A Life Cycle Analysis of the Effects of Medicare on Individual Health Incentives and Health Outcomes

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

Abstract

Medicare is the largest health insurance program in the U.S. This paper analyzes the effects of Medicare on individual health incentives and outcomes resulting from the dynamic trade-off induced by insurance between moral hazard in health related choices and mortality risk. 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. Model simulations imply that Medicare benefits elderly through insurance against medical expenditures, small improvements in health status and reduced mortality. However Medicare induces increased medical utilization, and small levels of moral hazard in alcohol consumption, smoking and exercise among elderly. Medicare also has a small effect in counteracting moral hazard in alcohol consumption, smoking and exercise for younger individuals by inducing self-protection. Overall, Medicare improves *life time* utility, with individuals willing to pay between $1,530.36 and $3,574.76 (in 1991 dollars) annually for its coverage. Additional simulations show that reforming Medicare by increasing the eligible age will lower insurance against medical expenditures, medical utilization, and health outcomes over the life cycle, and vice versa.

Keywords: Dynamic discrete choice, Medicare, moral hazard, mortality, human capital.

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

*I am grateful to Mike Keane for his numerous suggestions, and to Roger Feldman for an observation that led to this paper. This paper has also benefited from the comments of Victor Aguirregabiria, Will Dow, Ron Gallant, Donna Gilleskie, Han Hong, Tom Mroz, John Rust, Jon Skinner, Frank Sloan, Wilber van der Klaauw and Ken Wolpin, and seminar participants at the Econometric Society Winter Meetings, International Conference on Public Economic Theory, Conference on Improving Social Insurance Programs, Fuqua Finance Brown Bag Seminar, NBER Health Care Program Meeting, SUNY-Stony Brook, Triangle Health Economics Workshop, University of Maryland, and Federal Reserve Board. I acknowledge the use of the HRS data and the help of its staff and the support of the Minnesota Super Computer Institute. All errors are mine alone.

We hope that never again will our aged be threatened with economic ruin whenever illness strikes them. Medical care will free millions from their miseries. It will signal a deep and lasting change in the American way of life. It will take its place beside Social Security, and together they will form the twin pillars of protection upon which all our people can safely build their lives and their hopes.

–President Lyndon B. Johnson, remarks to the National Council of Senior Citizens, June 3rd, 1966.

1 Introduction

Medicare is the largest health insurance program in the U.S.[1] According to the Congressional Budget Office the expenditures on Medicare benefits were $212 million in 1999, which was 12.4% of federal outlays or 2.4% of GDP. In 2000 Medicare covered about 39.9 million or 13.9% of the U.S. population of which about 34.4 million were aged 65 or older. The primary aims of Medicare are to insure the elderly against medical expenditure risk and increase their access to medical care to improve health outcomes. However provision of insurance involves a trade-off between its benefits and distorted incentives (e.g., Arrow 1963, Pauly 1968, Manning et al 1987, Manning and Marquis 1996).[2] Moreover, individual decisions about medical utilization and health related behaviors, and the consequent health outcomes (Grossman 1972) are inter-related over the life cycle. This implies that the trade-off induced by insurance is dynamic and exists over the life cycle (Philipson and Becker 1998), e.g., an increase in unhealthy behaviors like smoking due to moral hazard from insurance coverage might elevate future mortality risk, which might in turn affect current health behaviors (e.g., Adda and Lechene 2001). Hence the effects of Medicare on health behaviors and outcomes are potentially dynamic, of a life cycle nature, and may undermine its desired benefits.

This paper analyzes the effects of the Medicare program on individual health incentives and outcomes resulting from the dynamic trade-off induced by insurance between moral hazard in health related choices and mortality risk. Such an analysis of Medicare also sheds light on the dynamic effects of contracts on individual incentives and outcomes, which is a subject of considerable interest to economists with little

[1] Medicare was enacted in 1965 as Title XVIII of the Social Security Act designated "Health Insurance for the Aged and Disabled" and introduced in 1966. It covers individuals who are 65 and older and those with certain disabilities, and consists of three parts. Hospital Insurance or part A is provided automatically and free of premiums to all eligible individuals. Any person or spouse of a person who is 65 or older, a citizen or permanent resident of the U.S., and has been employed in a Medicare covered job for 10 years is eligible. All individuals who are entitled to part A coverage are eligible to enroll for Supplementary Medical Insurance or part B on a voluntary basis. Part B covers certain physician and medical services not covered under part A. In 1997 under the Balanced Budget Act, Medicare + Choice or part C was established and it provided options for Medicare beneficiaries to enroll in certain private sector health care plans. In 1973 individuals entitled to Social Security or Railroad Retirement disability cash benefits for at least 24 months, and most individuals with end stage renal disease became eligible for Medicare coverage. In 2000 individuals with Amyotrophic Lateral Sclerosis or Lou Gehrig's disease were allowed to waive the 24 month waiting period. On 8th December 2003 the Medicare Prescription Drug, Improvement and Modernization Act was passed by the U.S. Congress. Under this act the Medicare + Choice HMO plans have been renamed Medicare Advantage. Detailed information about benefits is available at www.medicare.gov.

[2] An excellent review of this literature is found in Cutler and Zeckhauser (2000).

accompanying empirical research (Salanié 2002). 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). Simulations from the model are used to infer the effects of Medicare on medical utilization, out of pocket medical expenditures, health related behaviors and health outcomes by comparing outcomes under its coverage to those in a counter-factual situation in which individuals lack access to Medicare. Given the adverse financial implications of imminent demographic changes[3] there have been various proposals for Medicare reform.[4] Additional simulations are used to infer the effects of one particular reform proposal, i.e., changing the age of eligibility for Medicare.[5]

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.[6] Thus it may alter survival incentives for those under 65 like it alters employment incentives (Rust and Phelan 1997). For example, 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. Individuals may also be led to "stockpile" medical care prior to age 65, e.g., delay expensive treatments like coronary artery bypass graft. Alternatively, 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). 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 life time welfare effects of the program. This may be different from

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

[4]The call for Medicare reform goes back to the 1983 National Commission on Social Security Reform, often called the Greenspan Commission. Cutler (2000), Fuchs (2000), McClellan (2000), Reinhardt (2000) and Saving (2000) provide an excellent discussion of the proposed reforms.

[5]This idea has also been discussed in the popular press, e.g., "New Retirement Age Is Needed To Head Off Fiscal Train Wreck," Alan Murray, Wall Street Journal, July 8th, 2003.

[6]See footnote 1 for the exceptions.

the welfare effects on the elderly per se because Medicare is an inter-generational transfer of resources.

Medicare is administered by the federal government and implemented almost uniformly across the U.S. with little changes in its coverage since its inception.[7] Thus its effects cannot be analyzed by exploiting cross state and inter-temporal variation as often done for Medicaid.[8] Hence a simulation based analysis is employed. McClellan and Skinner (1999) use a similar strategy to 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 effects on health outcomes or health related behaviors, nor do they look at 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 Klaauw and Wolpin (2005). Due to the objectives of the Social Security program the analysis has focussed on examining the effects on employment and savings decisions. However given Medicare's objectives it can be argued that its effects on health behaviors and outcomes are of equal if not more importance and worthy of examination in their own right, as is the focus of this paper.

A small recent literature examines the effects of Medicare on health behaviors and outcomes. Skinner, Fisher and Wennberg (2001) estimate the effect of intensity of Medicare expenditures on survival.[9] Dow (2001) and Lichtenberg (2002) examine the effects on utilization, morbidity and mortality. Card, Dobkin and Maestas (2004) analyze the effect on medical utilization, self-reported health, mortality, smoking, exercise and obesity. Decker (2005) examines effects on medical utilization among women, especially mammography and the chances of early detection of breast cancer. Finkelstein and McKnight (2005) examine effects on mortality and medical expenditures. Though insightful, this literature is limited by its reliance on aggregate or repeated cross section data, e.g., it cannot account for individual specific unobserved heterogeneity in estimating the effects of Medicare.

This paper extends current research in five 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 behaviors and health outcomes (in particular endogenous mortality) in assessing the effects of

[7]Sometimes Medicare is called a state of the art insurance program for 1965. The largest expansion of Medicare enacted in December 2003, primarily to provide prescription drug benefits, would be difficult to justify as an *exogenous* policy change. In any case the changes have not been completely implemented, so data is not yet available to analyze their effects.

[8]Medicaid is a means based welfare program that is regulated by states. There is considerable cross state variation in its benefits and it has also undergone various expansions since its inception which can be used to estimate its effects, e.g., see Currie and Gruber (1996).

[9]Feldstein (1971) also used aggregate pooled data for the period 1st July 1966 - 30th June 1968 to examine the economic efficiency of Medicare.

Medicare on health incentives and outcomes, (ii) it controls for individual specific unobserved heterogeneity in the analysis, (iii) it examines the role of Medicare in insuring against medical expenditure risk, (iv) it examines the effects of Medicare on health related behaviors other than medical utilization that affect health outcomes, and which could undermine Medicare's objectives, and (v) it calculates the life time willingness to pay for Medicare accounting for the trade-offs inherent in insurance provision, and the differences in insurance costs and mortality in the absence of Medicare.

Model simulations imply that Medicare benefits the elderly through insurance against medical expenditure risk, small increases in health status and reductions in mortality. However Medicare induces increases in medical utilization, and small levels of moral hazard in alcohol consumption, smoking and exercise among the elderly. Medicare also has a small effect in counteracting moral hazard in alcohol consumption, smoking and exercise for the younger individuals by inducing self-protection (Ehrlich and Becker 1972) at younger ages. Medicare improves *life time* utility accounting for insurance costs and mortality. On average, individuals are willing to pay between \$1,530.36 and \$3,574.76 (in 1991 dollars) annually for Medicare coverage. Simulations show that reforming Medicare by increasing the eligible age will lower insurance against medical expenditures, medical utilization, and health outcomes with a consequent reduction in welfare, and vice versa. 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 examines the effects of Medicare including its life time welfare impact. Section 7 analyzes the effects of changing the age of eligibility and section 8 concludes.

2 The Model

2.1 Structure and Specification

For the purpose of analyzing the effects of Medicare 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, 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) 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,

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

subject to

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

$$X_t = Y_t - P_I - OOP_t \tag{3}$$

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

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

$$s_t = s(H_t, A_t, \epsilon_t^s) \tag{6}$$

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

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

The model accounts for the moral hazard induced by Medicare 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").[10]

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. 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 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."}

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

These alternatives are based on the empirical distribution of insurance choices in the HRS.[11] 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 coverage[12] 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 can not be uninsured, while plans 2 and 3 are defined as a mix of Medicare and other coverage as stated.

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. 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. 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 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[13] 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_t, age, A_t, and lagged variables in the utility function (1) are explained later in section 2.3.

The budget constraint (eq. 3) determines the consumption of the composite commodity through the difference between income, Y_t, and the insurance premium,[14] P_I, and the out of pocket (OOP) medical

[11] 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.

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

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

[14] The restrictive assumption that insurance premiums are constant over time is adopted because premium information is

costs, OOP_t.[15] Income is measured as the sum of wage and non-wage income. Income (eq. 4) is determined endogenously with the law of motion, $\log(Y_t) = \overline{Y_t} + \epsilon_t^y$, where $\epsilon_t^y \overset{\text{IID}}{\sim} N(0, \sigma_y^2)$ and $\overline{Y}_t$ 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.[16]

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, ϵ_t^H. Empirically health stock, H_t, is defined using two measures of health, i.e., (a) self-reported health status[17] (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 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 time t, $t = 1, \ldots, T$, given by,

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

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.[18] 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 respectively, $\eta_{3,q}$ 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., $s_{t-1} = 3$). 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

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

[16]The econometric specification is not presented, here and elsewhere, when it can easily be inferred through the tables presenting the results, i.e., tables 16- 22.

[17]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.

[18]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.

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 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 affect 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[19] 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 linear function of insurance status, medical utilization, health, household size and age, i.e., if $OOP_t > 0$, $\log(OOP_t) = \overline{OOP_t} + \epsilon_t^{oop}$ where $\epsilon_t^{oop} \overset{\text{IID}}{\sim} N(0, \sigma_{oop}^2)$. It is also assumed that the random components of OOP 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. 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 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.[21] Individuals can however save implicitly by investing in their health stock, which is the only asset in the model.

[19]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 2.3.

[20]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.

[21]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 7 in section 5.2).

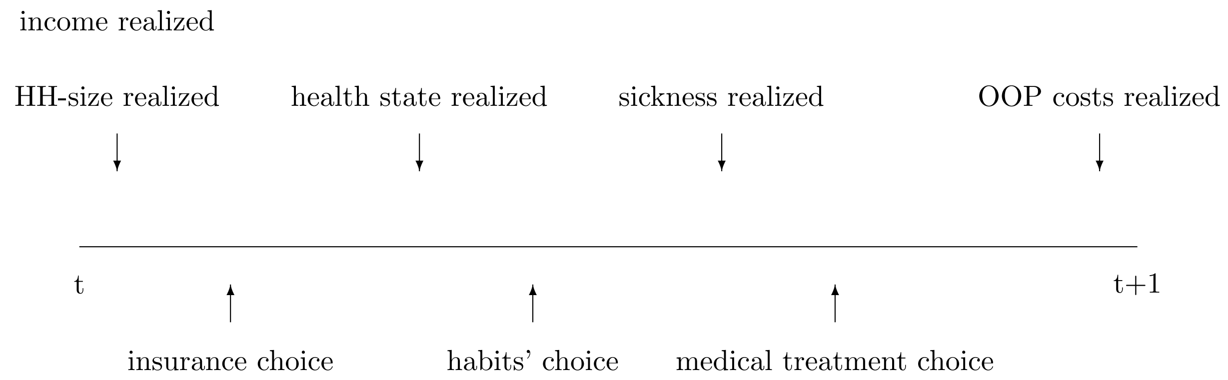

Figure 1: Timing of Per-period Decisions

2.2 The Timing Convention

The decisions of the individuals are modelled from ages 22 to 80.[22] 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 (fig. 1).

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 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 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.

2.3 Per Period Utility Functions

In the first sub-period an individual makes an insurance choice $i \in INS$.[23] 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

[22]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.

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

deterministic component[24] of the utility function takes the form

$$U_{(i)}(I_t, A_t) = \alpha_1 \cdot 1\{I_{t-1} \neq I_t^i\} \cdot 1\{A_t < 65\} + \alpha_2 \cdot 1\{I_{t-1} \neq I_t^i\} \cdot 1\{A_t \geq 65\} \ . \tag{9}$$

The term I_t^i 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).[25] The ζ_t^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 $(\rho_1\gamma, \frac{\pi^2\rho_1^2}{6})$, 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. 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., $U^*_{(j)}(H_t, a_t^j, c_t^j, e_t^j, A_t, \zeta_t^j) = U_{(j)}(H_t, a_t^j, c_t^j, e_t^j, A_t) + \zeta_t^j$. 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]

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

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

[24] 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.

[25] These are the only utility parameters that vary by age.

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

[27] The notation (a_t^j, c_t^j, e_t^j) 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.

[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.

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_t^j 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, $U^*_{(k)}(Y_t, P_I, OOP_t, s_t, m_t^k, HHS_t, \zeta_t^k) = U_{(k)}(Y_t, P_I, OOP_t, s_t, m_t^k, HHS_t) + \zeta_t^k$, 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,

$$\begin{aligned} U_{(k)}(Y_t, P_I, OOP_t, s_t, m_t^k, HHS_t) \quad = \quad & \frac{\alpha_{13}\cdot(Y_t - P_I - OOP_t)}{(HHS_t)^{\alpha_{19}}} + \frac{\alpha_{14}\cdot(Y_t - P_I - OOP_t)^2}{(HHS_t)^{2\alpha_{19}}} \\ & +\alpha_{15}\cdot 1\{m_t^k > 1\}\cdot(m_t^k)^2 + \Big[\alpha_{16}\cdot(m_t^k) - \alpha_{17}\Big]\cdot[1\{s_t = 2\}] \\ & +\alpha_{18}\cdot\Big[\alpha_{16}\cdot(m_t^k) - \alpha_{17}\Big]\cdot[1\{s_t = 3\}] \quad . \end{aligned} \tag{11}$$

The function is quadratic in consumption, with the coefficients α_{13} and α_{14}, to allow for risk aversion.[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

[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.

[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 5.2 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. The primary reason for this was that Quadrature as opposed to Monte Carlo integration was employed for the numerical solution of the model as given the model specification the former was found to be faster. 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.

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. 11 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 mitigated by the amount $[\alpha_{18} \cdot \alpha_{16}]$ in the case of high sickness. The term 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 $(\rho_3\gamma, \frac{\pi^2\rho_3^2}{6})$.

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[34] in period T and defined as,

$$U^*_{T+1}(H_T, s_T) = \alpha_{20} \cdot (H_T) + \alpha_{21} \cdot (s_T) + \alpha_{22} \cdot (H_T/s_T). \tag{12}$$

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).[35] The dynamic programming and the associated Bellman equations are described for a related model in Khwaja (2001). There is no closed form representation of the solution to the dynamic programming problem for the model. 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.[36] 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.[37] Without the

[34]The terminal value function was parsimoniously defined to include just three arguments to avoid over-fitting the model.
[35]The model does not suffer from the usual Independence of Irrelevant Alternatives (IIA) limitation because of its dynamic structure (Rust 1994a, p. 139).
[36]Blacks, Hispanics, and residents of the state of Florida were over sampled.
[37]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/.

Table 1: Data Summary (*two year* values in 1991 dollars)

Variable	Obs. (col. %)	Mean (Std. Dev.)
Age	9,321	59.11 (4.91)
Education[a]	9,321	12.45 (3.23)
Race	9,321	1.29 (0.63)
1-“white”	7,439 (79.81%)	
2-“black”	993 (10.65%)	
3-“other”	889 (9.54%)	
Health	9,321	4.44 (1.18)
1-“dead”	193 (2.07%)	
2-“poor”	407 (4.37%)	
3-“fair”	1,158 (12.42%)	
4-“good”	2,768 (29.70%)	
5-“very good”	2,916 (31.28%)	
6-“excellent”	1,879 (20.16%)	
Alcohol	9,128	1.86 (0.73)
1-“none”	3,170 (34.73%)	
2-“drink/day $\leqq 1$ ”	4,034 (44.19%)	
3-“drink/day > 1”	1,924 (21.08%)	
Smoking	9,128	1.41 (0.76)
1-“none”	6,938 (76.01%)	
2-“pack/day < 1 ”	669 (7.33%)	
3-“pack/day $\geqq 1$”	1,521 (16.66%)	
Exercise	9,321	1.31 (0.46)
1-“no”	6,263 (68.61%)	
2-“yes”	2,865 (31.39%)	
Medical care	9,128	1.82 (0.8)
1-“low”	3,878 (42.48%)	
2-“moderate”	2,976 (32.60%)	
3-“high”	2,274 (24.91%)	
Household (HH) size	9,321	2.33 (0.84)
HH-income ($)	9,128	10,6643.6 (10,5525.1)
HH-OOP costs ($)	4,276	3,069.039 (9,607.85)
Insurance (age $\leqq 64$)	8,062	2.18 (0.97)
1-“none”	1,041 (12.91%)	
2-“group”	5,870 (72.81%)	
3-“personal”	468 (5.81%)	
4-“VA/Champus”	247 (3.06%)	
5-“group/VA/Champus” (%)	244 (3.03%)	
6-“group/personal”	192 (2.38%)	
Insurance (age $\geqq 65$)	1,143	2.25 (0.90)
1-“Medicare”	352 (30.80%)	
2-“Medicare/Medigap/personal”	147 (12.86%)	
3-“Medicare/group”	644 (56.34%)	
Insurance cost reported by uninsured ($)	629	215.0 (1086.77)
Insurance premiums (age $\leqq 64$)		
Plan 2 (group) ($)	3,651	2,072.0 (3,661.7)
Plan 3 (personal) ($)	305	8,561.27 (6,888.16)
Plan 4 (VA/Champus) ($)	169	677.74 (1,613.64)
Plan 5 (group/VA/Champus) ($)	144	2,209.33 (6,566.13)
Plan 6 (group/personal) ($)	134	4,564.5 (6,543.99)
Insurance premiums (age $\geqq 65$)		
Plan 1 (Medicare) ($)	361	1,861.23 (2,544.51)
Plan 2 (Medicare/Medigap/personal) ($)	115	5,016.19 (5,055.49)
Plan 3 (Medicare/group) ($)	515	2,025.61 (2,532.37)

[a] Education-years of formal schooling

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.[38]

The sample ranges between ages 50 and 75.[39] The physical exercise variable includes activities such as sports, gardening, physical labor at work, household cleaning and maintenance etc. As stated earlier, the medical utilization variable is created by classifying an individual's total visits over two years to a physician or medical facility at *each* wave[40] in to one of three discrete categories, i.e., low, moderate or high. One advantage of wave specific categorization is that it accounts for time specific effects in demand, e.g., due to an epidemic.

The OOP expenditure data is limited to waves 2 and 3.[41] 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 \$8,561.27 and the VA/Champus plans are the least costly at \$677.74; (b) for individuals older than 65, Medicare/Medigap/personal plans are the most expensive at \$5,016.19 and the plain Medicare plan is least costly at \$1,861.23. 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.

[38] It is imperative that future research examine the effects of Medicare on females.

[39] 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.

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

[41] I am grateful to Dan Hill at HRS for making this data available to me.

Given the two year sampling period, the respondents are observed either at a sequence of odd or even numbered ages. Since 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 likelihood function that nests the solution to the dynamic programming problem (see e.g., Rust 1987, Gilleskie 1998). To illustrate, assume for simplicity that there are no sub-periods and the pre-determined state space at time t is given by $Z_t \equiv \{HHS_{t-1}, I_{t-1}, H_{t-1}, a_{t-1}, c_{t-1}, e_{t-1}, s_{t-1}, m_{t-1}\}$.[42] 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 $\{z_{n,t}\}_{t=t_{n,0}}^{T_n}$ except sickness[43], where $t_{n,0}$ is the initial observation and $T_n \leq T$ is the last observation for the individual, and (ii) the sequence of choices for insurance, alcohol consumption, smoking, exercise and medical treatment, and outcomes related to health, household size, income and OOP costs, i.e., $\{I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, H_{n,t}, HHS_{n,t}, Y_{n,t}, OOP_{n,t}\}_{t=t_{n,0}}^{T_n}$. 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. Using these the sample likelihood can be written as

$$\mathcal{L} = \prod_{n=1}^{N} \prod_{t=t_{n,0}}^{T_n} \Bigg(P\left[I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, H_{n,t}, HHS_{n,t}, Y_{n,t}, OOP_{n,t} \mid z_{n,t}, \Theta\right] \cdot P\left[z_{n,t} \mid \Theta\right] \Bigg). \tag{13}$$

Maximum likelihood estimation does not impose the requirement 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 1987, p. 588-589). In particular if OOP costs were observed regardless of treatment for all the *four* waves then the parameters of its distribution could be consistently estimated directly from

[42] 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.

[43] Sickness is treated as a latent variable in estimating the model and is integrated out from the likelihood function.

the data. However the OOP cost data suffer from the (standard) problems of selection and censoring.[44] Thus the OOP cost parameters are estimated jointly with the other parameters of the model. Simulations from the model provide a means of correcting for selection and censoring in the estimation procedure (Heckman 1979).[45] Hence even in the absence of data on OOP costs of all individuals [46] the parameters of the OOP cost distribution can be estimated consistently. To account for selection 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 unobservable 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 π_l. The estimation allows for unobservable differences in the health technology, i.e., $\eta_{10,l}$, the preferences for health, i.e., $\alpha_{3,l}$ and the income earning ability, i.e., $\kappa_{1,l}$.[47] 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 technology 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 sample likelihood as follows,

$$\mathcal{L} = \prod_{n=1}^{N} \sum_{l=1}^{L} \pi_l \left(\prod_{t=t_{n,0}}^{T_n} P\left[I_{n,t}, a_{n,t}, c_{n,t}, e_{n,t}, m_{n,t}, H_{n,t}, HHS_{n,t}, Y_{n,t}, OOP_{n,t} \mid z_{n,t}, \Theta_l, l\right] \cdot P\left[z_{n,t} \mid \Theta_l, l\right] \right). \tag{14}$$

It is assumed that the probability that an individual is a particular type, π_l, is a multinomial logit in education and race with parameters λ_l. Therefore the modified sample likelihood (eq. 14) is maximized with respect to the parameters $\overline{\Theta}_l$, where $\overline{\Theta}_l \equiv \{\Theta_l, \lambda_l\}_{l=1}^{L}$. The estimation also allows for additive measurement errors in income $\epsilon_\kappa \overset{\text{IID}}{\sim} N(0, \sigma_\kappa^2)$ (see e.g., Wolpin 1987) and in the OOP expenditures $\epsilon_\mu \overset{\text{IID}}{\sim} N(0, \sigma_\mu^2)$.

[44] 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 depending on the insurance coverage.

[45] 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.

[46] There are 4276 observations on OOP costs suggesting no cause for concern about the size of the sub-sample (table 1).

[47] This is the constant in the income function

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 correlated with his history, and in particular with his initial period choices and outcomes. Thus there is dynamic selection in to the behaviors and consequently in survival leading to an initial conditions problem.[48]

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).[49] To illustrate, assume that all the individuals are first observed at the age of 50 or $t = 15$.[50] The likelihood is modified to

$$\mathcal{L} = \prod_{n=1}^{N} \sum_{l=1}^{L} \pi_l \left(\prod_{t=15}^{T_n} P\left[I_{n,15}, a_{n,15}, c_{n,15}, e_{n,15}, m_{n,15}, H_{n,15}, HHS_{n,15}, Y_{n,15}, OOP_{n,15} \mid z_{n,15}, \overline{\Theta}_l, l\right] \cdot P\left[z_{n,15} \mid \overline{\Theta}_l, l\right] \right) \tag{15}$$

The likelihood calculation includes the term $P\left[z_{n,15} \mid \overline{\Theta}_l, l\right]$ which corrects for the initial conditions, dynamic selection and 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. 8) solely through the intercept term which represents the unobserved health production type of each individual, i.e., $\eta_{10,q}$, $q = 1, \ldots, 6$. 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, $P^*\left[z_{n,15} \mid \overline{\Theta}_l, l\right]$, where "$*$" denotes that these simulated probabilities may not be "continuous," i.e., there is a possibility that a given hyper-cell 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.[51] The "smoothened" sample likelihood is maximized using a Newton-Raphson hill climbing algorithm that employs the BHHH 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) for additional details for a related model. The asymptotic standard errors are also calculated using the BHHH method.

[48] 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.

[49] Wooldridge (2000) provides an alternative solution.

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

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

4.4 Identification

The model is identified[52] through a combination of economic and statistical assumptions, and exogenous variation in the data.[53] One important source of identification is the exogenous change in the set of insurance choices through eligibility for Medicare at age 65.[54] The identification assumption is that eligibility for Medicare at age 65 is independent of past health outcomes or health related behaviors of an individual.[55] 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 ages.[56] 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 elements of the model (e.g., preference shocks) are IID controlling for time invariant unobserved heterogeneity (Heckman and Singer 1984). Identification also comes from the economic assumptions of the model, e.g., the timing convention (section 2.2) provides a series of dynamic exclusion restrictions. To illustrate, current income affects current health but current health affects future income.

Another way to examine identification is to analyze the changes in the predictions of the model as the parameter values change (see e.g., Erdem, Imai and Keane 2003).[57] The parameters would be identified if the choice and outcome distributions were sensitive to the parameter values. In particular if parameter values affected distributions of the various choices and outcomes differently. Consider the parameters associated with medical care in the utility function (eq. 11). The life-cycle effects of changing these parameters is reported in table 2. The effect of increasing the absolute value of the transactions cost for seeking medical care, α_{15}, is to decrease the demand for medical care and OOP costs with a corresponding fall in the health outcomes. Individuals under 65 are also more likely be uninsured. Individuals older than 65 opt in greater numbers for plan 1 (i.e., only Medicare) and participate less in plan 3 (i.e., Medicare/group plan). There is also a decrease in alcohol consumption and smoking rates, while exercise rates increase.

[52] Formally the conditions for identification are the same as those 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) other than to examine the invertibility of the Hessian.

[53] See Rust (1994a, 1994b) for an excellent discussion of the assumptions required for identification of dynamic programming models.

[54] 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.

[55] The exceptions to this eligibility criterion are described in footnote 1. These do not adversely affect the identification assumption.

[56] 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. 9). Insurance status parameters in the OOP cost processes are shown in tables 16 and 17, 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.

[57] I am grateful to Mike Keane for bringing this method to examine identification to my attention.

Table 2: Changes in Average Life-cycle Choices and Outcomes Due to an Increase in Parameter Value

Parameter	Medical-care m_t	OOP-cost OOP_t	Health H_t	Insurance (Uninsured) $I_t < 65$	Insurance (Plan 1) $I_t \geq 65$	Alcohol a_t	Smoking c_t	Exercise e_t
Transactions cost (α_{15})	$--$	$--$	$-$	$++$	$+$	$-$	$-$	Small $+$
Palliative care (α_{16})	$++$	$++$	$+$	$--$	$-$	$-$	Small $+$	Small $+$
Curative care ($\eta_{4,\cdot}$)	$++$	$++$	Small $+$	$--$	Small $-$	$+$	$++$	$-$
Preventive care ($\eta_{9,\cdot}$)	$+$	$+$	Small $+$	$-$	Small $-$	Small $+$	Small $+$	Small $-$

Notation: "+" and "−" denote increase and decrease respectively; "++" and "−−" denote a sharp increase and decrease in a particular dimension relative to the other variables respectively.

An increase in the palliative effect of treatment, α_{16}, in equation (eq. 11) leads to an increase in the demand for medical care, OOP costs and health outcomes. Individuals under 65 are also less likely to be uninsured. Individuals older than 65 are less likely to be in plan 1 and more likely to be in plan 3. Alcohol consumption decreases while smoking and exercise increase. Thus the transactions cost and the palliative care parameters affect the predicted choices and outcomes differently and should be identifiable from each other through the observed variation in choices and outcomes in the data.

In the health production function (eq. 8) the effect of increasing the curative component of medical care, $\eta_{4,\cdot}$, leads to an increase in the demand for medical care, OOP costs, and health outcomes (table 2). The level of the uninsured under 65 and the demand for plan 1 over the age of 65 decreases. The demand for plan 3 over the age of 65 increases. There is an increase in alcohol consumption and smoking, while exercise rates decrease. The effect of increasing the preventive component of medical care, $\eta_{9,\cdot}$, is a much smaller increase (relative to the effects of the change in the curative component) in the demand for medical care, OOP costs and health outcomes. Similarly compared to the curative component case there is a relatively smaller decrease in the uninsured under 65, the demand for plan 1 over 65 and a milder increase in the demand for plan 3 over 65. Alcohol consumption and smoking increase while exercise decreases in smaller amounts relative to the curative component context. The differences between the changes in the choices and outcomes suggest that the two medical care parameters in the health production function should be identified from each other, and from those in the utility function. The identification of the other parameters of the model may be examined analogously but is not done due to space constraints.

5 The Estimation Results and Model Fit

5.1 The Estimates

The marginal utility of the composite consumption commodity at $X = 0$, α_{13}, is normalized to be 1 (see table 3). 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, α_1, of switching insurance plans is \$11,564.30. This is about $1/10^{\text{th}}$ the average two-year household income of \$106,643.60. It is about 3.2 times the average premium of \$3,617 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, α_2, is \$7,394.94. 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, α_4, is common to all types, and is estimated to be -15,118.50. The utility parameter linear in health is estimated to be different for each type of individual, for type 1 ($\alpha_{3,1}$) it is 149,745.23, for type 2 ($\alpha_{3,2}$) it is 166,328.98 and for type 3 ($\alpha_{3,3}$) it is 123,804.29. 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, α_5, is \$344.21. The switching cost for exercise, α_6, is \$1,161.36, implying the model will predict persistence in exercise choices. The utility from smoking, α_7, is \$72.00. The switching cost, α_8, is \$409.46. The start-up cost parameter, α_9, is \$75.30. 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, α_{10}, of \$103.71. The switching cost, α_{11}, is \$786.95. The start-up cost parameter, α_{12}, is \$68.01. 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, α_{15}, is \$1,534.03.

Table 3: Utility Parameters and Other Variables

Parameter	Estimate (Asy. S.E.)
Insurance choices:	
Switching cost-under 65 (α_1)	-11564.30 (83.50)
Switching cost-65 and older (α_2)	-7394.94 (499.28)
Habit choices:	
Utility of health-type 1 ($\alpha_{3,1}$)	149745.23 (683.29)
Utility of health-type 2 ($\alpha_{3,2}$)	166328.98 (416.73)
Utility of health-type 3 ($\alpha_{3,3}$)	123804.29 (489.21)
Utility of health-quadratic term (α_4)	-15118.50 (56.25)
Utility of exercise (α_5)	-344.21 (0.43)
Switching cost of exercise (α_6)	-1161.36 (2.64)
Utility of smoking (α_7)	72.00 (1.66)
Switching cost of smoking(α_8)	-409.46 (1.87)
Start-up cost of smoking (α_9)	-75.30 (0.56)
Utility of alcohol consumption (α_{10})	103.71 (4.75)
Switching cost of alcohol consumption (α_{11})	-786.95 (8.15)
Start-up cost of alcohol consumption (α_{12})	-68.01 (1.06)
Medical Care choices:	
Utility of consumption (α_{13}) (not estimated)	1.0
Utility of consumption-quadratic term (α_{14})	-1.1×10^{-6} (5.0×10^{-9})
Transactions cost of medical care (α_{15})	-1534.03 (5.34)
Utility of mitigative medical care-moderate sickness (α_{16})	1325.19 (3.07)
Utility of moderate sickness (α_{17})	-3279.46 (119.28)
Coefficient on multiplicative effects of high sickness (α_{18})	3.29 (0.019)
Coefficient on "equivalence factor" of household size (α_{19})	0.613 (6.8×10^{-3})
Terminal Value Function:	
Health (α_{20})	10690.71 (3402.00)
Sickness (α_{21})	-20750.20 (5627.22)
Health÷Sickness (α_{22})	44520.47 (14245.74)
Other variables:	
Annual discount rate (β) (not estimated)	0.95
Number of types ($\bar{L}$)	3
Likelihood	-102198.215

The disutility of moderate sickness, α_{17}, is \$3,279.46. The mitigative utility of medical care, α_{16}, is \$1,325.19. Medical care does not totally mitigate the disutility of sickness. A "moderately" sick person who seeks "moderate" medical treatment has a net disutility of \$3,488.03, inclusive of the transactions cost of seeking care.[58] Due to the convexity of the transactions costs, with "moderate" sickness and "high" level of medical care the net disutility is \$8,090.39. 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, α_{20}, is \$10,690.71 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, α_{21}, is \$20,750.20 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, α_{22}, is estimated to be \$44,520.47. 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 15). 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 60 below for details of the simulated sample). The mortality (survival) odds are calculated as $\left[\frac{P_{t,1}}{P_{t,0}}\right]$, where $P_{t,1}$ ($P_{t,0}$) 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.[59] 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 16 - 22).

[58] In this calculation the utility parameters are normalized so that the lowest level of care and sickness provide zero utility.

[59] Survival odds are computed for medical utilization. Mortality odds of "no exercise" to "exercise" are computed, i.e., the inverse of the ratio above.

Table 4: Within sample fit: Distribution of Health Outcomes and Choices by Age.

		HRS Data (*Baseline Simulations*)				
Age	Total Obs. [Survivors]	Mean health[a], [c]	Mean medical-care[b]	Mean alcohol-consumption[b]	Mean smoking[b]	Mean exercise[b]
50	134 [134]	4.761 (*4.576*)	1.672 (*1.704*)	1.955 (*1.960*)	1.672 (*1.607*)	1.194 (*1.318*)
52	724 [722]	4.691 (*4.524*)	1.693 (*1.740*)	1.921 (*1.918*)	1.536 (*1.545*)	1.242 (*1.331*)
54	1231 [1220]	4.642 (*4.488*)	1.736 (*1.792*)	1.929 (*1.916*)	1.507 (*1.521*)	1.258 (*1.340*)
56	1425 [1410]	4.549 (*4.459*)	1.765 (*1.800*)	1.903 (*1.892*)	1.440 (*1.462*)	1.320 (*1.344*)
58	1459 [1435]	4.478 (*4.407*)	1.751 (*1.818*)	1.862 (*1.869*)	1.404 (*1.421*)	1.315 (*1.356*)
60	1292 [1273]	4.404 (*4.352*)	1.833 (*1.819*)	1.851 (*1.848*)	1.382 (*1.390*)	1.309 (*1.354*)
62	1054 [1022]	4.380 (*4.335*)	1.843 (*1.879*)	1.853 (*1.848*)	1.351 (*1.362*)	1.326 (*1.354*)
64	797 [769]	4.281 (*4.299*)	1.941 (*1.787*)	1.811 (*1.817*)	1.307 (*1.332*)	1.389 (*1.343*)
66	531 [511]	4.209 (*4.168*)	1.971 (*2.262*)	1.802 (*1.791*)	1.286 (*1.303*)	1.382 (*1.349*)
68	339 [324]	4.127 (*4.131*)	2.117 (*2.276*)	1.725 (*1.794*)	1.302 (*1.274*)	1.327 (*1.353*)
70	188 [173]	3.952 (*4.086*)	2.208 (*2.267*)	1.775 (*1.803*)	1.301 (*1.259*)	1.370 (*1.333*)
72	105 [101]	3.819 (*4.071*)	2.208 (*2.241*)	1.683 (*1.781*)	1.317 (*1.248*)	1.396 (*1.338*)
74	42 [34]	3.500(*4.014*)	2.353 (*2.249*)	1.559 (*1.761*)	1.147 (*1.243*)	1.471 (*1.329*)

[a]The relevant data distribution is "Total Obs."
[b]The relevant data distribution is "Survivors."
[c]The measure of health, H_t, is a composite of SRHS and mortality.

5.2 Within Sample Fit

The estimated model is used to simulate a sample of 3671 individuals.[60] The baseline simulations replicate the health transitions (table 5) and the age distribution of mean health well, though the mean health is slightly under-predicted at younger ages and slightly over-predicted health at older ages (table 4).[61] 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

Table 5: Within Sample Fit: Health State Transitions (row %)

	HRS Sample (*Baseline Simulations*) Current Health (H_t)					
Lagged Health (H_{t-1})	Dead	Poor	Fair	Good	Very good	Excellent
Poor	22.67 (*24.0*)	41.70 (*49.2*)	23.89 (*22.4*)	6.88 (*3.5*)	2.83 (*0.9*)	2.02 (*0.1*)
Fair	7.01 (*9.2*)	11.34 (*13.3*)	43.17 (*38.5*)	28.01 (*29.8*)	8.61 (*8.7*)	1.78 (*0.5*)
Good	2.67 (*0.8*)	2.61 (*1.3*)	14.01 (*18.2*)	50.50 (*53.8*)	23.80 (*22.9*)	6.41 (*3.1*)
Very good	1.51 (*0.0*)	1.13 (*0.0*)	5.38 (*2.0*)	27.22 (*37.5*)	49.70 (*43.4*)	15.06 (*17.0*)
Excellent	0.95 (*0.0*)	0.71 (*0.0*)	2.22 (*0.0*)	13.03 (*7.6*)	32.09 (*37.6*)	50.99 (*54.7*)

[60]The simulation sample corresponds to each person in the estimation sample, i.e., 3562 persons, and 109 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 39). The sample size is limited to 3671 because of a significant computational cost in simulating from the model. Khwaja (2001) provides details for a related model.

[61]Tables for the frequencies are omitted in the interest of brevity. The relevant sample for the health and household size outcomes includes observations on dead individuals as this information is available in the event of death.

Table 6: Within Sample Fit: Smoking Choice Transitions (row %)

HRS Data (*Baseline Simulations*)

Lagged Smoking (c_{t-1})	Current Smoking (c_t) None	Pack/Day < 1	Pack/Day $\geqq 1$
None	97.78 (*96.5*)	1.09 (*2.0*)	1.14 (*1.4*)
Pack/Day < 1	26.14 (*22.6*)	56.09 (*54.7*)	17.77 (*22.8*)
Pack/Day $\geqq 1$	11.18 (*7.9*)	12.84 (*45.9*)	75.98 (*46.1*)

lag in increase in utilization in the data may be due to frictions 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 model is able to match the persistence in alcohol consumption (not shown) and smoking (see table 6) as well as the quitting rates for those consuming alcohol and smoking.[62] However for those smoking "a pack or more/day" the model tends to over-predict reduction of smoking to "less than a pack/day." A potential reason is that the switching cost for smoking is not allowed to differ by consumption level. This was done to keep the specification parsimonious.

The mean log-income and log-OOP cost profiles are matched well, especially the decline in income with

Table 7: Within Sample Fit: Distribution of Outcomes by Age

HRS Data (*Baseline Simulations*)

Age	Total Obs. [`Survivors`]	Mean (Log)-Income[a]	Mean (Log)-OOP-costs [`Obs.`][b]
50	134 [134]	11.520 (*11.405*)	7.015 (*6.630*) [1]
52	724 [722]	11.455 (*11.409*)	6.272 (*6.523*) [105]
54	1231 [1220]	11.391 (*11.333*)	6.352 (*6.654*) [556]
56	1425 [1410]	11.375 (*11.326*)	6.738 (*6.734*) [734]
58	1459 [1435]	11.330 (*11.253*)	6.764 (*6.817*) [692]
60	1292 [1273]	11.201 (*11.207*)	6.915 (*6.758*) [644]
62	1054 [1022]	11.153 (*11.239*)	6.792 (*6.793*) [519]
64	797 [769]	11.124 (*11.124*)	7.087 (*6.867*) [436]
66	531 [511]	11.118 (*11.057*)	6.927 (*6.891*) [262]
68	339 [324]	11.051 (*10.964*)	6.903 (*6.827*) [161]
70	188 [173]	10.955 (*10.944*)	6.850 (*6.936*) [104]
72	105 [101]	10.746 (*10.909*)	7.216 (*7.059*) [51]
74	42 [34]	10.731 (*10.839*)	7.129 (*6.859*) [11]

[a] The relevant data distribution is "`Survivors`."
[b] The data on OOP-costs is limited to individuals from waves 2 and 3.

[62] The transition matrices for the medical care, alcohol consumption, exercise and household size variables are not presented due to space limitations. These replicate the data well and are available from the author on request.

Table 8: Within Sample Fit: Insurance Choice (age$\leqq$ 64) Transitions, (row %)

Lagged Insurance ($I_{t-1},\ age \leqq 64$)	HRS Data (*Baseline Simulations*) Current Insurance ($I_t,\ age \leqq 64$) None	Group	Personal	VA/Ch.	Grp./VA/Ch.	Grp./Pers.
None	65.22 (*64.0*)	20.76 (*18.4*)	9.17 (*4.4*)	3.98 (*4.4*)	0.17 (*4.6*)	0.69 (*4.1*)
Group	3.15 (*3.1*)	90.18 (*88.9*)	2.78 (*2.0*)	0.34 (*2.7*)	0.74 (*1.7*)	2.81 (*1.6*)
Personal	11.24 (*7.6*)	33.72 (*23.9*)	51.55 (*53.4*)	0.00 (*6.2*)	0.39 (*5.0*)	3.10 (*3.9*)
VA/Ch.	8.0 (*6.5*)	5.60 (*21.6*)	0.80 (*5.7*)	76.00 (*58.4*)	9.60 (*4.3*)	0.00 (*3.5*)
Grp./VA/Ch.	0.00 (*8.1*)	19.44 (*24.4*)	0.00 (*4.8*)	13.89 (*6.6*)	63.89 (*53.5*)	2.78 (*2.6*)
Grp./Pers.	3.06 (*11.1*)	67.35 (*26.3*)	4.08 (*6.6*)	2.04 (*6.5*)	5.10 (*5.1*)	18.37 (*44.4*)

age (table 7). The insurance choice transitions under the age of 65 are replicated reasonably well (table 8). An exception is the transitions from the group/personal plan (plan 6) in to other plans. In particular the transitions in to the group plan (plan 2) are under-predicted and the transitions in to plan 6 itself are over-predicted. In general the insurance transitions at ages 65 and older are fit reasonably well (table 9). However the transitions from the Medicare with Medigap or personal plans (plan 2) in to plan 1 (Medicare) are under-predicted and those in to plan 3 (Medicare with group insurance) are over-predicted. A possible explanation is that the switching cost does not change with insurance status. This was done to keep the specification sparse to maintain policy invariance.

For individuals younger than 65, the model does well in replicating the frequency distribution of the uninsured and those on the group plan (table 10). These two insurance choices include 86% of the individuals (see table 1) so the frequencies of the other four choices in this age group are not shown in the interest of saving space. However those choices too are replicated well. The frequency distribution for individuals 65 or older who have only Medicare (plan 1) is flatter than the data. The model does better at replicating the data for those 65 or older who have a combination of Medicare and Medigap and/or personal insurance (plan 2). A potential reason for the predicted insurance choices at ages older than 70 is the small sample

Table 9: Within Sample Fit: Insurance Choice (age $\geqq$ 65) Transitions, (row %)

Lagged Insurance ($I_{t-1},\ age \geqq 65$)	HRS Data (*Baseline Simulations*) Current Insurance ($I_t,\ age \geqq 65$) Medicare	Medicare/Medigap/Personal	Medicare/Group
Medicare	78.91 (*70.0*)	7.81 (*8.6*)	13.28 (*21.4*)
M-care/M-gap/Pers.	48.65 (*19.4*)	39.19 (*52.0*)	12.16 (*28.7*)
Medicare/Group	19.40 (*13.5*)	7.36 (*9.1*)	73.24 (*77.4*)

Table 10: Within Sample Fit: Distribution of Insurance Choices by Age

HRS Data (*Baseline Simulations*)

Age	Total Obs. [Survivors]	Uninsured %- age < 65[a]	Group % (Plan 2, age ≦ 65[a])	Medicare % (Plan 1, age ≧ 65[a])	Medicare+ Medigap+Pers. % (Plan 2, age ≧65[a])
50	134 [134]	10.45 (*11.2*)	78.36 (*64.7*)		
52	724 [722]	13.43 (*10.5*)	76.73 (*67.3*)		
54	1231 [1220]	12.13 (*9.9*)	76.48 (*68.7*)		
56	1425 [1410]	12.06 (*8.7*)	75.18 (*70.1*)		
58	1459 [1435]	11.78 (*9.2*)	73.94 (*71.1*)		
60	1292 [1273]	12.25 (*10.6*)	71.01 (*70.0*)		
62	1054 [1022]	11.84 (*10.9*)	70.65 (*66.7*)		
64	797 [769]	11.57 (*12.8*)	69.05 (*57.9*)		
66	531 [511]			22.9 (*36.2*)	11.74 (*19.3*)
68	339 [324]			31.48 (*34.8*)	12.65 (*17.5*)
70	188 [173]			38.15 (*34.3*)	17.34 (*16.8*)
72	105 [101]			46.53 (*33.4*)	15.84 (*16.8*)
74	42 [34]			58.82 (*34.0*)	0.0 (*14.9*)

[a]The relevant data distribution is "Survivors."

size at those ages, e.g., (i) 101 and 34 observations respectively at ages 72 and 74, and (ii) no data on individuals in insurance plan 2 at age 74.

5.3 Out of Sample Fit

To assess the out of sample fit of the model data from the National Health Survey (NHS) 1962-63 are compared to the predicted distribution of uninsured individuals in the U.S. population under a counterfactual assumption that Medicare does not exist (table 11). 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

Table 11: Out of Sample Analysis: Uninsured Males in the US Population in the Absence of Medicare.

Uninsured Males %

Age[a]	Simulations- No Medicare- good[b] case	NHS 1962-63 Data- No hospital- insurance
25-34	26.72	23.4
35-44	18.9	21.1
45-54	11.12	21.7
55-64	7.56	25.2
65-74	31.06	39.0
75≤	52.43[c]	57.5

Data source: National Health Survey 1962-63, "Health Insurance Coverage," NCHS Series 10, no. 11, August 1964.
[a]The age categories replicate those reported in the NHS 1962-63.
[b]Individuals age ≥ 65 can purchase insurance available to those age < 65.
[c]The simulations are limited to the age range 75-80. The data includes all individuals 75 or 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 6).

Though the model under-predicts the percentage of uninsured in the middle part of the life cycle relative to the NHS 1962-63 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 NHS 1962-63 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 NHS 1962-63 data qualitatively is a strong validation of its predictability out of sample.

6 The Effects of Medicare on Health Behaviors and Outcomes

The effects of Medicare on individual health incentives and outcomes are examined by comparing medical utilization, OOP expenditures, health outcomes and health related behaviors in the baseline simulations with Medicare existent, to counter-factual simulations assuming the non-existence of Medicare. Such a comparison under different insurance environments should provide a measure of the effects of Medicare. It should be emphasized that the counter-factual simulations do not rest on the premise that Medicare will or should be disbanded.

In the absence of Medicare two extreme opposite insurance environments are assumed. In the first or *good* case 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.[63] In the second or *bad* case it is assumed that those 65 and older are priced out of the health insurance market and do not have any coverage in the absence of Medicare.[64] In fact a rationale for the introduction of Medicare was that significant proportions of elderly individuals lacked access to insurance. To make the good and bad cases symmetric it is assumed that there is no switching cost in dropping insurance coverage in the good case. Thus the good case is simply an expansion of the insurance choice set available in the bad case. As before a sample of 3671 individuals is simulated in each case. These two extreme environments help to bound the estimated effects of Medicare.[65]

[63] These choices are: 1-"none," 2-"group," 3-"personal," 4-"VA/Champus," 5-"group/VA/Champus," 6-"group/personal."

[64] I am grateful to Doug Staiger and Pat Bayer for independently suggesting this simulation.

[65] There is some empirical basis for these assumptions as revealed in the testimony before Congress in 1964 of Edwin Daly, MD, of the Group Health Association of America. Dr. Daly said,"The present HIP (Health Insurance Plan of Greater New York) annual premium rate for its medical care program for a person enrolled prior to 65, is $57.80 and the Blue Cross Insurance is $55.80. If that person then remains in HIP-Blue Cross, he (continues to pay those premiums after age 65 for) both medical and hospital coverage...These are community rates, and reflect the average costs of young people, middle age people, and old people...But the present Blue Cross rates for a person coming in after he is 65 jump up to $129.60 a year.

Table 12: Effects of Disbanding Medicare: % Differences in Means in the Absence of Medicare Relative to its Presence.

Age	Health[a]		Medical care		OOP costs[b]		Alcohol-consumption		Smoking		Exercise	
	good[c] case	bad[d] case	good case	bad case	good case	bad case	good case	bad case	good case	bad case	good case	bad case
22-42	0.001	0.001	0.085	0.061	0.355	0.337	0.009	0.035	0.009	0.023	0.000	0.000
44-64	0.001	0.000	1.584	-0.070	0.507	0.091	0.039	0.077	0.007	0.058	-0.009	-0.031
66-80	-0.183	-0.560	-32.50	-54.031	7.186	93.739	-0.100	-0.242	-0.479	-0.493	0.016	0.029

[a]The measure of health, H_t, is a composite of SRHS and mortality.
[b]The percentage differences in the OOP costs in levels are reported.
[c]Individuals older than 65 can purchase insurance available to those under the age of 65.
[d]No insurance markets exist for those over 65.

Table 12 reports the simulation results. Mean health is worse in the absence of Medicare between ages 66 to 80 with negligible changes at younger ages. For those aged between 66 and 80 it drops on average by 0.183% in the good and 0.560% in the bad case. The utility from health in each period evaluated at the mean health for this age group is $315,363.23 for type 1, $365,736.81 for type 2 and $236,566.95 for type 3 individuals.[66] The average per period loss in utility due to decrease in health between ages 66 and 80 in the good case is $420.49, $540.71, and $232.45 for type 1, 2 and 3 individuals respectively. In the bad case the average per period loss in utility is $1,302.49, $1,673.38, and $722.32 for the three types respectively. The value of the losses is small relative to the utility from health. Conversely this suggests that the Medicare improves health status by small amounts relative to the utility of health in this age group.[67] This is similar to the finding by Card, Dobkin and Maestas (2004) that Medicare coverage provides a small improvement in self-reported health status for people slightly older than 65. The mortality effects are presented in table 13. The last two columns present the distribution of survivors in the absence of Medicare. There is almost no difference in survival under the age of 70 but between ages 70 and 80 the difference increases. At age 80 the decline in survivorship is 0.32% in the good case and 0.82% in the bad case. Thus for cohorts born after 1944, assuming 3 million live births a year,[68] Medicare would annually save between approximately 9,600 and 24,600 lives of those aged 80 by the year 2024.[69] This is consistent with the conclusions of Dow

That is Blue Cross (hospital coverage) alone and is available only if that person is found medically acceptable. Blue Cross does not take them all, they take a few..." (U.S. Congress 1964). This testimony came to my attention through McClellan and Skinner (1997).

[66]The utility of health is $[\alpha_3(H_t - 1) - \alpha_4(H_t - 1)^2]$. The one is subtracted from H_t to normalize utility to zero in case of death. The mean health in the age group 66-80 is 4.038 (see table 4).

[67]This finding 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 differences in variable definitions.

[68]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).

[69]The simulations assume that Medicare exists from the age of 22 onwards for a cohort.

Table 13: Mortality Effects of Medicare on the US Population.

	Number of Survivors		
Age	Baseline Simulations	No Medicare-good^a case	No Medicare-bad^b case
20	100000	100000	100000
30	98992	98992	98992
40	96541	96541	96541
50	92454	92454	92454
60	85045	85045	85045
70	73631	73631	73522
80	60174	59984	59684

Data source: National Vital Statistics Reports 2000
[a]Individuals older than 65 can purchase insurance available to those under the age of 65.
[b]No insurance markets exist for those over 65.

(2001)[70] and Lichtenberg (2002) that Medicare has a significant effect in reducing mortality, and those of Card, Dobkin and Maestas (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.[71] In sum, the simulations imply that Medicare leads to small improvements in health status and reductions in mortality of the elderly but has no health effects on the young.

Mean OOP costs (in levels) between ages 66 and 80 increase by 7.186% and 93.78% in the good and bad case respectively. This implies that Medicare reduces OOP costs and provides insurance against medical expenditure risk. Interestingly, Finkelstein and McKnight (2005) reach a similar conclusion using very different methods and data.[72] The mean medical utilization between ages 66 and 80 declines by 32.50% and 54.03% in the good and bad case respectively. This implies that Medicare induces large increases in medical utilization by the elderly, as also concluded by Dow (2001), Lichtenberg (2002), and Card, Dobkin and Maestas (2004). The dynamic incentives that arise because of the life cycle nature of health production lead to some small changes in utilization for those younger than 65 as well. In the good case, for the age group 22-64, medical utilization shows a slight increase as a slightly larger number of individual's are insured relative to the baseline simulations in this age group. More individuals between 22 and 64 choose

[70]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.

[71]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.

[72]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.

to be insured in the good case relative to the baseline simulations for two reasons. First, there is a (fixed) switching cost of acquiring coverage after age 65. If insurance coverage is acquired early and maintained subsequently it reduces the expected average cost of insurance coverage by spreading the switching cost over a longer horizon. Second, insurance coverage is not guaranteed after the age of 65 which in turn reinforces the first incentive.[73]

The reason medical utilization changes by large amounts in absolute terms (i.e., between 32.50% to 54.03%) but the corresponding changes in mean health (i.e., between 0.183% and 0.560%) are smaller is that a large component of medical care is mitigative, i.e., purely for consumption benefits. Simulations shows that medical utilization would only decline by 17.4% over the life cycle if medical care was purely mitigative and had no curative or preventive components.[74] This conforms with the conclusion of Skinner, Fisher and Wennberg (2001) that $26 billion (in 1996 dollars) or 20% of Medicare expenditures have no benefit in terms of reducing mortality. Davis (1998) also finds a similar large non-pecuniary effect of Medicare and Medicaid coverage for elderly women.

The model incorporates moral hazard as individuals have an incentive to increase their alcohol consumption and smoking, and decrease exercise when they are insured against the risk of medical expenditures. Moreover, the moral hazard is dynamic because future expected changes in outcomes (e.g., health, medical expenditures) affect current consumption decisions. In addition the life cycle nature of health production introduces other dynamic effects in the choices regarding health behaviors. For example, the decision to consume alcohol involves a trade-off between the utility of current consumption and future expected costs arising from worse health, lower income and higher medical expenditures due to the detrimental effects of current consumption on future health. The model also allows for the reverse effect of health on alcohol consumption, i.e., individuals in better health may consume more alcohol. The model also permits rational addiction so current alcohol consumption depends on past and future expected consumption. Hence alcohol consumption in each period is determined by the interaction of multiple factors, and as their relative impact changes over the life cycle it introduces dynamic variation in the consumption decision. The decisions about smoking and exercise are similarly determined over the life cycle.

Average alcohol consumption for those over 65 decreases by 0.1% in the good and 0.242% in the bad case. The decrease in insurance coverage, and the associated decrease in medical utilization and health outcomes raises the "full cost" of alcohol consumption sufficiently to counteract the current utility of consumption for those 65 and older. Conversely, Medicare induces small levels of moral hazard in alcohol

[73]In the bad case the medical utilization is slightly lower in the age group 22-64, in fact it falls for those 44-64, as slightly fewer individuals are insured (82.26%) relative to the baseline simulations.

[74]This simulation is done under the assumption that $\alpha_{16} > 0$, $\eta_{4,\cdot} = 0$, and $\eta_{9,\cdot} = 0$.

consumption for the elderly. Alcohol consumption is increased slightly at ages younger than 65 in the absence of Medicare as the current utility of consumption at these ages outweighs the future expected cost. This is because individual anticipate lower medical utilization, and health, and increased mortality (table 13) and medical expenditures in their old age without Medicare. This lowers the future expected utility thereby reducing the incentive for investing in health and prolonging the endogenously determined life span (Philipson and Becker 1998). This indicates that Medicare provides a small incentive for self-protection (Ehrlich and Becker 1972) for individuals younger than 65 which counteracts moral hazard and lowers alcohol consumption. On average smoking shows a small decrease for individuals 65 and older, and a small increase for those under 65 in the absence of Medicare. The explanation for these effects is similar to that for alcohol consumption. The effects are reversed in the case of exercise. There is on average a small increase in exercise for individuals 65 and older, and a small decrease for those under 65. Since exercise has a positive impact on health outcomes it changes in opposite direction compared to alcohol consumption and smoking when Medicare is absent.[75] Thus Medicare generates small levels of moral hazard in alcohol consumption, smoking and exercise for the elderly. On the other hand the self-protection that it induces at younger ages leads to slightly lower levels of alcohol consumption and smoking, and slightly increased exercise at these ages.

Medicare is an inter-generational transfer of resources which could have different welfare consequences for those younger than 65 compared to the elderly. Hence the life time impact is computed rather than the welfare impact on the elderly per se. The welfare impact is evaluated by computing the compensating variation, W_q, where q is an index for the good and bad cases without Medicare. The compensating variation is calculated as the average annual amount by which the income for all individuals would have to change accounting for differences in insurance premiums and mortality in the absence of Medicare to provide the same level of aggregate present discounted value of life time utility as under Medicare coverage. All individuals are weighted equally in this calculation. More precisely, W_q, for $q \in \{\text{good, bad }\}$ is calculated using the equality,

$$\sum_{n=1}^{N}\sum_{t=1}^{T}\beta^{t-1}[U_{M,t,n}(Y_{M,t,n};P_{M,t})]^{1(H_{M,t,n}>1)} = \sum_{n=1}^{N}\sum_{t=1}^{T}\beta^{t-1}[U_{q,t,n}(Y_{q,t,n}+W_q;P_{q,t})]^{1(H_{q,t,n}>1)} \tag{16}$$

where M, indicates the availability of Medicare coverage, and n and t index individuals and time respectively. For an individual n at time t and the insurance environment $r \in \{\text{M, good, bad}\}$, $U_{r,t,n}(\cdot)$ is the monetary equivalent of the indirect utility, and $Y_{r,t,n}$, $P_{r,t}$ and $H_{r,t,n}$ are respectively measures of income, insurance premiums and health. A notable feature of the calculation in (16) is that it accounts for the dif-

[75] Card, Dobkin and Maestas (2004) find no changes in smoking or exercise at age 65. Though not shown, the model also predicts very small changes in smoking and exercise for individuals *at* age 65.

ferences in mortality rates under the different insurance environments.[76] The welfare impact of Medicare accounting for insurance costs and mortality on life time utility of individuals is positive. On average individuals would be willing to pay $1,530.36 (in 1991 dollars) in the good and $3,574.76 in the bad case annually for Medicare coverage.[77] The willingness to pay is lower in the good case as expected due to it being the better alternative in the absence of Medicare. To provide some context to these numbers, 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 $2,124.00.[78] Hence the total annual Medicare premium for an ineligible (see footnote 1 for eligibility rules) individual was $2,482.80, and well within the range of the computed willingness to pay, adding plausibility to these calculations. In sum, Medicare improves *life time* utility by providing small improvements in health status, decreases in mortality late in the life cycle, insuring against medical expenditure risk and by improving access to medical care that has a *mitigative* value to individuals.

Similar welfare calculations for the life cycle effects of Medicare do not exist so direct comparisons are difficult. However Finkelstein and McKnight (2005) reach a very similar conclusion. 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 (1997) 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.[79]

[76]Even though taxes were not explicitly included in the budget constraint, in performing the calculations in the absence of Medicare, individuals are reimbursed 1.45% of their income. This compensates consumption by the amount that would otherwise have been deducted as Medicare payroll tax.

[77]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 unfeasible due to computational limitations.

[78]For details see the "2004 Annual Report of the Board of Trustees of the Federal Hospital Insurance and Federal Supplementary Medical Insurance Trust Funds" available online at http://www.cms.hhs.gov/publications/trusteesreport/.

[79]Manning and Marquis (1996) use data from the RAND health insurance experiment (Manning et al. 1987) to evaluate the the welfare effects of health insurance coverage in a static framework. They examine insurance contracts that vary in coinsurance rates and stop-loss to balance the trade-off between risk pooling and increased utilization of medical care. Their calculations are not comparable but they find that insurance with 30% copayment would be close to optimal.

Table 14: Effects of Changing Age of Eligibility for Medicare to 55 or 75: Percentage Differences in Means with Eligibility for Medicare at 55 (or 75) Relative to Age 65.

Age	Mean health[a]		Mean medical care		Mean OOP costs[b] (in levels)	
	Eligible age		Eligible age		Eligible age	
	55	75	55	75	55	75
22-54	-0.003	0.000	-1.202	0.059	-0.140	0.012
56-64	0.042	0.003	19.119	3.391	-13.597	0.894
66-74	0.163	-0.060	0.002	-18.079	-0.169	17.372
76-80	0.065	-0.250	0.018	-0.313	-0.530	2.559
Avg. (22-80)	0.039	-0.035	2.508	-2.446	-2.427	3.307

[a]The measure of health, H_t, is a composite of SRHS and mortality.
[b]The percentage differences in the OOP costs in levels are reported.

7 The Impact of Changing the Age of Eligibility for Medicare

Reductions in benefits–through changes to the age for receiving full retirement benefits or through reforms to slow the growth of Medicare spending or through other means–can affect retirement, the labor force, and saving behavior. In addition, policies that link increases in longevity over time to the eligibility age for social security, and perhaps Medicare, may need to be considered. Such linkages would help protect the financial and, hence, the economic viability of these programs.

–Chairman Alan Greenspan before the Special Committee on Aging, U.S. Senate, February 27$^{\text{th}}$, 2003.

Alan Greenspan in his capacity as Chairman of the Board of Governors of the Federal Reserve System[80] drew attention on multiple occasions to the expected demographic shifts in the U.S. and the consequent implications for the financial stability of the Medicare program. To deal with the anticipated financial burden one proposal that he, among others, made is that the age of eligibility be modified as has been done for the Social Security program.[81] Hence the impact of changing the age of eligibility on life cycle medical utilization, OOP costs and health outcomes is examined. A sample is simulated with the Medicare eligible age decreased (increased) to 55 (75).[82] These ages were chosen so that the consequent changes in outcomes would be sufficiently large to be observable. The results are reported in table 14. 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 younger and older than 65 in the

[80]See e.g., testimony of Chairman Alan Greenspan before the Committee on the Budget, U.S. House of Representatives September 8, 2004, http://www.federalreserve.gov/boarddocs/testimony/2004/200409082/default.htm.
[81]The eligible age for Social Security is legislated to increase to 67 by 2027.
[82]The choice of the alternative ages for eligibility is not meant to suggest that such changes should be implemented. The simulations can be repeated for any particular change.

baseline simulations. As previously a sample of 3671 individuals is simulated in each instance.

The consequences of increasing the eligible age to 75 are discussed below. The effects of decreasing the eligible age to 55 are the converse and not discussed for sake of brevity. Increasing the eligible age would reduce mean health over the life cycle by 0.035%. The changes mainly occur after the age of 65, with the largest after 75, as increasing the eligible age worsens insurance coverage that in turn decreases access to medical care with consequent effects on health.[83] Increasing the eligible age decreases medical utilization over the life cycle by 2.45% due to the delayed availability of Medicare insurance coverage. The mean OOP costs (in levels) increase over the life cycle by 3.31% when the eligible age is increased because of the decrease in insurance coverage at later ages. For both medical utilization and OOP costs, the largest changes are in the age range most affected by the change in coverage, i.e., 66-74. These results coupled with the welfare analysis in section 6 suggest that postponing the age of eligibility will reduce life time individual welfare, with the elderly being the most affected.

8 Conclusion

Medicare coverage is found to have a considerable effect in insuring the elderly against medical expenditure risk. It is also found that Medicare leads to small increases in self-reported health status and reductions in mortality for the elderly. However it also induces large increases in medical utilization, and generates small levels of moral hazard in alcohol consumption, smoking and exercise among the elderly. Medicare also plays a small role in counteracting moral hazard in alcohol consumption, smoking and exercise by inducing self-protection for individuals at younger ages. The *life time* welfare impact of Medicare accounting for insurance costs and mortality is positive. On average, individuals are willing to pay between $1,530.36 and $3,574.76 (in 1991 dollars) annually for Medicare coverage.

Given the expected demographic changes in the U.S. and the consequent implications for the financial viability of the Medicare program, the model is used to predict the consequences of changing the eligible age for Medicare. It is found that increasing the eligible age will decrease utilization, increase OOP costs and worsen health outcomes with a consequent reduction in welfare for the elderly as well as over the life cycle, and vice-versa on decreasing the eligible age.

The results in this paper warrant a caveat that the model ignores the cost to society of subsidizing Medicare, i.e., the financing of medical services. However this would require a general equilibrium analysis that is beyond the scope of the paper. Inclusion of such costs may reduce the computed welfare benefits of Medicare. The model abstracts from savings or retirement decisions for reasons of computational

[83]The results for mortality per se were qualitatively similar and are not reported due to space limitations.

tractability. On the other hand, inclusion of such decisions in the model would likely lead to larger estimated welfare benefits of Medicare. Additionally, the effects of Medicare on innovation in medical technology (Weisbrod 1991, Newhouse 1992), which are not the focus of this research, are also potentially welfare enhancing.[84] Future research that examines the role of these factors in assessing the impact of Medicare is warranted. Nevertheless, on the basis of this examination it can be concluded that Medicare lives up to its intended role as government provided health insurance, and from the point of view of its present and future beneficiaries is worth the cost.

References

Adda, J., and V. Lechene (2001): "Smoking and Endogenous Mortality: Does Heterogeneity in Life Expectancy Explain Differences in Smoking Behavior?" Working Paper, University College London.

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

Arcidiacono, P., H. Sieg and F. Sloan (2001): "Living Rationally Under the Volcano? An Empirical Analysis for Heavy Drinking and Smoking," Working Paper No. 8602, NBER.

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

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.

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.

Blundell, R. W., and S. R. Bond (1998): "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics* 87, 115–143.

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

Card, D., C. Dobkin and N. Maestas (2004): "The Impact of Nearly Universal Insurance Coverage on Health Care Utilization and Health: Evidence From Medicare," Working paper No. 10365, NBER.

Currie, J. and J. Gruber (1996): "Health Insurance Eligibility, Utilization of Medical Care, and Child Health," *Quarterly Journal of Economics* 111, 431–466.

Cutler, D. M. (2000): "Walking the Tightrope on Medicare Reform," *Journal of Economic Perspectives* 14, 45–56.

Cutler, D. M. and R. J. Zeckhauser (2000): "The Anatomy of Health Insurance," in A. J. Culyer and J. P. Newhouse, eds., *Handbook of Health Economics*, Vol 1A, Elsevier, Amsterdam, 563–643.

Davis, M. A. (1998): "The Health and Financial Decisions of the Elderly," Ph.D. Dissertation, University of Pennsylvania

[84] The HRS data is not well suited to assess the impact of Medicare on changes in technology.

Deaton, A. S. and C. H. Paxson (1998): "Health, Income, and Inequality over the life cycle," in David Wise, ed., *Frontiers in the Economics of Aging,* Vol. 7, Chicago, University of Chicago Press, 431-457.

Decker, S. L. (2005): "Medicare and the Health of Women with Breast Cancer," *Journal of Human Resources*, forthcoming.

Dow, W. (2001): "The Introduction of Medicare: Effects on Elderly Health," Working Paper, Department of Health Policy and Administration, University of North Carolina, Chapel Hill.

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.

Ehrlich, I. and G. S. Becker (1972): "Market Insurance, Self-Insurance, and Self-Protection," *Journal of Political Economy* 80, 623–648.

Erdem, T., S. Imai and M. P. Keane (2003): "Brand and Quantity Choice Dynamics Under Price Uncertainty," *Quantitative Marketing and Economics* 1, 5–64.

Feldstein, M. S. (1971): "An Econometric Model of the Medicare System," *Quarterly Journal of Economics* 85, 1–20.

Finkelstein, A., and R. McKnight (2005): "What Did Medicare Do (And Was It Worth It)?" Working Paper, Society of Fellows, Harvard University.

French, E. (2005): "The Effects of Health, Wealth and Wages on Labour Supply and Retirement Behavior," *The Review of Economic Studies* 72, 395–427.

Fuchs, V. R. (2000): "Medicare Reform: The Larger Picture," *Journal of Economic Perspectives* 14, 57–70.

Juster, F. T. and R. Suzman (1995): "An Overview of the Health and Retirement Study," *Journal of Human Resources* 30, S7–S56.

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. (1979): "Sample Selection as a Specification Error," *Econometrica* 47, 153–162.

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

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

Hodgson, T. (1992): "Cigarette Smoking and Lifetime Medical Expenditures," *The Milbank Quarterly* 70, 81–125.

Hubbard, R. G., J. Skinner, and S. Zeldes (1995): "Precautionary Saving and Social Insurance," *Journal of Political Economy* 103, 360–399.

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.

Khwaja, A. (2001): “Health Insurance, Habits and Health Outcomes: A Dynamic Stochastic Model of Investment in Health,” Ph.D. Dissertation, University of Minnesota.

Lichtenberg, F. (2002): “”The Effects of Medicare on Health Care Utilization and Outcomes,” in Alan Garber, ed., *Frontiers in Health Policy Research*, Vol.5, Cambridge, MA: MIT Press.

Lleras-Muney, A. (2005): “The Relationship Between Education and Adult Mortality in the United States,” *The Review of Economic Studies* 72, 189–221.

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

Manning, W. and M. S. Marquis (1996): “Health Insurance: The Tradeoff Between Risk Pooling and Moral Hazard,” *Journal of Health Economics* 15, 609–639.

McClellan, M. (2000): “Medicare Reform: Fundamental Problems, Incremental Steps,” *Journal of Economic Perspectives* 14, 21–44.

McClellan, M. and J. Newhouse (1997): “The Marginal Cost-Effectiveness of Medical Technology: A Panel Instrumental-Variables Approach,” *Journal of Econometrics* 77, 39–64.

McClellan, M. and J. Skinner (1997): “The Incidence of Medicare,” Working paper No. 6013, NBER.

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.

Newey, W. K., and D. L. McFadden (1994): “Large Sample Estimation and Hypothesis Testing,” in R. F. Engle and D. L. McFadden, eds., *Handbook of Econometrics,* Vol. 4, Amsterdam: Elsevier, 2111–2245.

Newhouse, J. (1992): “Medical Care Costs: How Much Welfare Loss?,” *Journal of Economic Perspectives* 6, 3–21.

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

Orphanides, A. and D. Zervos (1995): “Rational Addiction with Learning and Regret,” *Journal of Political Economy* 103, 739–758.

Palumbo, M. G. (1999): “Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle,” *The Review of Economic Studies* 66, 395–421.

Pauly, M. V. (1968): “The Economics of Moral Hazard,” *American Economic Review* 58, 231–237.

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.

Reinhardt, U. (2000): “Health Care for the Aging Baby Boom: Lessons from Abroad,” *Journal of Economic Perspectives* 14, 271–84.

Richards, H. (1999) *Life and Worklife Expectancies.* Lawyers and Judges Publishing Company, Inc., Tucson, AZ.

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.

Rothschild M., and J. E. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay on the Economics of Imperfect Information," *Quarterly Journal of Economics* 90, 630–649.

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.

Salanié, B. (2002): *The Economics of Contracts,* Cambridge, MA: MIT Press.

Saving, T. R. (2000): "Making the Transition to Prepaid Medicare," *Journal of Economic Perspectives* 14, 85–99.

Skinner, J., E. Fisher and J. Wennberg (2001): "The Efficiency of Medicare," Working Paper No. 8395, NBER.

Weisbrod, B (1991): "The Health Care Quadrilemma: An Essay on Technological Change, Insurance, Quality of Care, and Cost Containment," *Journal of Economic Literature* 29, 523–552.

Wennberg, J. E. and M. M. Cooper (1999) *The Quality of Medical Care in the United States: A Report on the Medicare Program. The Dartmouth Atlas of Health Care in the United States.* American Health Association Press, Chicago.

Wooldridge, J. M. (2000): "A Framework of Estimating Dynamic, Unobserved Effects Panel Data Models With Possible Feed to Future Explanatory Variables," *Economics Letters* 68, 245–250.

Table 15: Health Transition Parameters

	Estimate (Asy. S.E.)					
			Health Status			
Parameter	Dead ($q = 1$)	Poor ($q = 2$)	Fair ($q = 3$)	Good ($q = 4$)	Very Good ($q = 5$)	Excellent ($q = 6$)
Lagged health ($\eta_{1,q}$)	-1.35 (2×10^{-3})	0.53 (3×10^{-3})	-0.36 (8×10^{-4})	-1.43 (8×10^{-4})	-1.62 (3×10^{-3})	-2.09 (1×10^{-3})
Lagged health-quadratic term ($\eta_{2,q}$)	-0.61 (4×10^{-4})	-0.42 (3×10^{-3})	-0.56 (0.003)	-0.62 (7×10^{-4})	-0.41 (0.001)	-0.28 (0.017)
Moderate sickness ($\eta_{3,q}$)	0	-1.21 (0.04)	-0.57 (0.03)	-0.07 (1×10^{-4})	-0.19 (4×10^{-5})	-0.06 (7×10^{-4})
Curative medical-care-moderate-sickness ($\eta_{4,q}$)	0	-0.044 (0.0019)	0.019 (0.0009)	-0.0012 (2×10^{-5})	0.0009 (2×10^{-5})	0.0014 (1×10^{-5})
Multiplicative effects-high sickness ($\eta_{5,q}$)	0	0.03 (0.006)	0.07 (0.006)	36.33 (0.14)	23.12 (0.11)	56.19 (0.30)
Alcohol-consumption ($\eta_{6,q}$)	0	0.0007 (3×10^{-4})	-0.0025 (3×10^{-4})	-0.0037 (2×10^{-4})	-0.0061 (5×10^{-5})	-0.0078 (9×10^{-5})
Smoking ($\eta_{7,q}$)	0	-0.022 (3×10^{-4})	-0.022 (6×10^{-5})	-0.024 (5×10^{-5})	-0.026 (8×10^{-5})	-0.027 (3×10^{-5})
Exercise ($\eta_{8,q}$)	0	0.0008 (0.0012)	0.0013 (6×10^{-5})	0.003 (1×10^{-5})	0.0046 (1×10^{-5})	0.0012 (4×10^{-6})
Preventive medical-care ($\eta_{9,q}$)	0	0.0002 (3×10^{-4})	0.0026 (5×10^{-4})	0.0008 (5×10^{-4})	0.0055 (0.001)	0.0012 (3×10^{-4})
Constant:						
Type 1 ($\eta_{10,1,q}$)	0	1.83 (0.086)	2.05 (0.059)	4.64 (0.009)	6.53 (0.022)	7.59 (0.030)
Type 2 ($\eta_{10,2,q}$)	0	1.96 (0.065)	1.71 (0.051)	4.66 (0.0043)	6.78 (0.013)	7.64 (0.019)
Type 3 ($\eta_{10,3,q}$)	0	2.07 (0.10)	1.97 (0.087)	4.66 (0.069)	6.50 (0.033)	7.32 (0.038)

Table 16: Log-OOP Process Parameters

Parameter	Estimate (Asy. S.E.)
Insurance choices under age 65:	
No insurance (μ_1)	3.213 (0.0048)
Group plan (μ_2)	1.023 (0.0095)
Personal plan (μ_3)	1.293 (0.047)
VA/Champus plan (μ_4)	3.194 (0.0099)
Group/VA/Champus plan (μ_5)	3.158 (0.011)
Group/Personal plan(μ_6)	2.979 (0.024)
Multiplicative factor- high treatment (μ_7)	1.086 (0.0025)
Insurance choices at 65 and over:	
Medicare (μ_8)	0.881 (0.041)
Medicare/Medi- gap/Personal plan (μ_9)	0.677 (0.077)
Medicare/Group- VA/Champus plan (μ_{10})	0.414 (0.047)
Multiplicative factor- high treatment (μ_{11})	1.286 (0.048)
Other variables:	
Health (μ_{12})	-0.122 (0.0015)
Household Size (μ_{13})	0.124 (0.0031)
Age (μ_{14})	1.768 (0.0021)

Table 17: Probability of Zero OOP Process Parameters

Parameter	Estimate (Asy. S.E.)
Constant (γ_1)	-2.297 (0.050)
Health (γ_2)	0.035 (0.013)
Household Size (γ_3)	-0.096 (0.016)
Insurance choices under age 65:	
Group plan (γ_4)	0.721 (0.021)
Personal plan (γ_5)	0.604 (0.104)
VA/Champus plan (γ_6)	0.449 (0.161)
Group/VA/Champus plan (γ_7)	0.385 (0.140)
Group/Personal plan (γ_8)	0.547 (0.171)
Multiplicative factor- high treatment (γ_9)	0.492 (0.055)
Insurance choices at 65 and over:	
Medicare (γ_{10})	0.948 (0.064)
Medicare/Medi- gap/Personal plan (γ_{11})	0.357 (0.123)
Medicare/Group/- VA/Champus plan (γ_{12})	0.407 (0.084)
Multiplicative factor- high treatment (γ_{13})	0.742 (0.123)

Table 18: Variance and Kernel Smoothing Aitchison-Aitken Bandwidth Parameters

Parameter	Estimate (Asy. S.E.)
Variance:	
Insurance choice (ρ_1)	5069.43 (15.37)
Habits choice (ρ_2)	988.63 (1.84)
Medical care choice (ρ_3)	1209.04 (9.8)
Income process (σ_y)	0.860 (0.017)
OOP costs process (σ_{oop})	0.704 (0.006)
Measurement error in income (σ_κ)	0.204 (0.071)
Measurement error in OOP costs (σ_μ)	0.224 (0.019)
A-A Bandwidth for the initial distribution of:	
Health (ω_1)	0.466 (0.022)
Insurance (ω_2)	0.023 (0.123)
Household Size (ω_3)	0.205 (0.041)
Exercise (ω_4)	2.6 $\times 10^{-6}$ (2.3 $\times 10^{-6}$)
Smoking (ω_5)	0.026 (0.116)
Alcohol Consumption (ω_6)	0.468 (0.034)

Table 19: Income Process Parameters

Parameter	Estimate (Asy. S.E.)
Lagged Health (κ_2)	0.181 (0.0017)
Age (κ_3)	0.036 (1.4×10^{-4})
Age-quadratic term (κ_4)	-0.0005 (3.5×10^{-7})
Constant:	
Type 1 ($\kappa_{1,1}$)	10.120 (0.019)
Type 2 ($\kappa_{1,2}$)	10.525 (0.007)
Type 3 ($\kappa_{1,3}$)	9.901 (0.019)

Table 20: Household Size Transition Parameters

	Estimate (Asy. S.E.)			
	Household Size (HHS)			
Parameter	"One member" ($q = 1$)	"Two members" ($q = 2$)	"Three members" ($q = 3$)	"Four or more - members" ($q = 4$)
Constant ($\psi_{1,q}$)	0	-1.359 (0.0046)	-2.309 (0.0062)	-2.826 (0.014)
Lagged HHS ($\psi_{2,q}$)	-0.626 (0.038)	-4.955 (0.0098)	-3.548 (0.0099)	-0.539 (0.0099)
Lagged HHS-quadratic term ($\psi_{3,q}$)	0.083 (0.015)	-1.243 (0.021)	-2.034 (0.0015)	-2.295 (0.0073)
Age ($\psi_{4,q}$)	0	0.230 (0.0001)	0.410 (0.0001)	0.620 (0.0002)
Age-quadratic-term ($\psi_{5,q}$)	0	-0.0022 (4×10^{-6})	-0.0039 (4×10^{-6})	-0.0068 (7×10^{-6})

Table 21: Sickness Transition Parameters

Parameter	Estimate (Asy. S.E.)		
	Sickness States		
	Low ($q = 1$)	Moderate ($q = 2$)	High ($q = 3$)
Constant ($\phi_{1,q}$)	0	-4.185 (1.4×10^{-3})	-7.155 (1.4×10^{-3})
Health ($\phi_{2,q}$)	0	-0.092 (8.3×10^{-4})	-0.142 (7.1×10^{-4})
Age ($\phi_{3,q}$)	0	0.297 (1.4×10^{-4})	0.351 (7.7×10^{-5})

Table 22: Unobserved Heterogeneity Distribution Parameters

Parameter	Estimate (Asy. S.E.)	
	Individual Type	
	"type 2" ($l = 2$)	"type 3" ($l = 3$)
Constant ($\lambda_{1,l}$)	-2.430 (0.0038)	-0.120 (4.4×10^{-5})
Education ($\lambda_{2,l}$)	0.101 (3.5×10^{-5})	-0.079 (8.2×10^{-6})
Race ($\lambda_{3,l}$)	0.188 (0.068)	0.174 (0.079)