 5). However the lagged OOP costs may affect the current insurance status. Also conditional on the insurance status, the current medical treatment choice affects the current *realized* OOP costs but not vice versa. However *expected* OOP costs may affect the current medical treatment decision i.e.,

[57] The latter assumption is used in forming the terms in the brackets on the third line of the likelihood function (25).
[58] The constant in the income function varies by unobserved type.

the individual knows the OOP cost distribution but makes the treatment decision without any information about the current period's realization, ϵ_t^{oop} (see section 2.2). The sickness distribution parameters are identified through variation in pre-determined health and age variables based on an exclusion restriction in the form of a conditional independence assumption between H_t and s_t. That is,

$$Pr[H_t, s_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] = Pr[s_t \mid H_t, A_t] \cdot Pr[H_t \mid H_{t-1}, s_{t-1}, e_{t-1}, c_{t-1}, a_{t-1}, m_{t-1}] \ . \quad (26)$$

This implies that the current sickness does not affect current health but conversely current health affects current sickness. Current sickness is however assumed to affect future health states.

6 The Estimation Results and Model Fit

6.1 The Estimates

In the estimation I try to match the sample of 3671 males from the HRS data. The model has been estimated using a gradient method to maximize the likelihood that incorporates the BHHH algorithm for approximating the hessian, and calculating standard errors. The estimates are provided in Tables 3 - 12. The estimates have been obtained under the assumption of three types of individuals. The type 1 individual is a moderate type, with moderate demand for health, moderate ability to produce health and a moderate income earning ability. The type 2 individual is a high type defined analogous to the type 1. The type 3 individual is a low type as defined in similar fashion to a type 1 or 2 individual. It should be noted that the parameter α_{13} is not estimated. It is a scale parameter and is set to 1. In fact this helps provide an interpretation to the utility parameters as being monetary equivalents of psychological costs or benefits.

It is seen from Figures 3 - 14 that the model does well in matching the means and frequencies[59] for all the choices and the health states. It needs to emphasized that only 221 of the 9650 (or approximately 2.2%) data points are under the age of 50 or over the age of 75 (see Figure 2). This is by design of the HRS data. Thus the model is does a poor job of fitting the data at the tails of the empirical distribution. In the figures I concentrate on establishing fit for individuals in the age range of 50-75.

The model does well in matching the persistence in the health states and the insurance choices (Tables 18 - 35). It does not do so well in matching the transitions for those who move on to medicare and associated plans at the age of 66 from the different choice set that exists for the individuals under the age of 65 (Tables 22 and 23). This could be because the model is under-parameterized in this dimension. In fact there is no parameter that captures this transition so the reasonable success in this dimension is

[59] Tables for the frequencies are omitted in the interest of brevity.

a cause itself to put some faith in the model. It should be noted that the model tends to over predict transitions in to the uninsured state for the under 65 individuals (see Tables 20 and 21) but this could just be a side effect of over predicting the number of uninsured at the younger ages (figure not shown). Other evidence on transitions is found in Tables 24 - 35 and the model is seen to do reasonably well in matching the persistence in the other choices and states. In particular the quitting behavior for smoking and alcohol choices is matched fairly well. Thus the model is able to address a limitation of the rational addiction models in being unable to match quitting behavior (see Choo (2000)).

Given the large number of coefficients, I do not discuss the estimates (see Tables 3 - 14) individually. The model may appear to be highly parameterized with 199 estimated parameters. However as in Rust and Phelan (1997) a relatively few number of parameters are needed to specify individual preferences (19 parameters). On the other hand a very large number of parameters are needed to specify the beliefs and technology. Importantly, the model attempts to fit 5 decisions processes (insurance, alcohol consumption, smoking, exercise and medical treatment) and 5 outcome processes (health, income, OOP costs, sickness and household size), some of which are highly non-linear. If a combination of linear and discrete dependent variable regressions were estimated with the model's endogenous variables being regressed on the relevant state variables there would be more than 600 parameters to be estimated.[60] By this standard the model is relatively parsimonious.

6.2 Out of Sample Fit

To further test the fit of the model I did some out of sample exercises. I compared the mortality (actually the survival rates) of the US population and the survival rates from the model using the estimated parameters. The obtained results are reproduced in Table 36. It is found that the predicted survival rates (under the column heading "baseline simulations") are quite close to the data except at the very old ages. The discrepancy at the older ages could be because of the paucity of data on the aged as mentioned earlier. The model is currently unable to reproduce the large increase in death rates as the individuals approach the age of 80.

In another exercise I compared the predicted percentage of the uninsured males from the model to the percentage of uninsured males for the US population as reported in the National Health Interview Survey (NHIS) Sample Adult component for 1998. This comparison is illustrated in Figure 15. It is seen that the model does well in predicting the percentage of uninsured males in the US population. I did a similar comparison for the mean alcohol consumption for males as predicted by the model and the mean alcohol

[60] A simple calculation showed that assuming reduced form specifications analogous to those adopted in the model for each of the 10 regressions would yield 651 parameters.

consumption for the US male population as reported in the NHIS 1998. This comparison was done in terms of the discrete nomenclature adopted by me in estimating the model. As Figure 16 shows the model does favorably in predicting out of sample alcohol consumption.

In the last of such exercises I compared the predicted mean smoking rates for males as predicted by the model and mean smoking rates for the US male population as obtained in the NHIS 1998. It is seen in Figure 17 that the model matched the out of sample smoking rates well for those over the age of 50. However for males under the age of 50 the model tends to over predict the mean smoking rate. One reason for this over prediction at the younger ages (i.e. between 22 and 50) could be that the model parameters have been recovered using data (overwhelmingly) from the 1931-41 birth cohort. However the NHIS 1998 data contains a mixture of cohorts. Given the fact that there is evidence that smoking rates have declined in the last fifty years it seems plausible to reason that for the age range of 22 to 50, the individuals in the 1931-41 birth cohort would have had much higher smoking rates compared to the mixture of cohorts that are sampled in the NHIS 1998.

6.3 Counter Factual Policy Experiments

I used the estimates for two counter factual policy experiments (Figures 18-23). The first experiment was designed to simulate the effects of the introduction of an insurance that provides comprehensive coverage. I changed the insurance regime to be such that every individual was mandated to be on the same insurance plan. The sole available insurance plan was deemed to charge a premium of \$1000.0 per annum and cover all out of pocket costs. The charge of \$1000.0 per annum is in the range of the amounts charged by the Group, Group/VA, Medicare and Medicare/Group plans as found in the HRS data. I simulated a sample of a 3000 individuals (a 1000 individuals for each type) using the preliminary estimates. As figure 15 shows the demand for medical care rose at all ages under the new regime compared to the previous one. The increase in the proportion of individuals seeking medical care was as much as 49% higher at some ages. There is also evidence that alcohol consumption in the new regime decreased slightly (up to 0.05 %) in the younger years and increased marginally (up to 0.4%) during the later years. As for smoking, there is a rise of up to 0.3% in the proportion of individuals smoking (see figures 16 - 18). However there was not any significant change in the average health over the life-cycle (figure 19). The important thing to note is that under the regime of subsidized medical care though there was a substantial increase in the demand for medical treatment there was *no large increase* in the risky habits. This goes against the grain of common wisdom that argues about the existence of moral hazard associated with the provision of subsidized medical treatment and unhealthy habits like smoking.

The intuition for this result is as follows. Since the individuals have life long access to subsidized medical care they do not hesitate to seek treatment whenever they are sick. Thus as the individuals age and get increasingly sicker they increase their demand for medical care. The increase in medical treatment is used primarily for care that mitigates the disutility from increased sickness that results from aging. Consequently though the life span of the individuals is not increased much their well being is improved considerably compared to the status quo. Thus the individuals lead more comfortable lives when they are older than they would have if they did not have access to subsidized medical care. Since the individuals are forward looking they realize when they are younger that under the regime of subsidized medical care they will be leading comfortable lives if they lived to be older. Thus there is an incentive to survive to be older and derive the benefits of a longer life. This incentive leads the individuals to abstain from the practice of harmful habits like smoking and alcohol consumption when they are younger. This is in anticipation of the increase in survival rates that is brought about by staying away from these habits. In other words the individuals are willing to trade off the instantaneous pleasure associated with smoking and alcohol consumption for the future benefits that the curtailment of these habits provides in terms of an increased life expectancy. This result provides empirical evidence that contrary to common wisdom there is no (or little) moral hazard associated with the provision of subsidized medical treatment and habits like smoking and alcohol consumption. It should be noted that spirit of this result is like that of measuring the wage elasticity of labor supply. Theoretically there are two countervailing effects of a habit like smoking, i.e. the instantaneous utility that smoking provides and the expected future cost in terms of the expected drop in future health outcomes. The estimates suggest that empirically the two effects are such that the moral hazard problem is non existent. Interestingly this result has the flavor of the result in Philipson and Becker (1998) that says that back loading the rewards in the life-cycle can affect decisions at younger ages.

The intuition behind the lack of change of the average health profile over the life-cycle (see Figure 22) even though there was drastic change in the amount of medical treatment demanded is the following. Most of the extra medical treatment demanded under the regime that was more generous in terms of medical care subsidy was for palliative purposes. Only small amounts of the additional treatment demanded played the role of investment in health. This fits well with the intuition that other than in the most exceptional of cases medical treatment that is imperative is sought (or received) by most people. However people do tend to forego some sorts of medical treatment that can have significant effects in improving their well-being but that which they perceive as unnecessary (and out of their financial reach), e.g. dental treatments, eye exams or health screening visits to doctors.

To further test this intuition I calculated the average consumption over the life cycle under the new regime and compared it to the average consumption in the status quo. I found that the rise in average consumption is as much as 7% (see Figure 23) under the new policy. Thus there is evidence that the savings on the out of pocket costs are at least being channeled towards the composite consumption commodity and not being spent on alcohol and cigarettes. This is also partial evidence of improvements in well being of the average individual under a policy of comprehensive health insurance. This is partial evidence because what one really needs to do to determine the answer to the question of welfare changes is to calculate the life time utility of individuals under the two regimes and compare these. However I have only compared one component of the life time utility, i.e. consumption of the composite good. This comparison is kept for future research. However this examination tends to substantiate the intuition provided for the absence of the moral hazard associated with the habits in the presence of subsidized medical insurance.

In the second experiment I tried to simulate the effects of a total withdrawal of the provision of subsidized medical care. Thus everyone in the economy was denied health insurance and had to pay for any medical care sought through their own pocket. As expected there was a fall in the proportion of individuals seeking medical care. The fall was most dramatic for the elderly, being as much as 95% for those over 70 years of age. It was also found there was an increase in the proportion of individuals smoking compared to the status quo at the younger ages (upto 0.17%), and at the later ages there was a fall in smoking (up to 0.6%). There was also an increase in the number of individuals consuming alcohol (as much as 0.3%) at the younger ages and a decrease (as much as 0.35%) at the older ages (graphs are omitted for brevity).

The intuition for this result is analogous to that provided for the first experiment. As the individuals age they get sicker. However since they have to pay for all their medical treatment out of their own pocket it is not rational for them to seek medical treatment as they approach their final years. This is because at the later ages they find it more worthwhile to consume their incomes directly rather than wait for the (expected) benefits of medical treatment to take effect in the future. This is natural since they do not have enough of a life left to recoup the (expected) benefits of the treatment. As far as the short run palliative aspect of medical care is concerned the elderly in this world of no health insurance prefer to consume their incomes and use that to compensate for their afflictions directly rather than seek solace through medical treatment. Given that the individuals are modeled to be forward looking, at younger ages the individuals foresee that they will be uncomfortable if they live to be old. Thus there is little incentive for the individuals to live a long life in this policy regime where there is complete absence of subsidized medical care. Hence the smoking rates and alcohol consumption go up under this regime as the individuals trade off the future benefits associated with curtailing these activities with the instantaneous pleasures that these activities provide. It is almost as if at the younger ages the individuals in this no subsidy world indulge in activities that are detrimental to themselves because they see no rationale to prolong their lives. However the individuals who have survived in this world try to do their utmost (by curtailing smoking and alcohol consumption) to improve their well being[61].

By doing a comparison between the predicted survival rates (and average health outcomes) from the two experiments (see table 35 and figure 20) it possible to say that the two policies provide examples of two extreme policies with potentially two different life expectancy implications for the elderly. Under the regime of complete subsidy for medical treatment the survival rates are predicted to rise marginally at the older ages. However under the policy of no subsidy for medical treatment the survival rates are predicted to exhibit a small drop in the later years. It remains a question for future research to calculate the economic implications of this for the population in terms of the life time expected utility accumulation.

Before concluding it might be worthwhile to provide an example of another experiment that could be done using the estimated model. The model can be used to examine the effects of barriers to health insurance coverage that individual health insurance companies may create using genetic information about individuals. If the objective is one of trying to simulate the effects of barriers to health insurance to the individuals with poorer quality genetic material then one can simply create barriers to health insurance coverage for the type 3 (i.e. low type) individuals (or deny them coverage altogether) in the model and simulate the effects on the individuals of the new policy in terms of their health outcomes and the other states and habits. In this and in many other ways the estimates of the model can be gainfully employed.

7 Conclusions

In conclusion, I have constructed a dynamic stochastic model of health insurance, exercise, smoking, alcohol and medical treatment choices. The model has been estimated using data from 3671 males from the Health and Retirement Study. By and large the model does well in matching the data. The model is also displays a good out of sample fit. It is successful at matching out of sample survival rates (by age) for the US population. Additionally the model predicts the proportion of the uninsured males for the US population fairly well.

I have also provided two examples of the kinds of experiments that can be performed using the estimated parameters of the model. In the first counter factual experiment where there is only one type of insurance

[61]This is much like the story of Mickey Mantle. He came from a family that was afflicted with a hereditary illness that had shortened the life span of all the males (no male in the family was said to have survived beyond 45). Micky Mantle led an extremely careless life as far as smoking and alcohol consumption were concerned. However he survived to a relatively old age. In his later years the following words have been attributed to him, "*If I had known I was going to live this long I would have taken better care of myself.*".

available in the economy that covers all out of pocket costs and is mandated upon every individual. This leads to an increased demand for medical care however smoking rates and alcohol consumption do not rise much, in fact the alcohol consumption exhibits a slight decline at the younger ages. In the second case where individuals are denied health insurance coverage there is a sharp fall in the demand for medical care, especially in the later years. Further there is a small rise in the smoking rates at the younger ages and a similar fall at later ages. Similarly there is a slight increase in the alcohol consumption in the younger years and a small decrease in the older years. The two experiments taken together suggest that there is negligible moral hazard associated with the provision of subsidized medical treatment in regards with smoking and alcohol consumption. The experiments also provide evidence that the consequences for individual well-being and welfare may be very different under the two radically contrasting health insurance regimes. The model also provides a vehicle for conducting various kinds of health insurance policy experiments and to determine the outcome of counter-factual health insurance policies on not just the demand for medical care but also individual habits.

References

Aitchison, J., and C. Aitken (1976): "Multivariate Binary Discrimination by the Kernel Method," *Biometrika* 63, 413–420.

Barnett, W. A. (1977): "Pollak and Wachter on the Household Production Approach," *Journal of Political Economy* 85, 1073–1082.

Becker, G. S., and K. Murphy (1988): "A Theory of Rational Addiction," *Journal of Political Economy* 96, 675–700.

Becker, G. S., M. Grossman and K. Murphy (1994): "An Empirical Analysis of Cigarette Addiction," *American Economic Review* 84, 396–418.

Benitez-Silva, H., M. Buchinsky, H-M. Chan, J. Rust and S. Sheidvasser (1999): "An Empirical Analysis of the Social Security Disability Application, Appeal and Award Process," *Labour Economics* 6, 147–178.

Berndt, E., B. Hall, E. Hall and J. Hausman (1974): "Estimation and Inference in Nonlinear Structural Models," *Annals of Economic and Social Measurement* 3, 653–665.

Butler, J. S., and R. Moffit (1982): "A Computationally Efficient Quadrature Procedure for the One-Factor Multinomial Probit Model," *Econometrica* 50, 761–764.

Cameron, A., P. Trived, F. Milne and J. Piggot (1988): "A Microeconomic Model of the Demand for Health Care and Health Insurance in Australia," *Review of Economic Studies* 55, 85–106.

Cameron, S. and J. Heckman (2001): "The Dynamics of Educational Attainment for Black, Hispanic, and White Males," *Journal of Political Economy* 109, 455–499.

Choo, E. (2000): " Estimating a Rational Model of Addictive Consumption: The Case of Cigarette Smoking," mimeo, Yale University.

Duan, N., W. Manning, C. Morris and J. Newhouse (1983): "A Comparison of Alternative Models for the Demand for Medical Care," *Journal of Business and Economics Statistics* 1, 115–126.

Duan, N., W. Manning, C. Morris and J. Newhouse (1984): "Choosing Between the Sample-Selection Model and the Multi-part Model," *Journal of Business and Economics Statistics* 2, 283–289.

Eckstein, Z., and K. Wolpin (1989): "The Specification and Estimation of Dynamic Stochastic Discrete Choice Models: A Survey," *Journal of Human Resources* 24, 562–598.

Fries, J. F. (1980): "Aging, Natural Death and the Compression of Morbidity," *The New England Journal of Medicine* 303, 130–135 .

Gilleskie, D. B. (1994): "A Dynamic Stochastic Model of Medical Care Use and Work Absence," Ph.D. Dissertation, University of Minnesota.

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.

Hahn, J, P. Todd and W. van der Klaauw (2001): "Identification and Estimation of Treatment Effects With A Regression-Discontinuity Design," *Econometrica* 69, 201–209.

Heckman, J. (1981): "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time - Discrete Data Stochastic Process," in *Structural Analysis of Discrete Data with Econometric Applications,* edited by C. Manski and D. McFadden. Cambridge, MA: MIT Press, 179-195.

Hubbard, R. G., J. Skinner, and S. Zeldes (1994): "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving," *Carnegie Rochester Conference Series on Public Policy* 40, 59–125.

Hurwicz, L. (1950): "Generalization of the Concept of Identification," in *Statistical Inference in Dynamic Economic Models,* edited by T. Koopmans. New York: John Wiley, 245–257.

Keane, M. P., and K. I. Wolpin (1997): "The Career Decisions of Young Men," *Journal of Political Economy* 105, 473–522.

Keeler, E., J. Newhouse and C. 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.

Lillard, L. A., and Y. Weiss (1997): "Uncertain Health and Survival: Effects on End-of-Life Consumption," *Journal of Business and Economics Statistics* 15, 254–268.

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

McFadden, D. (1981): "Econometric Models of Probabilistic Choice," in C. Manski and D. McFadden, eds., *Structural Analysis of Discrete Data with Econometric Applications,* Cambridge, MA: MIT Press, 198-272.

Mundlak, Y. (1963): "Estimation of Production and Behavioral Functions from a Combination of Cross-Section and Time Series Data," in *Measurement in Economics: Studies in Mathematical Economics and Econometrics in Memory of Yehuda Grunfeld,* edited by C. F. Christ et. al., Stanford, California: Stanford University Press, 139–166.

Muurinen, J-M. (1982): "Demand for Health : A Generalized Grossman Model," *Journal of Health Economics* 1, 5–28.

Newhouse, J. P. (1980): "On Having Your Cake and Eating It Too," *Journal of Econometrics* 13, 365–390.

Newhouse, J. P. (1987): "Health Economics and Econometrics," *American Economic Review* 77, 269–274.

Newhouse, J. P. (1993): *Free for All?* Cambridge, MA: Harvard University Press.

Olley, G. S., and A. Pakes (1996): "The Dynamics of Productivity in the Telecommunications Equipment Industry," *Econometrica* 64, 1263–1297.

Phelps, C. (1973): *Demand for Health Insurance : A Theoretical and Empirical Investigation.* Prepared for the Office of Economic Opportunity, R-1054-OEO, Santa Monica, CA: RAND.

Philipson, T. J., and G. S. Becker (1998): "Old-Age Longevity and Mortality-Contingent Claims," *Journal of Political Economy* 106, 551–573.

Pollak, R., and M. Wachter (1975): "The Relevance of the Household Production Function and Its Implications for the Allocation of Time," *Journal of Political Economy* 83, 255–278.

Pollak, R., and M. Wachter (1977): "Reply: 'Pollak and Wachter on the Household Production Approach'," *Journal of Political Economy* 85, 1083–1086.

Rosenzweig, M., and T. Paul Schultz (1983): "Estimating a Household Production Function: Heterogeneity, the Demand for Health Inputs, and Their Effects on Birth Weight," *Journal of Political Economy* 91, 723–746.

Rosenzweig, M., and K. Wolpin (1995): "Sisters, Siblings, and Mothers: The Effect of Teen-age Childbearing on Birth Outcomes in a Dynamic Family Context," *Econometrica* 63, 303–326.

Rust, J. (1987): "Optimal Replacement of GMC Bus Engines: An Empirical Model of Harold Zurcher," *Econometrica* 55, 999–1035.

Rust, J. (1994a): "Estimation of Dynamic Structural Models, Problems and Prospects: Discrete Decision Processes," in C. Sims, ed., *Advances in Econometrics: Sixth World Congress,* Vol. 2, Cambridge: Cambridge University Press, 119–170.

Rust, J. (1994b): "Structural Estimation of Markov Decision Processes," in R. F. Engle and D. L. McFadden, eds., *Handbook of Econometrics,* Vol. 4, Amsterdam: Elsevier, 3082–3143.

Rust, J., and C. Phelan (1997): "How Social Security and Medicare Affect Retirement Behaviour in a World of Incomplete Markets," *Econometrica* 65, 781–831.

Sickles, R., and A. Yazbeck (1998): "On the Dynamics of Demand for Leisure and the Production of Health," *Journal of Business and Economic Statistics* 16, 187–197.

Vita, A., R. B. Terry, H. B. Hubert and J. F. Fries (1998): "Aging, Health Risks and Cumulative Disability," *The New England Journal of Medicine* 338, 1035–1041.

Wagstaff, A. (1986): "The Demand for Health : Some new empirical evidence," *Journal of Health Economics* 5, 195–233.

Wagstaff, A. (1989): "Econometric Studies in Health Economics: A Survey of the British Literature" *Journal of Health Economics* 8, 1–51.

Wolpin, K. (1984): "An Estimable Dynamic Stochastic Model of Fertility and Child Mortality," *Journal of Political Economy* 92, 852–874.

Variable	Value
β	0.9025 (≡.95*.95)
$\overline{L}$ (# of types)	3

Table 3: General Variables (Not Estimated)

Parameter	ω_1	ω_2	ω_3	ω_4	ω_5	ω_6
Value	0.05	0.05	0.15	0.01	0.01	0.1

Table 4: Aitchison-Aitken Bandwidths for the Kernel Smoothing

Variable	Estimate	(Asy. S.E.)
ρ_1	5071.4	(10.4)
ρ_2	993.6	(0.9)
ρ_3	1012.1	(3.8)
σ_y	0.86	(0.02)
σ_{oop}	0.75	(0.003)
σ_κ	0.2	(0.08)
σ_μ	0.01	(0.2)

Table 5: Variance Parameters

Parameter	α_1	α_2	$\alpha_{3,1}$	$\alpha_{3,2}$	$\alpha_{3,3}$	α_4
Estimate	-10478.8	-9027.2	150131.1	169986.6	129969.8	-15103.3
(Asy. S.E.)	(53.2)	(268.5)	(258.8)	(267.06)	(341.9)	(28.04)
Parameter	α_5	α_6	α_7	α_8	α_9	α_{10}
Estimate	-344.9	-1153.2	191.7	-412.3	-111.1	95.8
(Asy. S.E.)	(0.2)	(1.3)	(0.3)	(0.8)	(0.2)	(0.8)
Parameter	α_{11}	α_{12}	α_{13}	α_{14}	α_{15}	α_{16}
Estimate	-322.6	-73.9	1.0	-9×10^{-7}	-3239.8	4201.2
(Asy. S.E.)	(3.6)	(0.3)		(4×10^{-9})	(35.2)	(20.7)
Parameter	α_{17}	α_{18}	α_{19}			
Estimate	4132.9	4.11	0.52			
(Asy. S.E.)	(43.2)	(0.04)				

Table 6: Utility Parameters

Parameter	$\eta_{1,1}$	$\eta_{2,1}$	$\eta_{3,1}$	$\eta_{4,1}$	$\eta_{5,1}$
Estimate	-0.4	-0.24	0	0	0
(Asy. S.E.)	(6×10^{-4})	(2×10^{-4})			
Parameter	$\eta_{1,2}$	$\eta_{2,2}$	$\eta_{3,2}$	$\eta_{4,2}$	$\eta_{5,2}$
Estimate	0.25	-0.4	-1.1	0.2	0.1
(Asy. S.E.)	(7×10^{-4})	(7×10^{-4})	(0.02)	(0.005)	(0.003)
Parameter	$\eta_{1,3}$	$\eta_{2,3}$	$\eta_{3,3}$	$\eta_{4,3}$	$\eta_{5,3}$
Estimate	-0.35	-0.6	-0.5	0.2	0.3
(Asy. S.E.)	(0.002)	(0.001)	(0.01)	(0.001)	(0.002)
Parameter	$\eta_{1,4}$	$\eta_{2,4}$	$\eta_{3,4}$	$\eta_{4,4}$	$\eta_{5,4}$
Estimate	-1.52	-0.8	-0.1	-0.001	27.2
(Asy. S.E.)	(0.002)	(0.001)	(5×10^{-4})	(2×10^{-5})	(0.07)
Parameter	$\eta_{1,5}$	$\eta_{2,5}$	$\eta_{3,5}$	$\eta_{4,5}$	$\eta_{5,5}$
Estimate	-1.54	-0.4	-0.23	0.002	17.1
(Asy. S.E.)	(0.004)	(0.002)	(6×10^{-4})	(6×10^{-5})	(0.05)
Parameter	$\eta_{1,6}$	$\eta_{2,6}$	$\eta_{3,6}$	$\eta_{4,6}$	$\eta_{5,6}$
Estimate	-1.57	-0.7	-0.1	0.005	47.4
(Asy. S.E.)	(0.002)	(6×10^{-4})	(3×10^{-4})	(5×10^{-5})	(0.1)

Table 7: Health Transition Parameters (a)

Parameter	$\eta_{6,1}$	$\eta_{7,1}$	$\eta_{8,1}$	$\eta_{9,1}$	$\eta_{10,1}$	$\eta_{11,1}$
Estimate	0	0	0	0	0	0
Parameter	$\eta_{6,2}$	$\eta_{7,2}$	$\eta_{8,2}$	$\eta_{9,2}$	$\eta_{10,2}$	$\eta_{11,2}$
Estimate	0.0008	-0.0002	-0.00016	0.0002	1.28	1.36
(Asy. S.E.)	(1×10^{-6})	(7×10^{-5})	-0.00016	(2×10^{-4})	(0.04)	(0.03)
Parameter	$\eta_{6,3}$	$\eta_{7,3}$	$\eta_{8,3}$	$\eta_{9,3}$	$\eta_{10,3}$	$\eta_{11,3}$
Estimate	0.0015	-0.014	-0.0017	0.0004	1.79	1.91
(Asy. S.E.)	(6×10^{-6})	(2×10^{-5})	(2×10^{-5})	(1×10^{-4})	(0.05)	(0.03)
Parameter	$\eta_{6,4}$	$\eta_{7,4}$	$\eta_{8,4}$	$\eta_{9,4}$	$\eta_{10,4}$	$\eta_{11,4}$
Estimate	0.003	-0.019	-0.0039	0.0006	4.66	4.83
(Asy. S.E.)	(1×10^{-5})	(2×10^{-5})	(5×10^{-5})	(5×10^{-4})	(0.04)	(0.02)
Parameter	$\eta_{6,5}$	$\eta_{7,5}$	$\eta_{8,5}$	$\eta_{9,5}$	$\eta_{10,5}$	$\eta_{11,5}$
Estimate	0.0045	-0.02	-0.0055	0.0008	6.52	6.58
(Asy. S.E.)	(8×10^{-6})	(3×10^{-5})	(7×10^{-5})	(6×10^{-4})	(0.01)	(0.006)
Parameter	$\eta_{6,6}$	$\eta_{7,6}$	$\eta_{8,6}$	$\eta_{9,6}$	$\eta_{10,6}$	$\eta_{11,6}$
Estimate	0.0012	-0.02	-0.0072	0.001	7.82	7.94
(Asy. S.E.)	(7×10^{-6})	(3×10^{-5})	(8×10^{-5})	(4×10^{-4})	(0.02)	(0.008)

Table 8: Health Transition Parameters (b)

Parameter	$\eta_{12,1}$	$\eta_{13,1}$	$\eta_{14,1}$	$\eta_{15,1}$
Estimate	0	0	0	0
Parameter	$\eta_{12,2}$	$\eta_{13,2}$	$\eta_{14,2}$	$\eta_{15,2}$
Estimate	1.21	1.13	0.98	1.17
(Asy. S.E.)	(0.05)	(2.8)	(10.3)	(16.9)
Parameter	$\eta_{12,3}$	$\eta_{13,3}$	$\eta_{14,3}$	$\eta_{15,3}$
Estimate	1.73	1.07	1.11	1.13
(Asy. S.E.)	(0.06)	(31.8)	(13.3)	(11.1)
Parameter	$\eta_{12,4}$	$\eta_{13,4}$	$\eta_{14,4}$	$\eta_{15,4}$
Estimate	4.54	0.97	1.03	1.12
(Asy. S.E.)	(0.05)	(345.8)	(182.1)	(16.5)
Parameter	$\eta_{12,5}$	$\eta_{13,5}$	$\eta_{14,5}$	$\eta_{15,5}$
Estimate	6.41	9.11	9.54	8.48
(Asy. S.E.)	(0.01)	(986.6)	(0.4)	(8.2)
Parameter	$\eta_{12,6}$	$\eta_{13,6}$	$\eta_{14,6}$	$\eta_{15,6}$
Estimate	7.69	9.52	10.55	8.07
(Asy. S.E.)	(0.02)	(230.6)	(1.1)	(7.6)

Table 9: Health Transition Parameters (c)

Parameter	$\phi_{1,1}$	$\phi_{2,1}$	$\phi_{3,1}$
Estimate	0	0	0
Parameter	$\phi_{1,2}$	$\phi_{2,2}$	$\phi_{3,2}$
Estimate	-3.98	-0.1	0.3
(Asy. S. E.)	(0.002)	(0.0004)	(0.00008)
Parameter	$\phi_{1,3}$	$\phi_{2,3}$	$\phi_{3,3}$
Estimate	-6.96	-0.15	0.35
(Asy. S. E.)	(0.002)	(0.0003)	(0.00008)

Table 10: Sickness Transition Parameters

Parameter	κ_1	κ_2	κ_3	κ_4	κ_5	κ_6
Estimate	0.185	0.0362	-0.0004	10.06	10.14	9.89
(Asy. S.E.)	(0.001)	(0.0001)	(2×10^{-6})	(0.01)	(0.01)	(0.02)

Table 11: Income Generating Process Parameters

Parameter	$\psi_{1,1}$	$\psi_{2,1}$	$\psi_{3,1}$	$\psi_{4,1}$	$\psi_{5,1}$	$\psi_{6,1}$
Estimate	0	-4.11	1.53	0	0	0
(Asy. S.E.)		(0.03)	(0.005)			
Parameter	$\psi_{1,2}$	$\psi_{2,2}$	$\psi_{3,2}$	$\psi_{4,2}$	$\psi_{5,2}$	$\psi_{6,2}$
Estimate	21.48	-19.2	-7.4	0.27	-0.005	2.52
(Asy. S.E.)	(0.003)	(0.008)	(0.005)	(6×10^{-5})	(4×10^{-6})	(108.7)
Parameter	$\psi_{1,3}$	$\psi_{2,3}$	$\psi_{3,3}$	$\psi_{4,3}$	$\psi_{5,3}$	$\psi_{6,3}$
Estimate	27.1	-5.59	-8.89	0.54	-0.0075	2.1
(Asy. S.E.)	(0.005)	(0.01)	(0.006)	(2×10^{-6})	(1×10^{-4})	(108.7)
Parameter	$\psi_{1,4}$	$\psi_{2,4}$	$\psi_{3,4}$	$\psi_{4,4}$	$\psi_{5,4}$	$\psi_{6,4}$
Estimate	18.23	9.64	-7.9	0.79	-0.01	1.2
(Asy. S.E.)	(0.009)	(0.01)	(0.006)	(1×10^{-4})	(4×10^{-6})	(0.85)

Table 12: Household Size Process Parameters

Parameter	μ_1	μ_2	μ_3	μ_4	μ_5	μ_6	μ_7
Estimate	2.4	0.79	0.8	3.21	3.27	2.58	1.11
(Asy. S.E.)	(0.004)	(0.004)	(0.04)	(0.003)	(0.003)	(0.006)	(0.001)
Parameter	μ_8	μ_9	μ_{10}	μ_{11}	μ_{12}	μ_{13}	μ_{14}
Estimate	0.89	0.97	0.85	0.8	-0.1	0.152	1.76
(Asy. S.E.)	(0.03)	(0.06)	(0.03)	(0.03)	(0.002)	(0.002)	(0.001)

Table 13: Out of Pocket Expenditure Process Parameters

Parameter	γ_1	γ_2	γ_3	γ_4	γ_5	γ_6	γ_7
Estimate	-2.35	0.0049	-0.113	0.15	0.3	0.01	0.05
(Asy. S.E.)	(0.03)	(0.009)	(0.008)	(0.02)	(0.07)	(0.03)	(0.03)
Parameter	γ_8	γ_9	γ_{10}	γ_{11}	γ_{12}	γ_{13}	
Estimate	0.2	1.1	0.2	0.3	-0.2	2.7	
(Asy. S.E.)	(0.01)	(0.2)	(0.04)	(0.07)	(0.04)	(0.3)	

Table 14: “Probability of Zero Out of Pocket Expenditure” Process Parameters

Insurance Type	Premium
Uninsured (1)	0
Group (2)	2072.0
Personal (3)	8561.27
VA/Champus (4)	677.74
Group/VA/Champus (5)	2209.33
Group/Personal(6)	4564.5

Table 15: Insurance premiums by insurance type for age ≤ 64

Insurance Type	Premium
Only Medicare (1)	1864.58
Medicare/Medigap/Personal (2)	5044.1
Medicare/Group/VA-Champus (3)	2032.35

Table 16: Insurance premiums by insurance type for age ≥ 65

Variable	Obs	Mean	Std. Dev.	Min	Max
Age	9627	59.03	5.47	26	80
Education	9627	12.42	3.26	0	17
Sex	9627	1	0	1	1
Race	9627	1.29	0.63	1	3
Health	9627	4.43	1.18	1	6
Exercise	9423	1.31	0.46	1	2
Smoking	9423	1.41	0.76	1	3
Alcohol	9423	1.86	0.73	1	3
Medical care	9423	1.82	0.8	1	3
Log of hh-income	9423	11.25	0.86	2.19	15.08
Log of hh-OOP costs	4421	6.78	1.61	0.58	12.33
Household size	9627	2.33	0.84	1	4
Insurance cost for the uninsured	629	215.0	1086.77	0	13572.87
Insurance premiums:					
Group plans	3651	2072.0	3661.7	0	67864.36
Personal plans	305	8561.27	6888.16	0	52097.14
VA/Champus plans	169	677.74	1613.64	0	7936.06
Group/VA/Champus plans	144	2209.33	6566.13	0	76912.94
Group/Personal. plans	134	4564.5	6543.99	0	40366.68
Medicare	361	1861.23	2544.51	0	13941.73
M-care/M-gap/Personal. plans	115	5016.19	5055.49	0	45824.3
M-care/Group plans	515	2025.61	2532.37	0	13747.29

Table 17: Data summary (all values are for 2 years and in 1991 dollars; Race:1-white, 2-black, 3-other; Education-years of formal schooling)

			$\mathbf{H}_t$			
$\mathbf{H}_{t-1}$	**1**	**2**	**3**	**4**	**5**	**6**
2	21.79	40.47	24.90	7.78	3.11	1.95
3	8.19	11.31	42.52	27.83	8.45	1.69
4	2.63	2.74	14.33	50.46	23.69	6.16
5	1.56	1.09	5.47	27.34	49.38	15.16
6	1.00	0.69	2.32	12.82	32.2	50.97

Table 18: Health State Transitions: Data (%)

			$\mathbf{H}_t$			
$\mathbf{H}_{t-1}$	**1**	**2**	**3**	**4**	**5**	**6**
2	22.00	41.10	33.10	1.50	2.20	0.00
3	7.90	12.80	40.60	25.40	13.20	0.00
4	1.20	1.30	13.80	54.50	27.30	1.90
5	0.10	0.10	1.20	30.50	44.70	23.40
6	0.00	0.00	0.00	3.20	35.60	61.10

Table 19: Health State Transitions: Baseline Simulations (%)

			$\mathbf{I}_t$ $(age$	$< 65)$		
$\mathbf{I}_{t-1}(age < 65)$	**1**	**2**	**3**	**4**	**5**	**6**
1	65.49	21.25	9.11	4.05	0.17	0.84
2	3.12	90.26	2.78	0.33	0.78	2.73
3	11.70	33.96	50.57	0	0.38	3.40
4	8.66	5.51	0.79	75.59	9.45	0.00
5	0.00	19.31	0.00	14.48	63.45	2.76
6	3.03	67.68	4.04	2.02	5.05	18.18

Table 20: Insurance Choice Transitions: Data (%)

			$\mathbf{I}_t$ $(age$	$< 65)$		
$\mathbf{I}_{t-1}(age < 65)$	**1**	**2**	**3**	**4**	**5**	**6**
1	62.20	18.00	3.80	7.00	5.00	4.00
2	4.00	85.40	2.40	3.60	2.70	1.80
3	10.30	25.90	45.40	8.40	5.30	4.60
4	9.00	18.80	4.30	59.10	5.40	3.30
5	10.10	22.50	5.40	8.60	49.30	4.20
6	11.50	25.20	5.90	9.90	8.00	39.70

Table 21: Insurance Choice Transitions: Baseline Simulations (%)

$\mathbf{I}_{t-1}(age = 64)$	$\mathbf{I}_t$ $(age = 66)$ **1**	**2**	**3**
1	51.28	33.33	15.38
2	18.32	7.69	73.99
3	62.50	29.17	8.33
4	37.50	12.50	50.00
5	10.00	10.00	80.00
6	37.50	0.00	62.50

Table 22: Insurance Choice Transitions: Data (%)

$\mathbf{I}_{t-1}(age = 64)$	$\mathbf{I}_t$ $(age = 66)$ **1**	**2**	**3**
1	47.50	17.40	35.10
2	45.90	18.50	35.50
3	44.60	12.40	43.10
4	43.20	18.50	38.20
5	44.50	13.90	41.60
6	49.60	16.80	33.60

Table 23: Insurance Choice Transitions: Baseline Simulations (%)

$\mathbf{I}_{t-1}(age > 66)$	$\mathbf{I}_t$ $(age > 66)$ **1**	**2**	**3**
1	80.26	6.58	13.16
2	50.00	39.36	10.64
3	20.00	7.69	72.31

Table 24: Insurance Choice Transitions: Data (%)

$\mathbf{I}_{t-1}(age > 66)$	$\mathbf{I}_t$ $(age > 66)$ **1**	**2**	**3**
1	78.30	7.40	14.30
2	26.80	50.10	23.10
3	21.50	7.90	70.60

Table 25: Insurance Choice Transitions: Baseline Simulations (%)

		$\mathbf{m}_t$	
$\mathbf{m}_{t-1}$	**1**	**2**	**3**
1	63.28	23.81	12.91
2	26.94	46.13	26.94
3	15.07	32.51	52.42

Table 26: Medical Choice Transitions: Data (%)

		$\mathbf{m}_t$	
$\mathbf{m}_{t-1}$	**1**	**2**	**3**
1	60.80	14.30	24.90
2	45.10	20.10	34.80
3	36.90	15.30	47.80

Table 27: Medical Choice Transitions: Baseline Simulations (%)

		$\mathbf{c}_t$	
$\mathbf{c}_{t-1}$	**1**	**2**	**3**
1	97.78	1.10	1.12
2	25.49	56.37	18.14
3	10.94	12.81	76.26

Table 28: Smoking Choice Transitions: Data (%)

		$\mathbf{c}_t$	
$\mathbf{c}_{t-1}$	**1**	**2**	**3**
1	92.10	4.70	3.30
2	21.70	54.40	23.90
3	7.40	44.70	47.90

Table 29: Smoking Choice Transitions: Baseline Simulations (%)

		$\mathbf{a}_t$	
$\mathbf{a}_{t-1}$	**1**	**2**	**3**
1	86.28	12.04	1.68
2	17.20	72.58	10.23
3	4.98	25.32	69.70

Table 30: Alcohol Choice Transitions: Data (%)

		$\mathbf{a}_t$	
$\mathbf{a}_{t-1}$	**1**	**2**	**3**
1	88.70	6.50	4.90
2	12.20	65.40	22.40
3	1.50	37.70	60.80

Table 31: Alcohol Choice Transitions: Baseline Simulations (%)

	$\mathbf{e}_t$	
$\mathbf{e}_{t-1}$	**1**	**2**
1	73.53	26.47
2	34.42	65.58

Table 32: Exercise Choice Transitions: Data (%)

	$\mathbf{e}_t$	
$\mathbf{e}_{t-1}$	**1**	**2**
1	84.50	15.50
2	34.50	65.90

Table 33: Exercise Choice Transitions: Baseline Simulations (%)

		HHS$_t$		
HHS$_{t-1}$	**1**	**2**	**3**	**4**
1	90.75	7.65	1.25	0.36
2	6.83	85.39	5.58	2.19
3	5.88	38.53	48.04	7.55
4	3.24	23.17	31.52	42.06

Table 34: Household Size Transitions: Data (%)

		HHS$_t$		
HHS$_{t-1}$	**1**	**2**	**3**	**4**
1	86.30	13.40	0.30	0.00
2	7.40	82.00	10.60	0.00
3	2.10	41.00	45.60	11.30
4	1.50	2.70	30.00	65.80

Table 35: Household Size Transitions: Baseline Simulations (%)

		Suvivors		
Age	**Data**	**Baseline Simulations**	**Mandatory Insurance**	**No Insurance**
20	1000	1000	1000	1000
30	990	987	987	987
40	976	962	962	962
50	948	922	922	921
60	886	853	855	852
70	752	749	751	747
80	512	633	635	626

Table 36: Out of Sample Analysis: Aggregate Mortality Rate of the US Population (***Survivors by Age***). Data source: CDC, Atlanta, GA

Figure 2: Frequency of Sample Observations

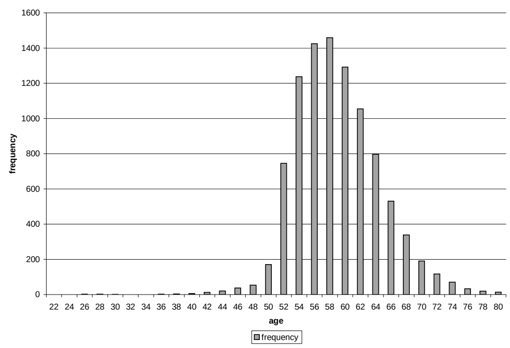


Figure 3: Mean Health by Age

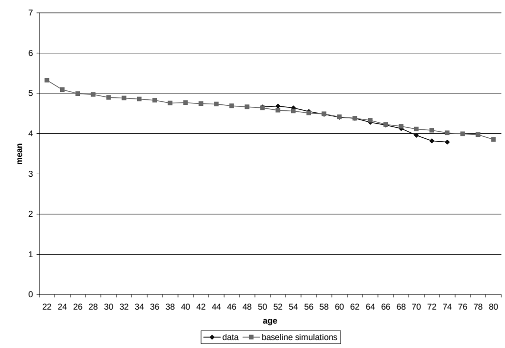


Figure 4: Mean Household Size by Age

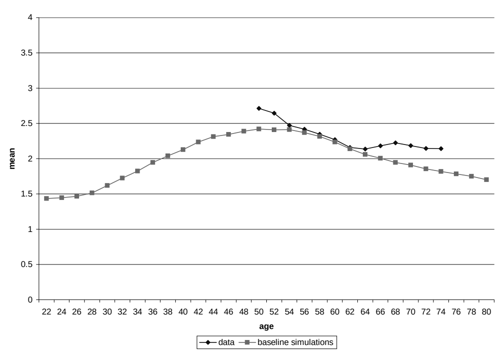


Figure 5: Mean Medical Care by Age

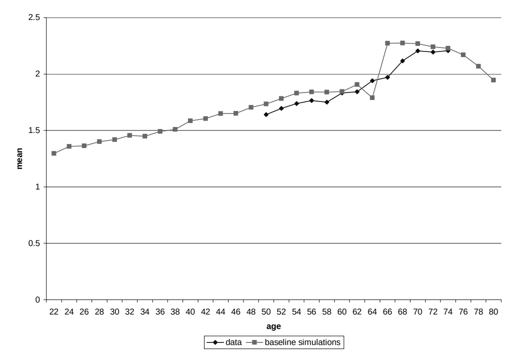


Figure 6: Mean Alcohol Consumption by Age

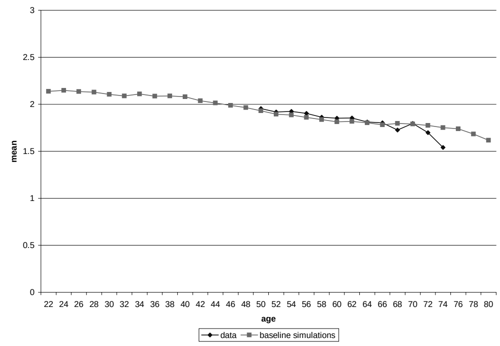


Figure 7: Mean Smoking by Age

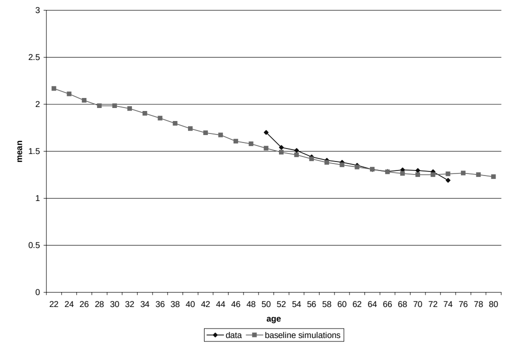


Figure 8: Mean Exercise by Age

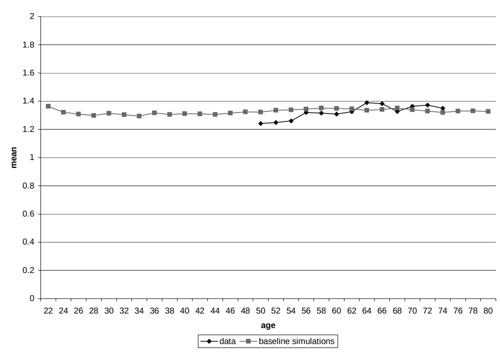


Figure 9: Mean Log-Income by Age

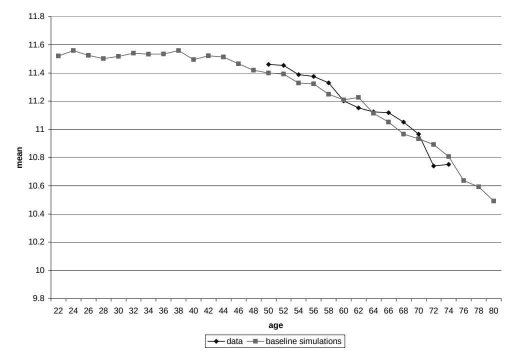


Figure 10: Mean Log-OOP Expenditures by Age

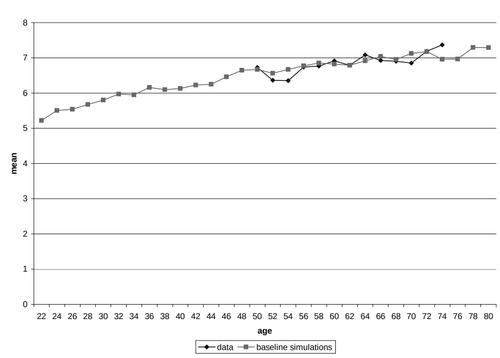


Figure 11: Frequency of Un-insured Individuals (Under 65) by Age

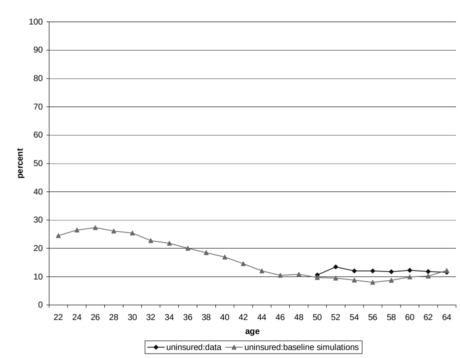


Figure 12: Frequency of Individuals with Group Insurance (Under 65) by Age

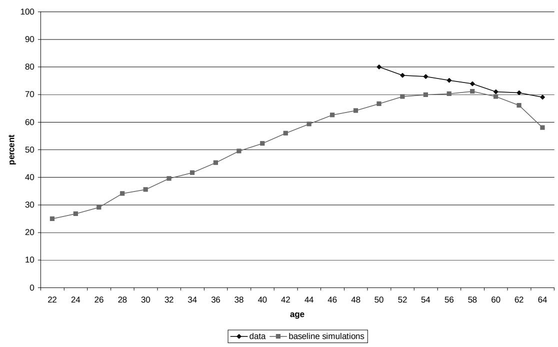


Figure 13: Frequency of Individuals with Only Medicare (Over 65) by Age

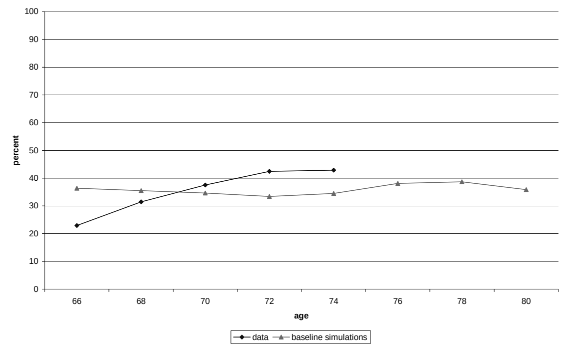


Figure 14: Frequency of Individuals with Medicare and Medigap Insurance (Over 65) by Age

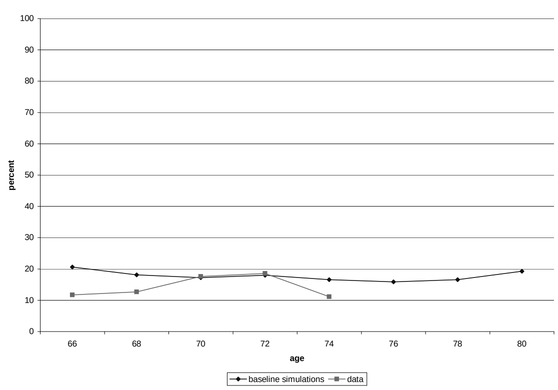


Figure 15: Out of Sample Analysis: Percentage of Un-insured Individuals

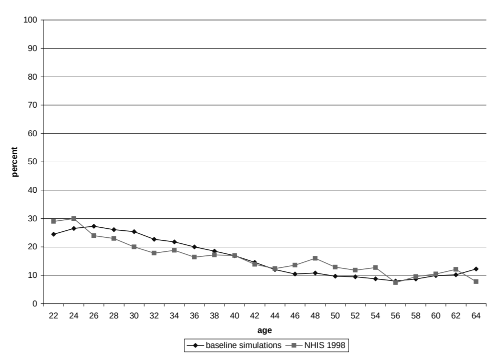


Figure 16: Out of Sample Analysis: Mean Alcohol Consumption

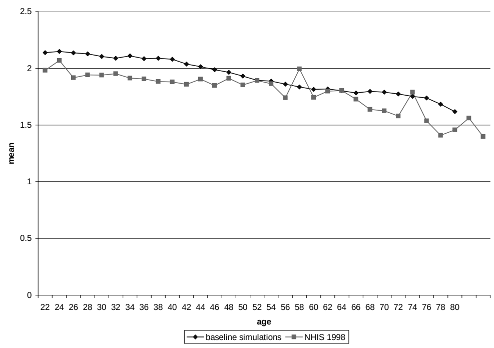


Figure 17: Out of Sample Analysis: Mean Smoking

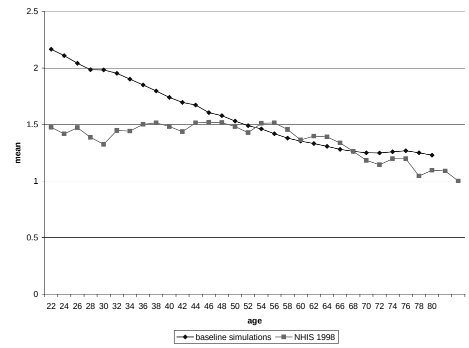


Figure 18: Proportion of Individuals Seeking Medical Treatment Under the Different Insurance Policy Environments

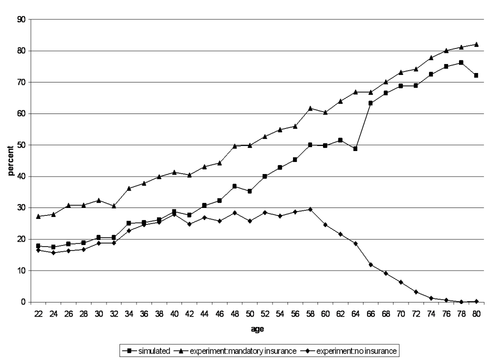


Figure 19: Proportion of Individuals Exercising Under the Different Insurance Policy Environments

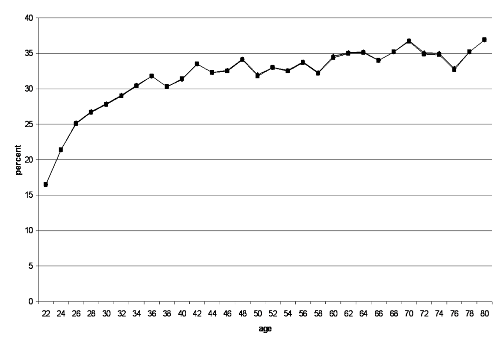
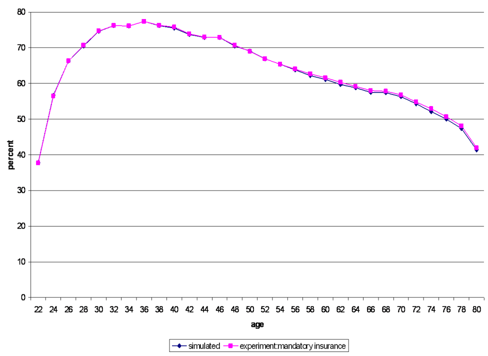


Figure 20: Proportion of Individuals Consuming Alcohol Under the Different Insurance Policy Environments

Figure 21: Proportion of Individuals Smoking Under the Different Insurance Policy Environments



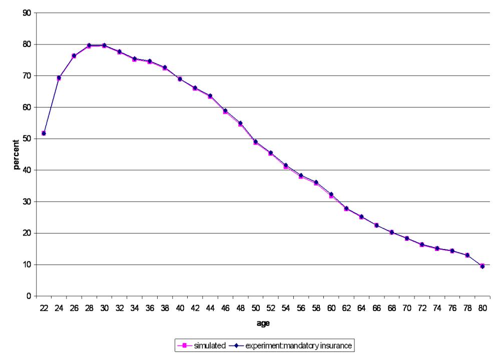


Figure 22: Proportion of Individuals in "Excellent" and "Very Good" Health Under the Different Insurance Policy Environments

Figure 23: Average Consumption of the Composite Consumption Commodity Under the Different Insurance Policy Environments