Estimating willingness to pay for medicare using a dynamic life-cycle model of demand for health insurance

Ahmed Khwaja

Fuqua School of Business, Duke University, One Towerview Drive, 27708 Durham, NC, United States

Keywords: Medicare Willingness to pay Life-cycle human capital Dynamic discrete choice

Abstract

Medicare is the largest health insurance program in the US. This paper uses a dynamic random utility model of demand for health insurance in a life-cycle human capital framework with endogenous production of health to calculate the individual willingness to pay (WTP) for Medicare. The model accounts for the feature that the demand for health insurance is derived through the demand for health, which is jointly determined with the production of health over the life-cycle. The WTP measure incorporates the effects of Medicare insurance on aggregate consumption through effects on medical expenditures and mortality, and consumption utility of health. The model is estimated using panel data from the Health and Retirement Study. The average WTP or change in lifetime expected utility resulting from delaying the age of eligibility to 67 is found to be $24,947 in 1991 dollars ($39,435 in 2008 dollars). However, there is considerable variation in the WTP, e.g., in 1991 dollars the WTP of individuals who have less than a high school education and are white is $28,347 ($44,810 in 2008 dollars), while the WTP of those with at least a college degree and who are neither white nor black is $15,584 ($24,635 in 2008 dollars). More generally, the less educated have a higher WTP to avoid a policy change that delays availability of Medicare benefits. Additional model simulations imply that the primary benefits of Medicare are insurance against medical expenditures with relatively smaller benefits in terms of improved health status and longevity. Medicare also leads to large increases in medical utilization due to deferring of medical care prior to eligibility.

1 Introduction

Medicare is the largest health insurance program in the US. According to the Congressional Budget Office the expenditures on Medicare benefits were $212 million in 1999, which was 12.4% of...
Thus it may alter survival incentives for those under 65 as it alters employment incentives (e.g., (Card et al., 2004; Benitez-Silva et al., 2005)). Hence, Medicare plays an important role in insuring the elderly against medical expenditure risk and improving their access to medical care. This paper uses a dynamic random utility model of demand for health insurance in a life-cycle human capital framework with endogenous production of health to calculate the individual willingness to pay (WTP) for Medicare. The model accounts for the feature that the demand for health insurance is derived through the demand for health. Moreover, individual decisions about health insurance, medical utilization and health related behaviors, and the consequent health outcomes (Grossman, 1972; Phelps, 1973) are inter-related over the life-cycle. Hence the welfare effects of Medicare need to be considered in a dynamic life-cycle framework.

The WTP measure computed in this paper incorporates the effects of Medicare insurance on aggregate consumption through the effects on medical expenditures and mortality, and the consumption utility of health. A life-cycle human capital model of endogenous decisions about health insurance, medical utilization, alcohol consumption, smoking and exercise is estimated using panel data from the Health and Retirement Study (HRS). In order to understand the effects of Medicare on individuals and their willingness to pay for its benefits, simulations from the model are used to infer its effects on medical utilization, out of pocket medical expenditures, and health outcomes by comparing outcomes under its coverage to those in a counter-factual situation in which the age of eligibility is delayed to 67. Given the adverse financial implications of imminent demographic changes this is one particular reform that has been proposed.

There are four important reasons for analyzing the effects of Medicare in a life-cycle framework. First, such a framework helps evaluate the dynamic impact of Medicare. Medicare is a “mortality contingent claim” (Philipson and Becker, 1998) because individuals are entitled to its benefits conditional on survival to age 65. Thus it may alter survival incentives for those under 65 as it alters employment incentives (Rust and Phelan, 1997). Better coverage and thus potentially improved health and higher utility in old age, might induce individuals younger than 65 to on the margin increase behaviors (e.g., exercise) that decrease mortality risk (Philipson and Becker, 1998). Individuals may also defer medical care until they are eligible for Medicare at age 65, e.g., delay expensive treatments like a coronary artery bypass graft. Similarly the anticipated availability of generous coverage for the elderly may induce individuals younger than 65 to on the margin increase behaviors (e.g., smoking) that raise the risk of future medical expenditures and adverse health events. Second, given the life-cycle nature of health production, changes in individual health incentives and behaviors at younger ages may in turn affect health behaviors and outcomes after age 65. Third, the life-cycle nature of health production also implies that there will be dynamic selection (e.g., (Rosenzweig and Wolpin, 1995; Cameron and Heckman, 2001)) in medical utilization, i.e., individuals whose past behaviors raise their current and future health risks will consume relatively more medical care over the life-cycle. In particular mortality will be endogenous to past behaviors causing selection through survivorship. Employing a life-cycle model provides a means to correct for dynamic selection in the empirical analysis in a manner similar to the method proposed by Heckman (1979). Fourth, such an analysis allows for an evaluation of the lifetime welfare effects of the program which for the reasons just outlined is important.

Medicare is administered by the federal government and implemented almost uniformly across the US with little changes in its coverage since its inception. Medicare used to be called a state of the art insurance program for 1965 until December 2003 when the largest expansion of Medicare (Medicare Part D) was enacted. This was done primarily to provide prescription drug benefits as part of the Medicare Prescription Drug, Improvement, and Modernization Act (MMA) that went into effect on January 1, 2006. Using the variation generated by this expansion it is possible to study via reduced form methods the effects of this change in Medicare policy on behaviors and outcomes (e.g., (Duggan and Scott Morton, 2008)). However, such an analysis cannot be used to compute WTP for changes in the Medicare program. Using methods similar to those in this paper, McClellan and Skinner (2006) examine the value of Medicare in completing the missing market for health insurance for the elderly. They compute the parameters of a dynamic model using micro data. Using simulations they evaluate the value of Medicare insurance to its beneficiaries, and also Medicare reforms like progressive premiums and government vouchers. However, they do not analyze its dynamic life-cycle implications. Simulation based methods have also been used to examine the effects of Social Security benefits, e.g., (French, 2005), and van der Klauw and Wolpin (2008). The benefits of Medicare are also analyzed by Lakdawalla and Bhattacharya (2005), who conclude in contrast to McClellan and Skinner (2006) that it is a very progressive program.


This paper extends current research in various ways. In a unified framework using a single source of individual level panel data (i) it accounts for the inter-related life-cycle nature of medical utilization, health related decisions and health outcomes (in particular endogenous mortality) in assessing the effects of Medicare, (ii) it controls for individual-specific unobserved heterogeneity in the analysis, (iii) it examines the role of Medicare in insuring against medical expenditure risk, and (iv) it calculates the lifetime willingness to pay for Medicare.

---

2 The Centers for Medicare & Medicaid Services estimates that the Medicare population will nearly double to 77.2 million or 22% of the US population in 2030, with the beneficiaries 65 or older comprising 68.6 million or 19.6% of the population.


4 See footnote 1 for the exceptions.

5 In principle, the method adopted in this paper can be used to compute specifically the WTP for Medicare Part D or to examine the effects of Medicare Part D on individual behaviors and outcomes, see, e.g., Yang et al. (2009).

6 Feldstein (1971) also used aggregate pooled data for the period 1st July 1966 – 30th June 1968 to examine the economic efficiency of Medicare.
The model is estimated using panel data from the Health and Retirement Study. The WTP or change in lifetime expected utility resulting from delaying the age of eligibility to 67 is found to be $25,539 (in 1991 dollars) on average. However, there is considerable variation in the WTP, e.g., the WTP of individuals who have less than a high school education and are white is $27,252, while the WTP of those with at least a college degree and whose race is neither white or black is $20,012. More generally, the less educated have a higher WTP to avoid a policy change that delays availability of Medicare benefits. Additional model simulations imply that the primary benefits of Medicare are insurance against medical expenditures with relatively smaller benefits in terms of improved health status and longevity. Medicare also leads to large increases in medical utilization due to 'stockpiling' of medical care prior to eligibility.

The rest of the paper is organized as follows. Section 2 describes the model and Section 3 the data. Section 4 discusses the estimation procedure. Section 5 presents estimation results and assesses model fit. Section 6 evaluates the WTP for Medicare benefits, and its impact on health related decisions and health outcomes, and Section 7 concludes.

2. Model

2.1. Structure

For the purpose of evaluating the WTP for Medicare and analyzing its effects a dynamic discrete choice (e.g., (Rust, 1987; Gilleskie, 1998)) stochastic life-cycle model in which health is both a consumption good and human capital (Grossman, 1972; Khwaja, 2001) is employed. 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, \( l_t \), 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, \( X \). The model adopts a random utility specification (McFadden, 1981) with a stochastic component, \( \xi_t \), for the preferences associated with each of the choices. Defining \( \beta \) as the discount rate and \( U() \) as the single period utility function over decisions and states described below, the maximization problem is represented as,

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

subject to

\[
H_t = h(H_{t-1}, a_{t-1}, c_{t-1}, e_{t-1}, s_{t-1}, m_{t-1}, e^{U}_{t-1})
\]

\[
X_t = Y_t - p_t - OOP_t
\]

\[
Y_t = y(H_{t-1}, A_t, e_t)
\]

\[
OOP_t = o(l_t, H_t, m_t, A_t, HHS_t, e^{opp}_t)
\]

\[
s_t = s(H_t, A_t, e_t)
\]

\[
HHS_t = f(HHS_{t-1}, A_t, e^{HHS}_{t-1})
\]

\[
H_0^* = I_0^1; I_0 = I_0^2; a_0 = a_0; c_0 = c_0; e_0 = e_0^*; m_0 = m_0; HHS_0 = HHS_0
\]

A detailed description of the model follows and for ease of reference a summary of the notation and variables in the model is presented at the end of this section in Table 1. In the model 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 of one of two values (1-“no”, 2-“yes”).

The model allows for insurance to protect 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. Hence the model allows for the Medicare generated trade-off between benefits of insurance and distortions in incentives for health behaviors. As specified in the model, insurance does not provide any utility per se but there is a monetary equivalent of the cost of switching insurance plans. Hence \( l_t \) 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,
\[
INS = \{1-“none”, 2-“group”, 3-“personal”, 4-“VA/Champus”, 5-“VA/Champus/other-personal”, 6-“group/personal”\}
\]

(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.8 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, i.e., VA/Champus coverage. Plan 5 represents a mix of group and VA/Champus coverage9 and plan 6 represents a mix of group and personal insurance coverage. Plan choices for those 65 and over are similarly defined. In this case individuals have at least basic Medicare coverage (plan 1) and cannot be uninsured, while plans 2 and 3 are defined as a mix of Medicare and other coverage as stated.

An advantage of modeling 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. Fully endogenizing supply of insurance would also make the model computationally intractable. Thus all individuals are assumed to have access to the same set of insurance choices conditional on age.10 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

---

8 There are only two choices for exercise due to data limitations in the HRS.
9 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 modeled.
10 An alternative approach would be to define classes of individuals depending on their type of insurance coverage, see e.g., (Rust and Phelan, 1997). This would relax the assumption of individuals having access to the same set of insurance choices but on the other hand insurance would no longer be an endogenous choice but be exogenously determined.
cost of seeking medical care in the utility function (1), e.g., cost of scheduling and waiting for an appointment. The medical care choices, \(m_c\), 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 1/3, or between 1/3 and 2/3, or more than 2/3 of the empirical distribution. The model also distinguishes between the stock of health and a flow variable measuring sickness, \(s_t\). This helps to account for consumption of medical care without any change in the underlying health stock, and also for the different aspects of medical utilization, e.g., mitigative, curative and preventive (these terms are defined below). The model allows for disutility from sickness. The sickness variable takes one of three values: 1—“none”, 2—“moderate”, 3—“high”. The roles of household size, \(HHS\), age, \(A\), and lagged variables in the utility function (1) are explained later in Section 2.2.

The budget constraint (Eq. (3)) determines the consumption of the composite commodity through the difference between income, \(Y_r\), and the insurance premium.\(^{12}\) \(P_i\), and the out of pocket (OOP) medical costs, \(OOP\).\(^{13}\) Income is measured as the sum of wage and non-wage income. Income (Eq. (4)) is determined endogenously with the law of motion, \(\log(Y_t) = \bar{\eta}_t + e_t\), where \(e_t \sim N(0, \sigma_t^2)\) and \(\bar{\eta}_t\) is mean income of an individual in period \(t\). Mean income is specified as a function of lagged health, \(H_{t-1}\), and current age.\(^{14}\)

To examine the dynamic effect of Medicare on the insurance trade-off through future health risks, especially mortality risk, the model includes 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, \(\epsilon_t\).\(^{11}\) Empirical health stock, \(H_t\), is defined using two measures of health, i.e., (a) self-reported health status\(^{15}\) (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”. 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 \(q\) of \(1, \ldots, 6,\) at time \(t, t = 1, \ldots, T\), given by,

\[
\begin{align*}
\frac{\eta_y}{\eta_5} \cdot \beta_t &= \eta_{q,1} \cdot (q - H_{t-1}) + \eta_{q,2} \cdot (q - H_{t-1})^2 + \\
& - \left[ \eta_{s,3} - \eta_{s,4} \cdot m_{t-1} \cdot 1 [s_{t-1} = 2] \right] - \\
& - \left[ \eta_{oq} \cdot \left[ \eta_{q,1} - \eta_{q,4} \cdot m_{t-1} \cdot 1 [s_{t-1} = 3] \right] \right] + \\
& + \eta_{oq,1} \cdot q_{t-1} \cdot \eta_{s,1} \cdot c_{t-1} + \eta_{oq,2} \cdot e_{1-t} \cdot \eta_{s,2} \cdot m_{t-1} + \eta_{oq,3} \cdot m_{t-1} + \eta_{oq,4} \cdot \eta_{s,3} \cdot (q - H_{t-1}) + \eta_{oq,5} \cdot m_{t-1} + \eta_{oq,6} \cdot \eta_{s,4} \cdot \eta_{s,5} \cdot (q - H_{t-1})^2 + \\
& \text{A multinomial specification is adopted rather than an ordered logit because the former (here and elsewhere below) allows for greater flexibility in replicating the inter-temporal transitions.}\(^{16}\)
\end{align*}
\]

The current health outcome depends on the lagged health outcome via the terms \(\eta_{s,1}\) and \(\eta_{s,2}\). The quadratic specification allows for the persistence in health transitions. The effects of lagged choices on current health are represented by \(\eta_{oq}\) (alcohol), \(\eta_{oq,1}\) (smoking), \(\eta_{oq,2}\) (exercise), and \(\eta_{oq,3}\) (medical treatment). The lagged sickness affects current health through respectively, \(\eta_{s,1}\) in case of moderate sickness (i.e., \(s_{t-1} = 2\)), and \(\eta_{s,2}\) in case of high sickness (i.e., \(s_{t-1} = 3\)). The \(\eta_{s,3}\) is a proportionality factor for high sickness. Utilization of medical care alleviates the effect of moderate sickness by the amount \(\eta_{oq,4}\). In case of high sickness, utilization alleviates the effect of sickness by the amount \(\eta_{oq,5} \cdot \eta_{s,4}\). Hence \(\eta_{oq,4}\) is the effect of medical care in alleviating the depression in health stock. It is referred to as the curative component of medical treatment. On the other hand \(\eta_{oq,5}\) is the net investment effect of medical care on future health. This is referred to as the preventive component of medical care.

The function describing the endogenous OOP costs (Eq. (5)) of the individual takes a flexible mixed-continuous form to allow for non-linearities in reimbursement. The OOP cost function is used to model the effect of insurance coverage on individual health behaviors, and consequent health outcomes, i.e., through the budget constraint (Eq. (3)). The model allows for the OOP expenditures, \(OOP_t\), to be zero conditional on medical utilization if coverage is sufficiently generous. 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\(^{17}\) and age. This probability is modeled as a logit. If the OOP costs are positive at time \(t\), the mean OOP expenditure, \(\bar{OOP}_t\), is a linear function of insurance status, medical utilization, health, household size and age, i.e., if \(OOP_t > 0\),

\[
\log(OOP_t) = \bar{OOP}_t + \epsilon_t \text{ where } \epsilon_t \sim N(0, \sigma_{\epsilon}^2),
\]

It is also assumed that the random components of OOP medical expenditures and income are uncorrelated.\(^{18}\)

The current sickness (Eq. (6)) depends on the current health, age and a random term, \(\epsilon_t\). It is specified to evolve as a multinomial logit process. The household size (Eq. (7)) depends on the lagged household size, \(HHS_{t-1}\), current age and a random element, \(\epsilon_t\).\(^{19}\) It is specified to take one of four values: 1—“one 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.\(^{19}\)

Additionaly, 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 (Khwaja, 2001, 2009) for details).

---

11\ Given the discrete nature of the utilization variable the cut-offs are not exact.

12\ The restrictive assumption that insurance premiums are constant over time is adopted because premium information is extremely limited in the HRS data.

13\ 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. Arcidiacono et al. (2001), among others, find that alcohol and cigarette prices are not significant determinants of these consumption decisions for the HRS sample.

14\ The econometric specification is not presented, here and elsewhere, when it can easily be inferred through the tables presenting the results, i.e., Tables 8–14.

15\ SRHS is a rough but good measure of life-cycle health that does well in predicting significant health events like mortality, see e.g., Deaton and Paxson (1998). Using a more complicated measure of health status would have made the model computationally intractable. For a recent critique of the SRHS model especially in the context of estimating models of labor supply see e.g., Lindboom and Kerkhofs (2009).

16\ 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.

17\ Household size is included because individuals purchasing insurance typically consider the future needs of their household members along with their own. They form expectations over not just their own but also the medical expenditures of the household. Non-wage 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 2.2.

18\ The model allows for current income to depend on lagged health and for current OOP expenditures to depend on current health. The model’s dynamics should capture the correlation between income and OOP expenditures through the joint effect of health on these variables as well as the reverse effect of income on health status.

19\ 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 (Khwaja, 2001, 2009) for details).
2.2. The timing of decisions and flow utilities

The decisions of the individuals are modeled from ages 22 to 80. A terminal value function that depends on health and sickness represents the future consequences of decisions at age 80. The HRS surveys respondents at two year intervals so each time period represents two calendar years. The choice set in each time period is very large. There are 324 potential choices for those under age 65 and 162 for those 65 and over. Hence each time period is divided into three sub-periods to reduce computational demands.

At the beginning of each time period \( t \), \( t = 1, \ldots, T \), a random draw from the income distribution is realized. The household size is also concurrently but independently determined. In the first sub-period, the individual makes a decision about insurance. Next 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. In this sub-period the individual also derives utility from the accumulated stock of health. Next the random draw for the sickness is realized. In the third sub-period the individual makes the medical treatment choice without knowledge of the actual costs. He also derives utility from the composite consumption good. At the end of the period the random draw for the OOP expenditures is realized and the actual OOP expenditures are determined.

In the first sub-period an individual makes an insurance choice \( i \in \text{INS,} \) 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, \( U_{it}(I_t, A_t, \xi) = U_{ii}(I_t, A_t) + \zeta_i \). The deterministic component of the utility function takes the form

\[
U_{ii}(I_t, A_t) = \alpha_t \cdot 1(I_{t-1} \neq I_t) \cdot 1(A_t < 65) \\
+ \alpha_t \cdot 1(I_{t-1} \neq I_t) \cdot 1(A_t \geq 65).
\]

The term \( I_t \) denotes that the choice at time \( t \) is \( i \in \text{INS} \) 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 be different for those who are younger than 65 (i.e., \( \alpha_1 \)) from those 65 or older (i.e., \( \alpha_2 \)). The \( \zeta_i \) is an additive period \( t \) and choice \( i \) specific stochastic component of the preferences for insurance. It is assumed to be IID and drawn from a multivariate extreme value distribution.

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

\[
U_{ij}(H_t, A_t, \xi) = U_{ii}(H_t, A_t, \xi) + \zeta_i.
\]

Inclusion of lagged choices allows for habit persistence and addiction in these behaviors (Becker et al., 1994). The deterministic component of the utility function has the form,

\[
U_{i0}(H_t, A_t, \xi) = \alpha_3 \cdot H_t + \alpha_4 \cdot H_t^2 + \alpha_5 \cdot e_t + \alpha_6 \cdot 1(e_{t-1} \neq e_t) \\
+ \alpha_9 \cdot 1(e_{t-1} \neq \xi) + \alpha_10 \cdot 1(e_{t-1} \neq \xi) + \alpha_11 \cdot 1(a_{t-1} \neq \xi)
\]

\[
+ \alpha_12 \cdot 1(a_{t-1} \neq \xi) \cdot 1(a_{t-1} = 1) \cdot A_t.
\]

This function is quadratic in the health stock, with the parameters \( \alpha_3 \) and \( \alpha_4 \), to allow for risk aversion with respect to health. The net utility from the behaviors is given by \( \alpha_{10} \) (alcohol), \( \alpha_{11} \) (smoking) and \( \alpha_{12} \) (exercise). There is a switching cost associated with alcohol consumption (\( \alpha_11 \)) and smoking (\( \alpha_8 \)) and exercise (\( \alpha_9 \)) 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 (\( \alpha_{10} \)) and smoking (\( \alpha_8 \)) 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. 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. The \( \zeta_i \) 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. 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, \( U_{i0}(Y_t, P_t, OOP_t, S_t, M_t, HHS, \xi) = U_{ii}(Y_t, P_t, OOP_t, S_t, M_t, HHS, \xi) + \zeta_i \) is the indirect utility function for each medical care choice \( k \in K \). 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. The deterministic component of the indirect

---

20 An avenue for future work would be to extend the model to allow for financial savings, and examine whether the results are robust to this extension.

21 The focus of the research is the adult life-cycle behaviors. Individuals typically leave college at age 22 and after that they usually first make independent decisions about health insurance and medical care. Therefore the starting age is assumed to be 22. The terminal age is assumed to be 80 to reduce the computational burden.

22 It should be noted that INS is 6 if age \( \leq 64 \) and INS is 3 if age \( \geq 65 \).

23 Strictly the notation should be \( U_{ii}(I_t, I_{t-1}, A_t) \). 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.

24 These are the only utility parameters that vary by age.

25 There are 3 choices for alcohol consumption, 3 for smoking and 2 for exercise and so \( 
\#J = 18 \).
utility function is,
\[ U_{t}(Y_{t}, P_{t}, OOP_{t}, s_{t}, m_{t}^{k}, HHS_{t}) = \frac{\alpha_{13} \cdot (Y_{t} - P_{t} - OOP_{t})}{(HHS_{t})^{2/3}} + \alpha_{14} \cdot (Y_{t} - P_{t} - OOP_{t})^2 \]
+ \frac{\alpha_{15} \cdot (m_{t}^{k} > 1) \cdot (m_{t}^{k})^2}{(HHS_{t})^{2/3}}
+ \frac{[\alpha_{16} \cdot (m_{t}^{k} - \alpha_{17}) \cdot [1(s_{t} = 2)]}{(HHS_{t})^{2/3}}
+ \frac{[\alpha_{18} \cdot (\alpha_{16} \cdot m_{t}^{k} - \alpha_{17}) \cdot [1(s_{t} = 3)]}{(HHS_{t})^{2/3}}.
\] (12)

The function is quadratic in discrete, with the coefficients \( \alpha_{13} \) and \( \alpha_{14} \), to allow for risk aversion. The individual consumption is the household consumption normalized non-linearly, via \( \alpha_{19}. \) 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, \( \alpha_{15} \), 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. (12) 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 \( \alpha_{16} \). Similarly the disutility of high sickness (i.e., \( s_{t} = 3 \)) is \( -\alpha_{18} \cdot \alpha_{17} \), where \( \alpha_{18} \) is a proportionality constant. This disutility is mitigated by the amount \( \alpha_{18} \cdot \alpha_{16} \) in the case of high sickness. The term \( \alpha_{16} \) represents the pure consumption or mitigative component of medical care. The \( \alpha_{18} \) 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.

If an individual survives to the final time period \( T \) then a terminal value function represents the utility from the remaining life span. This is parameterized as a function of health and sickness states\(^{35} \) in period \( T \) and defined as,
\[ U_{T}^{*}(H_{T}, S_{T}) = \alpha_{20} \cdot (H_{T}) + \alpha_{21} \cdot (S_{T}) + \alpha_{22} \cdot (H_{T}/S_{T}). \] (13)

The stochastic components of the preferences \( \xi_{t}, \xi_{t}^{k}, \) and \( \xi_{t}^{k} \) 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,b).\(^{36} \) The dynamic programming and the associated Bellman equations are described for a related model in Khwaja (2001, 2009). There is no closed form representation of the solution to the dynamic programming problem for the model. There is also a set of initial conditions (Eq. (8)). The model is solved numerically through backward recursion.

3. Data

The model is estimated using the first four waves of publicly available HRS data spanning 1991–98. The HRS sample is nationally representative and contains individuals aged 51–61 in 1991–92, and their spouses irrespective of age.\(^{37} \) The unique data set includes information on all variables of interest for the model, e.g., wage earnings, non-wage income, health insurance, medical utilization, medical expenditures, health related behaviors and health outcomes. See Juster and Suzman (1995) for additional details.\(^{38} \) Without the HRS data estimation of such a model would be impossible. The longitudinal estimation sample consists of 9321 observations on 3562 non-disabled males. The sample includes individuals for whom complete information is available regarding the variables of interest for at least two waves. The summary statistics are in Table 1. A sample of men is used because older males tend to be in poorer health compared to women, thus access to medical care through insurance would have a more pronounced effect on them. Males are also more extreme in their health related behaviors relative to women and the consequent effects are more likely to be observable in their case. Including females in the estimation sample would also add to the computational burden by expanding the state space to allow for differences by gender.\(^{39} \)

The sample ranges between ages 50 and 75.\(^{40} \) The lack of OOP cost data for all waves does not affect the estimation in any adverse way as explained in Section 4. The information on insurance premiums is highly limited in the HRS. The insurance premiums are defined to be the plan specific means for the estimation sample. This is a very rough estimate of the true premium paid by each individual but was adopted as a last resort given the data set. However as seen in Table 1 the plan specific means have plausible values, e.g., (a) for individuals younger than 65, personal insurance is the most expensive with a two-year premium of $ 8561 and the VA/Champus plans are the least costly at $ 678; (b) for individuals older than 65, Medicare/Medigap/personal plans are the most expensive at $ 5016 and the plain Medicare plans are the least costly at $ 1861. Further, the uninsured report that they incurred average insurance costs of $ 215 over two years. A potential explanation is that even though these individuals reported being uninsured at the time of the survey they may have been insured for any part of the preceding two years. There is no information in the HRS about the timing of the loss of insurance. Thus in estimating the model these individuals are considered to be uninsured for the entire two year duration, with no associated insurance costs.

Given the two year sampling period, the respondents are observed either at a sequence of odd or even numbered ages. Since

---

32 The terminal value function was parsimoniously defined to include just three arguments to avoid over-fitting the model.

33 The model does not suffer from the usual independence of Irrelevant Alternatives (IIA) limitation because of its dynamic structure ([Rust, 1994a], p. 139).

34 Blacks, Hispanics, and residents of the state of Florida were over-sampled.

35 Further information about the HRS may be found in The Journal of Human Resources, Vol. 30, 1995, “Special Issue on the Health and Retirement Study: Data Quality and Early Results” and at http://hrsonline.isr.umich.edu/.

36 It is imperative that future research examine the effects of Medicare on females.

37 There were 109 individuals younger than 50 or older than 75 for whom complete information was available. Given the small size of the sample at these ages, these individuals were excluded from the estimation sample as it would be highly unlikely that these individuals would be representative of the population at these ages, e.g., there was only one individual aged 26 in the sample.

38 For this reason the distribution of medical treatment choices over the 4 waves does not have 3 equal parts.

39 I am grateful to Dan Hill at HRS for making this data available to me.
Table 1
Notation and data summary (two year values in 1991 dollars).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Type</th>
<th>No. of Obs.</th>
<th>Col. (%)</th>
<th>Mean (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control variables</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$a_t$</td>
<td>Alcohol consumption</td>
<td>Discrete</td>
<td>9128</td>
<td>1.9 (0.7)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“none”</td>
<td></td>
<td>3170</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“drink/day ≤ 1”</td>
<td></td>
<td>4034</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“drink/day &gt; 1”</td>
<td></td>
<td>1924</td>
<td>21</td>
<td></td>
</tr>
<tr>
<td>$c_t$</td>
<td>Cigarette consumption</td>
<td>Discrete</td>
<td>9128</td>
<td>1.4 (0.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“none”</td>
<td></td>
<td>6938</td>
<td>76</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“pack/day &lt; 1”</td>
<td></td>
<td>669</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“pack/day ≥ 1”</td>
<td></td>
<td>1521</td>
<td>17</td>
<td></td>
</tr>
<tr>
<td>$e_t$</td>
<td>Exercise</td>
<td>Discrete</td>
<td>9321</td>
<td>1.3 (0.5)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“no”</td>
<td></td>
<td>6263</td>
<td>69</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“yes”</td>
<td></td>
<td>2865</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td>$I_t$</td>
<td>Health status</td>
<td>Discrete</td>
<td>9128</td>
<td>1.8 (0.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“Medicare”</td>
<td></td>
<td>352</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“Medicare/Medigap/personal”</td>
<td></td>
<td>147</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“Medicare/group”</td>
<td></td>
<td>644</td>
<td>56</td>
<td></td>
</tr>
<tr>
<td>$m_t$</td>
<td>Medical utilization</td>
<td>Discrete</td>
<td>9128</td>
<td>1.3 (0.8)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“low”</td>
<td></td>
<td>3878</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“moderate”</td>
<td></td>
<td>2976</td>
<td>33</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“high”</td>
<td></td>
<td>2274</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

State and Other Exogenous Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
<th>Type</th>
<th>No. of Obs.</th>
<th>Col. (%)</th>
<th>Mean (Std. Dev.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A_t$</td>
<td>Age</td>
<td>Discrete</td>
<td>9321</td>
<td>59 (4.9)</td>
<td></td>
</tr>
<tr>
<td>$E_t$</td>
<td>Education (years of schooling)</td>
<td>Discrete</td>
<td>9321</td>
<td>12.5 (3.2)</td>
<td></td>
</tr>
<tr>
<td>$H_t$</td>
<td>Health status</td>
<td>Discrete</td>
<td>9321</td>
<td>4.4 (1.2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“dead”</td>
<td></td>
<td>193</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“poor”</td>
<td></td>
<td>407</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“fair”</td>
<td></td>
<td>1158</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4-“good”</td>
<td></td>
<td>2768</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5-“very good”</td>
<td></td>
<td>2916</td>
<td>31</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6-“excellent”</td>
<td></td>
<td>1879</td>
<td>20</td>
<td></td>
</tr>
<tr>
<td>$HHSt$</td>
<td>Household size</td>
<td>Discrete</td>
<td>9321</td>
<td>2.3 (0.8)</td>
<td></td>
</tr>
<tr>
<td>$OOP_t$</td>
<td>Household OOP costs ($$)</td>
<td>Continuous</td>
<td>4276</td>
<td>3,069 (9,068)</td>
<td></td>
</tr>
<tr>
<td>$Pt$</td>
<td>Insurance premiums (age ≤ 64)</td>
<td>Continuous</td>
<td>3651</td>
<td>2,072 (3,662)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 2 (group) ($)</td>
<td></td>
<td>3651</td>
<td>2,072 (3,662)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 3 (personal) ($)</td>
<td></td>
<td>305</td>
<td>8,561 (6,888)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 4 (VA/Champus) ($)</td>
<td></td>
<td>169</td>
<td>678 (1,614)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 5 (group/VA/Champus) ($)</td>
<td></td>
<td>144</td>
<td>2,209 (6,566)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 6 (group/personal) ($)</td>
<td></td>
<td>134</td>
<td>4,565 (6,544)</td>
<td></td>
</tr>
<tr>
<td>$Pt$</td>
<td>Insurance premiums (age ≥ 65)</td>
<td>Continuous</td>
<td>361</td>
<td>1,861 (2,545)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 1 (Medicare) ($)</td>
<td></td>
<td>361</td>
<td>1,861 (2,545)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 2 (Medicare/Medigap/personal) ($)</td>
<td></td>
<td>115</td>
<td>5,016 (5,055)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Plan 3 (Medicare/group) ($)</td>
<td></td>
<td>515</td>
<td>2,026 (2,532)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Insurance cost reported by uninsured ($)</td>
<td></td>
<td>629</td>
<td>215 (1087)</td>
<td></td>
</tr>
<tr>
<td>$R_t$</td>
<td>Race</td>
<td>Discrete</td>
<td>9321</td>
<td>1.3 (0.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1-“white”</td>
<td></td>
<td>7439</td>
<td>80</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-“black”</td>
<td></td>
<td>993</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td></td>
<td>3-“other”</td>
<td></td>
<td>889</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>$Y_t$</td>
<td>Household income ($)</td>
<td>Continuous</td>
<td>9128</td>
<td>106,644 (10,5525)</td>
<td></td>
</tr>
</tbody>
</table>

Each time period $t$ represents two years; the model would need to be solved separately for those surveyed at odd and even numbered ages. To reduce the computational burden the ages are re-coded so that every individual’s age is an even number. The insurance states of the individuals close to the age of 65 were kept unchanged in this process, i.e., those aged 65 or older are not coded to an age below 65 and vice versa for those under 64. The most severe limitation of this re-coding is that the parameters associated with an age regressor do not disentangle the effect of a given even numbered age from that of an odd numbered age one year later. These effects however do not have any significant bearing on the fit of the model or the simulations.
4. Estimation

4.1. The likelihood function

The model’s parameters are estimated by maximizing a like-
lihood function that nests the solution to the dynamic program-
ing problem (see e.g., Rust, 1987; Gillekis, 1998). To illustrate,
assume for simplicity that there are no sub-periods and the pre-
determined state space at time \( t \) is given by \( Z_t = \{ \text{HHS}_t, \}
H_{t-1}, \text{Y}_{t-1}, a_{n,t-1}, c_{t-1}, e_{t-1}, s_{-1}, m_{t-1} \}. \) Define \( z_{t} \in Z_t \) as an
element of the state space. The HRS provides data for a sample of
\( n = 1, \ldots, N \) individuals on (i) all the pre-determined elements
\( \{ z_{n,t}, \text{H}_{t-1} \}, \) except sickness; \( 40 \) where \( \text{H}_{0,t} \) is the initial observation and \( T \)
is the last observation for the individual, and (ii) the sequence of choices for insurance, alcohol consump-
tion, smoking, exercise and medical treatment, and outcomes
related to health, household size, income and OOP costs, i.e.,
\( \{ I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, \text{HHS}_{n,t}, \text{Y}_{n,t}, \text{OOP}_{n,t} \}_{t = t_{0}}^{T} \). The solution
of the model provides the joint probability of observing the
choices and outcomes for each individual at \( n \) time \( t \) conditional
on the pre-determined state and the model’s parameters \( \Theta \).
Using these the sample likelihood can be written as
\[
\mathcal{L} = \prod_{n=1}^{N} \prod_{t=1}^{T} \left( \pi_{n,t} \quad \prod_{l=15}^{10} \prod_{L_{n,t}} \frac{P[I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, \text{HHS}_{n,t}, \text{Y}_{n,t}, \text{OOP}_{n,t}]}{\tilde{P}[z_{n,t} | \Theta_{l}]} \right). \quad (14)
\]
Maximum likelihood estimation does not impose the require-
ment that components associated with the outcomes of health,
household size, income and OOP costs be included in the likelihood
function (see e.g., Eckstein and Wolpin, 1989, p. 588–589). In partic-
ular if OOP costs were observed regardless of treatment for all
the four waves then the parameters of its distribution could be con-
sistently estimated directly from the data. However the OOP cost
data suffer from the (standard) problems of selection and censor-
ing. \( 42 \) Thus the OOP cost parameters are estimated jointly with the
other parameters of the model. Simulations from the model pro-
vide a means of correcting for selection and censoring in the esti-
mation procedure (Heckman, 1979). \( 43 \) Hence even in the absence
of data on OOP costs of all individuals \( 44 \) the parameters of the OOP
cost distribution can be estimated consistently. To account for selec-
tion and endogeneity in the other outcome variables the parameters
of those distributions are similarly estimated jointly with the
choice parameters.

4.2. Unobserved heterogeneity, serial correlation and measurement error

Individuals may differ in unobservable ways with the unob-
serveable components being serially correlated over time. Ignoring
this possibility could lead to biased estimates. Consequently the
method proposed by Heckman and Singer (1984) is used to control
for unobserved heterogeneity. Specifically, suppose that there are
\( l = 1, \ldots, L \) types of individuals with the probability of being the
\( l \)-th type given by \( \pi_{l} \). The estimation allows for unobservable differ-
ences in the health technology, i.e., \( \lambda_{n, l} \), the preferences for health,
income, \( \alpha_{n, l} \), and the income earning ability, i.e., \( \kappa_{n, l} \). \( 45 \) It is assumed that the
preferences and technology are common across the types, and
the individuals know their type. Consequently conditional on an
individual’s observed characteristics and the unobserved type, the
unobserved random components of the preferences, health tech-
nology and income process are assumed to be serially uncorrelated.
By treating the unobserved type as a random effect it is possible to
integrate out the probability of being a particular type in the sam-
ple likelihood as follows,
\[
\mathcal{L} = \prod_{n=1}^{N} \sum_{l=1}^{L} \pi_{l} \prod_{t=15}^{10} \prod_{L_{n,t}} \frac{P[I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, \text{HHS}_{n,t}, \text{Y}_{n,t}, \text{OOP}_{n,t}]}{\tilde{P}[z_{n,t} | \Theta_{l}]} . \quad (15)
\]
It is assumed that the probability that an individual is a particu-
lar type, \( \pi_{l} \), is a multinomial logit in education and race with
parameters \( \lambda_{l} \). Therefore the modified sample likelihood (Eq.
(15)) is maximized with respect to the parameters \( \Theta_{l} \), where \( \Theta_{l} = \{ \pi_{l}, \lambda_{l} \}_{l=1}^{L} \). The estimation also allows for additive measurement
errors in income \( e_{n, l} \sim N(0, \sigma_{n, l}^{2}) \) (see e.g., Wolpin, 1987) and in the OOP
expenditures \( e_{n, l} \sim N(0, \sigma_{n, l}^{2}) \).

4.3. Initial conditions, dynamic selection and survivorship bias

Incorporating unobserved heterogeneity in the estimation
implies that the choices and outcomes of each individual are
related with his history, and in particular with his initial
period choices and outcomes. Thus there is dynamic selection in
the behaviors and consequently in survival leading to an initial
conditions problem. \( 46 \) The estimation corrects for the initial conditions, dynamic
selection and survivorship by approximating the distribution for the
initial choices and outcomes conditional on the unobserved
effects while jointly specifying and estimating the distribution of
the unobserved effects (Heckman, 1981). \( 47 \) To illustrate, assume that
all the individuals are first observed at the age of 50 or \( t = 15 \). \( 48 \) The likelihood is modified to
\[
\mathcal{L} = \prod_{n=1}^{N} \sum_{l=1}^{L} \pi_{l} \prod_{t=15}^{10} \prod_{L_{n,t}} \frac{P[I_{n,15}, a_{n,15}, c_{n,15}, e_{n,15}, m_{n,15}, \text{H}_{n,15}, \text{Y}_{n,15}, \text{OOP}_{n,15}]}{\tilde{P}[z_{n,15} | \Theta_{l}]} . \quad (16)
\]
The likelihood calculation includes the term \( P[z_{n,15} | \Theta_{l}, I] \) which corrects for the initial conditions, dynamic selection and

\( 40 \) 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.
\( 41 \) Sickness is treated as a latent variable in estimating the model and is integrated
out from the likelihood function.
\( 42 \) The OOP cost data is limited to waves 2 and 3 and is available only if an
individual sought treatment. It could also be zero even in the case of treatment
de pending on the insurance coverage.
\( 43 \) The mixed-continuous distribution for the OOP costs in the model explicitly
allows for OOP costs to be unobserved if the individual did not obtain treatment,
and to be zero even when obtaining treatment conditional on insurance coverage.
In addition there is an explicit measurement error to capture any other discrepancies
in the data.
\( 44 \) There are 4276 observations on OOP costs suggesting no cause for concern
about the size of the sub-sample (Table 1).
\( 45 \) This is the constant in the income function.
\( 46 \) The HRS respondents had to be between the ages of 51 and 61 in 1991–92.
However spouses of any age are included in the data set.
\( 47 \) Wooldridge (2000) provides an alternative solution.
\( 48 \) The problem can be generalized easily to include different ages at which an
individual’s history becomes observable.
survivorship. Assuming that the model is correctly specified, it is possible to simulate a history of choices and outcomes for sample of individuals between the ages of 22 (i.e., \( t = 1 \)) and 50 (i.e., \( t = 15 \)) conditional on type. The simulations assume that the initial state of health at \( t = 1 \) is determined by the health production technology (Eq. (9)) solely through the intercept term which represents the unobserved health production type of each individual, i.e., \( q_{10}, q = 1, \ldots, 6 \). The initial household size is similarly determined 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 choices.

The simulated sample can be used to obtain, \( P^* \left[ z_{1, 15} \mid \mathcal{F}_t, f \right] \), where "\( * \)" denotes that these simulated probabilities may not be "continuous," i.e., there is a possibility that a given hypercell determined by the intersection of the elements of the state space at \( t = 15 \) may contain zero elements. To obtain continuous initial probabilities a kernel smoothing procedure (Aitchison and Aitken, 1976) is adopted.\(^{49}\) The "smoothed" sample likelihood is maximized using a Newton–Raphson hill climbing algorithm that employs the BHHT method (Berndt et al., 1974) to calculate the Hessian. The procedure involves jointly estimating the initial distribution of health and household size with the other parameters of the model through a process of "backcasting." See Khwaja (2001, 2009) for additional details for a related model. The asymptotic standard errors are also calculated using the BHHT method.

5. The estimation results and model fit

5.1. The estimates

The marginal utility of the composite consumption commodity at \( X = 0 \), \( \alpha_{15} \), is normalized to be 1 (see Table 2). Hence the utility parameters are in 1991 dollars and refer to a two year duration. For individuals under 65, the monetary equivalent of the cost, \( \alpha_{1} \), of switching insurance plans is $11,564. This is about 1/10th the average two-year household income of $106,644. It is about 3.2 times the average premium of $3617 for the five available insurance plans (the sixth category is "uninsured"). The switching cost includes the disutility of changing insurance plans, e.g., understanding new guidelines, dealing with new providers etc. It also represents the disutility of being uninsured and having to search for coverage when uninsured.

For individuals older than 65 the switching cost, \( \alpha_{2} \), is $7,395. This is about 2.5 times the average premium paid by this group. This switching cost is lower than that under 65 because the three plans available to this age group have greater similarities. Also with access to Medicare, these individuals are guaranteed insurance and do not face the prospect of seeking coverage once uninsured. In controlling for unobserved heterogeneity, individuals of three types are estimated. The utility parameter associated with the quadratic term in health, \( \alpha_{3} \), is common to all types, and is estimated to be \(-15.119\). The utility parameter linear in health is estimated to be different for each type of individual, for type 1 (\( \alpha_{4,1} \)) it is 149.745, for type 2 (\( \alpha_{4,2} \)) it is 166.329 and for type 3 (\( \alpha_{4,3} \)) it is 123.804. Thus type 2 individuals value health the most and type 3 the least.

The disutility and opportunity cost (e.g., value of time lost) of exercise, \( \alpha_{6} \), is $344. The switching cost for exercise, \( \alpha_{6} \), is $1161, implying the model will predict persistence in exercise choices. The utility from smoking, \( \alpha_{7} \), is $72. The switching cost, \( \alpha_{8} \), is $409. The start-up cost parameter, \( \alpha_{9} \), is $75. The actual start-up cost increases proportionately with age, i.e., \( \alpha_{9} \times A_t \). Hence (re)starting smoking is more costly than quitting or switching levels of smoking. This is an incentive not to quit for individuals who might expect to re-start smoking in the future. Thus the model would predict that only a small number of individuals would (i) quit or (ii) relapse. The predicted smoking choices will therefore exhibit high persistence. Alcohol consumption provides utility, \( \alpha_{10} \), of $104. The switching cost, \( \alpha_{11} \), is $787. The start-up cost parameter, \( \alpha_{12} \), is $68. The actual start-up cost increases proportionately with age, i.e., \( \alpha_{12} \times A_t \). It is much more costly to (re)start alcohol consumption than to quit or switch consumption levels. The predicted alcohol consumption will thus exhibit persistence similar to smoking.

The transactions cost of seeking medical care, e.g., scheduling and waiting time costs, \( \alpha_{13} \), is $1534. The disutility of moderate sickness, \( \alpha_{14} \), is $3279. The mitigative utility of medical care, \( \alpha_{15} \), is $1325. Medical care does not totally mitigate the disutility of sickness. A "moderately" sick person who seeks "moderate" medical treatment has a net disutility of $3488, inclusive of the transactions cost of seeking care.\(^{50}\) Due to the convexity of the transactions costs, with "moderate" sickness and "high" level of medical care the net disutility is $8090. Hence there is no incentive for a well person to seek medical treatment unless treatment has a preventive (or investment) benefit. The utility from health in the terminal value function, \( \alpha_{20} \), is $10,691 which is less than the utility prior to age 80 and represents the decreased utility of health after 80. The disutility of sickness in the terminal value function, \( \alpha_{21} \), is $20,750 which is about 6.3 times the disutility at ages below 80, indicating the lower ability to tolerate sickness after 80. The parameter on the ratio of health stock to sickness in the terminal value function, \( \alpha_{22} \), is estimated to be $44,520. This represents the (non-linear) value individuals place on the relative level of health stock to sickness conditional on the level of health stock and sickness per se.

The health production function has a multinomial logit specification so the parameters can not be interpreted directly (Table 7). The effects of health inputs may be inferred by calculating the two year mortality (or survival) odds for a simulated sample of individuals (see footnote 52 below for details of the simulated sample). The mortality (survival) odds are calculated as \( \frac{P_{1, t}}{P_{0, t}} \), where \( P_{1, t} \) (\( P_{0, t} \)) is the mean probability of two year mortality (survival) at time \( t \) with (without) the input conditional on the distribution of characteristics of the simulated sample.\(^{51}\) It is found that in general medical care and exercise have a positive impact on health production while alcohol consumption and smoking are detrimental to health. Seeking "high" relative to "low" level of medical care increases two year survival odds by 1.7% on average over the life-cycle. "High" alcohol consumption (more than a drink a day) relative to abstinence increases the mortality odds by 2.7% on average over the life-cycle. "High" smoking (a pack or more per day) relative to "no smoking" raises two year mortality odds by an average of 6.8% over the life-cycle. The two year mortality odds of "not exercising" relative to "exercising" average 2.0% over the life-cycle. Other production function parameters may be similarly examined. This is not done for space limitations, which is also the reason for not discussing the other parameters of the model (see Tables 8–14).

\(^{49}\) This procedure came to my attention through Gilleskie (1994).

\(^{50}\) In this calculation the utility parameters are normalized so that the lowest level of care and sickness provide zero utility.

\(^{51}\) Survival odds are computed for medical utilization. Mortality odds of "no exercise" to "exercise" are computed, i.e., the inverse of the ratio above.
The baseline simulations replicate the the
utility parameters and other variables. Table 2

distribution of health outcomes and choices by age. Table 3

sample correspond to each person in the estimation sample, i.e., 3562 persons, and
additional persons observed with complete information in at least two waves in
the data. The latter individuals were excluded from the estimation sample because
they are aged either less than 50 or more than 75 (see footnote 37).

5.2. Model fit

The estimated model is used to simulate a sample of individuals
to assess its fit. The baseline simulations replicate the the
age distribution of mean health well, though the mean health is slightly under-predicted at younger ages and slightly over-

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>(Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \alpha_1 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_3 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_4 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_5 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_6 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_7 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_8 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_9 )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_{10} )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_{11} )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \alpha_{12} )</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>( \beta )</td>
<td>0.95</td>
<td></td>
</tr>
<tr>
<td>( L )</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>Likelihood</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3
Within sample fit: distribution of health outcomes and choices by age.

<table>
<thead>
<tr>
<th>Age</th>
<th>Total obs. [Survivors]</th>
<th>Mean health</th>
<th>Mean medical-care</th>
<th>Mean alcohol-consumption</th>
<th>Mean smoking</th>
<th>Mean exercise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Data Baseline Simulations)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>134 [134]</td>
<td>4.76</td>
<td>4.56</td>
<td>1.67</td>
<td>1.71</td>
<td>1.96</td>
</tr>
<tr>
<td>52</td>
<td>724 [722]</td>
<td>4.69</td>
<td>4.53</td>
<td>1.69</td>
<td>1.76</td>
<td>1.92</td>
</tr>
<tr>
<td>54</td>
<td>1231 [1220]</td>
<td>4.64</td>
<td>4.48</td>
<td>1.74</td>
<td>1.80</td>
<td>1.93</td>
</tr>
<tr>
<td>56</td>
<td>1425 [1410]</td>
<td>4.55</td>
<td>4.46</td>
<td>1.77</td>
<td>1.82</td>
<td>1.90</td>
</tr>
<tr>
<td>58</td>
<td>1459 [1435]</td>
<td>4.48</td>
<td>4.42</td>
<td>1.75</td>
<td>1.85</td>
<td>1.86</td>
</tr>
<tr>
<td>60</td>
<td>1292 [1273]</td>
<td>4.40</td>
<td>4.37</td>
<td>1.83</td>
<td>1.84</td>
<td>1.85</td>
</tr>
<tr>
<td>62</td>
<td>1054 [1022]</td>
<td>4.38</td>
<td>4.33</td>
<td>1.84</td>
<td>1.83</td>
<td>1.85</td>
</tr>
<tr>
<td>64</td>
<td>797 [769]</td>
<td>4.28</td>
<td>4.31</td>
<td>1.94</td>
<td>1.76</td>
<td>1.81</td>
</tr>
<tr>
<td>66</td>
<td>531 [511]</td>
<td>4.21</td>
<td>4.28</td>
<td>1.97</td>
<td>2.23</td>
<td>1.80</td>
</tr>
<tr>
<td>68</td>
<td>339 [324]</td>
<td>4.13</td>
<td>4.23</td>
<td>2.12</td>
<td>2.25</td>
<td>1.71</td>
</tr>
<tr>
<td>70</td>
<td>188 [173]</td>
<td>3.95</td>
<td>4.17</td>
<td>2.21</td>
<td>2.25</td>
<td>1.78</td>
</tr>
<tr>
<td>72</td>
<td>105 [101]</td>
<td>3.82</td>
<td>4.11</td>
<td>2.21</td>
<td>2.26</td>
<td>1.68</td>
</tr>
<tr>
<td>74</td>
<td>42 [34]</td>
<td>3.50</td>
<td>4.10</td>
<td>2.35</td>
<td>2.24</td>
<td>1.56</td>
</tr>
</tbody>
</table>

a The first number in each column is the data followed by the baseline simulations.
b The relevant data distribution is “Total obs.”
c The relevant data distribution is “Survivors”.
d The measure of health, \( H_t \), is a composite of SRHS and mortality.
predicted health at older ages (Table 3). The mean medical utilization profile matches the data well, except that the model predicts an increase in utilization between the ages of 64 and 66 while the data shows an increase between ages 66 and 68. The lag in increase in utilization in the data may be due to friction created by factors such as learning about Medicare plans or habit persistence in utilization that are not in the model. The alcohol consumption and smoking profiles are matched well. In particular the decline in consumption with age is replicated by the baseline simulations. The exercise profile though matched closely does not replicate the slight increase in exercise with age.

The fit of the model is further assessed by comparing its simulated predictions out of sample to data from the National Health and Nutrition Examination Survey (NHANES) 1999–2000. The estimated model is used to simulate the predicted distribution of uninsured individuals in the US population under a counter-factual assumption that Medicare does not exist (Table 4). In simulating a sample of individuals in the absence of Medicare it is assumed that the same insurance choices are available to individuals 65 and older as those available under age 65, i.e., 6 choices listed in Table 1, including the choice to be uninsured (more details are provided in Section 5). In other words, it is assumed that absent Medicare persons 65 and older can continue to choose one of the six health insurance alternatives that are available to those under 65, and without any risk adjustment of insurance premiums.

Though the model under-predicts the percentage of uninsured in the middle part of the life-cycle relative to the NHANES 1999–2000 data, it does match the U-shape of the age distribution remarkably well. A reason for predicting a lower level of uninsured compared to the NHANES 1999–2000 data is that these data refer to hospital insurance which provided less generous coverage compared to the health insurance choices in the HRS data. In the HRS sample period relative to that in the 1960’s, the difference in insurance characteristics coupled with advances in medical technology and improvements in quality of medical care, would be expected to increase the demand for medical care, and in turn the demand for insurance. Hence the ability of the model to match the NHANES 1999–2000 data qualitatively is a strong validation of its predictability out of sample. More discussion on the within sample fit of a related model, which is very good, can be found in Khwaja (2001, 2009).

6. Evaluating willingness to pay for Medicare and its effects on behaviors and outcomes

The willingness to pay (WTP) is computed by comparing expected lifetime utility for individuals under the status quo with what would result when if the age of eligibility was delayed to 67 from the current 65. The WTP is computed based on a compensating variation measure that calculates the amount of additional lump-sum income that would be required at t = 0 to make an individual indifferent under the two alternative eligibility ages.

In particular, the WTP to avoid a delay in the availability of Medicare benefits for two years is computed in the following way. For individuals of type \( i = 1, \ldots, N_t \), the amount of (one-time) lump-sum income at \( t = 0 \), \( WTP_i \), that would on average equate lifetime expected utility from delaying the age of eligibility to 67 to that when the age of eligibility is 65 is computed (17). This calculation provides a measure of mean WTP for individuals conditional on type, i.e., \( WTP_i \), because it averages over all individuals \( n = 1, \ldots, N_t \) of type \( i \).

\[
1 \left( \frac{1}{N_t} \sum_{n=1}^{N_t} \left\{ E \left[ \sum_{t=1}^{T} U_t(Y_{n,t,L,M}, P_{n,M}) \right] \right\} \right) = 1 \left( \frac{1}{N_t} \sum_{n=1}^{N_t} \left\{ E \left[ \sum_{t=1}^{T} U_t(Y_{n,t,L,M} + WTP_i P_{n,M}) \right] \right\} \right)
\]

(17)

The following notation is used in (17). The subscript \( M \) represents utility under the current Medicare policy of eligibility at age 65, and \( \Delta M \) refers to a changed Medicare policy of eligibility at 67. For an individual in the age of type \( i \) at time \( t \) and in the insurance environment \( q \in \{ M, \Delta M \} \), \( Y_{n,t,L,M}, P_{n,M} \) and \( H_{n,t,L,M} \) are respectively measures of income, insurance premiums and health, and \( U_t() \) represents the flow utility functions described in Section 2.2. It should be noted that subscripts \( n \) and \( i \) do not appear in the insurance premiums as these do not vary across individuals or types, however, income and health are individual and type specific and the calculation accounts for these differences, as it also does for the heterogeneity in utility for health across types. An even more noteworthy feature of the calculation in (17) is that it accounts for the differences in mortality rates under the different insurance environments through the indicator function \( 1(H_{n,t,L,M} > 1) \). This indicator function accounts for whether an individual is alive (recall, death is represented by a health status equal to 1, i.e., \( H_{n,t,L,M} = 1 \)).

A sample is simulated with the Medicare eligible age increased to 67. The simulations assume that before and after attaining the changed eligible age the individuals have access to the same menu of insurance choices that is available respectively to individuals.
It is observed that the highest WTP ($28,347 in 1991 dollars or $44,810 in 2008 dollars) to avoid a delay in age of eligibility is for those individuals who have less than a high school education and are white. The lowest WTP ($15,584 in 1991 dollars or $24,635 in 2008 dollars) to avoid a postponement of eligibility is for college graduates who are neither white nor black. More generally, it is apparent that individuals with a lower amounts of education have a higher WTP to avoid a policy change that would increase the age of eligibility. This implies that less educated individuals value Medicare benefits more than those with higher education. To provide some context to these numbers, it should be noted that the median household income for households with members 65 and older in 1998 (the last year of the estimation sample) was $19,527 (in 1991 dollars).55 Furthermore, in 1991 the annual Medicare part B premium was $358.80 and individuals ineligible for part A coverage could purchase this for an additional annual premium of $212.4.56 Hence the total annual Medicare premium for an ineligible (see footnote 1 for eligibility rules) individual was $2,482.80, which is about 1/10th the average WTP. Further context is provided below after discussion of the effects of delay in eligibility on out of pocket medical expenditures.

Similar welfare calculations for the life-cycle effects of Medicare do not exist so direct comparisons are difficult. However, a very similar conclusion is reached by Lakdawalla and Bhatattacharya (2005), who find that Medicare is a progressive program with larger benefits accruing to the less educated. These results are also consistent with the findings of Finkelstein and McKnight (2005). They evaluate the insurance benefits of Medicare by examining changes in the risk premium of beneficiaries due to a decrease in exposure to medical expenditure risk in a static framework. They find that the per beneficiary annual welfare benefit of Medicare in the first ten years after its inception was $519 (in 2000 dollars). They argue that under conservative assumptions this benefit covers between forty five and seventy five percent of the cost of Medicare’s provision, and can be considered to be its major benefit irrespective of its effect on health. McClellan and Skinner (2006) evaluate the role of Medicare in smoothing household consumption for its beneficiaries given uncertainty about health outcomes and medical expenditures using a two period model. They find that the greatest benefit is to low income beneficiaries. Beneficiaries with income below $15,000 (in 1990 dollars) value Medicare benefits at $240 or 17% more than their dollar amount, while those with income above $50,000 value the coverage by $670 or 23% less than its dollar amount.61

In order to better understand the variation in WTP across demographic categories, the effects on changing the age of eligibility on OOP expenditures across these demographic categories is computed. The changes in OOP expenditures are computed using (20)–(22) in the same way as the WTP measures are computed. On average the OOP costs increase by $447 but again there is variation by demographic categories. This calculation may also be used

### Table 5
Willingness to pay to avoid delaying age of eligibility for Medicare to 67 (in 1991$).

<table>
<thead>
<tr>
<th>Education</th>
<th>White</th>
<th>Black</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than high school (&lt; 11 yrs)</td>
<td>28,347</td>
<td>25,235</td>
<td>21,925</td>
</tr>
<tr>
<td>High school graduate (12 yrs &lt; &amp; 16 yrs)</td>
<td>26,759</td>
<td>23,498</td>
<td>19,721</td>
</tr>
<tr>
<td>College graduate (16 yrs &lt;)</td>
<td>22,819</td>
<td>19,262</td>
<td>15,584</td>
</tr>
</tbody>
</table>

As before, a sample of 11013 individuals is simulated, i.e., 3 simulated individuals for every individual in the estimation sample of 3562 persons plus the 109 additional persons observed with complete information in at least two waves in the data (see footnote 52).


59 It is likely that the WTP would be calculated to be even higher if the analysis accounted for the benefits of Medicare on savings and retirement behavior. As mentioned this is impossible due to computational limitations.
to provide additional context to the WTP numbers reported above. First, recall that the counterfactual experiment is under the assumption that individuals would have had the same insurance plans until age 67 as are available in the status quo. If one makes the assumption that on average 20% of total medical expenditures are out of pocket and 80% are covered by Medicare then the savings on medical expenditures per capita to Medicare by delaying benefits by two years would be $1788. This is far less than the average WTP (note also that the WTP calculations account for the out of pocket medical expenditures).

\[
\Delta OOP_t = \frac{1}{N_t} \sum_{i=1}^{N_t} \left\{ \sum_{t=1}^{T} \left[ OOP(Y_{i,t,1,LM}, P_{i,t,1}) \right] 1_{H_{i,LM} > 1} \right\} - \sum_{i=1}^{N_t} \left[ OOP(Y_{i,t,1,LM}, P_{i,t,1}) \right] 1_{H_{i,LM} > 1} \right\} (20)
\]

\[
\Delta OOP = \sum_{t=1}^{T} Pr[I \mid E] \cdot \Delta OOP_t (21)
\]

\[
\Delta OOP_{E,R} = \sum_{t=1}^{T} Pr[I \mid E, R] \cdot \Delta OOP_t (22)
\]

The results are presented in Table 6, and imply that the less educated see their OOP medical expenditures go up the most when the age of eligibility is increased. This is one reason that they value avoiding the delay in eligibility for Medicare benefits the most. It should be noted that the WTP numbers in Table 5 are much larger than those in Table 6 as the former includes the value of cumulative utility from aggregate consumption, health, and other behaviors, and not just the effects of OOP expenditures through the budget constraint on aggregate consumption. This disparity is particularly large because of the big value placed on utility from health (see discussion in Section 5).

Similar values for changes in level of health status and consequently income (recall, that income in the model depends on health) are computed but not reported in the interest of brevity. The numbers are computed using analogues of the OOP calculations in (20)–(22). The average changes in health status and income over the life-cycle due to a change in the age of eligibility are small. Income is higher under the current Medicare policy by $98. The level of health status is higher by 1.08 × 10^{-4} (recall average health status in the sample is 4.4, see Table 1). This is also similar to the finding by Card et al. (2004) that Medicare coverage provides a small improvement in self-reported health status for people slightly older than 65.

In simulation results not reported it is found that for a cohort living its entire life under the current Medicare policy the survival would be 0.02% higher at age 80 compared to a cohort living under the altered policy with eligibility at age 67. The difference in mortality would occur as expected in the later stages of the life-cycle, and would impact individuals of type 1 and 2 the most. To provide some interpretation, assuming 3 million live births a year for cohorts born after 1944, Medicare in its current form would annually save approximately 600 lives of those aged 80 for this cohort by the year 2024 compared to the number of lives it would have saved if the eligible age was 67. This is consistent with the conclusions of Dow (2001) and Lichtenberg (2002) that Medicare has a significant effect in reducing mortality. Nor is it inconsistent with the findings of Card et al. (2004) that Medicare has no effect on mortality for those just older than 65. Finkelstein and McKnight (2005) find that Medicare had no significant effect in reducing elderly mortality in the first ten years of its implementation. In sum, the simulations imply that Medicare leads to small improvements in health status and reductions in mortality for the elderly.

The effects of Medicare on medical utilization and OOP expenditures over the life-cycle can be examined in more detail. Fig. 1 reports the change in proportion of individuals seeking ‘moderate’ and ‘high’ amounts of medical care when the age of eligibility is 67 compared to the status quo (note, this calculation reverses the order of the terms compared to the calculations earlier in (17) and (20)). It can be seen that at age 66 there is large drop in medical utilization. The proportion of individuals seeking either ‘high’ or ‘moderate’ levels of medical care reduces by 29% when the age of eligibility is delayed to 67 compared to the eligibility being at 65. Conversely, this suggests Medicare induces large increases in medical utilization. Furthermore, it implies that there is deferring of medical care just before individuals become Medicare eligible. This is further substantiated because the the changes in medical utilization are much smaller over the rest of the life-cycle. A similar conclusion is reached by Dow (2001), Lichtenberg (2002), and Card et al. (2004).

Fig. 1 also shows that there is an increase in the difference in OOP expenditures around the age of 66 between the policy with eligibility at 67 compared to the status quo. This spike (of 18%) corresponds to the ages when there would be no Medicare coverage if the age of eligibility were delayed. Conversely, this implies that Medicare reduces OOP costs and hence, provides insurance against medical expenditure risk. Interestingly, Finkelstein and McKnight (2005) reach a similar conclusion using very different methods and data.

Finally, the simulations also shed light on the aspect that medical utilization changes by large amounts but the corresponding changes in the level of health status are much smaller in absolute terms. A reason for this is that a large component of medical care is mitigative, i.e., purely for consumption benefits. Simultaneously, a small component is proactive, i.e., purely for consumption benefits.

Table 6 reports the negative of the numbers calculated in (20)–(22).

The finding about small effects on health status cannot be directly compared to the results of Dow (2001) that Medicare has little effect in reducing sick bed days, or those of Lichtenberg (2002) that Medicare reduces morbidity due to basic differences in how the health variable was measured in those studies in contrast to this one.

As always, appropriate caution must be exercised in making this comparison as health status is a categorical variable.

63 Table 6 reports the negative of the numbers calculated in (20)–(22).
64 The finding about small effects on health status cannot be directly compared to the results of Dow (2001) that Medicare has little effect in reducing sick bed days, or those of Lichtenberg (2002) that Medicare reduces morbidity due to basic differences in how the health variable was measured in those studies in contrast to this one.
65 As always, appropriate caution must be exercised in making this comparison as health status is a categorical variable.
66 There were 2.86 million live births in 1945. Live births peaked at 4.27 million in 1961 and were about 4.02 million 2002 (source: 2002 Vital Statistics).
67 The simulations assume that Medicare exists from the age of 22 onwards for a cohort.
68 The findings in this paper are consistent with Dow’s findings for males and blacks, for whom he estimates reductions in mortality across various specifications.
69 Two potential reasons, among others, for this could be that the results of this paper are based on (i) a sample of males, who are assumed to have Medicare coverage from the age of 22 onwards, and (ii) data from the 1990s. This is not the case for the analysis by Finkelstein and McKnight.
70 Finkelstein and McKnight estimate the changes in the distribution of out of pocket medical expenditures of the elderly pre and post medicare from the Survey of Health Service Utilization and Expenditures (SHSUE) data from 1963 and 1970 using a “differences in differences” approach.
Fig. 1. Effect of delaying Medicare eligibility to 67 on medical utilization and OOP expenditures.

71 Not reported here but available from the author on request.
72 This simulation is done under the assumption that $\sigma_{16} > 0$, $\eta_{4} = 0$, and $\eta_{9} = 0$.
73 The HRS data is not well suited to assess the impact of Medicare on changes in technology.
Table 7

Health transition parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Health Status</th>
<th>Dead (q = 1)</th>
<th>Poor (q = 2)</th>
<th>Fair (q = 3)</th>
<th>Good (q = 4)</th>
<th>Very good (q = 5)</th>
<th>Excellent (q = 6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\eta_{1,q}$</td>
<td>Lagged health</td>
<td>$-1.4 \times 10^{-3}$</td>
<td>$0.5 \times 10^{-3}$</td>
<td>$-0.4 \times 10^{-4}$</td>
<td>$-1.4 \times 10^{-4}$</td>
<td>$-1.6 \times 10^{-3}$</td>
<td>$-2.1 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\eta_{2,q}$</td>
<td>Lagged health-quadratic term</td>
<td>$-0.6 \times 10^{-4}$</td>
<td>$-0.4 \times 10^{-4}$</td>
<td>$-0.6 \times 10^{-4}$</td>
<td>$-0.6 \times 10^{-4}$</td>
<td>$-0.4 \times 10^{-3}$</td>
<td>$-0.3 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\eta_{3,q}$</td>
<td>Moderate sickness</td>
<td>0</td>
<td>$-0.2 \times 10^{-3}$</td>
<td>$-0.6 \times 10^{-3}$</td>
<td>$-0.07 \times 10^{-4}$</td>
<td>$-0.2 \times 10^{-3}$</td>
<td>$-0.06 \times 10^{-3}$</td>
</tr>
<tr>
<td>$\eta_{4,q}$</td>
<td>Curative medical-care-moderate-sickness</td>
<td>0</td>
<td>$-0.04 (0.002)$</td>
<td>$0.02 (0.0009)$</td>
<td>$-0.001 (2 \times 10^{-5})$</td>
<td>$0.001 (2 \times 10^{-5})$</td>
<td>$0.001 (1 \times 10^{-5})$</td>
</tr>
<tr>
<td>$\eta_{5,q}$</td>
<td>Multiplicative effects-high sickness</td>
<td>0</td>
<td>$0.0007 (3 \times 10^{-4})$</td>
<td>$-0.003 (3 \times 10^{-4})$</td>
<td>$-0.004 (2 \times 10^{-4})$</td>
<td>$-0.006 (5 \times 10^{-5})$</td>
<td>$-0.008 (9 \times 10^{-5})$</td>
</tr>
<tr>
<td>$\eta_{6,q}$</td>
<td>Alcohol-consumption</td>
<td>0</td>
<td>$-0.02 (3 \times 10^{-4})$</td>
<td>$-0.02 (6 \times 10^{-4})$</td>
<td>$-0.02 (5 \times 10^{-5})$</td>
<td>$-0.03 (8 \times 10^{-5})$</td>
<td>$-0.03 (3 \times 10^{-5})$</td>
</tr>
<tr>
<td>$\eta_{7,q}$</td>
<td>Smoking</td>
<td>0</td>
<td>$0.001 (0.0012)$</td>
<td>$0.001 (6 \times 10^{-5})$</td>
<td>$0.003 (1 \times 10^{-5})$</td>
<td>$0.005 (1 \times 10^{-5})$</td>
<td>$0.001 (4 \times 10^{-5})$</td>
</tr>
<tr>
<td>$\eta_{8,q}$</td>
<td>Exercise</td>
<td>0</td>
<td>$0.0002 (3 \times 10^{-4})$</td>
<td>$0.003 (5 \times 10^{-4})$</td>
<td>$0.001 (5 \times 10^{-4})$</td>
<td>$0.006 (0.001)$</td>
<td>$0.001 (3 \times 10^{-4})$</td>
</tr>
<tr>
<td>$\eta_{9,q}$</td>
<td>Preventive medical care</td>
<td>0</td>
<td>$1.8 (0.09)$</td>
<td>$2.1 (0.06)$</td>
<td>$4.6 (0.01)$</td>
<td>$6.5 (0.02)$</td>
<td>$7.6 (0.03)$</td>
</tr>
<tr>
<td>$\eta_{10,q}$</td>
<td>Type 1</td>
<td>0</td>
<td>$2.0 (0.07)$</td>
<td>$1.7 (0.05)$</td>
<td>$4.7 (0.004)$</td>
<td>$6.8 (0.01)$</td>
<td>$7.6 (0.02)$</td>
</tr>
<tr>
<td>$\eta_{11,q}$</td>
<td>Type 2</td>
<td>0</td>
<td>$2.1 (0.1)$</td>
<td>$2.0 (0.09)$</td>
<td>$4.7 (0.07)$</td>
<td>$6.5 (0.02)$</td>
<td>$7.3 (0.04)$</td>
</tr>
</tbody>
</table>
### Table 8
Log-OOP process parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>(Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu_1$ (No insurance)</td>
<td>3.21</td>
<td>(0.005)</td>
</tr>
<tr>
<td>$\mu_2$ (Group plan)</td>
<td>1.02</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\mu_3$ (Personal plan)</td>
<td>1.29</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\mu_4$ (VA/Champus plan)</td>
<td>3.19</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\mu_5$ (Group/VA/Champus plan)</td>
<td>3.16</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\mu_6$ (Group/Personal plan)</td>
<td>2.98</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\mu_7$ (Multiplicative factor-high treatment)</td>
<td>1.09</td>
<td>(0.003)</td>
</tr>
</tbody>
</table>

Insurance choices at 65 and over:

| $\mu_8$ (Medicare) | 0.88 | (0.04) |
| $\mu_9$ (Medicare/Medigap/Personal plan) | 0.68 | (0.08) |
| $\mu_{10}$ (Medicare/Group/VA/Champus plan) | 0.41 | (0.05) |
| $\mu_{11}$ (Group/VA/Champus plan) | 1.29 | (0.05) |

Other variables:

| $\mu_{12}$ (Health) | −0.12 | (0.002) |
| $\mu_{13}$ (Household size) | 0.12 | (0.003) |
| $\mu_{14}$ (Age) | 1.77 | (0.002) |

### Table 9
Probability of zero OOP process parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>(Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_1$ (Constant)</td>
<td>−2.3</td>
<td>(0.05)</td>
</tr>
<tr>
<td>$\gamma_2$ (Health)</td>
<td>0.04</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\gamma_3$ (Household size)</td>
<td>−0.1</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

Insurance choices under age 65:

| $\gamma_4$ (Group plan) | 0.72 | (0.02) |
| $\gamma_5$ (Personal plan) | 0.6 | (0.1) |
| $\gamma_6$ (VA/Champus plan) | 0.45 | (0.16) |
| $\gamma_7$ (Group/VA/Champus plan) | 0.39 | (0.14) |
| $\gamma_8$ (Group/Personal plan) | 0.55 | (0.17) |
| $\gamma_9$ (Multiplicative factor-high treatment) | 0.49 | (0.06) |

Insurance choices at 65 and over:

| $\gamma_{10}$ (Medicare) | 0.95 | (0.06) |
| $\gamma_{11}$ (Medicare/Medigap/Personal plan) | 0.36 | (0.12) |
| $\gamma_{12}$ (Medicare/Group/VA/Champus plan) | 0.41 | (0.08) |
| $\gamma_{13}$ (Multiplicative factor high treatment) | 0.74 | (0.12) |

### Table 10
Variance and kernel smoothing aitchison-aitken bandwidth parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>(Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_1$ (Insurance choice)</td>
<td>5069</td>
<td>(15.4)</td>
</tr>
<tr>
<td>$\rho_2$ (Habits choice)</td>
<td>889</td>
<td>(1.8)</td>
</tr>
<tr>
<td>$\rho_3$ (Medical care choice)</td>
<td>1209</td>
<td>(9.8)</td>
</tr>
<tr>
<td>$\sigma_y$ (Income process)</td>
<td>0.86</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\sigma_{oop}$ (OOP costs process)</td>
<td>0.7</td>
<td>(0.006)</td>
</tr>
<tr>
<td>$\sigma_\epsilon$ (Measurement error in income)</td>
<td>0.2</td>
<td>(0.07)</td>
</tr>
<tr>
<td>$\sigma_\mu$ (Measurement error in OOP costs)</td>
<td>0.22</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>

A-A Bandwidth for the initial distribution of:

| $\omega_1$ (Health) | 0.47 | (0.02) |
| $\omega_2$ (Insurance) | 0.02 | (0.12) |
| $\omega_3$ (Household size) | 0.21 | (0.04) |
| $\omega_4$ (Exercise) | 2.61 × 10^{-6} | (2.3 × 10^{-6}) |
| $\omega_5$ (Smoking) | 0.03 | (0.12) |
| $\omega_6$ (Alcohol consumption) | 0.47 | (0.03) |

### Table 11
Income process parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate</th>
<th>(Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\kappa_2$ (Lagged Health)</td>
<td>0.18</td>
<td>(0.002)</td>
</tr>
<tr>
<td>$\kappa_3$ (Age)</td>
<td>0.04</td>
<td>(1.4 × 10^{-4})</td>
</tr>
<tr>
<td>$\kappa_4$ (Age-quadratic term)</td>
<td>−0.001</td>
<td>(3.5 × 10^{-7})</td>
</tr>
<tr>
<td>$\kappa_{1,1}$ (Constant)</td>
<td>10.1</td>
<td>(0.02)</td>
</tr>
<tr>
<td>$\kappa_{1,2}$ (Type 1)</td>
<td>10.5</td>
<td>(0.01)</td>
</tr>
<tr>
<td>$\kappa_{1,3}$ (Type 2)</td>
<td>9.9</td>
<td>(0.02)</td>
</tr>
</tbody>
</table>
Table 12
Household size transition parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“One member” (q = 1)</td>
</tr>
<tr>
<td>$\psi_{1,e}$</td>
<td>Constant</td>
</tr>
<tr>
<td>$\psi_{1,e}$</td>
<td>Lagged HHS</td>
</tr>
<tr>
<td>$\psi_{1,e}$</td>
<td>Lagged HHS quadratic term</td>
</tr>
<tr>
<td>$\psi_{1,e}$</td>
<td>Age</td>
</tr>
<tr>
<td>$\psi_{1,e}$</td>
<td>Age-quadratic term</td>
</tr>
</tbody>
</table>

Table 13
Sickness transition parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Low (q = 1)</td>
</tr>
<tr>
<td>$\phi_{1,e}$</td>
<td>Constant</td>
</tr>
<tr>
<td>$\phi_{2,e}$</td>
<td>Health</td>
</tr>
<tr>
<td>$\phi_{3,e}$</td>
<td>Age</td>
</tr>
</tbody>
</table>

Table 14
Unobserved heterogeneity distribution parameters.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Estimate (Asy. S.E.)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>“Type 1” (I = 1)</td>
</tr>
<tr>
<td>$\lambda_{1,J}$</td>
<td>Constant</td>
</tr>
<tr>
<td>$\lambda_{2,J}$</td>
<td>Education</td>
</tr>
<tr>
<td>$\lambda_{3,J}$</td>
<td>Race</td>
</tr>
</tbody>
</table>

References

Finkelstein, A., McKnight, R. 2005. What did medicare do (and was it worth it)? Working Paper, Society of Fellows, Harvard University.


