The Effect of changing the Medicare Eligibility Age on the Health of the Near-Retirement Population

Peyman Firouzi Naeim *

This Version: July 19, 2019
[Latest Version]

Abstract

Raising the eligibility age for Medicare, the third largest program in the federal budget, could lead to a large reduction in the federal budget deficit; however, the effect of this change on the welfare and the health of the near-retirement population is unclear. Using Health and Retirement Study (HRS) dataset, I measure the effect of a change in Medicare eligibility age on the welfare of the elderly population by estimating a dynamic discrete choice model of health and retirement that endogenizes health investment decisions. The empirical model allows for tracking the health behavior, labor supply and health status among the other key variables. Using Forward Simulation and Conditional Choice Probability estimator (CCP), I incorporate a large, multidimensional state space consisting fixed unobserved heterogeneity that serves as a measure to better identification of the effects. I find that labor supply, life expectancy, and mental health will be affected positively in response to an increase in the Medicare eligibility age. The welfare effect, however, is negative and there is some evidence of cost transfers from Medicare to the Social Security Program.

*Department of Economics, Andrew Young School of Policy Studies, (✉ pfirouzinaeim1@gsu.edu , ☎ +1(252)378-7882, 🏠 www.pfirouzi.com). I am extremely thankful to my advisers Tom Mroz, Daniel Kreisman, Pierre Nguimkeu and Melinda Pitts, for their guidance on this project, as well as Barry Hirsch, John Gibson, Garth Heutel, Rusty Tchernis, and Spencer Banzhaf, who provided invaluable advice throughout its development. I also benefited from the helpful comments of the participants in the GSU Brown Bag Health Economics Workshop, GSU Ph.D. seminar series and the IHS Summer 2018 Graduate Research Colloquium.

1 Introduction

Medicare, the universal health insurance program for elderly people initiated in 1965, has undergone extensive changes in coverage and compensation methods. Surprisingly, one aspect that has been unchanged is the age at which people become eligible. The Medicare eligibility age (MEA), was initially set at 65 in accordance with the Full Retirement Age (FRA)[1]. While the FRA was changed gradually from age 65 to 67 in response to demographic changes and the increasing pressure on the federal budget, the MEA has remained unchanged.

As baby boomers are approaching the retirement age, the aggregate utilization of Medicare is rising. Improvements in the medical industry, have led to an increase in life expectancy and a rise in Medical expenditures. As a result, people are using Medicare insurance for longer and more expensive services. Both factors have an adverse effect on the federal budget. One policy suggestion is to raise the MEA and reduce the burden on the federal budget[2]. The key concerns are the potential effects on the physical as well as mental health of the elderly population and change in welfare.

The MEA affects the access to the medical care and possibly alters people incentives to pursue a healthier lifestyle. As a result, changes in the MEA can potentially affect health outcomes. The direction of this effect is, however, ambiguous. I address two main questions in this paper. First, *"what is the effect of an increase in the MEA on both physical and mental health of the near-retirement and elderly people"*. And second, *"what is the welfare effect of this change in policy?"*

Because of the lack of exogenous treatments (changes in MEA), researchers cannot employ simple evaluation methods such as randomized controlled trial. This limitation necessitates the use of structural estimation techniques. Also structural estimation helps to reduce the ambiguity in the direction of the health effects by providing detailed mechanisms indicating how different dimensions of health and health investment affect other behaviors such as labor supply, consumption and ultimately welfare.

In determining the effect of a change in MEA on health outcomes, both labor supply and health investment decisions should be taken into account. A change in the MEA might affect health outcome negatively, by limiting access to health insurance, and consequently health care utilization. But, people also have other options to improve their health, such as pursuing a healthier lifestyle by altering their smoking habits or exercising patterns. Ruhm (2000) discusses that health behaviors can compensate for the negative health effects

[1] The age at which a person may first become entitled to full or unreduced retirement benefits.

[2] Fontaine et al. (2011)

from a reduction in income or losing employer provided health insurance, the channel that accentuates the role of a budget constraint in health investment decisions. Ex-ante moral hazard, as tracked by health behaviors, is another channel that affects health behavior and ultimately health [Dhaval and Kaestner (2009), Stanciole (2008), and Kelly and Markowitz (2009)]. People can change their health behaviors as a response to a change in the price of medical services. A rise in the MEA, consecutively can trigger such behaviors since it delays the access to the inexpensive Medicare health insurance and leaves more expensive health insurance plans as the only feasible options. While budget constraint and medical services price are both possibly altered as a result of a change in MEA, it is necessary to incorporate these channels into analysis where the goal is to evaluate the effect of a change in MEA on health.

Although the existence of an endogenous change in health behaviors is well documented, the limited studies which investigate the change in the MEA disregard endogeneity problem. This happens because of the curse of dimensionality that limits the inclusion of more factors in a structural estimation model. Blau and Gilleskie (2008), for example show that raising MEA to 67 has a minor effect on the labor supply when the effect of the health insurance can only be transmitted through the budget constraint, production of health, and aversion to risk. French and Jones (2011) provide a model of retirement and health insurance and analyze the effect of a change in the age of Medicare eligibility on the labor supply. They find that raising the age of Medicare eligibility from 65 to 67 increases labor supply modestly between age 60 to 69. Their main focus, however, is not health status; they do not model health behavior as a decision and their only measure of health is self-reported health. Scholz and Seshadri (2013) endogenize the total medical expenditure and exercise as decisions which affect the production of health. They find that the availability of the postretirement health insurance affects the household's retirement decisions and as a result they retire one quarter earlier. Pelgrin and St-Amour (2016) propose a stochastic life-cycle framework with rational agents who maximize lifetime utility over consumption, leisure, and health investments. In addressing the retirement decision, they do not allow for a reversible retirement decision (returning to the job market after declaring retirement). As a result, their model is not flexible enough to explain partial retirement and returns to the labor force. All of these phenomena are observed frequently in recent data, and they are important determinants of health behaviors and health outcomes.

This study improves upon the available models of health insurance and retirement by incorporating health behavior channels as decisions; this addresses the ex-ante moral hazard in the health insurance literature, as well as the possibility of the return to work after retirement that affects monetary budget constraint. I introduce a model of health investment and labor supply with endogenous and explicitly modeled health investment decisions about smoking and exercise. Using the longitudinal data from the Health and Retirement Study (HRS) from 2000 to 2014, I look at both physical and mental health outcomes.

I measure mental health, using the objective index of the CES-D scale[3]. To avoid the subjectivity problem of the measure of physical health, I produce a new physical health index, using the various objective aspects of health, such as Activities of Daily Living (ADLs), Functional Limitations and different health conditions[4] using Item Response Theory (IRT) estimated by Bayesian methods. The dynamic optimization model that I provide is capable of measuring the effect of the change in MEA on the physical and mental health outcomes of the elderly people while taking into account endogenous health behaviors.

I use Conditional Choice Probability Estimator of Arcidiacono and Miller (2011), to reduce the computational burden in the structural estimation of my model. This allows for a more detailed state space and decision set in the model, as well as incorporating unobserved fixed heterogeneity types.

In particular, this paper:

- Studies the impact of a change in Medicare eligibility age from 65 to 69 on the welfare, physical, and mental health of the elderly population in the U.S.
- Provides a structural model of Health investment and Labor Supply that accounts for the endogenous health investments.
- Applies Conditional Choice Probability in the health and retirement models.

I show that raising MEA from 65 to 69 improves physical and mental health status. As a result of this policy change, men work 151 hours more annually and live for about 0.46 years longer; Also, from the federal budget point of view, I find some evidence of cost transfer from the Medicare program to the Social Security program. Mean lifetime social security benefit collection increases by \$ 12,688. That is due to the increase in labor supply, postponing the social security benefits claim, raising monthly benefits, and the increase in the life expectancy. Welfare will be affected negatively as a result of the change in MEA. People, on average, would ask for about \$ 151,000 to accept the change as of age 50.

I begin with a background discussion of the relevant aspects of the Medicare program and related literature. Section two introduces the model where I discuss timing, per-period decisions, transition probabilities, state variables and the specification of the model. Section

[3]Center for Epidemiologic Studies Depression Scale (CES-D)

[4]Different health conditions are: Blood Pressure, Diabetes, Cancer, Lung Disease, Heart Problem, Stroke, Psychological Problem and, Arthritis

three describes the data and construction of the variables. Section four discusses the estimation method. Section five provides the results for the model performance and counterfactual policy effects, and section six concludes.

1.1 Medicare and Backgroung

Passed in 1965 and implemented in 1966, Medicare provides nationwide and mandatory health insurance for individuals aged 65 and older. It subsequently expanded to include people who are permanently disabled for at least two years and those with the end-stage renal disease who need kidney dialysis treatments or kidney transplant. Medicare has two main parts which were incorporated into the program from the beginning and insured individuals against hospital services expenditures and physician services costs (Part A and Part B). Through time the Medicare program became more comprehensive. In what follows, I summarize a background on the aspects of Medicare which are more relevant to the analysis in the current study:

> *Part A (Hospital Services):* It pays for services provided by hospitals. Enrollment in part A was mandatory for every individual older than 65 or receiving social security benefit. Medicare part A is paid by the Medicare trust fund, which is a government account funded by income tax and Medicare tax which adds to worker's Social Security tax. In the initial design, Medicare Part A is not paying for hospital care of the first day of the hospital visit. Then Medicare pays 100 % of hospital charges for days 2 to 60. From day 61 to 90, Medicare pays for 25 %, then 50 % for days 91-150. Medicare does not compensate costs for days 151 and beyond.
>
> *Part B (Physician Services):* The second part is part B that covers the physician services and was called Supplemental Medical Insurance. Part B is voluntary health insurance, although the premiums were so low that almost all of the Part A Medicare beneficiaries were participating in Part B as well. The premium for Part B has not remained constant through time. Starting from $ 3 per month, Centers for Medicare and Medicaid Services (CMS) set the basic Part B premium at $96.4 in 2008. Then in 2010, CMS introduced the program that is called income-related monthly adjustment amount (IRMAA) and set the premium to be higher for individuals with more than $ 85,000 annual income with a sliding income scale. This scheme reaches the highest premium of up to $ 396.1 per month. Medicare Part B is funded by the combination of the premiums and revenues respectively. Law specifies that the premiums for Medicare Part B should cover 25 % of the program's costs.

Part D (Drugs): In 2006, a new Part D added prescription drug insurance. The growing importance of the new drugs in curing diseases which previously required hospitalization, and increasing demand for the new drugs, led the CMS to introduce Medicare Part D. The program offers catastrophic protection to those with large expenses, but not to those with intermediate drug expenses. The program does not pay for the first $ 320. From $ 320 to $ 2,930, it pays for 75 % of the expenses. Then from $ 2,930 to $ 6,657.5, there is a so-called "donut hole" where the program does not pay any amount on drugs. For the expenses which exceed $ 6,657.5 program almost fully compensates the costs by paying for 95 % of the costs.

Medigap: While Part B is supplemental insurance, there is another type of supplemental insurance that is less common. Private Supplemental Insurance or Medigap had about 12 million enrollees in 2015 where the number of Medicare beneficiaries reached a total of 55 million. Medicare coverage leaves some people uninsured or under-insured. Medicare Part A is not insuring against large health shocks when patients need a long period of hospitalization. Donut hole in Part D, leaves individuals with intermediate to large drug expenses under-insured. Medigap fills in the gap through private health insurance contracts. The annual premiums of the Medigap policies vary according to the plan chosen and location. The average across the nation in 2005 was between $ 1,150 to $ 1,750 with a significant variation depending on the location.

Medicare is the third largest program in the federal budget. In 2016, total Medicare expenditures were $ 679 billion and about 3.64 % of GDP [The Boards of Trustees (2017)]. The Boards of Trustees (2017) predicts that by the current increasing trend in the Medicare expenditures, trust fund becomes depleted in 2029. Fiscal pressure on Medicare steams the discussion of the different cost-saving measures. Most proposals focus on measures to limit the growth rate by altering the provider payment methods, expanding the cost-sharing aspect of the Medicare, and a less discussed measure of raising the MEA to reduce the number of beneficiaries.

One of the consequences of a rise in MEA is the change in retirement patterns as a result of the interaction between MEA and Full Retirement Age (FRA), as discussed by French (2005). FRA is defined as the age in which individuals with sufficient employment history become eligible to claim to receive Social Security Benefits, and was raised from 65 to 67 to address the recent changes in U.S. demography. Since the initiation of the Medicare program in 1965, the life expectancy for a 65-year-old man has risen by more than five years. This improvement in life expectancy is about four years for a 65-year-old woman.

There are abundant studies on the cost-saving aspects of a rise in MEA and its impact on healthcare utilization, the effect on health outcome, however, is less explored. Congressional Budget Office (2016) estimate that raising MEA from 65 to 67, by two months each year, and starting in 2020 produces a net saving of $ 14.9 billion between 2020 to 2026. Lichtenberg (2002) analyzes the effect of the healthcare utilization and documents a sharp increase in ambulatory care at age 65 and a smaller increase in inpatient care. He also finds some evidence of fewer hospital days and a lower probability of death than would be predicted based on pre-65 trends. Finkelstein (2007) shows that the introduction of Medicare was associated with an increase in health care utilization and spending. Card et al. (2008) find that the Medicare eligibility threshold at age 65 is associated with an increase in health insurance coverage as well as an increase in medical care services. Assuming that Medicare is designed to fill the gap in the health insurance coverage, resulting from the termination of employer health insurance, Lichtenberg (2002) argues that it is expected to see no difference in the health care utilization before and after age 65. Hence, increase in utilization can be interpreted as a sign of postponing of the healthcare utilization as individuals approach MEA.

The effect of the Medicare on health is less obvious. Card et al. (2008), using regression-discontinuity design around MEA (at age 65) and find that Medicare eligibility reduces 7-day to 9 months mortality by about .8 to 1 percentage points. Finkelstein and McKnight (2008) on the other hand, use age-based identification strategy and exploiting the geographical variation in the increase in health insurance coverage, resulted from introducing of Medicare and show there is, at best, very modest health benefit from Medicare. Although, they show that the introduction of Medicare had a positive welfare effect by reducing the risk exposure of the large health shocks. This study finds that the current policy with MEA at 65 is not optimum and compared with the policy with MEA at 69, hurts both mental and physical health as measured by depression and an objective measure of physical health, respectively.

2 Model

I develop a dynamic model in which agents are maximizing their discounted lifetime utility by deciding over consumption, leisure, individual health investments and purchase of insurance policies. Because of the data structure, every period in the model corresponds with two years in the data.

2.1 Variables

The variables can be separated into three distinctive groups: state variables, temporary variables, and decision variables. State variables (S) are those variables that individual requires to be informed about their values at the beginning of each period in order to update his current period decisions and state. Temporary variables (N) are not transferred from one period to the next and instead are calculated in every period and incorporate in the process of shaping the state variables. Lastly, the decision variables (D) are chosen by individual in every period, using utility maximization rationale.

2.1.1 State Variables

This group contains seven different sets of variables, including macro shock, attributes, employment history, health, reactionary healthcare, monetary variables, and unobserved heterogeneity.

Macro shock: Economic shock (σ^e) determines if economy is in the bad state or not;

Attributes: Age(A), education(E), full retirement age(f), gender(g), marital status(m);

Employment History: Tenure(x^1), experience(x^2), involuntary job loss(σ^l);

Health: Physical Health(h^p), mental health(h^m), number of health conditions(h^c);

Reactionary Healthcare: Number of nights in hospital(k^1), number of doctor visits(k^2);

Monetary variables: Social security benefit amount(SSB), total medical expenditure(ME), assets(a), medicare(i^m)

Unobserved Heterogeneity Type: Determines the fixed and unobserved heterogeneities which can be one of the three types (δ)

Hence the vector of the state variables is defined as:

$$\begin{aligned}\overrightarrow{S}_t = \quad & \{\sigma^e_t, A_t, E, f, g, m_t, x^1_t, x^2_t, \sigma^l_t, \\ & h^p_t, h^m_t, h^c_t, k^1_t, k^2_t, SSB_t, ME_t, a_t, i^m_t, \delta\}\end{aligned}$$

2.1.2 Temporary Variables

These variables are produced in every period and will be used to calculate the state variables. Since state variables convey all the necessary information for the next period decision-making process and transitions, there is no need to pass the temporary variables to the next period. The model consists of the eight temporary variables:

$$\overrightarrow{N}_t = \{y_t, OOP_t, ip_t^{emp}, ip_t^{prv}, ip_t^{m}, ssb_t^{earl}, ssb_t^{full}, ssb_t^{late}\}$$

Where y_t is income in period t, OOP_t is the out-of-pocket medical expenditure, ip_t^{emp} is the premium to pay if one acquires health insurance from the employer; ip_t^{prv} is the premium for the private health insurance; ip_t^{m} is the health insurance premium for the Medicare; ssb_t^{earl} is the social security benefit amount if person claims the social security at the earliest time possible[5] conditional on being in that age; ssb_t^{full} is the social security benefit amount at being made around the full retirement age; and ssb_t^{late} is the social security benefit amount, if a person delays her social security benefit.

2.1.3 Decision Variables

In every period the agent makes decisions regarding consumption level (c_t), labor supply [6], buying private health insurance (i_t^{prv}), buying health insurance from employer (i_t^{emp}), exercise (e_t) and, smoking (s_t). This makes the decision vector D_t:

$$\overrightarrow{D}_t = \{c_t, L_t, i_t^{emp}, i_t^{prv}, e_t, s_t\}$$

I assume that the job search cost is zero and an agent can decide to work unconditionally. However, income is a deterministic function of the state and decision variables and imposes restrictions on the budget constraint. I also assume that employer-provided health insurance is available, and it is only up to the agent's decision whether to purchase it or not The insurance premium does affect the agent's choice through the budget constraint.

2.2 Timing

Beginning each period t, agent is in state $\overrightarrow{S}_{t-1}$, which is the state at the end of the previous period. Variables such as gender and full retirement age are fixed through time, I also assume that education is fixed and since the study population contains only people older than 50 years, this assumption is not restrictive. Furthermore, I allow for fixed, unobserved heterogeneity. As a result, four of the state variables (gender, full retirement age, education, and unobserved fixed heterogeneity) are constant for each individual over time. However,

[5] Early retirement age is 62.

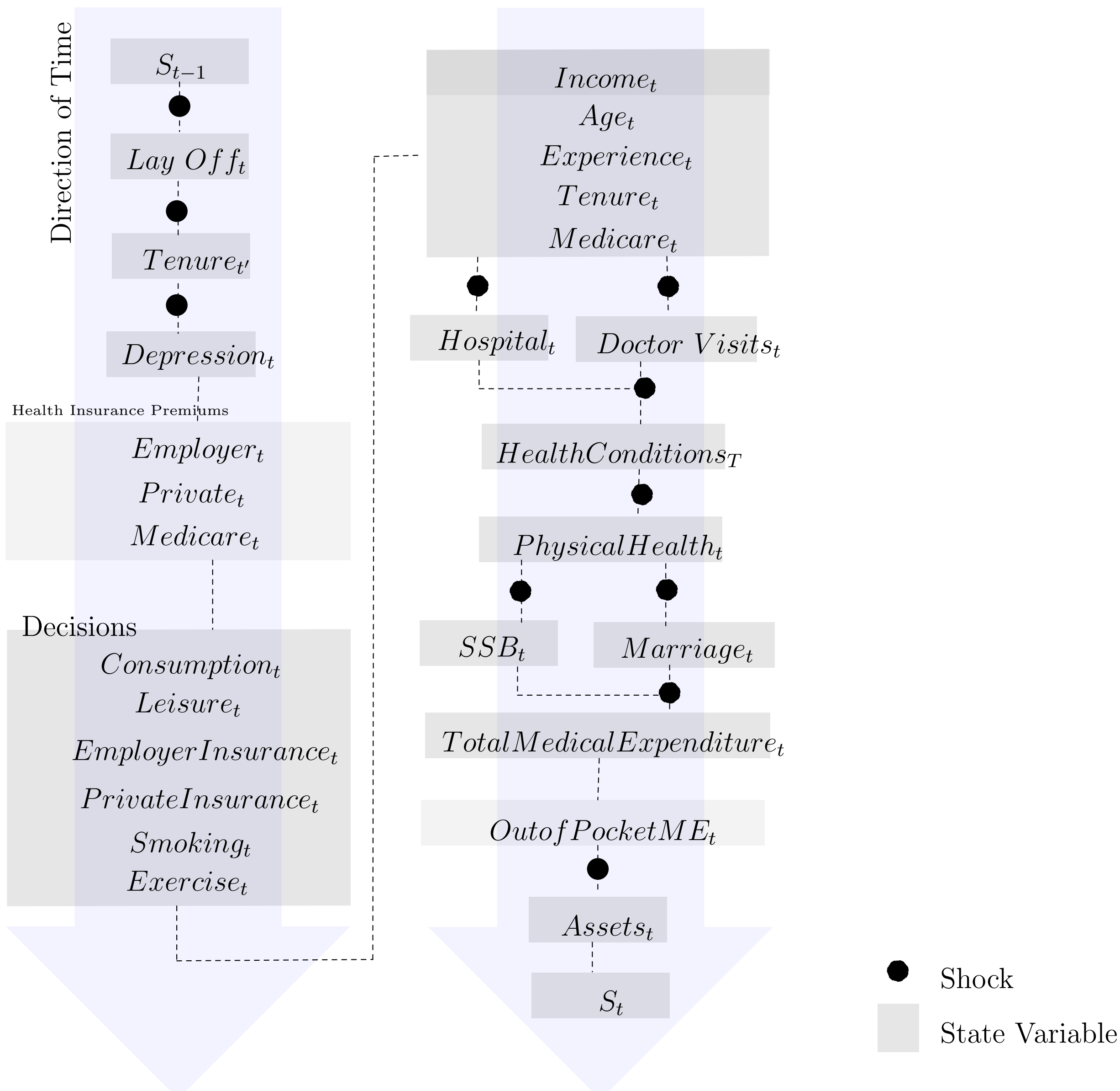
[6] Employment (e_t) or leisure (L_t)

other variables are changing through time. I implement a step by step updating procedure to better understand the model.

As depicted in figure 2.2, at the beginning of period t an economic shock hits the economy. Economic shock is a dummy that is one if a bad economic shock hits, and zero otherwise. The combination of the economic shock (σ_t^e) and S_{t-1} affects the probability of receiving a layoff shock for each individual, and this determines layoffs for individuals who were working in period $t-1$ (σ_t^l) and their tenure (x_t^1). Next, mental health will be updated stochastically knowing the occurrence of layoff (σ_t^l) and subsequently, health insurance premiums for three insurance types [7] will be revealed (ip_t^{emp} for employer-provided health insurance, i_t^{prv} for the private health insurance and $\imath_t^m$ for Medicare health insurance).

[7]Employer health insurance (i_t^{emp}), Private health insurance (i_t^{prv}) and Medicare (i_t^m) if accessible.

Figure 1: The updating sequence of the variables in the model.



In every period, agent inherits the state from the end of the previous period S_{t-1}. Showing shocks by the black circles, the updating sequence follows the direction which is shown by the two arrows.

Next, individual makes simultaneous decisions about the consumption level (c_t), leisure (l_t), buying health insurance from employer ($\imath_t^{emp}$), buying private health insurance ($\imath_t^{prv}$), smoking (s_t) and exercise (e_t). After the decisions, income from working (y_t), Age (A_t), tenure (x_t^1), experience (x_t^2) and receiving Medicare (i_t^m) are determined deterministically.

Then, both of the reactionary healthcare variables (number of nights in hospital and number of doctor visits) will be renewed, knowing the updated types of the health insurance, decisions, income, age as well as variables which have not yet been updated. In the next step, health shock is revealed and as a result, physical health (h_t^p) and the number of health conditions (h_t^c) can be determined, stochastically. Then, a marriage shock hits the individual and will determine marriage status (m_t). Social security benefit (SSB_t) and total health care expenditures (ME_t) are next to be renewed.

If the individual already had applied for the social security benefit, its amount will be unchanged, but if he had not been claimed it, conditional on reaching early retirement age, the individual would receive social security benefit shock that determines if he will claim it in this period or not. Then, the social security benefit amount is a deterministic function of the individual's updated state (S) and particularly work experience and income. Knowing the health insurance type, I deterministically set the health insurance co-payment levels and hence can determine the out-of-pocket medical expenditure (OOP_t). Finally, knowing assets level from previous period (a_{t-1}), consumption level (c_t), income (y_t), social security benefit (SSB_t,), health insurance premium and out-of-pocket medical expenditure, assets will be updated stochastically conditional on the interest rate (r_t). The timing is shown in the figure 2.2 and is explored in more details in the **Appendix** 6.

2.3 Utility Specification

The utility flow in each period is function of , consumption , health behaviors, health status, leisure and different types of the health insurance. I assume a semi CES preferences that is flexible and allows for unconstrained optimization of the likelihood. Moreover, the parameters will be estimated for three different types of heterogeneity:

$$
\begin{aligned}
&u(c, L, e, s, i^{prv}, i^{emp}, h^p, A, m, a) = \\
&\sum_{t=1}^{T} \beta_\delta^{t-1} \Bigg\{ \\
&I[h_t^p \neq 6] \times \Bigg\{ \Bigg(\Bigg(\exp(\bar{\phi} + \phi_\delta^0 + \phi_\delta^A.A_t + \phi_\delta^{h^p}.h_t^p) L^{1/\gamma_\delta^L} + \theta_\delta^e.\exp^{(\phi^e A_t)}.e_t + 1 \Bigg) \\
&.\Bigg(\theta_\delta^c.\Big(\frac{c_t}{(1+m_t)^{.7}} \Big)^{1/\gamma_\delta^c} + \theta_\delta^s.\exp^{(\phi^s.A_t)}.s_t + 1 \Bigg) \Bigg) + \Big(\theta_\delta^{prv}.\exp^{(\phi^{prv}.A_t)}.i_t^{prv} \Big) + \Big(\theta_\delta^{emp}.i_t^{emp} \Big) \Bigg\} \\
&+ [(h_t^p = 6] \times \Big\{ \lambda_\delta.a_t^{1/\gamma_\delta^c} \Big\} \Bigg\}
\end{aligned}
\tag{1}
$$

Where, A_t is age at time t, h_t^p is physical health, L_t is the leisure, c_t is consumption level, s_t is smoking dummy which is one if smoking, e_t is exercise dummy, i_t^{emp} is dummy that indicates the purchase of employer provided health insurance and i_t^{prv} is dummy variable for the private health insurance, ϵ_t is a choice specific unobserved (to the researcher) preference shock. The index δ indicates that the parameter is estimated separately for each unobserved heterogeneity types ($\delta \in \{1, 2, 3\}$). I follow Scholz and Seshadri (2013) and French et al. (2016) among others to use an equivalence scale for consumption, so that couples are consuming less than twice of the amount of separate individuals. Also I use a bequest function that allows the individuals to enjoy their bequest in the last period they appear in the model.

2.3.1 Assets law of motion

Assets evolve, by the following constraint:

$$
\begin{aligned}
a_{t+1} = \;& y_t + a_t(1 + r_t) \\
+ \;& SSB_t - OOP_t - c_t - ip_t^{emp}.i_t^{emp} - ip_t^{prv}.i_t^{prv} - ip_t^m.i_t^m
\end{aligned}
\tag{2}
$$

I do not model the cigarette price explicitly and assume that cigarette expenditure is negligible comparing to the other components of asset.

2.3.2 Health Insurance Premiums

Health insurance premiums for three types of insurance in the model are deterministic functions of updated states. More specifically I assume the covariate vector:

$$
\overrightarrow{X}_{ip} = [1, m_{t-1}, h_{t-1}^m, {h_{t-1}^m}^2, A_{t-1}, {A_{t-1}}^2, h_{t-1}^c, {h_{t-1}^c}^2, k_{t-1}^1, {k_{t-1}^1}^2, k_{t-1}^2, {k_{t-1}^2}^2, i_{t-1}^{emp}, L_{t-1}, s_{t-1}, g, f, E, E^2]
$$

and use three linear transformations, using pre-estimated coefficients $\overrightarrow{\beta}_{ip}$ to determine the premia;

$$ip_t^{emp} = \overrightarrow{X}_{ip} \times \overrightarrow{\beta'}_{emp} \quad (3)$$

$$ip_t^{prv} = \overrightarrow{X}_{ip} \times \overrightarrow{\beta'}_{prv} \quad (4)$$

$$ip_t^{m} = \overrightarrow{X}_{ip} \times \overrightarrow{\beta'}_{m} \quad (5)$$

2.3.3 Income

After individual makes decisions for consumption, labor supply, healthy behaviors and different health insurance types, income will set deterministically using a linear transformation:

$$y_t = \overrightarrow{X}_y \times \overrightarrow{\beta'}_y \quad (6)$$

Where

$$\overrightarrow{X}_y = [1, m_{t-1}, A_{t-1}, {A_{t-1}}^2, h_{t-1}^p, {h_{t-1}^p}^2, h_{t-1}^m, {h_{t-1}^m}^2, k_{t-1}^1, {k_{t-1}^1}^2, g, E, E^2, x_t^1, {x_t^1}^2, i_t^{emp}, L_t, s_t]$$

2.3.4 Out-of-Pocket Medical Expenditure

Out-of-pocket medical expenditure, depends on the health insurance type. Since, the different types and details of the insurance can affect the co-payment rates, and due to lack of data, I do not explicitly model all the details. I employ other attributes of the state space to better predict co-payments and hence the out-of-pocket medical expenditure. I predict the out-of-pocket medical expenditures, using the linear transformation :

$$OOP_t = \overrightarrow{X}_{OOP} \times \overrightarrow{\beta'}_{OOP} \quad (7)$$

Where

$$\overrightarrow{X}_{OOP} = [1, m_t, A_t, {A_t}^2, h_t^m, (h_t^m)^2, h_t^c, (h_t^c)^2, k_t^1, (k_t^1)^2, k_t^2, (k_t^2)^2, g, E, E^2 ... \\ , e_t, s_t, i_t^{emp}, i_t^{prv}, i_t^m, (i_t^{prv}.i_t^m), (i_t^{emp}.i_t^{prv}), (i_t^m.i_t^{emp}), (i_t^{emp}.i_t^{prv}.i_t^m)]$$

Where, $\overrightarrow{\beta}_{OOP}$ is a 1×24 vector of coefficients.

Social Security Benefits

The amount of the Social Security Benefits is determined using the transformation and shock structure which is discussed in details in **Appendix** 6.

2.4 Dynamic Optimization

Assuming forward looking agents and by utilizing Bellman's principle of optimality, the best possible value of the present discounted payoff is defined recursively:

$$V_t(\overrightarrow{S}_t, \epsilon_t) = \max_{\overrightarrow{D}_t} \Big(u(\overrightarrow{S}_t, \overrightarrow{D}_t, \epsilon_t(\overrightarrow{D}_t)) + \beta E(V_{t+1}(\overrightarrow{S}_{t+1}, \epsilon_{t+1}(\overrightarrow{D}_{t+1}) | \overrightarrow{S}_t, \overrightarrow{D}_t)) \Big)$$

where $\overrightarrow{S}$ is vector of state variables and $\overrightarrow{D}$ is vector of decisions. Assuming Conditional Independence (CI) and Additive Separability (AS), ex ante value function can be defined as:

$$\bar{V}_t(\overrightarrow{S}_t) = \int V_t((\overrightarrow{S}_t, \epsilon_t) g(\epsilon_t) d\epsilon_t = E \max_{\overrightarrow{D}} \Big(\nu(\overrightarrow{D}, \overrightarrow{S}_t) + \epsilon_t \Big) \tag{8}$$

where $\nu(\overrightarrow{D}, \overrightarrow{S}_t)$, is choice specific value function and can be written as:

$$\nu(\overrightarrow{D}_t, \overrightarrow{S}_t) = u(\overrightarrow{D}_t, \overrightarrow{S}_t) + \beta \int V(\overrightarrow{S}_t) dF(\overrightarrow{S}_{t+1} | \overrightarrow{D}_t, \overrightarrow{S}_t) \tag{9}$$

More specifically I define the choice specific value function as:

$$\begin{aligned} \nu(\overrightarrow{S}_t, \overrightarrow{D}_t) =& u(\overrightarrow{S}_t, \overrightarrow{D}_t) + \\ &+ \beta . \Bigg\{ \sum_{S_t} \Bigg(\pi(x_{2,t+1} | D_t) . \pi(A_{t+1}) . \pi(i^m_{t+1}) . \pi(k^1_{t+1} | \overrightarrow{X}_{k^1}) . \pi(k^2_{t+1} | \overrightarrow{X}_{k^2}) \\ & . \pi(h^p_{t+1} | \overrightarrow{X}_{h^p}) . \pi(h^c_{t+1} | \overrightarrow{X}_{h^c}) . \pi(m_{t+1} | \overrightarrow{X}_m) . \pi(SSB_{t+1} | \overrightarrow{X}_{SSB}) . \pi(ME_{t+1} | \overrightarrow{X}_{ME}) \\ & . \pi(r_{t+1} | \sigma^e) . \pi(\sigma^l_{t+1} | \overrightarrow{X}_l) . \pi(x_{1,t+1} | \overrightarrow{X}_{x_1}) . \pi(h^m_{t+1} | \overrightarrow{X}_{h^m}) \Bigg) . \bar{V}(\overrightarrow{S}_{t+1}) \Bigg\} \end{aligned} \tag{10}$$

where $\pi(.)$'s are the transition probabilities conditional on the most recently updated subspace of the state space and decisions, the vector $\overrightarrow{X}$ is defined as follows:

$$\overrightarrow{X}_{k^1} = [D_t, m_{t-1}, h^p_{t-1}, A_t, y_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$
$$\overrightarrow{X}_{k^2} = \overrightarrow{X}_{k^1}$$
$$\overrightarrow{X}_{h^p} = [D_t, m_{t-1}, h^p_{t-1}, h^c_{t-1}, a_{t-1}, g, E, A_t, L_t, y_t, k^1_t, k^2_t, e_t, s_t]$$
$$\overrightarrow{X}_{h^c} = [D_t, m_{t-1}, h^c_{t-1}, h^p_t, A_t, y_t, k^1_t, k^2_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$
$$\overrightarrow{X}_{m} = [D_t, m_{t-1}, h^p_t, h^c_t, A_t, y_t, k^1_t, k^2_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$
$$\overrightarrow{X}_{SSB} = [D_t, a_{t-1}, g, E, m_t, h^m_t, h^p_t, h^c_t, A_t, L_t, y_t, k^1_t, k^2_t, e_t, s_t, \sigma^e_t, c_t]$$
$$\overrightarrow{X}_{ssb} = [D_t, m_t, ssb_{t-1}, y_t, g, x^1_t, x^2_t]$$
$$\overrightarrow{X}_{ME} = [D_t, m_t, h^p_t, h^c_t, A_t, k^1_t, k^2_t, g, E]$$
$$\overrightarrow{X}_{l} = [D_t, m_{t-1}, x^1_{t-1}, y_{t-1}, \sigma^e_t]$$
$$\overrightarrow{X}_{x_1} = [D_t, L_{t-1}]$$
$$\overrightarrow{X}_{h^m} = [D_t, m_{t-1}, h^m_{t-1}, A_{t-1}, h^p_{t-1}, L_{t-1}, y_{t-1}, k^1_{t-1}, k^2_{t-1}, h^c_{t-1}, a_{t-1}, c_{t-1}, g, E, \sigma^e_t]$$

3 Data

I am using the Health and Retirement Study (HRS) and HRS-RAND (version P) database as the main source, to estimate the model. HRS is a nationally representative, biannual, longitudinal survey that interviews people older than 50 years old. I employ the waves 5-12 which matches the years 2000 to 2014. I also use Current Population Survey (CPS) monthly data from January 1998 to January 2014 to construct an instrument of the variable for involuntary layoff. In this section, I explain the construction of all the variables which were previously discussed in section 2.

I start with $181,033$ person-year observations and drop all the persons with missing values for the variables I use in estimation. I also drop all individuals who receive Medicaid, are not financially respondent of the household and observations with wealth greater than \$ 2,000,000. This leaves $115,488$ observations of $22,355$ unique individuals. In what follows in this section, I discuss the construction of the estimation variables. Variables are separated into four different groups: Attributes, employment, health, and monetary variables.

3.1 Attributes

Attributes include age (A), education (E), full retirement age (f), gender (G), and marital status (m). Since there is no ambiguity about the gender, I do not discuss it further. To construct age, I drop all the interviewees who are younger than 50 years. Then age bins will be constructed by leveling age at the even numbers, beginning from 50. As a result, both individuals who are 50 and 51, are coded as 50 in my estimation sample. This procedure is done to reduce the size of the state space while simultaneously keeping track of the dynamics in the biennial database of HRS.

HRS reports each person's years of education. I reduce it to seven categories. If an individual has less than nine years of education, it is coded as 0. Years of education between 9 and eleven is coded as 9 and reflects some high school. 12 is coded unchanged and is associated with a high school diploma. Then, anything between 13 and 15 years of education is considered 13 and depicts some college. 16 is coded as itself and is the college degree. Finally, anything more than 16 is coded as 17 and is regarded as graduate school experience.

Full retirement age is gradually changed from 65 to 66 by two months increments for every one year increase in the birth year beginning from 1938. **Table 1** shows changes in details. I set the full retirement age to 65 for people whose birth year is before 1943. All people who are born in 1943 or later are assigned full retirement age of 66.

Table 1: Full Retirement Age variation of the different cohorts

Year of Birth	<1938	1938	1939	1940	1941	1942	1943-1954
Full Retirement Age	65	65 and 2 months	65 and 4 months	65 and 6 months	65 and 8 months	65 and 10 months	66
Year of FRA	<2003	2003-4	2004-5	2005-6	2006-7	2008-9	2009-2020

3.2 Employment

Employment variables are leisure(work), tenure in the current job, experience, and involuntary layoff. Leisure takes three values of 0, 1/2, and one, if a person works full-time, part-time, and not working, respectively. Following the RAND database definition, I assume that person is working full-time if working 35+ hours per week, 36+ weeks per year. Any amount less than this (1260 hours) is considered as part-time. Full-time job status reflects 1800 hours annual work. I assume people who reported their labor force status as unemployed, retired or not in labor force are all not working and set their leisure to 1.

Tenure reports the tenure years in the current job. I code it as a discrete variable with two years increments, starting from 0. As a result, an individual's tenure is coded as zero if either he reports zero or one. I top code the individuals with more than 20 years of tenure as 20 to set the limit for the state space at the expense of the loss of outliers.

Experience keeps track of total years of work for each individual. Similar to tenure, I code experience by two years increments and to code it at 40 years of experience. As a result, experience takes one of the 20 possible values increasing by two years increments starting from 0.

Involuntary job loss is a dummy variable that is one if job loss is due to layoff or business closure and zero otherwise. Thus, all other job losses are coded as a quit and not considered involuntary layoff. Using the CPS monthly, short panel data, I calculate Job Separation and Job Entrance rates for each industry-census region. Then I use these measures as instruments for involuntary layoff to address endogeneity concerns caused by the mental and physical health status effects on the probability of receiving layoff shock. In the CPS I can observe each individual for four consecutive months. The individual is excluded from the survey for eight months, then included again for an additional four months. If individual reports that he is not working in the industry that was reported in a previous month, I add one job separation for current month-industry-region. The denominator of the Job Separation ratio is the last month's total number of people in that industry-region. The same logic is used when creating the job entrance ratio.

3.3 Health

Measuring health can be a subtle task due to the ambiguity in the definition of health. The word "Health," in contrast with its everyday life usage, does not have a precise definition. While cancer and feeling of pain both affect "health," they affect different decisions. I separate the analysis of health into the three subcategories of physical health, mental health, and health conditions.

HRS provides self-reported health status, as well as a rich set of questions which are designed to objectively report different dimensions of an individual's health. While it is convenient to use a uni-dimensional measure of health such as self-reported health, as a measure of physical health, the subjectivity of this measure and the problem of potential misreporting, can contaminate self-reported health as a proxy for physical health and hence any result which is made based upon it. On the other hand, relying upon objective, multi-dimensional sets of variables to summarize a person's health, inherits the problem of revealing the correct weight for each dimension. I address this issue using Item Response Theory (IRT) as discussed by Firouzi Naeim (2017). I assume there is an underlying physical health status that governs an individual's response to the different items. Items are sets of objective questions; each is measuring a particular aspect of physical health. I inform the impulse response function in the IRT process by a set of objective variables which are designed to measure functional limitations and health dummy that takes value one if individual reports that her health takes a value greater than or equal to 3. HRS asks individuals three sets of questions under a general category of Functional Limitations: Activities of daily living (ADL), Lower Body Mobility (LBM) and Upper Body Agility (UBA). The underlying variables in each group and their associated survey question is summarized in **Table** 2. I assume that the probability of a binary response to each item follows a Probit model and estimate two parameters for each item: Sensitivity and scale. Sensitivity provides information on the exactness of the item and its sensitivity to the true underlying physical health. The scale shows how likely is it that individuals respond to an item either negatively or positively. I employ a Bayesian approach, more specifically the Gibbs sampling method, to uncover the underlying physical health (a posterior distribution) of each individual in the sample at each date.

To provide aspects of health that self-reported health does not represent, HRS asks individuals if they have ever been told by a doctor that they have high Blood Pressure, Diabetes, Cancer, Lung Disease, Heart Problems, Stroke, Psychological Problems or Arthritis. I make a new variable (number of health conditions) by summing up the different health conditions. This variable takes integers from 0 to 8.

For mental health, I use CES-D[8] to track the depression patterns as a proxy for mental health. CES-D is an index that is made by aggregating different objective mental health-related measures in the survey and takes integer values from 0 to 8.

3.4 Monetary Variables

This group includes income, assets, consumption, and social security benefits. I adjust all variables for inflation, using the Consumer Price Index (CPI) and the base year 2000. As a result, all the variables are in real terms. I drop observations with $ 200,000 or more in real income as high earners. Then, I discretize income into five bins, with increments of 40,000 $. Assets are the sum of all wealth components after subtracting by all debts. (Value of primary residence + Net value of real estate + Net value of vehicles + Net value of businesses + Net value of IRA, Keogh accounts + Net value of stocks, mutual funds, and investment trusts + Value of checking, savings, or money market accounts + Value of CD, government savings bonds, and T-bills + Net value of bonds and bond funds + Net value of all other savings) - (Value of all mortgages/land contracts (primary residence) + Value of other home loans (primary residence) + Value of other debt)

Finally, to measure consumption, I use spending data from RAND-CAMS version D,

[8] Center for Epidemiologic Studies Depression Scale

Table 2: Functional limitations

Variable	Question
	Lower Body Mobility
RwWALKS	Because of a health problem do you have any difficulty with walking several blocks?
RwJOG	(Because of a health problem do you have any difficulty) with running or jogging about a mile?
RwSIT	(Because of a health problem do you have any difficulty) with sitting for about two hours?
RwCHAIR	(Because of a health problem do you have any difficulty) with getting up from a chair after sitting for long periods?
RwCLIMBS	(Because of a health problem do you have any difficulty) with climbing several flights of stairs without resting?
	Upper Body Agility
RwSTOOP	(Because of a health problem do you have any difficulty) with stooping, kneeling, or crouching?
RwARMS	(Because of a health problem do you have any difficulty) with reaching or extending your arms above shoulder level?
RwPUSH	(Because of a health problem do you have any difficulty) with pulling or pushing large objects like a living room chair?
RwLIFT	(Because of a health problem do you have any difficulty) with lifting or carrying weights over 10 pounds, like a heavy bag of groceries?
RwDIME	(Because of a health problem do you have any difficulty) with picking up a dime from a table?
	Activities of daily living (ADL)
RwBATH	(Because of a health or memory problem do you have any difficulty with) bathing or showering?
RwEAT	(Because of a health or memory problem do you have any difficulty with) eating, such as cutting up your food?
RwBED	(Because of a health or memory problem do you have any difficulty with) getting in or out of bed?
RwDRESS	(Because of a health or memory problem, do you have any difficulty) dressing?
RwWALKR	(Because of a health or memory problem, do you have) walking across a room?
RwTOILT	(Because of a health or memory problem, do you have) any difficulty using the toilet?

Name of the variable column shows the names as recorded in RAND database.

which is based on the Consumption and Activities Mail Survey (CAMS) supplement to the HRS data. Its main focus is to measure total household spending. It covers about 20% of the HRS respondents and asks detailed information on durables, non-durables, transportation, and housing. I employ a flexible, semi-parametric method using income, wealth, change in wealth, education, size of the household, physical health, marriage status, and gender, to calculate the household consumption for the observations with missing values.

4 Estimation

Estimation of the dynamic discrete choice models, like the one suggested in this study, involves finding the optimum path for each possible state of the world and requires the evaluation of all the possible paths for each point of the state space to find the best one. Such evaluation usually can be achieved by solving a Bellman equation for the problem and is called full solution method. However, the computational cost of the full solution method is high and can be infeasible when the model incorporates a high dimensional, large state space. This problem is called the curse of dimensionality and refers to the fact that by adding one new dimension to the state space the paths that researcher should evaluate for each point of the state space is multiplied by the number of possible values of the new dimension. This problem is even more severe in dynamic optimization. Namely, if the size of the different combinations which should be evaluated to find the optimum path for each state is $O(S_0^T)$, where S_0 is state space before adding a new dimension and T is number of periods that researcher allows the agents to look into the future, adding new dimension with s_1 new possible points increases the problem size to $O((S_0 * s_1)^T)$.

To prevent the curse of dimensionality, the number of non-full solution methods were developed. Rust (1987) assumes a Markovian process with the conditional independence of the state space (S_t) and unobserved preference shocks (ϵ_t), where flow utilities are additively separable. This allows for choice specific value functions of the form shown by equation 9.

For the estimation purpose, I rely on the Conditional Choice Probability estimator (CCP), suggested by Arcidiacono and Miller (2011). Utilizing the computational advantage of the CCP, compared with the full solution methods, I can estimate a complex model with a high dimensional state space. Also, CCP can be combined with the Expectation Maximization (EM) algorithm that allows for the fixed unobserved heterogeneity in the model. Both of the computational inexpensiveness and the incorporation of the unobserved heterogeneity are appealing in the models of health insurance and retirement. However, the computational advantage of the CCP is not costless. CCP is a data-intensive method and in the lack of a database that covers the state space, relies on the smoothing techniques. Hotz et al. (1994) show that when we rely on the nonparametric estimates of the choice probabilities, smoothed estimators can reduce the bias. In this study, I utilize a large number of observations, as well as smoothing techniques to address this concern.

CCP is a 2-stage maximum likelihood estimator and hence separates the coefficients of the model into first stage coefficients which are the coefficients of the transition probabilities and production functions, and second stage coefficients which appear in the utility function. In what follows in this section, I explain the estimation method, as well as estimation results, for each of the two stages.

4.1 Identification

To be able to predict the effect of a counterfactual policy (raising the Medicare Eligibility Age in this study), structural estimations rely on the policy-invariant estimation of the structural parameters in the model. The first problem to address is the Lucas critique in the Brunner, K. and Meltzer, A. (1983) that questions the validity of the marginal effect that is based on the primitives which are themselves a function of policy intervention. The treatment understudied in this paper (receiving Medicare health insurance) is not a function of states (such as health status) and decisions (Labor supply and health behaviors) of individuals and hence estimating the utility parameters and beliefs are independent of the treatment. In other words, As a result, the choice of treatment serves as an identification assumption as discussed by Khwaja, A. (2010).

While the policy-invariant estimate of the marginal effect is the necessary condition of the validity of the counterfactual analysis, it is not sufficient. Another problem of so-called *the identification problem* in the structural estimation argues the sensitivity of the marginal effects to the parameters of the model. The identification problem occurs when the observational likelihood of the two different sets of structural parameters (primitives) are equal, or in other words, two sets of primitives are doing equally well in fitting the data and parameter sets cannot be ranked by evaluating their goodness of fit. This problem is vital noting that two different sets of parameters with the same performance in the in-sample goodness of fit might produce different out-of-sample predictions.

To address the problem of identification, I control for two sources of endogeneity: First is the model endogeneity that emerges from the relationship between contemporaneous variables. The second source of endogeneity is the endogeneity from the initial conditions that refer to the unobserved decisions individuals already have made through life which have led them to the states that are observed by the researcher for the first time in the sample when people are in their midlife.

4.1.1 Model Endogeneity

Functional specification of the relationship between different contemporaneous state variables along with the detailed timing (explained in Section 2.2 and Figure 2.2), are imposed for better identification. However, while the timing convention reduces the identification problem for the variables with the intertemporal relationship, it is not useful when the endogeneity is derived from the contemporaneous relationships. To address the contemporaneous endogeneity, I rely on the functional assumptions and two sources of exogenous variations in my model. I utilize the business cycle shocks as the first source of exogenous variation that affects the stochastic interest rate and the probability of involuntary layoff and introduces exogenous variations in the time dimension. For the business cycles, I use different periods of recessions in the U.S. as reported by the NBER's Business Cycle Dating Committee. The second source of exogenous variation is through the involuntary layoff. HRS dataset indicates involuntary layoff for an individual if the person is not working in the same job as used to work in the previous wave, and reports the reason as "involuntary layoff" or "business closure." However, the relationships between mental health and physical health and the involuntary layoff might still suffer from reverse causality, i.e., while involuntary layoff affects the probability of being in the poor mental and physical health, person's health also affects the probability of receiving an involuntary layoff shock. I utilize the Current Population Survey data (CPS) to calculate the industry level job separation rate in each census region for every period, and employ this rate as an instrument for the probability of involuntary job loss for individual at time t, who works in particular industry j and lives in a census region cr, to isolate the effect of involuntary layoff on the mental and physical health. This procedure works as the inter-state exclusion restriction and improves the model identification by detaching the link from different health dimensions to the labor supply decisions.

While I use two sources of exogenous variation, the identification power which is introduced by them can be more than what we would expect in a static model. As discussed in the Proposition 1 of Bhargava (1991), utilizing the cross-equation restrictions when exogenous time-varying variables are introduced into the model in every period can provide additional identification conditions and less demanding identification sufficient condition, as opposed with the sufficient condition stating the number of time-invariant exogenous variables should be greater than time-varying endogenous variables. In the long panels, exogenous variations of the time-varying instruments are transmitted to the next period when the evolution of the endogenous variables follows a Markovian process and as a result instruments in the different periods have separate effects on the endogenous variables, and it leads to the multiplicity of instruments. In this study business cycles interacts with the region and industry that individual is associated with them and leads to the much more exclusion restriction than involuntary layoff alone.

Also, Mroz and Savage (1998) argue the non-linearities in the dynamics of the contemporaneous variables can provide even further moment conditions than what is expected from the lagged variables in the Markovian process.

4.1.2 Initial Conditions

Without any adjustments, initial conditions[9] introduce bias in the estimation of the structural parameters, if heterogeneities which are not modeled (unobserved heterogeneities) interacting with the decisions which were made through life cycle, before the observed periods and determine the initial conditions. In the Health and Retirement Study data which I use in this paper, individuals are interviewed for the first time when they are at least 50 years old. As a result, observed initial characteristics of the people in the sample are not exogenous. The initial observed states are based on the decisions that people made in life before entering into the sample, and as a result, not addressing the initial condition problem produces the biased estimates of the parameters and hence the marginal effects. To resolve this problem, I follow Van der Klaauw and Wolpin (2008) and estimate the unobserved heterogeneities conditional on the initial conditions.

4.2 Estimation Method

4.2.1 Simulated Value Function (SV)

Assuming the Additive Separability (AS) of the flow utility and the preference shocks, and Conditional Independence (CI) between the state (S_t) and preference shocks, the full information maximum likelihood estimator θ can be defined as:

$$\hat{\theta} = \underset{\theta}{\text{argmax}}\, l(\theta) = \prod_{a=1}^{A}\prod_{t=1}^{T_a} Pr(D_t^a|S_t^a, \theta_f, \theta_u) f(S_t^a|S_{t-1}^a, D_{t-1}^a, \theta_f)$$

This likelihood function is separable and can be estimated in two separate steps. Calling it 2-stage maximum likelihood, the parameters which govern the transition probabilities (θ_f) can be estimated at first, and then given the estimated parameters of the first stage ($\hat{\theta}_f$), the parameters in the utility function (θ_u) will be estimated using a likelihood function that utilizes the choice probabilities, $Pr(D_t^a|S_t^a, \theta)$. As discussed by McFadden (1981) and Rust (1987), If preference shock is an IID extreme value distribution, Conditional choice probabilities can be written as follows:

$$Pr(D_t|X_t) = \frac{1}{\sum_{D_t} exp(\nu(S_t, D_t) - \nu(S_t, d_t))}$$

As a result, reaching the consistent estimate of θ_u involves calculating of the choice specific value functions as described in equation 9. However, even by utilizing the Bellman's prin-

[9] The conditions which are observed for the first time when individual enters the data set

ciple of optimality, this process is computationally expensive. Hotz and Miller (1993) show that if preference shock is an IID extreme value distribution and CI and AS are satisfied, the difference between choice specific value functions can be defined as a function of the conditional choice probabilities $P(D_t|S_t)$. More specifically, the difference between the value of choosing decision D and an anchor decision 1 can be written as:

$$\Delta\hat{\nu}(S,D) = log\{\frac{P(D|S)}{P(1|S)}\}$$

This mapping between decision space and preference is helpful to reveal the difference between the choice specific value functions without pursuing the backward induction procedure which is required in the full solution method of estimation. Hotz et al. (1994) introduce the simulated value (SV) function estimator that is a smooth function of the structural parameters θ. The idea behind SV estimator is to use a non-parametric estimation to get consistent estimate of Conditional Choice Probabilities and then inverting them to obtain the value function which is normalized by an anchor choice ($\Delta\ \widehat{\nu}(\mathrm{x},\mathrm{d}) = \nu(x,d) - \nu(x,d_o)$).

Ideally, we can use a bin estimator of $p(D_t|S_t)$ as a completely non-parametric estimate that gives us the probability of each choice in every state. But data limitations force us to employ some smoothing techniques. For this purpose, I smooth the surface of conditional choice probabilities by employing a flexible functional form including quadratic terms and interactions of state variables in a logistic platform. For transition probability of stochastic state variables $f(S_t|S_{t-1},D_{t-1})$, I employ sets of the Multinomial Logistic models with lagged variables, and other controls as explanatory variables. The algorithm to apply the SV estimator of Hotz et al. (1994) as described by Rust (1994) is based on knowing the transition probabilities of stochastic state variables or individual beliefs, $f(S_t|S_{t-1},D_{t-1})$ and CCP estimates of $\hat{P}(D\ |.\ S)$. Then the Simulation Algorithm is defined as follows:

1. Calculating $\Delta\ \widehat{\nu}(\mathrm{x},\mathrm{d})=\ \log\{\frac{\hat{P}(d|x)}{\hat{P}(1|x)}\}$ [Hotz and Miller (1993)]
2. Then for each state, choice pair (data points) and each person in each period, use (x_t^n, d_t^n) as the initial point and:
   (a) Given (x_{t-1},d_{t-1}) draw x_t from previously estimated $\pi(x_t|x_{t-1},d_{t-1})$
   (b) Given x_t from previous step draw ϵ_t from assumed $q(\epsilon_t\ |.\ x_t)$ [EV1] and keep them as $\widetilde{\epsilon_t}$
   (c) Calculate $d_t = \widehat{\delta}(x,\epsilon) = \mathrm{argmax}_{d\in D(x)}\ [\Delta\ \widehat{\nu}(\mathrm{x},\mathrm{d}) + \epsilon(d)]$
   (d) Repeat until reaching terminal period (or until time preference makes the next period's effect infinitesimal).

3. Now that we have simulated values of $(\widetilde{x_t},\ \widetilde{d_t})$, using each initial value from step 2, compute the simulated value function for each θ:

$$\widetilde{\nu_\theta}(x_0, d_0) = \sum_{t=0}^{T} \beta^t \{u_\theta[\widetilde{x_t}, \widehat{\delta}(\widetilde{x_t}, \widetilde{\epsilon_t})] + \widetilde{\epsilon_t}\}$$

4. Knowing the simulated choice specific values, $\nu_\theta(S_0, D_0)$, I can form the conditional choice probabilities as a smooth function of θ_u:

$$\tilde{Pr}(D_t|X_t, \theta_u) = \frac{1}{\sum_{D_t} exp(\Delta\tilde{\nu}(S_0, D_0))}$$

And estimate the second stage likelihood estimator of θ_u:

$$\widetilde{\hat{\theta}}_u = \underset{\theta_u}{\operatorname{argmax}}\, L(\theta_u) = \sum_{a=1}^{A}\sum_{t=1}^{T_a} Log[\tilde{Pr}(D^a_{t+1}|S^a_{t+1}, \theta_u, \hat{\theta}_f)]$$

Since the simulated paths are sensitive to the simulated preference shocks $\tilde{\epsilon}_t$, I need to repeat the process above for each observation a to acquire a consistent estimator of θ_u. In this study, I expand each observation to 10 observations. Thus the maximum likelihood estimator averages over the $\tilde{\epsilon}_t$ and corrects for the possible bias.

The estimation method used in this paper is based on the 2-stage maximum likelihood and allows for incorporating fixed and unobserved heterogeneity types. Next section provides the adjustment in the likelihood function that is needed for this purpose.

4.2.2 Incorporating the Unobserved Heterogeneity

Following Arcidiacono and Miller (2011) I use the modified Expectation Maximization (EM) algorithm to address the fixed, unobserved heterogeneity. EM iterates over two steps. In the Expectation step, the probabilities of choosing each alternative conditional on the observed and unobserved states are updated. The maximization step can be done assuming the unobserved states are observed and using the probabilities of each alternative conditional on being in particular unobserved state as weights. The full explanation of the estimation method, consisting the Simulation Value function estimator and Conditional Choice Probability that allows for the unobserved heterogeneity in the model are discussed in details in the Appendix 6.

4.2.3 Policy Sensitive Value Functions

The value functions calculated by utilizing the CCP estimation method, are not sensitive to the change in policy. It is because they are non-parametric functions of the data and in particular related to the relative frequencies of the observed decisions. Since the sample is fixed under the current policy, value functions cannot be recalculated for the counterfactual policy.

This problem can be solved by solving the Bellman's equations recursively for the estimated values of the structural parameters and assigning the new value to each point in the state space, under each counterfactual policy. Starting from the terminal period T, each state receives a vector of values where the elements are associated with each decision available at the state, determined by the current utility (assuming there is no period ahead). Then the best decision is chosen based on the maximum value that can be reached. This optimum value then is assigned as the state and is called the value of state at T. Then moving one step back to the period $T-1$, the expected value of each decision will be calculated for each state by employing the transition probability matrix and the values calculated for the period T. These expected utilities need to be added to the current period's utility associated with each decision. Then we proceed by finding the maximum value that can be acquired from different available decisions. This process should be repeated until the value of the state is not changing by going backward in time. By repeating this backward induction for each counterfactual policy, I can assign a policy specific value function for each state.

Since I have already estimated the structural parameters by utilizing the non-full-solution estimation technique (CCP), it is not required to search over the parameter space to find the parameters which are optimally explaining the sample. However, I need to solve the Bellman's equation once for each counterfactual policy. Note that solving Bellman's equation by backward induction involves the curse of the dimensionality problem which was discussed earlier and adds exponentially to the computational expense. Considering the number of dimensions that the model in this study introduces, I have a large state space that makes the direct solution to the Bellman's equation infeasible.

To reduce the computational complexity and to be able to solve the Bellman's equation, I follow the interpolation based approximation technique suggested by Keane and Wolpin (1994). The technique involves the estimation based interpolation at every step of the backward induction. Starting at the terminal period T, approximation involves taking J draws from the joint distribution of the decision specific preference shocks and calculating the maximum value of the value functions over the decision space. Utilizing the Monte Carlo integration, it is possible to reduce the computational expense even more, by simulating the choice of specific value functions. Interpolation is based on the ordinary least squares regression of the ex-ante value functions and choice specific value functions. Then the values of the non-chosen points in the state space can be estimated.

While the interpolation technique makes it possible to assign values to each point of the state space, it involves the approximation at every step of the backward induction problem and increases the probability of the identification problem, by assigning same values to the points in the state space for a different set of the estimated parameters. This problem is more severe if the interpolation function does not support enough curvature, and is directly related to the number of the points one chooses for the exact calculation of the value functions. I alleviate this problem by avoiding approximation at every step of the backward induction, and use the interpolation only when I reach the time zero in the backward induction.

I avoid the curse of dimensionality by only limiting state space size to the states which are observed in the data and the states which can be reached highly likely from the observed states. I assume the observed states are those who are highly likely to be reached from the other states, and those states that are not observed are not likely to be reached in the society that I study. Adding the other highly likely states to the set of states I use in the backward induction process, is to increase the flexibility of the simulations when I study the marginal effects of the counterfactual policies. It is important to note that even after directly calculating the values of all observed and highly likely states, the Keane and Wolpin (1994) interpolation technique is crucial to the simulation because it provides the value to the states that might be reached in the forward simulation of each counterfactual policy.

4.3 Structural Estimation Results

4.3.1 First Stage: Transitions and Production Functions

Estimation of the transition probabilities and production functions follows the timing of the model that was discussed in **section** 2.2. The probability of transitioning from one point in a particular dimension of the state space to another point is modeled as multinomial logistic regression. Then assuming independence between different dimensions of the state space, probability of transitioning from one point of the state space to another can be calculated by multiplying the estimated probabilities in each dimension [10]. I avoid imposing the theory to the regression functions unless the channels are clear either by the nature of the transition

[10] Knowing that individual has Medicare and is 60 years old, the probability of transitioning to the state with Medicare ($P[i_{t+1}^{Med} = 1 | i_t^{Med} = 1)] = 1$) and being 58 ($P[A_{t+1} = 58 | A_t = 60] = 0$) is zero:

$$P[i_{t+1}^{Med} = 1, A_{t+1} = 58 | i_t^{Med} = 1, A_t = 60] = P(i_{t+1}^{Med} = 1 | P(i_t^{Med} = 1)).P(A_{t+1} = 58 | A_t = 60)$$

or by abundant evidence in the literature. However, I impose a detailed structure into the timing of the model as discussed previously. The specification of each logistic regression and its index function and the estimated probabilities are discussed in the Appendices 6 and 6, respectively.

4.3.2 Second Stage: Utility Parameters

Table 3 shows the estimation results, for the parameters of the utility function 1 which are estimated in the second stage.

Propensity for Leisure: There are three types of unobserved heterogeneity. All three types show the positive trend in the propensity for leisure as individuals grow older and their health deteriorates. However, the positive association between increase in age and health deterioration and propensity for leisure weakens as the unobserved heterogeneity type is changing from type one to type tow and three respectively, meaning leisure is becoming more desirable for type one in older ages and worse physical health, than type two and type three. The effect of this association cannot be interpreted directly without taking into account the dynamic and highly non-linear nature of the problem, but one can roughly conclude that individuals with the type one unobserved heterogeneity tend to work less as they are getting older and their health deteriorates, compared with the type two and type three. For type three, deterioration in health has a positive effect on the propensity for leisure. However, the effect is not statistically different from zero. Considering the fact that leisure takes three values: 0, 1/2 and 1, small values for γ^L in all three unobserved heterogeneity types indicate a large difference in utility, between working full-time job ($L = 1$) and two other possible decisions toward labor supple (part-time job and retirement). In another words, individuals do not distinguish between retirement and part-time job as much as they distinguish between full-time job and part-time job. As a result, there is a jump in utility level, as individual changes his choice from full-time job to any of the other two choices. This gap in utility is smaller for type one unobserved heterogeneity than the other two types.

Propensity to Consume: Consumption has positive utility for all three unobserved heterogeneity types, as expected by the consumer theory. Determined by θ^c (slope) and γ^c (curvature) in Table 3, individual with type one unobserved heterogeneity enjoys consumption less than type two and type two enjoys it less than type three. However, the increments in the consumption are more distinct for those with type one unobserved heterogeneity than type two or three. The combination of the two competing effects of slope and curvature shows that marginal propensity to consume is higher for the individual with type one unobserved heterogeneity. Changing the consumption from lowest possible normalized value ($\frac{1}{6}$) to the maximum consumption level (1) increases the utility by 3.55 times for the type one unobserved heterogeneity compared with 2.47 for the type three unobserved heterogeneity. Disregarding the complex dynamic in the decision making process where the present discounted value of the future utilities determines individual's decision, it is possible to say that individuals with the type one unobserved heterogeneity are more sensitive toward changes in consumption than those with type two or type three unobserved heterogeneity. However, individuals with type three and type two unobserved heterogeneity are more likely to favor decisions which promote more consumption on average (determined by the slope parameter θ^c).

Health behaviors: Health behaviors are represented in the model using smoking (s) and exercise (e). Shown by ϕ^s and ϕ^e, both smoking and exercise gain relative importance in individual's decision making process. For example, individual who distastes smoking is more hesitant to smoking as he gets older. Individuals show different attitudes toward health behaviors depending on the unobserved heterogeneity types. θ^s and θ^e indicate this attitude. Those individuals with the type one unobserved heterogeneity dislike smoking while indifferent about exercise and can be categorized as healthy type. Those with type two unobserved heterogeneity enjoys smoking and dislikes exercise and I call them unhealthy type. Individuals with the third unobserved heterogeneity type dislikes both smoking and exercise.

Bequest Preference and Discount Factor: Unobserved heterogeneity type determines individual's tendency toward bequest. Tracked by λ, agents with type one unobserved heterogeneity appreciate positive bequest values, while the other types do not enjoy the bequest. The discount factor (inter-temporal time preference) is about the same for all types of unobserved heterogeneity and is 0.856, 0.86 , and 0.857 for unobserved heterogeneity type one, two ,and three, respectively. In another words, discounted present value of one unit of utility in 20 periods (40 years) is translated to 0.044, 0.49, and 0.045 for the three unobserved heterogeneity types.

5 Results

I discuss the results for men with 12 years of education. In what follows, I evaluate the in-sample prediction power of the estimated model further more, by comparing the simulated

Table 3: Utility Parameters Estimates

Parameter	$\delta = 1$	$\delta = 2$	$\delta = 3$
$\bar{\phi}$	-1.235	-1.235	-1.235
	(.103)	(.103)	(.103)
ϕ_δ^0	-1.177	-0.185	0.703
	(.177)	(.093)	(.107)
ϕ_δ^A	2.238	1.692	0.985
	(.194)	(.14)	(.301)
$\phi_\delta^{h^p}$	1.012	0.373	0.006
	(.164)	(.194)	(.143)
γ_δ^L	0.183	0.174	0.192
	(.072)	(.183)	(.124)
θ_δ^c	0.083	0.104	0.185
	(.043)	(.029)	(.043)
θ_δ^s	-1.046	0.468	-0.347
	(.002)	(.07)	(.105)
θ_δ^e	0.009	-0.152	-0.321
	(.173)	(.042)	(.009)
γ_δ^c	1.380	1.162	2.070
	(.146)	(.043)	(.215)
θ_δ^{prv}	-1.268	-1.277	-1.931
	(.013)	(.22)	(.416)
θ_δ^{emp}	0.951	0.170	0.324
	(.007)	(.063)	(.01)
β_δ	0.856	0.860	0.857
	(.018)	(.077)	(.106)
λ_δ	0.092	0.001	0.001
	(.013)	(.03)	(.122)
ϕ^{prv}	0.572	0.572	0.572
	(.328)	(.328)	(.328)
ϕ^s	0.899	0.899	0.899
	(.145)	(.145)	(.145)
ϕ^e	0.727	0.727	0.727
	(.021)	(.021)	(.021)
N	26,110	13,102	26,562

Standard Deviations in parantheses
Utility parametrs estimates.

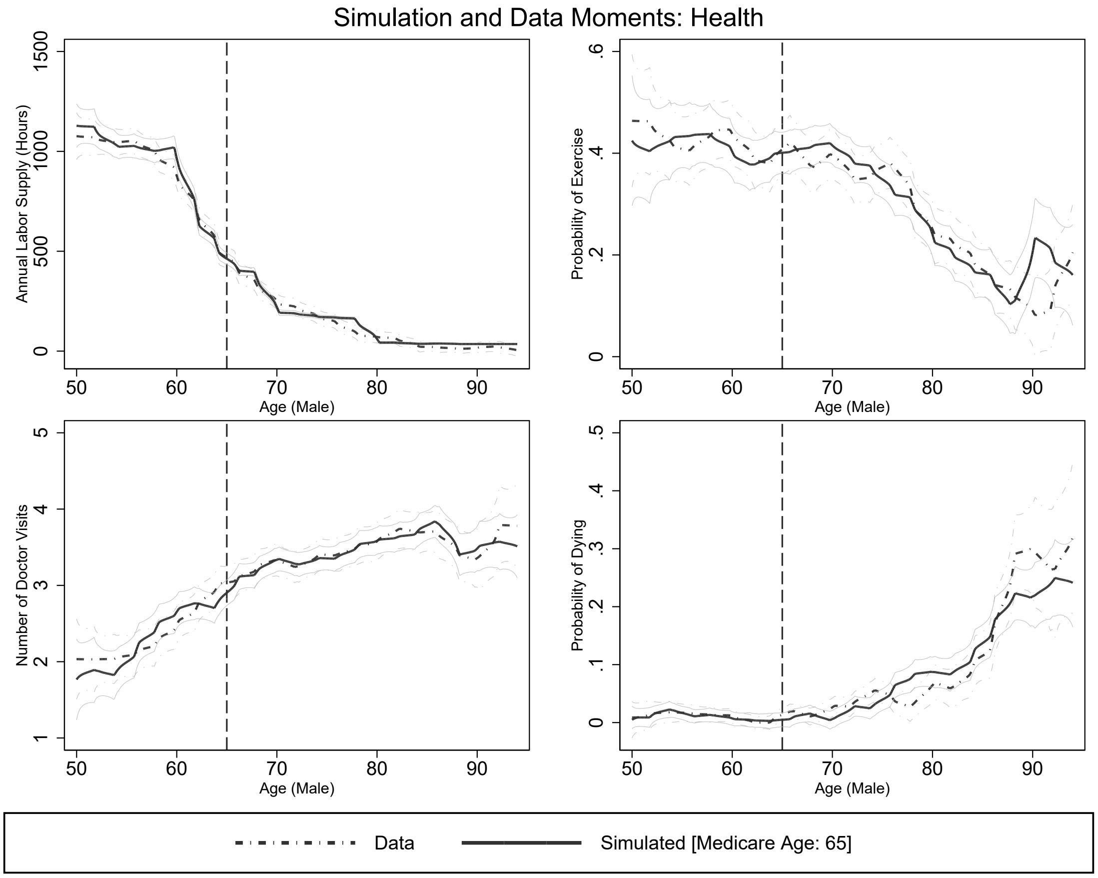
data and the data from HRS. This is done by looking at the moments of the two “samples”. After verifying the in-sample credibility of the estimated model, I analyze the marginal effects of the change in the policy by changing the Medicare Eligibility Age from 65 to 69.

5.1 Evaluating the Moments

To evaluate how well the simulation matches the data I compare the simulated results based on the current policy (Medicare Eligibility age at 65) and data from HRS using the unconditional means by age. I take the wave 5 of the HRS data that corresponds with the year 2000 and simulate it forward using the estimated utility function parameters evaluated for using the current policy. As a result, I can compare the real panel data from HRS to the simulated panel.

The moments are defined for each variable as the age-specific average of the observations. For example, I calculate the average health status at age 54 for the HRS data and the simulated data. To compare the two averages, I utilize the t-statistic for the null hypothesis that states the two averages are equal. A summary of the results for the main outcomes are provided in the Table 17 and in Figure 2. As shown, the model is replicating the observed data closely. The full evaluation of the moments is discussed in the **Appendix** 6.

Figure 2: First moment comparison of the real data and simulated data under the current policy.

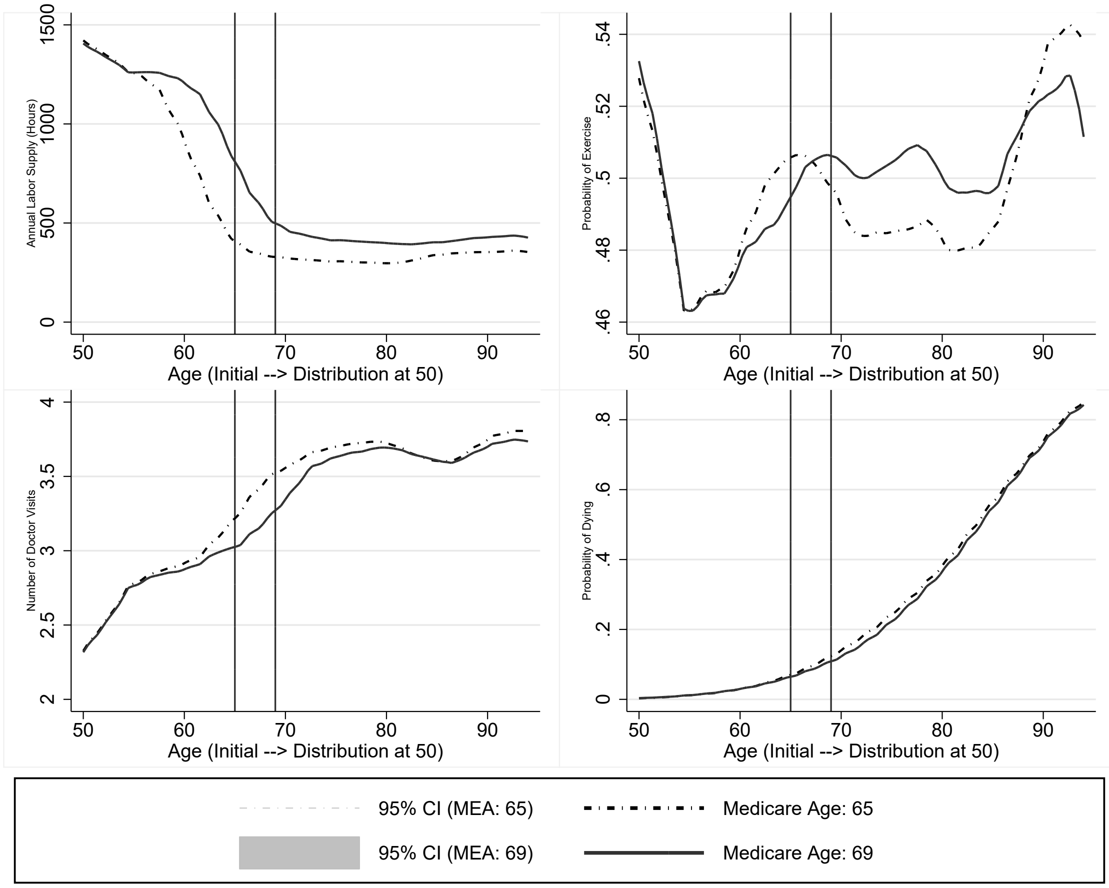


5.2 Counterfactual Policy Analysis

After confirming the model's performance in replicating the current policy (MEA at 65), I run a simulation under the assumption that the counterfactual policy is in place and analyze the effect of the counterfactual policy (MEA at 69). For this purpose, I employ a synthetic sample of people who are 50 years old and weight this sample so that it represents the pool of individuals who are 50 years old in the data. Calling this sample the weighted average sample, I simulate it for 44 years, so that individuals in the sample will be 94 years old if they stay alive in the simulation process. Note that, for age 85 and older there is not enough data in the HRS dataset and hence simulating the model for more than 85 does not provide reliable inference, especially since the estimation method I use to reveal the structural parameters is a data-intensive method and has a tendency to the sample average.

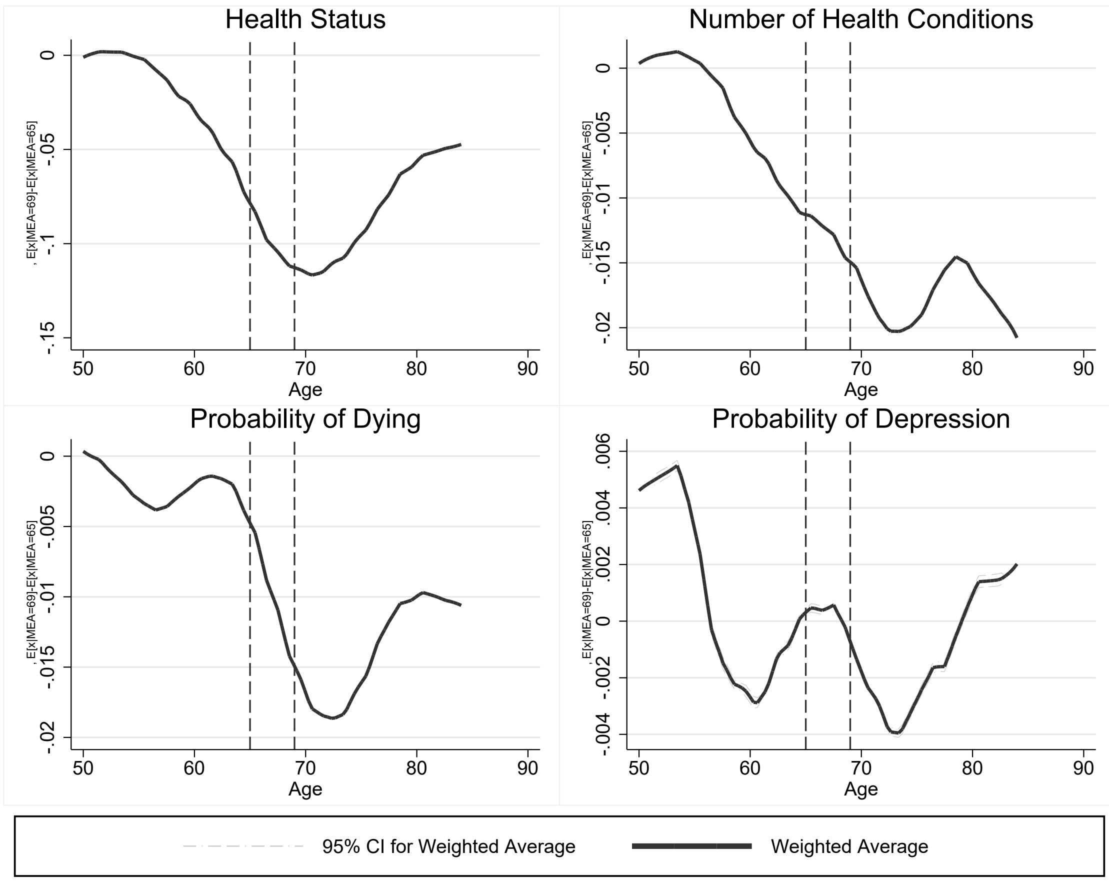
The effect of a change in the Medicare Eligibility Age (MEA) from 65 to 69 is analyzed through different measures. Table 5 shows the expected values at age 50 while considers the survival probability for every future period. Table 4 on the other hand, shows the mean outcome for different age groups conditional on being alive during the reported age group. Figure 3 shows the resulted lifetime path of the annual labor supply, probability of exercise, number of doctor visits and the probability of dying, for both current and counterfactual policies.

Figure 3: The trajectory of the selected variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).



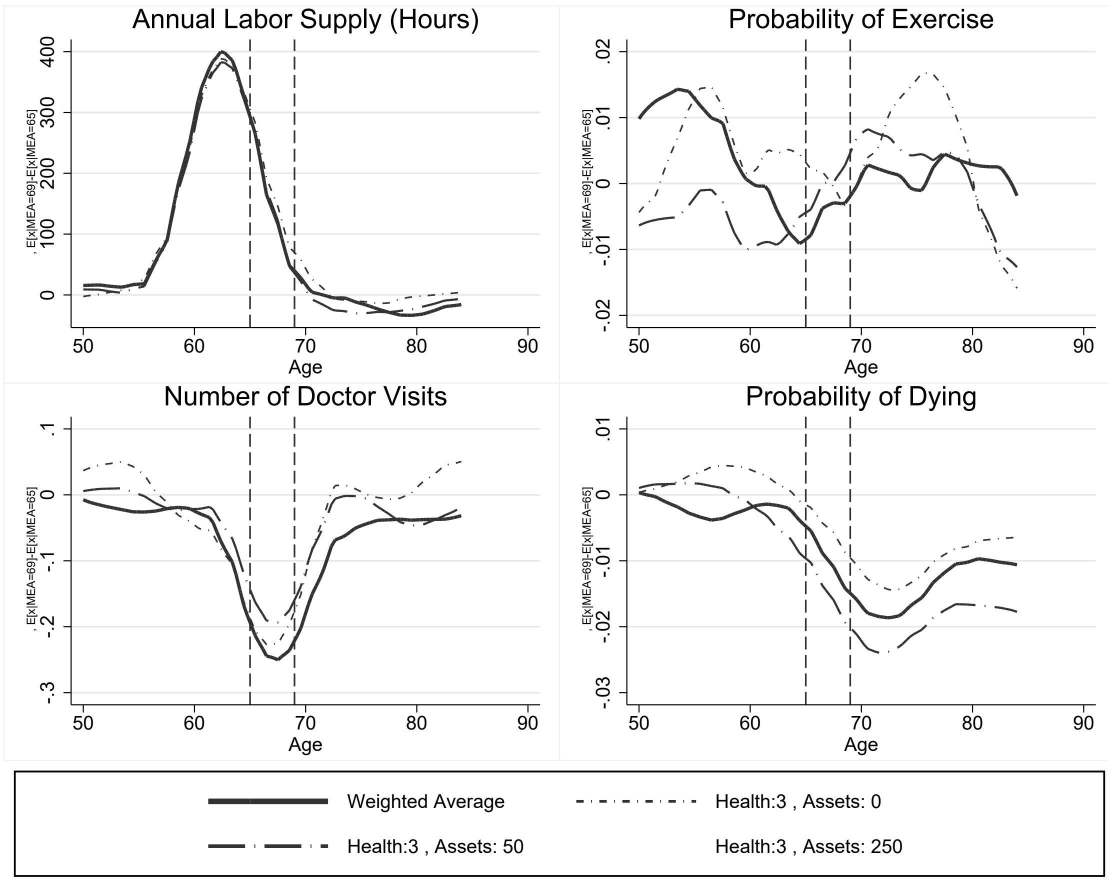
To show the effect of a change in policy, the difference between the current policy trajectory and counterfactual policy trajectory is shown in figure 4. The difference is augmented by the 95 % confidence interval and can be used to evaluate the effect of a change in policy at different ages. As a result of the hypothetical change in policy, annual labor supply is increasing. Individuals in the simulation pool on average tend to work more hours annually, beginning at age 54. The gap between the labor supply under the two policies is increasing by age and is maximum at age 64. The first row in the table 4 compares the expected average total hours of work at age 50 under the two policies and indicates that under the counterfactual policy men expected to work about 5000 hours more. This result will be followed by an increase in consumption of about $ 28,500 and leads to higher assets accumulation of about $ 10,100. Probability of exercise shows the same trend. Under the counterfactual policy, individuals tend to be more active and exercise more for a large portion of their life. The number of the doctor visits, on the other hand, is decreasing under the new policy for the years in which, Medicare is delayed. This is consistent with the ex-post moral hazard, as, given the existence of the ex-post moral hazard, we expect the reduction in the marginal price of the medical services leads to increase in medical services utilization. Row six of the table 4 shows the expected lifetime medical expenditure at age 50 under two regimes. Expected lifetime medical expenditure is lower ($ about 4,800) under the counterfactual policy. The overall effect of the change in behaviors determines the health measures and among them, the hazard of dying (probability of dying given individual is alive at a certain age) and average health status. Under the counterfactual policy, the hazard of dying is decreasing. Shown in row 9 of the table 4 men at age 50, are expected to live 0.5 year longer under the counterfactual policy. This also affects the total social security benefits amount that is expected to be collected by men at age 50. Rows 10 and 11 of he table 4 show the number of years receiving social security benefits and its total amount. While the number of years is lower under the new policy by 0.39 year, total social security benefits are increased by $ 12,688 . This is because of the higher labor supply and delay in the social security age claim under the counterfactual policy.

Figure 4: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on the selected variables



The whole analysis then is repeated for another simulation pool in which individuals are set to have low assets and poor health. Figure 5 shows the marginal effects for the two simulation pool of the weighted average sample and poor with poor health sample. By utilizing the new simulation pool of poor individuals with poor health, the marginal effect of the change in policy on the annual labor supply is shown to be unchanged. The hazard of dying is decreased for poor individuals with poor health, but the effect is smaller compared with the weighted average sample.
More detailed analysis of the results is provided in what follows.

Figure 5: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on selected variables



5.2.1 Effect on Health Investment

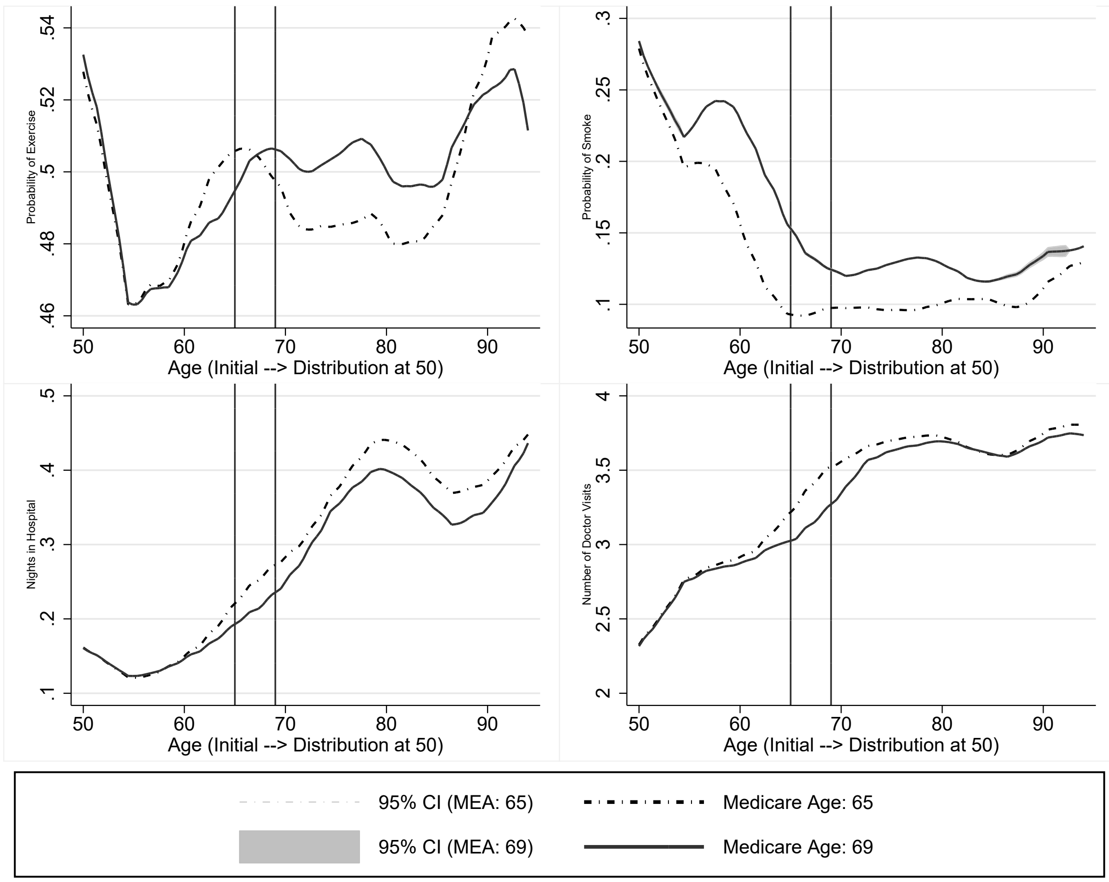
This study investigates the effect of the change in the policy on the four health investment variables. The probability of the exercise, probability of the smoking, number of night stays in hospital and number of doctor visits.

- **Probability of Exercise:** The trajectories of the probability of the exercise (vigorous activity) under the current and counterfactual policies are shown in the top-left graph of the Figure 6. For the individuals in the simulation pool, the counterfactual policy of MEA at 69 alters their optimal behavior toward pursuing an active lifestyle, and as a result, the probability of exercise is higher for them under the alternative policy comparing to the current policy. It can be seen that as a response to not having access to Medicare health insurance, the probability of exercise increases from .42 at age 50 to

.5 at the age of 80. While the counterfactual policy induces people to be more active as they are getting older, under the current policy probability of exercise does not change by age. The top-left panel of the Figure 7 shows that the difference between the probability of exercise under two regimes is statistically significant and is increasing by age. The gap is widening at a higher pace between age 65 to 72 and shows that the immediate impact of the change in policy on the optimal behavior is higher than the indirect effect that affects the optimal behavior through the forward-looking channel.

- **Probability of Smoking:** Decision to smoke is modeled directly in the model. The top-right panel in the figure 6 provides the probability of smoking under two regimes for different ages. For people older than 54, It is more likely to choose to smoke under the counterfactual policy than the current one. After Age 66, this relationship is stabilizes. However, the difference between the optimal behavior under the two policies is not significantly different from zero after age 66, as depicted in the top-left panel of the figure 7

Figure 6: The trajectory of the **Health Investment** variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).



- **Number of nights in Hospital:** Unlike the exercise and smoking, number of nights in the hospital is not modeled as the direct decision in this study. However, the dynamic structure of the model enables the agents to affect the probability of staying a particular number of nights in the hospital. As shown in column 5 of the Table 4 and the bottom-left graph of the Figure 6 number of nights in the hospital is affected by the change in policy. Starting at age 65, individuals in the simulation pool begin to utilize more of the hospital services under the current policy that reduces the cost of accessing the medical services, than the counterfactual policy. The bottom-left panel of the figure 7 provides a clearer view of the difference between the number of nights in hospital under two policies. While there is no statistically significant difference between two regimes until age 60, as the effect of a change in the policy, agents are utilizing fewer hospital services under the new regime.

- **Number of Doctor Visits:** The bottom-right panel of the Figure 6 shows the number of doctor visits for different ages under the counterfactual and current policies. Like the number of nights in hospital, the change in policy induces people to utilize lower amounts of the medical services. However, unlike the number of nights in the hospital, the gap which is initiated by the change in policy will be closed after the immediate effect of policy disappears. This temporary effect of the change in policy is shown more exactly in the bottom-right panel of the figure 7. While there is no difference between the number of doctor visits before age 60, starting at age 62, agents in the simulation pool are visiting doctor less frequently. The gap between the two regime is reaching its maximum at age 70 and will be closer completely at age 76. Column 4 of the Table 4 also shows the difference in the two policies for different age groups conditional on being alive. The expected number of doctor visits for a representative agent of the simulation pool at age 50 is shown in the column 4 of the Table 5. Representative agent is expected to visit doctor 106.02 times under the current policy in his life after age 50, while this number increases to the 105.15 under the counterfactual policy. This increase in the number of doctor visits can be explained by the increase in life expectancy under the counterfactual regime. Note that expected number of Doctor visits, uses the survival probabilities to provide expected values and hence is not equal to the number of doctor visits one can acquire by adding them through simulated life.

Figure 7: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different measures of **Health Investments**

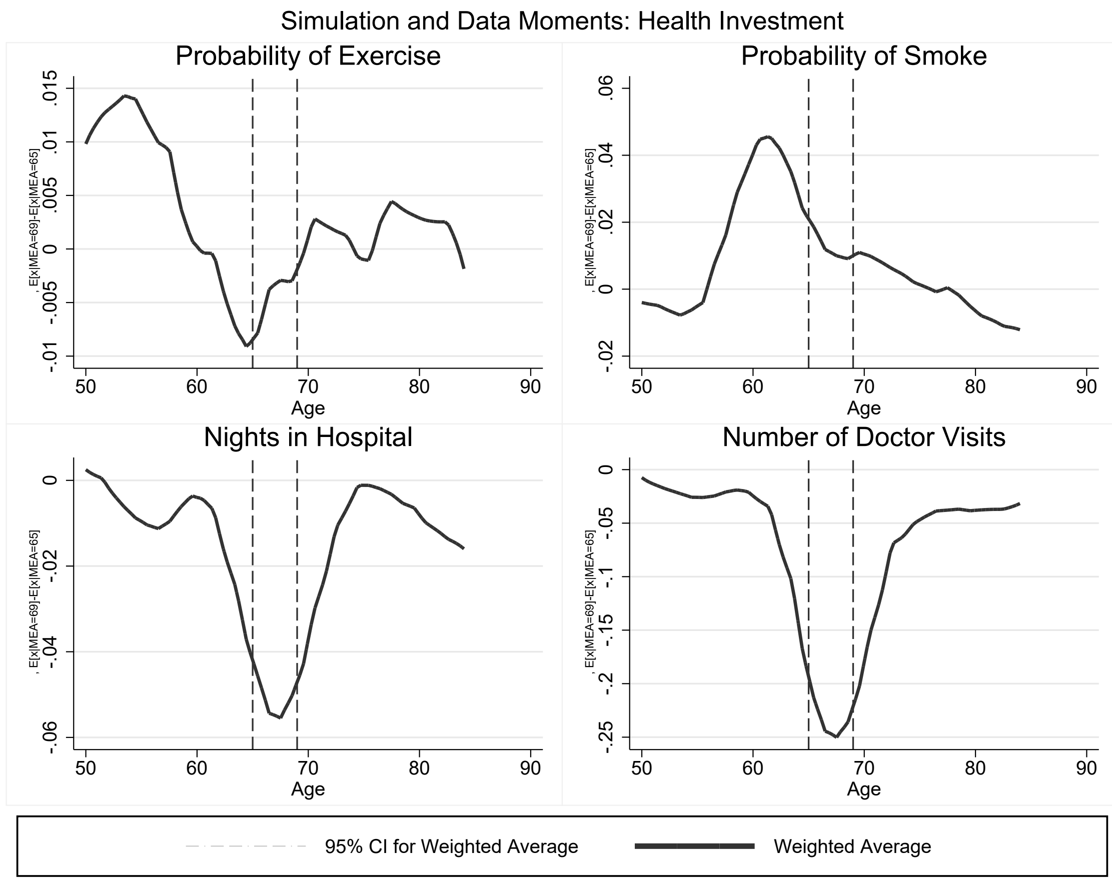


Table 4: Average effects of change in the policy

		50-54	55-59	60-64	65-69	70-74	75-79	80-84
Hours of Work	MEA : 65	1,369.307	1,237.003	653.902	325.312	306.780	315.946	291.935
	MEA : 69	1,358.392	1,256.101	1,120.242	550.000	431.546	414.087	370.643
		-(00.86)	(00.91)	(29.04)	(08.36)	(05.71)	(03.22)	(03.14)
Assets	MEA : 65	65.745	65.745	70.589	70.589	106.309	106.309	118.449
	MEA : 69	65.802	65.802	76.062	76.062	122.096	122.096	136.366
		(00.02)	(00.07)	(01.54)	(02.80)	(03.05)	(02.35)	(02.22)
Consumption	MEA : 65	45.746	45.011	45.591	45.148	44.697	45.167	42.479
	MEA : 69	45.985	45.239	45.239	45.883	45.197	45.845	42.734
		(00.62)	(00.37)	-(00.82)	(01.13)	(01.11)	(00.90)	(00.41)
Doctor Visits	MEA : 65	2.468	2.815	2.961	3.411	3.562	3.674	3.414
	MEA : 69	2.457	2.797	2.917	3.063	3.471	3.606	3.429
		-(00.20)	-(00.24)	-(00.83)	-(04.68)	-(01.65)	-(00.84)	(00.22)
Hospital	MEA : 65	0.148	0.121	0.167	0.258	0.295	0.407	0.392
	MEA : 69	0.149	0.124	0.157	0.216	0.274	0.376	0.350
		(00.07)	(00.23)	-(00.95)	-(02.35)	-(01.34)	-(01.09)	-(01.87)
Total Medical Expenditure	MEA : 65	6.159	7.350	9.141	10.839	11.217	14.243	12.575
	MEA : 69	6.181	7.409	8.931	10.624	10.758	13.779	11.952
		(00.10)	(00.20)	-(00.81)	-(00.46)	-(01.45)	-(00.78)	-(01.43)
Health	MEA : 65	1.745	1.797	2.012	2.337	2.808	3.394	4.023
	MEA : 69	1.735	1.797	1.991	2.247	2.676	3.268	3.911
		-(00.40)	-(00.02)	-(00.60)	-(01.61)	-(02.55)	-(01.70)	-(01.82)
CES-D	MEA : 65	0.958	0.726	0.689	0.686	0.577	0.588	0.518
	MEA : 69	0.943	0.733	0.661	0.637	0.541	0.541	0.477
		-(00.41)	(00.16)	-(01.06)	-(01.11)	-(01.34)	-(01.19)	-(01.25)
Life Expectancy	MEA : 65	1.00	1.00	0.99	0.99	0.98	0.98	0.93
	MEA : 69	1.00	1.00	0.99	1.00	0.98	0.98	0.94
		(00.00)	(00.00)	(00.84)	(00.87)	(00.54)	-(00.16)	(00.74)
Probability of Receiving SSB	MEA : 65	0.000	0.000	0.373	0.870	0.948	0.968	0.929
	MEA : 69	0.000	0.000	0.295	0.781	0.940	0.968	0.935
				-(07.70)	-(06.02)	-(01.20)	-(00.07)	(00.72)
SSB	MEA : 65	0.000	0.000	3.915	6.490	6.480	6.525	6.212
	MEA : 69	0.000	0.000	4.068	6.972	6.981	6.990	6.696
				(04.67)	(06.14)	(07.02)	(05.00)	(05.12)
Probability of Being Married	MEA : 65	0.598	0.602	0.576	0.542	0.462	0.397	0.307
	MEA : 69	0.597	0.599	0.574	0.516	0.442	0.374	0.297
		-(00.08)	-(00.15)	-(00.13)	-(01.20)	-(01.11)	-(00.95)	-(00.53)

t-statistics in parentheses
For each outcome, the first two rows are the outcome under the current (MEA: 65) or counterfactual policy (MEA: 69). Values in parentheses are t-statistics of the null hypothesis, $H_0 : E(X_{\text{MEA }:65}|Age) = E(X_{\text{MEA }:69}|Age)$. The effects under the current and the counterfactual policies are simulated by employing the pool of individuals at age 50 and weight for other characteristics to reflect the population of 50 years old in the sample. The pool size is 3000 and the simulation is continued for 40 years.

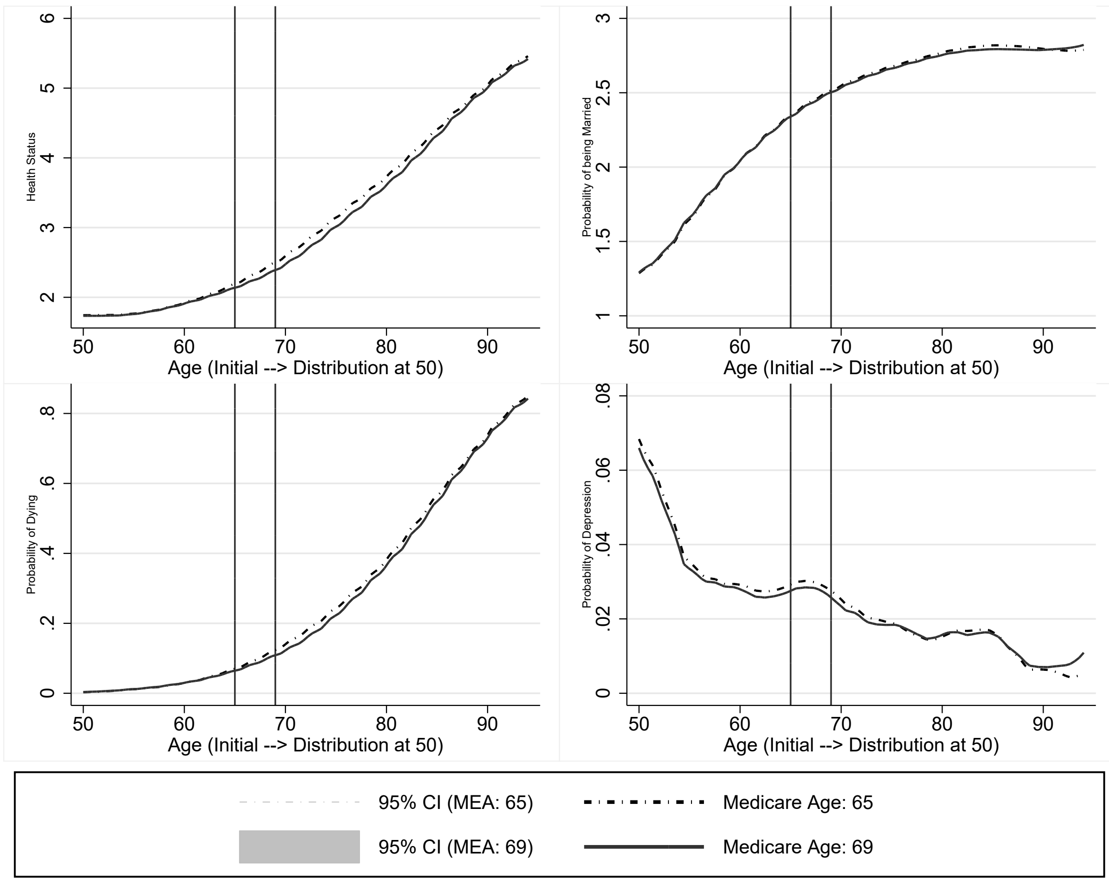
5.2.2 Effect on Health

For the health-related outcomes, I look at the physical and mental health status, the number of medical conditions and probability of dying.

- **Physical Health:** Physical health is measured using the index which is informed by the objective binary responses to the functional limitation questions and varies between 1 (excellent health) to 5 (poor health). The top-left panel in the Figure 8 shows the trajectories of the physical health under two policies. For the agents in the simulation pool, living under the counterfactual policy regime induces better physical health compared to the current policy. The top-right panel of the Figure 9 shows that this gap is increasing by age and is statistically different from zero in 95 % confidence level. Better physical health can be explained by the more active lifestyle in the counterfactual regime. Column 7 of the Table 4 shows the average physical health under two regimes for different age groups. While the health status is better under the counterfactual policy than the current one, for the individuals at age 50, the average expected lifetime health status is not large between the two regimes. This result is shown in column 7 of the Table 5.

- **Mental Health:** Mental health is measured using the CES-D index and also reported here by the probability of depression. The bottom-right graph in the figure 8 shows the probability of depression under the two policies and indicates that the probability of depression is lower under the counterfactual regime than the current regime. The gap between the depression probability under two regimes is initiated at age 54 and widens until age 84. Only attempts to closing the gap between age 54 and 84 is at the time of implementation of the new policy, between age 65 and 69. Then, again the gap persists. It might reflect the important tole that health insurance plays to provides financial stability and reduction of stress related diseases like depression. The gap is shown in the bottom-right panel of the figure 9. While the probability of depression is different under the two policies, the difference in the CES-D measure under the two regimes is not statistically different than zero. This is shown in the column 8 of the Table 4 and 5

Figure 8: The trajectory of the **Health** variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).



- **Number of Health Conditions:** Number of health conditions varies between 0 and 8 and provides another health status measure that covers more severe health problems which are not necessarily affect the functional limitation of the person. The number of health conditions under the two policies is shown in the top-right panel of the Figure 8. People are experiencing a slightly fewer number of health conditions under the counterfactual regime than the current one. There s no difference in the number of health conditions until age 62. Starting at age 64 the difference in the number of health conditions under the different policies is significantly different from zero and the gap is widening until age 72 and is stable after age 74, as shown in the top-right panel of the Figure 9.

- **Probability of Dying:** Probability of dying is an objective measure of health that integrates all the other health measures and also affects the aggregate level of the other important variables of the model through the life expectancy. The bottom-right graph in the Figure 8 shows the probability of dying for both current and counterfactual policies. Individuals in the simulation pool have the average probability of dying 0 at age 50. The probability increases under both regimes as individuals are getting older and reach around .5, for the current and alternative policies respectively. The bottom-right panel of the Figure 9 indicates that the gap between the effect of two policies appears at age 64 and grows until age 84. The probability of dying is .01 percentage point lower for the people at age 84 who are living under the counterfactual regime (MEA at 69) comparing to the current regime (MEA at 65). The probability of dying can be translated to the life expectancy. The life expectancy under the two different regimes is shown in column 9 of the Table 5 and indicates that people are on average living .46 years longer under the counterfactual regime than the current regime. This difference in the life expectancy explains the puzzling differences in the medical expenditure, number of doctor visits, and physical health status under the different policies.

Figure 9: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different measures of **Health**

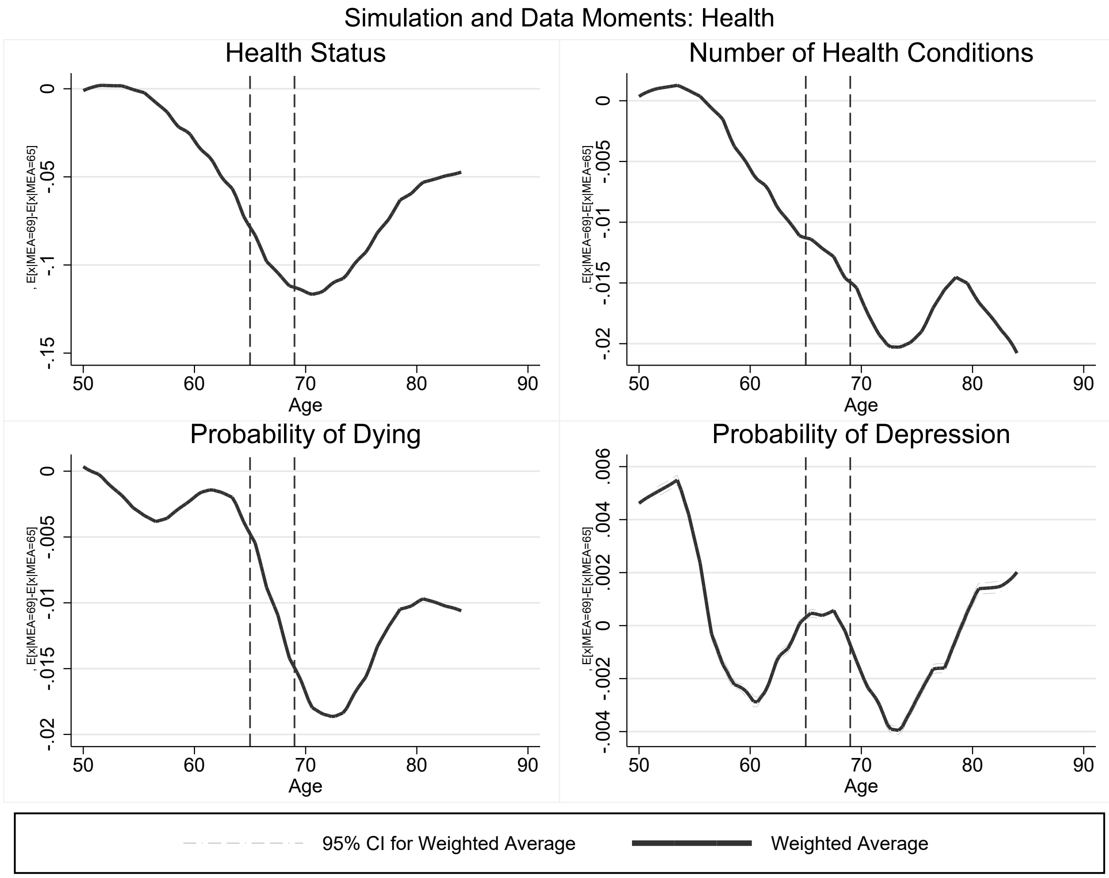


Table 5: Expected value of the outcome at age 50

Variable	MEA : 65	MEA : 69	Difference
Total Hours of Work	22,420.17	27,500.87	5,080.70
	(208.42)	(262.41)	(34.28)
Assets	97,807.99	107,930.57	10,122.58
	(116.44)	(125.91)	(9.46)
Total Consumption	1,506,692.17	1,535,095.64	28,403.47
	(363.46)	(379.49)	(4.97)
Total Doctor Visits	106.02	105.15	-0.87
	(296.34)	(296.66)	-(1.73)
Total Hospital	8.47	7.97	-0.51
	(150.52)	(147.96)	-(6.67)
Total Medical Expenditure	342,173.91	337,321.74	-4,852.17
	(223.78)	(230.17)	-(2.34)
Health	5.45	5.28	-0.17
	(508.35)	(496.66)	-(6.22)
CES-D	0.68	0.65	-0.03
	(195.39)	(190.)	-(4.36)
Life Expectancy	33.01	33.46	0.46
	(385.09)	(394.88)	(3.83)
Number of Years Receiving SSB	18.84	18.46	-0.39
	(224.55)	(224.31)	-(3.32)
Total SSB	137,342.61	150,031.30	12,688.70
	(216.62)	(225.69)	(13.5)
Number of years being married	16.17	16.02	-0.15
	(151.08)	(151.93)	-(1.61)

t-statistics in parentheses
Values in parantheses are t-statistics of the null hypothesis, $H_0 : E(X_{Data}|Age) = E(X_{Simulatied}|Age)$

5.2.3 Effect on Labor Supply

Labor supply is measured by the annual hours of work, the probability of working full time, the probability of working part-time and the probability of not working. The Annual hours of work for both policies are shown in the top-left graph of the Figure 10, where continuous line shows the current policy of MEA at age 65 and dashed line is the counterfactual policy of MEA at age 69. The divergence in the annual hours of work begins at age 54. People are working for more hours annually under the counterfactual policy regime and the gap is higher as individuals in the simulation pool are getting older. Most of the effect is caused by the extensive margin where people stay in the full time job instead of getting retired (stop working). Individuals tend to work more in the part-time job under the counterfactual policy regime before age 70, and after age 70 this relation is stabilized. The top-left graph in Figure 11 shows that the difference is significant in 95 % confidence for hours of work, the probability of working full time and the probability of not working, while the probability of working part-time job is not statistically different from zero. First column in Table 4 shows the average amount of hours of work for different 5-year age groups for those individuals in the simulation pool who are alive in each age group. First column of the Table 5 shows the expected lifetime hours of work for people at age 50. This result incorporates the probability of being alive in the future periods and as a result, reflects the health level as well. While under the current policy 50-year-old men are expected to work for 22,420 hours in their remaining portion of the lifetime, the counterfactual policy induces them to work for more 5,080 hours which can be interpreted as three years of full-time work. Most of the difference is due to the work during the transition period between current and counterfactual policy, meaning age 60 to 69.

Figure 10: The trajectory of the **Labor Supply** variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).

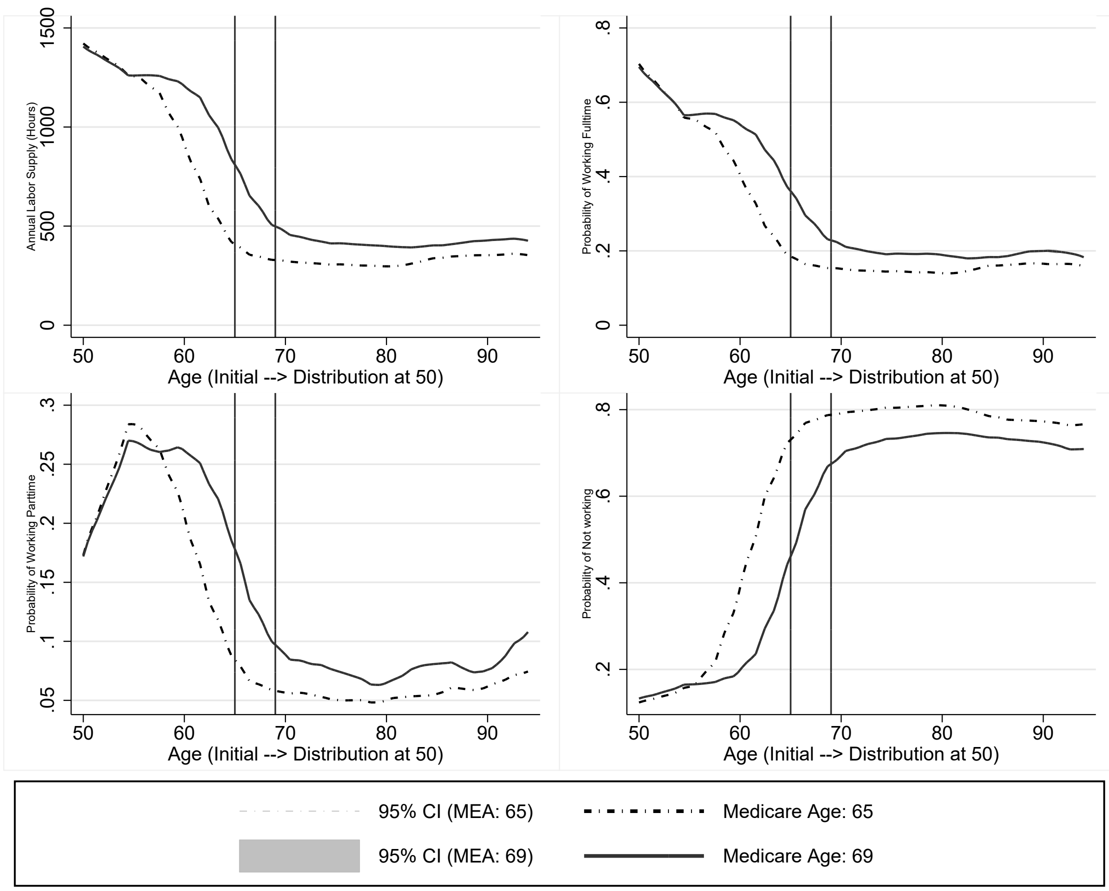
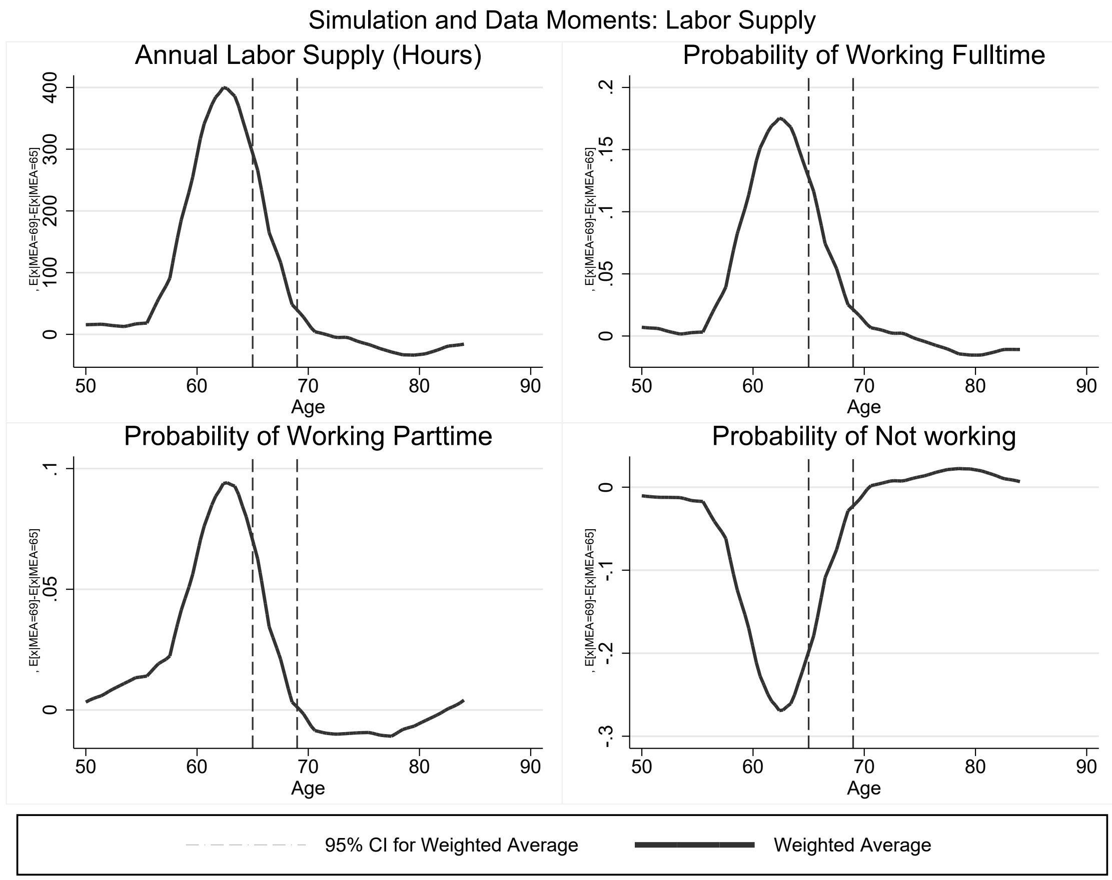


Figure 11: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on **Labor Supply**



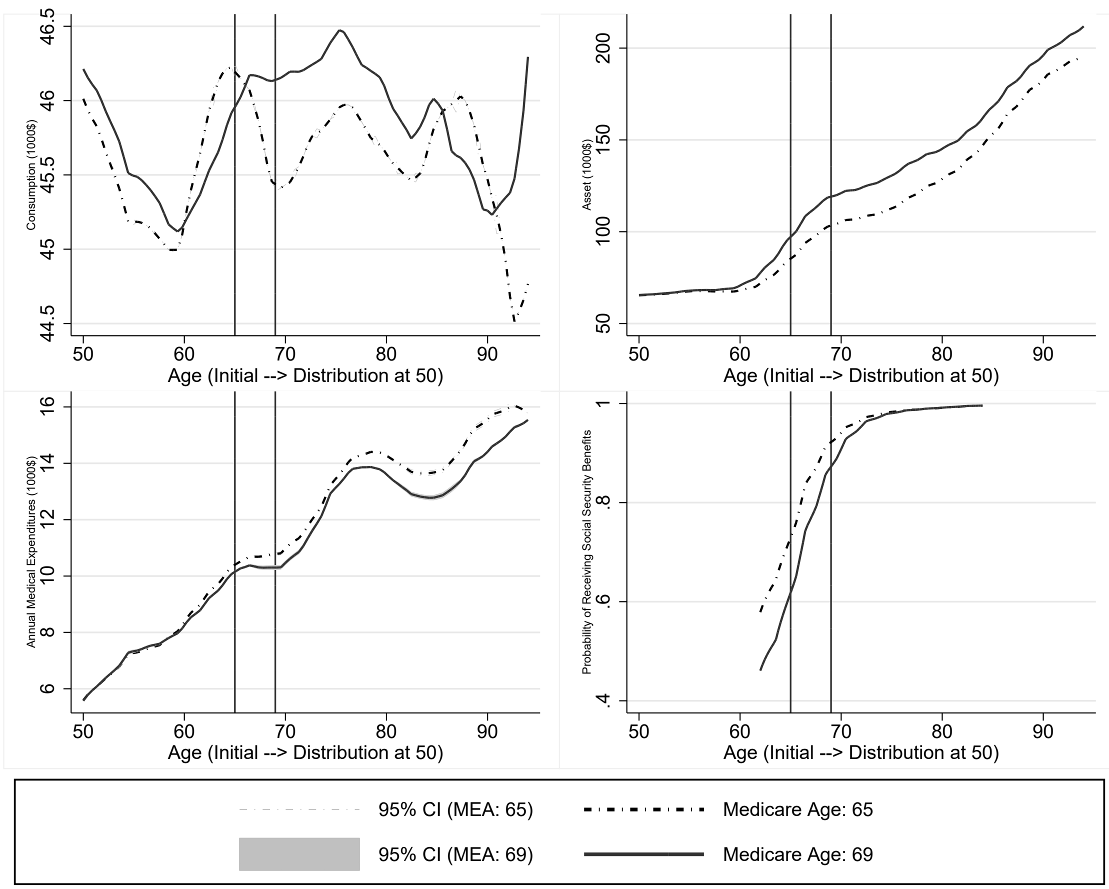
5.2.4 Effect on the Monetary Variables

Consumption, Assets (wealth), Annual medical expenditure and the social security benefits are the four monetary variables that we analyze.

- **Consumption:** Consumption level under the both regimes is shown in the top-left graph of the Figure 12. Simulation results indicate that consumption is relatively stabilized under both regimes and is not changing as the simulation pool is getting older, however under the counterfactual regime consumption is higher for a considerable portion of time, than the current policy regime. This accession in the consumption begins at age 66 and lasts until age 80. The top-left graph in the Figure 13 shows that the gap is at it's maximum at age 62 and stays almost constant until age 84. It should be noted that the difference is not statistically significant before age 70 as is shown in column 3 of the Table 4. The expected lifetime difference in consumption under the two regimes for the average individual in the simulation pool at age 50 is $ 28,403 which means that under the counterfactual regime, individuals are expected to consume about one year worth of consumption more, comparing to the current policy regime.

- **Assets:** Next monetary variable is the level of assets. The simulated trajectory of the assets under two regimes is shown in the top-right part of the Figure 12. Individuals in the simulation pool are accumulating lower levels of assets under the current regime. The top-right graph in the Figure 13 shows that the gap begins at age 60 and is expanding until age 69. Column 2 of the Table 5 shows that average level of assets for the representative person in the simulation pool is 97,807 $ under the current regime and 107,930 $ under the alternative regime.

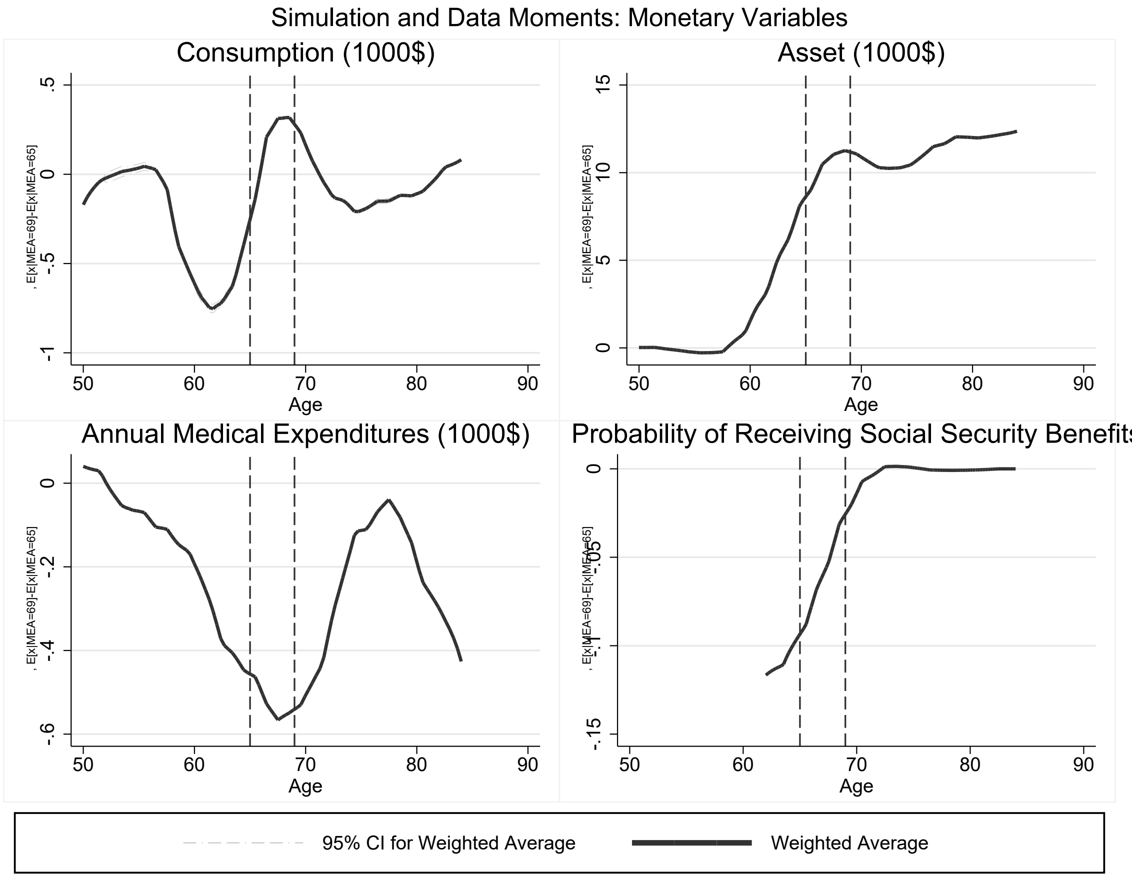
Figure 12: The trajectory of the **Monetary** variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).



- **Medical Expenditure:** Annual medical expenditure is increasing as individuals are getting older, and it shares the same trajectory under both regimes until age 60. As shown in the bottom-left part of the Figure 12, the gap in medical expenditure begins to grow at age 60. Medical expenditure is lower for individuals who live under the counterfactual regime in which they do not have access to the Medicare health insurance before age 69, comparing to the same individuals who live under the current regime. The bottom-left graph in the Figure 13, provides the difference in the medical expenditure under the two regimes and shows that the gap which initially began to appear at age 60, is growing through the lifetime and is statistically significant after age 68. Column 6 of the Table 5 shows that the expected lifetime medical expenditure for individuals at age 50 is 4,852 less, under the counterfactual regime compared to the current regime.

- **Social Security Benefits:** Social security benefit is the last monetary variable. The bottom-right graph of the Figure 12 shows that the amount of social security benefit that the average individual in the simulation pool receives under the alternative policy is higher compared to the current regime. Starting at age 62, which is the early retirement age, two policies induce different levels of social security benefits. The reason can be seen in both working histories and postponing the social security benefits claim. It is discussed under the labor supply title that counterfactual policy motivates people to work more. Thus, individuals have more average income which will affect the social security benefit amount they receive. Provided in column 10 of the Table 4 that shows the probability of receiving social security benefit under the two regimes and in the different age groups, the probability of receiving social security benefits is lower under the counterfactual policy. As a result, under the alternative policy, people have higher average income in their recent work history at each age and also they delay their social security claim, both of which increases the expected social security benefit that they receive

Figure 13: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different **Monetary Variables**

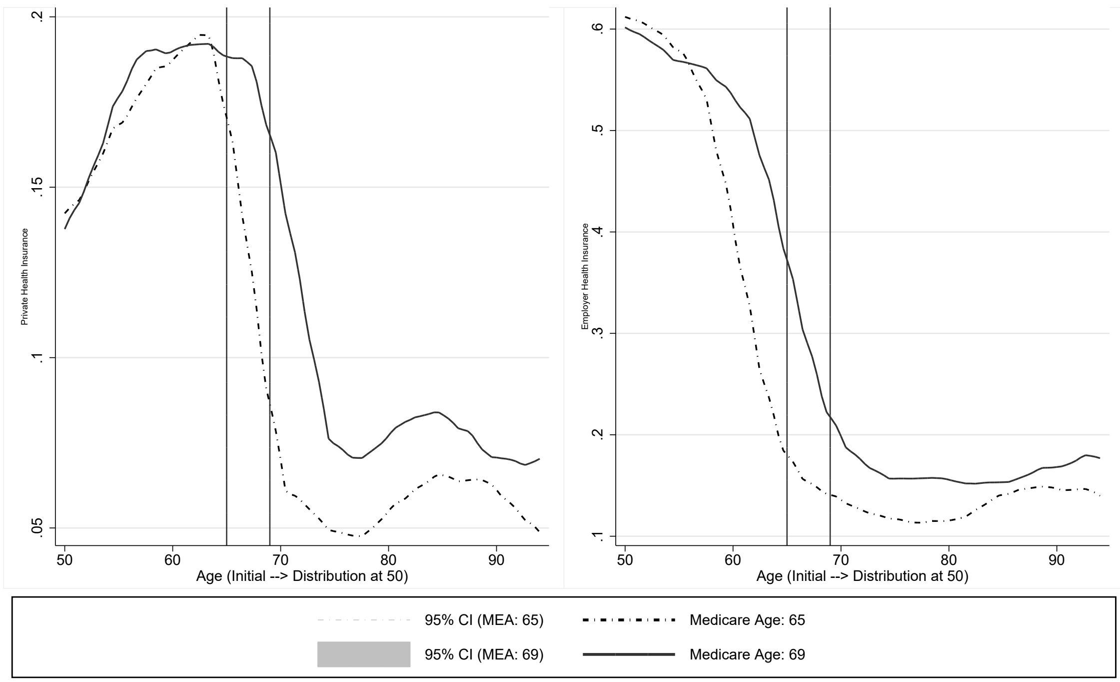


5.2.5 Effect on Health Insurance

The decisions to purchase private health insurance and employer health insurance are modeled directly as utility function variables. This study investigates the effect of the change in Medicare Eligibility Age from age 65 to age 69. Purchasing private and employer health insurance are the substitutes to the Medicare health insurance and should act as a buffer to minimize the harmful effects of this change in policy for the individuals. This section discusses the optimal response of the individuals to the change in policy through these two instruments.

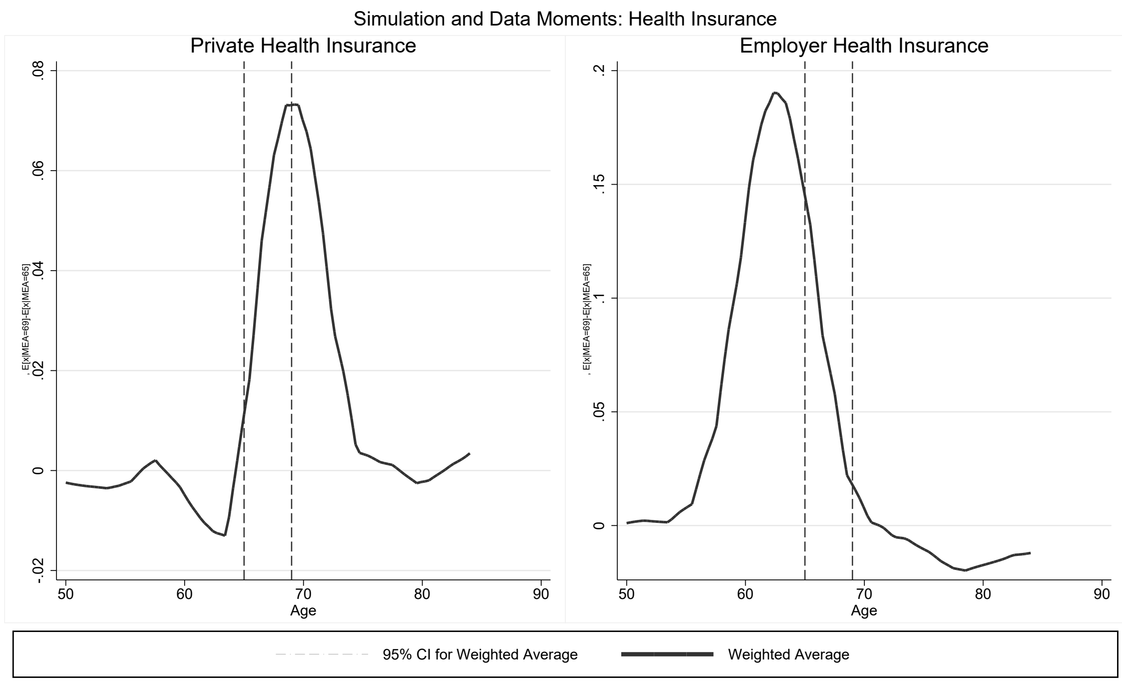
- **Purchasing Private Health Insurance:** The Left panel of the Figure 14 shows the probability of purchasing private health insurance under each policy. Individuals in the simulation pool are more likely to purchase private health insurance under the counterfactual policy after age 65, comparing to the current policy. The probability is statistically the same for both policies between age 50 to 65. The difference between the probabilities and the statistical significance is shown in the left panel of the figure 15.

Figure 14: The trajectory of the **Health Insurance** variables under current policy (MEA at 65) and counterfactual policy (MEA at 69).



- **Purchasing Employer Health Insurance:** Probability of purchasing the employer health insurance is linked to the decision to work and is decreasing under both policies as simulation pool is getting older. As shown in the right panel of the Figure 14, the reduction in the probability of purchasing the employer health insurance by age is occurring consistently at age of the implementation of Medicare policy. The right panel in the Figure 15 shows the difference in the probability under the two policies. The gap in the likelihood of purchasing the employer health insurance is appearing at age 64 and is widening steadily until age 69. At age 65 that is the Medicare Eligibility Age under the current regime, the gap between the two policies is about six percentage points. The gap reaches eight percentage point at around age 69.

Figure 15: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different **Health Insurance Types**



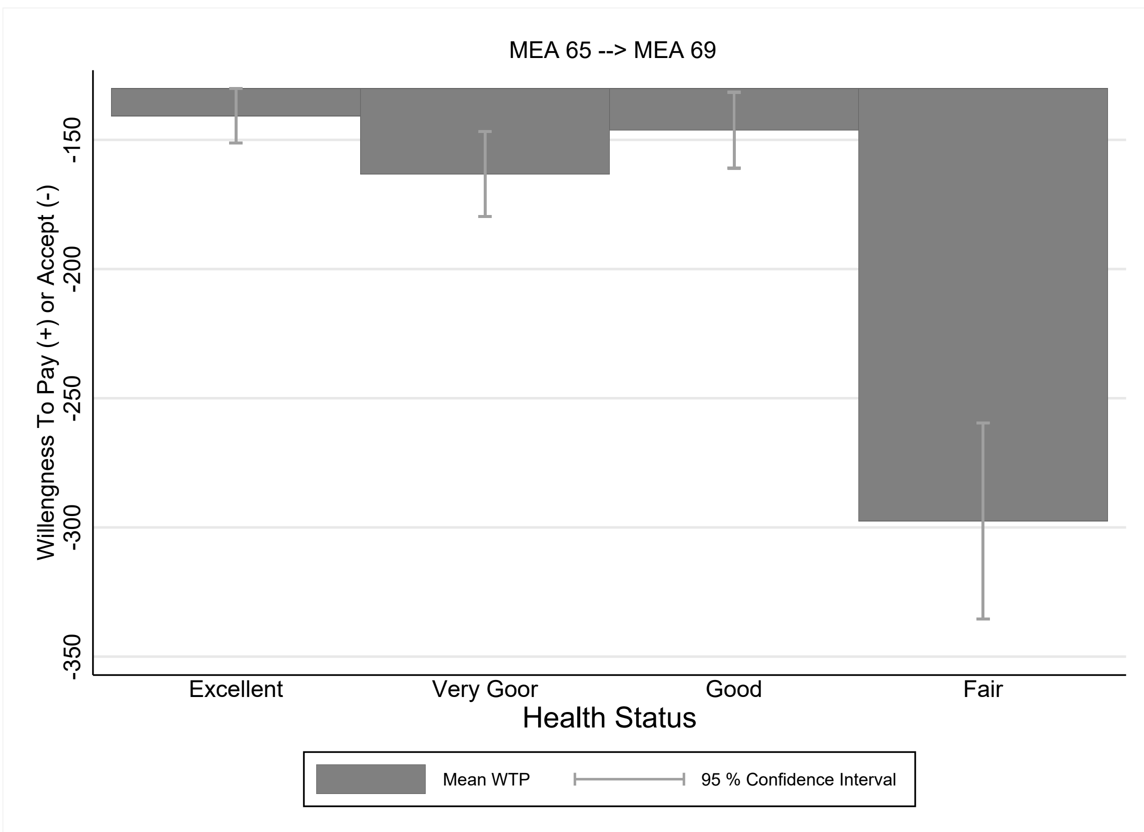
5.3 Welfare Effect

The change in the Medicare Eligibility Age induces individuals to work for more hours to compensate for the reduced access to Medicare health insurance. They also pursue a healthier lifestyle by being more active (exercise). They experience better mental health and consume more as health and consumption are complementary. However, losing the free health insurance and working for more years and more hours are not beneficial. To evaluate the welfare effect of the change in Medicare Eligibility age from age 65 to 69, I calculate the Willingness to Pay(WTP) of the change in policy.

I do this by calculating the willingness to pay for each individual between age 50 to 54 in the sample. This is done by assigning the ex-ante value function for each individual under the two different policies and searching for the change in the wealth that is required to make the Present Discounted Utility of the counterfactual policy equal to the one for the current policy. This change in wealth is interpreted as the willingness to pay and indicates the maximum amount that an individual pays to have the new policy. A negative willingness to pay means that individual prefers the current policy (MEA at 65) and to have him indifferent between current and counterfactual policy, policymaker should compensate him

(Willingness To Accept). Figure 16 shows the willingness to pay for the four different health groups: Excellent, very good, good ,and fair. Willingness to pay is negative for all the four groups and this means that individuals need to be compensated to become indifferent between current and the counterfactual policy. Otherwise, they prefer current policy. The Willingness to pay is more negative as the health status is worsening and indicates people with poor health dislike more the new policy. It can be explained by the value of the Medicare health insurance for the people with poor health who are more likely to utilize the medical care services. For people with excellent and very good health status, WTP is $ -140,733 and $ 1466,256 , respectively . WTP decreases to $ -163,235 for the people with good health status. For people with fair and poor health WTP is on average $ -297,500 .

Figure 16: Calculated Willingness To Pay (Accept) for the proposed policy of delaying Medicare Eligibility Age by four years.



6 Conclusion

Raising the Medicare Eligibility Age has been discussed as a policy to help balance the federal budget. The main concern in conducting such change in policy is the health and welfare of the near retirement population. This study utilizes the out of sample prediction power of the structural estimation models to address the effect of a raise in the Medicare Eligibility Age. Blau and Gilleskie (2008), Scholz and Seshadri (2013), French et al. (2016) and Pelgrin and St-Amour (2016) among others, have analyzed the effect of change in health insurance on labor supply and health, by developing structural models. While, these models of health and retirement can be utilized to address the effect of change in the Medicare Eligibility Age, the endogenous behavioral response of the individuals to such a change in policy should be incorporated in the models. This study models the labor supply and health investment behaviors of the men with 12 years of education by developing a model in which health behaviors such as smoking and exercise and different health insurance types are chosen by individuals and directly affect individual's utility. Incorporating three types of the unobserved fixed heterogeneity, the model under studied is flexible and sensitive to the desired changes in the health insurance policy. This study employs the Conditional Choice Probability (CCP) estimator suggested by Arcidiacono and Miller (2011), and hence is able to incorporate a large state space that allows for more detailed transition functions and timing.

I find that raising the Medicare Eligibility Age from 65 to 69 leads to a higher level of labor supply and higher life expectancy. However, the effect of the change in policy on the welfare is negative and people on average ask for about $ 151,000 to accept the change with those in poor health at age 50 needing nearly $ 297,500 to be compensated. I also find some evidence of cost transfer from Medicare program to the Social Security program. Men's lifetime social security benefits increase by $ 12,688 that is due to the increase in labor supply, postponing the social security benefits claim and, increase in the life expectancy.

References

Arcidiacono, P. and R. Miller (2011). Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity. *Econometrica*, 79(6):1823–1867.

Bhargava, A. (1991). Identification and panel data models with endogenous regressors. *The Review of Economic Studies*, 58(1):129–140.

Blau, D. and D. Gilleskie (2008). The role of retiree health insurance in the employment behavior of older men. *International Economic Review*, 49(2):475–514.

Brunner, K. and Meltzer, A. (1983). Econometric policy evaluation: A critique. *Theory, Policy, Institutions: Papers from the Carnegie-Rochester Conferences on Public Policy*, 1:257.

Card, D., C. Dobkin, and N. Maestas (2008). The impact of nearly universal insurance coverage on health care utilization: evidence from medicare. *American Economic Review*, 98(5):2242–58.

Congressional Budget Office (2016). Options for reducing the deficit: 2017 to 2026. *CONGRESSIONAL BUDGET OFFICE Report.*

Dempster, A., N. Laird, and D. Rubin (1977). Maximum likelihood from incomplete data via the em algorithm. *Journal of the royal statistical society*, B(methodological):1–38.

Dhaval, D. and R. Kaestner (2009). Health insurance and ex ante moral hazard: evidence from medicare. *International journal of health care finance and economics*, 9(4):367.

Finkelstein, A. (2007). The aggregate effects of health insurance: evidence from the introduction of medicare. *Journal of Economics*, 122(1):1–37.

Finkelstein, A. and R. McKnight (2008). What did medicare do? the initial impact of medicare on mortality and out of pocket medical spending. *Journal of public economics*, 92(7):1644–1668.

Firouzi Naeim, P. (2017). Measuring health in hrs dataset using bayesian approach. *Working Paper.*

Fontaine, P., T. Gullo, H. Harvey, T. Bradley, K. Crawley, J. Hearne, J. Holland, S. Jennings, and S. Papenfuss (2011). Reducing the deficit: Spending and revenue options. *CONGRESSIONAL BUDGET OFFICE (US CONGRESS) WASHINGTON DC.*

French, E. (2005). The effects of health, wealth, and wages on labour supply and retirement behaviour. *The Review of Economic Studies*, 72(2):395–427.

French, E. and J. Jones (2011). The effects of health insurance and self-insurance on retirement behavior. *Econometrica*, 79(3):693–732.

French, E., H. von Gaudecker, and J. Jones (2016). The effect of the affordable care act on the labor supply, savings, and social security of older americans. *SSRN*.

Hotz, V. and R. Miller (1993). Conditional choice probabilities and the estimation of dynamic models. *The Review of Economic Studies*, 60(3):497–529.

Hotz, V., R. Miller, S. Sanders, and J. Smith (1994). A simulation estimator for dynamic models of discrete choice. *The Review of Economic Studies*, 61(2):265–289.

Keane, M. and K. Wolpin (1994). The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte carlo evidence. *The review of economics and statistics*, Pp. 648–672.

Kelly, I. and S. Markowitz (2009). Incentives in obesity and health insurance. *Inquiry*, Pp. 418–432.

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

Lichtenberg, F. (2002). The effects of medicare on health care utilization and outcomes. *Forum for Health Economics and Policy*, 5(1).

McFadden, D. (1981). Econometric models of probabilistic choice. *Structural analysis of discrete data with econometric applications*, 198272.

Mroz, T. and T. Savage (1998). The long-term effects of youth unemployment. *Journal of Human Resources*, 41(2):259–293.

Pelgrin, F. and P. St-Amour (2016). Life cycle responses to health insurance status. *Journal of health economics*, 49:76–96.

Ruhm, C. (2000). Are recessions good for your health? *The Quarterly journal of economics*, 115(2):617–650.

Rust, J. (1987). Optimal replacement of gmc bus engines: An empirical model of harold zurcher. *Econometrica: Journal of the Econometric Society*, Pp. 999–1033.

Rust, J. (1994). Structural estimation of markov decision processes. *Handbook of econometrics*, 4:3081–3143.

Scholz, J. and A. Seshadri (2013). Health insurance and retirement decisions. *Ann Arbor MI: University of Michigan Retirement Research Center (MRRC) Working Paper*, 292:5.

Stanciole, A. (2008). Health insurance and lifestyle choices: Identifying ex ante moral hazard in the us market. *The Geneva Papers on Risk and Insurance-Issues and Practice*, 33(4):627–644.

The Boards of Trustees (2017). Annual report of the boards of trustees of the federal hospital insurance and federal supplementary medical insurance. *Annual Report.*

Van der Klaauw, W. and K. Wolpin (2008). Social security and the retirement and savings behavior of low-income households. *Journal of Econometrics*, 145(1-2):21–42.

Transitions and Production Functions

Involuntary Layoff

Probability of receiving layoff shock is logistic function of previous period marriage status (m_{t-1}), tenure (x^1_{t-1}), income (y_{t-1}) and current period economic environment (σ^e_t). Defining vector $\overrightarrow{X}_{\sigma_l}$:

$$\overrightarrow{X}_l = [1, m_{t-1}, x^1_{t-1}, y_{t-1}, \sigma^e_t]$$

Then define the probability of being in a specific mental health state as a logistic function of the linear transformation of index:

$$Pr(\sigma^l = a|\overrightarrow{X}_{\sigma^l}) = \frac{\exp(\overrightarrow{X}_{\sigma^l} \times \overrightarrow{\beta'}_{\sigma^l,a})}{1 + \sum_{b\neq 0}(\exp(\overrightarrow{X}_{\sigma^l} \times \overrightarrow{\beta'}_{\sigma^l,b}))} \quad \forall \quad a \in \{1\} \tag{11}$$

Where, $\overrightarrow{\beta}_{\sigma^l}$ is a 1×5 vector of coefficients and, 0 as the base category.

Tenure

Tenure can be an integer between zero and $\bar{X}_1$. I assume that tenure increases only if the agent keeps the previous job. It will be zero if the person decides not to work and will be two if the person starts working a new job. After layoff shock is revealed, If a layoff occurs, tenure resets to zero, and if there is no layoff, $x_{1,t}$ remains same at the beginning of period t and after the shocks. Finally, decision toward employment updates tenures value for the next period.

$$x_{1,t} = \begin{cases} 0, & \text{if } \sigma_{t-1} = 1 \\ x_{1,t-1}, & \text{otherwise} \end{cases} \tag{12}$$

Mental Health

Mental health is an integer between 0 and 8, with zero means excellent mental health and 8 means depressed. I assume logistic probability for the mental health shock. The probability of being in one of the states of the mental health is function of previous period marriage status, mental health, age, physical health, leisure, income, number of nights in hospital, number of doctor visits, number of health conditions, assets, consumption, gender, education and updated layoff shock . I define the index vector $\overrightarrow{X}_{h^m}$:

$$\overrightarrow{X}_{h^m} = [1, m_{t-1}, h^m_{t-1}, A_{t-1}, h^p_{t-1}, L_{t-1}, y_{t-1}, k^1_{t-1}, k^2_{t-1}, h^c_{t-1}, a_{t-1}, c_{t-1}, g, E, \sigma^e_t]$$

Then define the probability of being in a specific mental health state as a logistic function of the linear transformation of index:

$$Pr(h_t^m = a|\overrightarrow{X}_{h^m}) = \frac{\exp(\overrightarrow{X}_{h^m} \times \overrightarrow{\beta'}_{h^m,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{h^m} \times \overrightarrow{\beta'}_{h^m,b}))} \quad \forall \quad a \in \{1, ..., 8\} \tag{13}$$

Where, $\overrightarrow{\beta}_{h^m}$ is a 1×15 vector of coefficients and, 0 as the base category.

Experience and Tenure

If an individual decides to work (part-time or full-time), then both experience and tenure will be added by 2. If an individual chooses not to work, tenure resets to zero, while Experience will be unchanged.

Age and Medicare

Age increases by 2. Then Medicare will be one if updated age is greater than or equal the Medicare eligibility age, and zero otherwise.

Number of Nights in Hospital

After revealing insurance type, healthy behaviors, and income, I update the number of nights in the hospital, using a shock with logistic distribution and conditional on a subspace of state spate. Since data is concentrated at zero nights in the hospital, I set all the observations with one and two nights in the hospital to be one. Also, because there are not enough observations with more than three nights in the hospital, I set all observations with more than three nights, equal to 3.

$$\overrightarrow{X}_{k^1} = [1, m_{t-1}, h_{t-1}^p, A_t, y_t, e_t, s_t, i_t^{emp}, i_t^{prv}, i_t^m]$$

The probability of spending a specific number of nights in the hospital is a logistic function of the linear transformation of index:

$$Pr(k_t^1 = a|\overrightarrow{X}_{k^1}) = \frac{\exp(\overrightarrow{X}_{k^1} \times \overrightarrow{\beta'}_{k^1,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{k^1} \times \overrightarrow{\beta'}_{k^1,b}))} \quad \forall \quad a \in \{1, 3\} \tag{14}$$

Where, $\overrightarrow{\beta}_{k^1}$ is a 1×10 vector of coefficients and, 0 as the base category.

Number of Doctor Visits

Updating procedure of the number of doctor visits follows the same procedure as the number of nights in the hospital. Again, because of the data concentration around zero, I assume three bins for the number of doctor visits, 0, 1 and five which implies all observations with the number of doctor visits between 1 and four are set to 1 and all with the number of doctor visits greater and equal 5, set to 5.

$$\overrightarrow{X}_{k^2} = [1, m_{t-1}, h^p_{t-1}, A_t, y_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$

I model the structure of shock as conditional logistic distribution. Thus, the probability of the number of doctor visits is a logistic function of the linear transformation of index:

$$Pr(k^2_t = a|\overrightarrow{X}_{k^2}) = \frac{\exp(\overrightarrow{X}_{k^2} \times \overrightarrow{\beta'}_{k^2,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{k^2} \times \overrightarrow{\beta'}_{k^2,b}))} \quad \forall \quad a \in \{1, 5\} \tag{15}$$

Where, $\overrightarrow{\beta}_{k^2}$ is a 1×10 vector of coefficients and, 0 as the base category.

Physical Health

Knowing the reactionary health investments, healthy behaviors, age, leisure and income, individual receives the physical health shock. Using the same assumption for the structure of the shock, I can update physical health status. Physical health status is an integer that takes one of the values 1 through 6, with one as excellent health status and six as dead. The index that physical health status is based upon is defined as:

$$\overrightarrow{X}_{h^p} = [1, m_{t-1}, h^p_{t-1}, h^c_{t-1}, a_{t-1}, g, E, A_t, L_t, y_t, k^1_t, k^2_t, e_t, s_t]$$

Then the probability of being in each one of the six different values is:

$$Pr(h^p_t = a|\overrightarrow{X}_{h^p_t}) = \frac{\exp(\overrightarrow{X}_{h^p_t} \times \overrightarrow{\beta'}_{h^p_t,a})}{1 + \sum_{b \neq 1}(\exp(\overrightarrow{X}_{h^p_t} \times \overrightarrow{\beta'}_{h^p_t,b}))} \quad \forall \quad a \in \{2, 3, 4, 5, 6\} \tag{16}$$

Where, $\overrightarrow{\beta}_{h^p}$ is a 1×14 vector of coefficients and, 1 as the base category.

Number of Health Conditions

Individual also receives another health shock regarding the number of health conditions. I use three bins for the number of health conditions which are 0, 1 and 3. As the number of nights in the hospital and the number of doctor visits, the reason that I am using three bins for the number of health conditions is the data concentration around 0. For each individual, if the number of health conditions is 1 or 2, I set it to 1, and if it is greater than or equal 3, I set it to 3. Then the number of health condition shock is conditional on a vector of variables containing a subset of updated and un-updated state space:

$$\overrightarrow{X}_{h^c} = [1, m_{t-1}, h^c_{t-1}, h^p_t, A_t, y_t, k^1_t, k^2_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$

Then the probability of being in each one of the six different values is:

$$Pr(h^c_t = a|\overrightarrow{X}_{h^c_t}) = \frac{\exp(\overrightarrow{X}_{h^c_t} \times \overrightarrow{\beta'}_{h^c_t,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{h^c_t} \times \overrightarrow{\beta'}_{h^c_t,b}))} \qquad (17)$$
$$\forall \quad a \in \{1, 3\}$$

Where, $\overrightarrow{\beta}_{h^c}$ is a 1×13 vector of coefficients and, 0 as the base category.

Marriage Status

Marriage is not explicitly modeled as a decision in my model. Instead, I update marriage status by endogenous shock. An individual can affect the probability of being married by the decisions about labor supply and hence income, healthy behaviors such as smoking and exercise and health insurance type which all are important in explaining one's attractiveness and also need for marriage. Marriage status is the dummy that is one if an individual is married and zero otherwise. Endogenous marriage shock follows the logistic distribution with a conditional mean that is a function of the linear transformation of a subspace of state space. I define the index function of the logistic function as:

$$\overrightarrow{X}_m = [1, m_{t-1}, h^p_t, h^c_t, A_t, y_t, k^1_t, k^2_t, e_t, s_t, i^{emp}_t, i^{prv}_t, i^m_t]$$

I model the structure of marriage shock as conditional logistic distribution. Thus, the probability of the number of doctor visits is a logistic function of the linear transformation of index:

$$Pr(m_t = a|\overrightarrow{X}_m) = \frac{\exp(\overrightarrow{X}_m \times \overrightarrow{\beta'}_{m,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_m \times \overrightarrow{\beta'}_{m,b}))} \qquad (18)$$
$$\forall \quad a \in \{1\}$$

Where, $\overrightarrow{\beta}_m$ is a 1×13 vector of coefficients and, 0 as the base category.

Social Security Benefit

Like marriage status, I do not explicitly model the decision for claiming the social security benefits. I define a dummy variable for claiming social security benefit that takes value one if individual claims the social security benefit and zero otherwise. Receiving social security benefits is absorbing state, which means, if an individual has not started to receive the social security benefits, she can start claiming it in the current period by a probability that is defined by the shock structure. But if an individual already receives social security benefits, she continues receiving it. Receiving the social security variable gets an update by a logistic type shock with mean that is a function of the subset of state space. As a result, an individual can affect the probability of receiving social security benefits by affecting the state space through decisions which are explicitly modeled. Since an individual cannot claim social security benefits before reaching the early retirement age, I constrained the updating procedure to be started after early retirement age and set receiving social security to be zero for individuals younger than it. The vector of the relevant subset of the state space is:

$$\overrightarrow{X}_{SSB} = [1, a_{t-1}, g, E, m_t, h_t^m, h_t^p, h_t^c, A_t, L_t, y_t, k_t^1, k_t^2, e_t, s_t, \sigma_t^e, c_t] \quad for A_t \geq \underline{f}$$

Where $\underline{f}$ is the early retirement age. Social security benefit updating probability has a conditional logistic distribution:

$$Pr(SSB_t = a | \overrightarrow{X}_{SSB}) = \frac{\exp(\overrightarrow{X}_{SSB} \times \overrightarrow{\beta'}_{SSB,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{SSB} \times \overrightarrow{\beta'}_{SSB,b}))} \qquad \forall \quad a \in \{1\} \tag{19}$$

$\overrightarrow{\beta}_{SSB}$ is a 1×13 vector of coefficients and, 0 as the base category. If an individual happens to collect the social security benefits, its amount is determined deterministically, using a linear transformation of the following vector:

$$\overrightarrow{X}_{ssb} = [1, m_t, ssb_{t-1}, y_t, g, x_t^1, x_t^2] \quad for A_t \geq \underline{f}$$

Then depending A_t, I use one of the three sets of pre-estimated coefficients to determine the social security benefits amount. More specifically:

$$\beta_{ssb} = \begin{cases} \beta_{ssb,early}, & \text{if} \quad A_t < f \\ \beta_{ssb,full}, & \text{if} \quad A_t = f \\ \beta_{ssb,late}, & \text{if} \quad A_t > f \end{cases}$$

Knowing the relevant set of coefficients, I compute the social security amount as a linear transformation:

$$\beta_{ssb} = \begin{cases} \overrightarrow{X}_{ssb} \times \overrightarrow{\beta'}_{ssb}, & \text{if} \quad SSB_{t-1} = 0 \\ ssb_{t-1}, & \text{if} \quad SSB_{t-1} = 1 \end{cases}$$

Total Medical Expenditure

Total medical expenditure is a noisy measure, and it makes it hard to predict its amount. I define a medical expenditure shock that improves the simulation of the variable. I assume a conditional logistic distribution for the medical expenditure shock. This shock on top of the physical health shock and number of health conditions shock which affect physical health and number of health conditions respectively helps to better replication of the medical expenditure. I discretize the continuous medical expenditure to seven bins of 0, 5, 10, 20, 30, 40, 50. The probability of facing one of the medical expenditure bins follows a multinomial logistic formula. The index vector is:

$$\overrightarrow{X}_{ME} = [1, m_t, h_t^p, h_t^c, A_t, k_t^1, k_t^2, g, E]$$

Then, for each individual, the probability of being hit by one of the medical expenditure hits is a multinomial logistic:

$$Pr(ME_t = a|\overrightarrow{X}_{ME}) = \frac{\exp(\overrightarrow{X}_{ME} \times \overrightarrow{\beta'}_{ME,a})}{1 + \sum_{b \neq 0}(\exp(\overrightarrow{X}_{ME} \times \overrightarrow{\beta'}_{ME,b}))} \quad \forall \quad a \in \{5, 10, 20, 30, 40, 50\} \tag{20}$$

Where, $\overrightarrow{\beta}_{ME}$ is a 1×9 vector of coefficients and, 0 as the base category.

Interest Rate

I draw the interest rate (r_t), in every period conditional on the economic environment

$$r_t = \begin{cases} (-5\%, 1\%), & \text{if} \quad \sigma_t^e = 1 \\ (-1\%, 5\%), & \text{if} \quad \sigma_t^e = 0 \end{cases}$$

This range replicates the interest rates in different periods of crisis in the U.S. as reported by the NBER's Business Cycle Dating Committee.

Assets

Knowing components of asset and their transitioning probabilities, I can construct the asset at the end of the current period. Asset at the end of period t (a_t) is rebuilt by adding income from working, social security benefit income and dividend from previous periods accumulated asset that assumes to grow by the revealed interest rate (r_t). Then, consumption, out-of-pocket medical expenses, and different health insurance premiums are subtracted from asset:

$$\begin{aligned} a_{t+1} = \quad & y_t + a_t(1 + r_t) \\ + \quad & SSB_t - OOP_t - c_t - ip_t^{emp}.i_t^{emp} - ip_t^{prv}.i_t^{prv} - ip_t^m.i_t^m \end{aligned} \tag{21}$$

Simulated Value Function (SV)

Assuming Additive Separability (AS) of the preference shocks, agent i's discounted lifetime utility can be shown by:

$$\sum_{t=1}^{T} \beta^{t-1}.\big(u(D_{i,t}, S_{i,t}) + \epsilon(D_{i,t})\big) \tag{22}$$

Where β is discount factor, T is the terminal stage, ϵ is the choice specific preference shock and $D_{i,t}$ and $S_{i,t}$ are respectively, decision and state of the individual i at time t. Individual's maximization problem then can be written as::

$$\max_{\delta} \big\{ E\big[\sum_{t=1}^{T} \beta^{t-1} [u(S_t, D_t) + \epsilon(D_t)] \big] \big\}$$

Calling the future utilities, the **value function**, and defining it as the expected discounted lifetime utility following current state and preference shock we have:

$$V(S_t, \epsilon_t) = \max_{\delta} \big\{ E\big[\sum_{\tau=t+1}^{T} \beta^{\tau-1} [u(S_\tau, D_\tau) + \epsilon(D_\tau) | S_t, \epsilon(D_t)] \big] \big\}$$

Then Bellman's principle of optimality can define the value function as:

$$V(S_t, \epsilon_t) = \max_{D_t} \big[u(S_t, D_t) + \epsilon(D_t) + \beta E\big(V_{t+1}(S_{t+1}, \epsilon(D_{t+1}) | S_t, D_t) \big) \big]$$

or more specifically:

$$\begin{aligned} V(S_t, \epsilon_t) = & \sum_{D_t} I[\kappa_t = D_t] \big\{ u(S_t, D_t) + \epsilon(D_t) \\ & + \beta \sum_{S} \int_{\epsilon} \big(V_{t+1}(S_{t+1}, \epsilon(D_{t+1})) g(\epsilon(D_{t+1})) d\epsilon(D_{t+1}) \big) f(S_{t+1} | S_t, D_t) \big\} \end{aligned}$$

Where κ is the decision rule. Next, we define *ex ante* value function by integrating out the current period preference shock. Ex-ante value function defines the value of each state before the preference shock is revealed:

$$\bar{V}(S_t) = \int_{\epsilon} V_t(S_t, \epsilon(D_t)).g(\epsilon(D_t)).d\epsilon(D_t)$$

Having the ex-ante value function, we can define the choice specific value function $\nu(S_t, D_t)$, that defines the value of choosing decision D_t while being in state S_t and acting optimally for the future periods:

$$\nu(D_t, S_t) = u(D_t, S_t) + \beta \sum_{S_t} \bar{V}(S_{t+1}) f(S_{t+1}|D_t, S_t)$$

McFadden (1981) defines the social surplus function that is simply expected utility of choosing a specific alternative::

$$G[\{\nu(S_t, D_t)|S_t\}] = \int \max_{D_t} [\{\nu(S_t, D_t)|S_t\}] g(\epsilon(D_t))$$

Assuming AS and CI:

$$Pr(dS_{t+1}, D_{t+1}) = Pr(D_{t+1}|S_{t+1}).f(dS_{t+1}|S_t, D_t)$$

Where conditional choice probability is:

$$Pr(D_{t+1}|S_{t+1}) = \frac{\sigma G[\{\nu(S_t, D_t)|S_t\}]}{\sigma\{\nu(S_t, D_t)|S_t\}}$$

Substituting the choice specific value functions, and given the observed panel, the full information maximum likelihood estimator θ can be defined as:

$$\hat{\theta} = \underset{\theta}{\operatorname{argmax}}\, l(\theta) = \prod_{a=1}^{A} \prod_{t=1}^{T_a} Pr(D^a_{t+1}|S^a_{t+1}, \theta) f(S^a_t|S^a_{t-1}, D^a_{t-1}, \theta)$$

We can reduce the number of parameters which require to be estimated in likelihood function by employing asymptotically normal and inefficient 2-stage procedure that separates parameters into those which are appeared only in transition function f and others which only appear in u, β and g. Hence the first stage and the second stage estimator can be estimated separately. Where the first stage estimator θ_f is:

$$\hat{\theta}_f = \underset{\theta_f}{\operatorname{argmax}}\, l(\theta_f) = \prod_{a=1}^{A} \prod_{t=1}^{T_a} f(S^a_t|S^a_{t-1}, D^a_{t-1}, \theta_f)$$

And the second stage estimator θ_u, is based on the assumption that the first stage estimated parameters $(\hat{\theta}_f)$ are the truth:

$$\hat{\theta}_u = \underset{\theta_u}{\operatorname{argmax}}\, l(\theta_u) = \prod_{a=1}^{A}\prod_{t=1}^{T_a} Pr(D_{t+1}^a | S_{t+1}^a, \theta_u, \hat{\theta}_f)$$

Taking log from the above formula:

$$\hat{\theta}_u = \underset{\theta_u}{\operatorname{argmax}}\, L(\theta_u) = \sum_{a=1}^{A}\sum_{t=1}^{T_a} Log[Pr(D_{t+1}^a | S_{t+1}^a, \theta_u, \hat{\theta}_f)]$$

Reaching the consistent estimate of θ_u involves calculating of the choice specific value functions by the solving the Bellman's equation which is computationally expensive. As shown by Rust (1987), If preference shock is an IID extreme value distribution, Conditional choice probabilities can be written as follows:

$$Pr(D_t|X_t) = \frac{1}{\sum_{D_t} exp(\nu(S_t, D_t) - \nu(S_t, d_t))}$$

Hotz and Miller (1993) show that if preference shock is an IID extreme value distribution and CI and As are satisfied, the difference between choice specific value functions can be defined as a function of the conditional choice probabilities $P(D_t|S_t)$. More specifically, the difference between the value of choosing decision D and an anchor decision 1 can be written as:

$$\Delta\hat{\nu}(S, D) = log\{\frac{P(D|S)}{P(1|S)}\}$$

Based on Hotz and Miller (1993), Hotz et al. (1994) introduce the simulated value (SV) function estimator that is a smooth function of the structural parameters θ. The idea behind SV estimator is to use a non-parametric estimation to get consistent estimate of Conditional Choice Probabilities and then inverting them to obtain the value function which is normalized by an anchor choice ($\Delta\, \widehat{\nu}(\mathrm{x}, \mathrm{d}) = \nu(x, d) - \nu(x, d_o)$). Ideally, we can use a bin estimator of $p(D_t|S_t)$ as a completely non-parametric estimate that gives us the probability of each choice in every state. But data limitations force us to employ some smoothing techniques. For this purpose, I smooth the surface of conditional choice probabilities by employing a flexible functional form including quadratic terms and interactions of state variables in a logit platform. For transition probability of stochastic state variables $f(S_t|S_{t-1}, D_{t-1})$, I employ sets of the Multinomial Logistic models with lagged variables, and other controls as explanatory variables. The algorithm to apply the SV estimator of Hotz et al. (1994) as described by Rust (1994) is based on knowing transition probability of stochastic state variables or individual beliefs, $f(S_t|S_{t-1}, D_{t-1})$ and CCP estimates of $\hat{P}(D\ |.\ S)$. Then the Simulation Algorithm is defined as follows:

1. Calculating $\Delta\ \widehat{\nu}(\text{x}, \text{d}) = \ \log\{\frac{\hat{P}(d|x)}{\hat{P}(1|x)}\}$ [Hotz and Miller (1993)]

2. Then for each state, choice pair (data points) and each person in each period, use (x_t^n, d_t^n) as the initial point and:

   (a) Given (x_{t-1}, d_{t-1}) draw x_t from previously estimated $\pi(x_t|x_{t-1}, d_{t-1})$

   (b) Given x_t from previous step draw ϵ_t from assumed $q(\epsilon_t\ |.\ x_t)$ [EV1] and keep them as $\widetilde{\epsilon_t}$

   (c) Calculate $d_t = \widehat{\delta}(x, \epsilon) = \text{argmax}_{d\in D(x)}\ \ [\Delta\ \widehat{\nu}(\text{x}, \text{d}) + \epsilon(d)]$

   (d) Repeat until reaching terminal period (or until time preference makes the next period's effect infinitesimal).

3. Now that we have simulated values of $(\widetilde{x_t},\ \widetilde{d_t})$, using each initial value from step 2, compute the simulated value function for each θ:

$$\widetilde{\nu_\theta}(x_0, d_0) = \sum_{t=0}^{T} \beta^t \{u_\theta[\widetilde{x_t}, \widehat{\delta}(\widetilde{x_t}, \widetilde{\epsilon_t})] + \widetilde{\epsilon_t}\}$$

4. Knowing the simulated choice specific values, $\nu_\theta(S_0, D_0)$, I can form the conditional choice probabilities as a smooth function of θ_u:

$$\tilde{Pr}(D_t|X_t, \theta_u) = \frac{1}{\sum_{D_t} exp(\Delta\tilde{\nu}(S_0, D_0))}$$

   And estimate the second stage likelihood estimator of θ_u:

$$\widetilde{\hat{\theta}}_u = \underset{\theta_u}{\text{argmax}}\, L(\theta_u) = \sum_{a=1}^{A}\sum_{t=1}^{T_a} Log[\tilde{Pr}(D_{t+1}^a|S_{t+1}^a, \theta_u, \hat{\theta}_f)]$$

   Since the simulated paths are sensitive to the simulated preference shocks $\tilde{\epsilon}_t$, I need to repeat the process above for each observation a to acquire a consistent estimator of θ_u. In this study, I expand each observation to 10 observations. Thus the maximum likelihood estimator averages over the $\tilde{\epsilon}_t$ and corrects for the possible bias.

The estimation method provided here is based on the 2-stage maximum likelihood and allows for incorporating fixed and unobserved heterogeneity types. Next section provides the adjustment in the likelihood function that is needed for this purpose.

Incorporating the Unobserved Heterogeneity

Following Arcidiacono and Miller (2011) I use the modified Expectation Maximization (EM) algorithm to address the fixed, unobserved heterogeneity. EM iterates over two steps. In the Expectation step, the probabilities of choosing each alternative conditional on the observed and unobserved states are updated. The maximization step can be done assuming the unobserved states are observed and use the probabilities of each alternative conditional on being in particular unobserved state as weights. The Likelihood function for observation n at time t, when unobserved heterogeneity is introduced into the model can be written as:

$$L_{t,n}(D_{nt}, S_{n,t}|S_{nt}, \delta; \theta) = Pr(D_{nt}|S_{nt}, \delta; \theta_u).f_{tn}(S_{n,t}|D_{n,t-1}, Sn, t-1, \delta; \theta_f)$$

Where $\theta = \{\theta_u, \theta_f\}$ and, θ_f and θ_u are respectively first and second stage parameters, S is observed heterogeneity and δ is unobserved heterogeneity in state space. Integrating out unobserved heterogeneity the likelihood function and estimation problem are:

$$[\widehat{\theta}, \widehat{\zeta}] = \underset{\theta,\zeta}{\operatorname{argmax}} \sum_{n=1}^{N} \log(\sum_{s=1}^{S} \zeta(\delta|S_{n1}) \prod_{t=1}^{T} L_t(D_{nt}, S_{n,t+1}|.S_{nt}, \delta; \theta))$$

Where ζ is the distribution of the unobserved heterogeneity. Note that the likelihood function here, cannot be estimated by the familiar 2-stage procedure, because it is not a separable likelihood. Arcidiacono and Miller (2011) show that the likelihood can be replaced by a separable likelihood function suggested by Dempster et al. (1977) that shares the same First Order Conditions. Defining q as the conditional probability of person n being in unobserved state s given data and parameters $(\theta,\ \pi)$:

$$q(\delta|D_n, S_n; \theta, \zeta) = \frac{\zeta(\delta|S_{n1}) \prod_t (L_t(D_{nt}, S_{n,t+1}|S_{nt}, \delta; \theta))}{\sum_{\delta=1}^{3} \zeta(\delta|S_{n1}) \prod_{t=1}^{T} L_t(D_{nt}, S_{n,t+1}|S_{nt}, \delta; \theta)}$$

And ζ, the Average over all individuals in one observed state and gives the estimation of $\widehat{\zeta}(\delta|S_1)$:

$$\widehat{\zeta}(\delta|S_1) = \frac{\sum_n q(\delta|D_n, S_n;\ \widehat{\theta}, \widehat{zeta}) I(S_{1n} = S_1)}{\sum_n I(S_{1n} = S_1)}$$

The new separable likelihood function is defined as:

$$\widehat{\theta} = \underset{\theta}{\operatorname{argmax}} \sum_n \sum_t q(s\ |.\ d_n, x_n; \widehat{\theta}, \widehat{\pi}) Log(L_t(d_{nt}, x_{nt}, s, \theta))$$

Note that, this maximization problem is separable in contrast with what we had before. Also, it uses $q(\delta|D_n, S_n; \widehat{\theta}, \widehat{\zeta})$ as population weight. To reveal the unknown q we assume it as given and utilize EM algorithm iterations to reveal it.
Following Arcidiacono and Miller (2011) and using CI feature, let j be the choice indicator, given (Starting arbitrarily) ζ^1 and θ_f^1 we can calculate $m+1$ value of $q(\delta|D_n, S_n; \theta_f, \zeta)$:

1. Expectation Step

$$q^{m+1}(\delta|D_n, S_n; \theta_f^m, \zeta^m) = \frac{\zeta^m(\delta|S_{n1}) \prod_t \left(L_t(D_{nt}, S_{n,t+1}|S_{nt}, \delta, p^m; \theta_f^m)\right)}{\sum_{\delta'=1}^{3} \zeta^m(\delta'|S_{n1}) \prod_{t=1}^{T} L_t(D_{nt}, S_{n,t+1}|S_{nt}, \delta'; \theta_f^m)}$$

   But, replacing the conditional choice probability by non-parametric, flexible estimations, we do not need conditioning over θ_u^m:

$$\begin{aligned} &L_t(D_{nt}, S_{n,t+1}|S_{nt}, \delta, p^m; \theta^m) \\ &= p_t(D_{nt}|S_{nt}, \delta; \theta_u^m) f_t(S_{nt+1}|D_{nt}, S_{nt}, \delta; \theta_f^m) \\ &= p_t^m(D_{nt}|S_{nt}, \delta) f_t(S_{nt+1}|D_{nt}, S_{nt}, \delta; \theta_f^m) \\ &= \frac{\sum_{n=1}^{N} \left(q^m(\delta|D_n, S_n) D_{njt} I(S_{nt} = S_t)\right)}{\sum_{n=1}^{N} q^m(\delta|D_n, S_n) I(S_{nt} = S_t)} \\ &\quad .f_t(S_{nt+1}|D_{nt}, S_{nt}, \delta; \theta_f^m) \end{aligned}$$

   Then we have:

$$q(\delta|D_n, S_n; \theta_f^m, \zeta^m)\ \forall\ n \to \widehat{\zeta}^{m+1}(\delta|S_1) = \frac{\sum_n q^{m+1}\ (\delta|D_n, S_n; \widehat{\zeta^M}) I(S_{1n} = S_1)}{\sum_n I(S_{1n} = S_1)}$$

$$p_t^{m+1}(j|S_t, s) = \frac{\sum_{n=1}^{N} \left(q^{m+1}(\delta|D_n, S_n) D_{njt} I(S_{nt} = S_t)\right)}{\sum_{n=1}^{N} q^{m+1}(\delta|D_n, S_n) I(S_{nt} = S_t)}$$

2. Maximization Step: Now we can assume q^{m+1} as given and hence will be able to use a separable maximization problem. Maximize with respect to θ_f using:

$$\theta_f^{m+1} = \underset{\theta_f}{\operatorname{argmax}} \sum_n \sum_t \sum_\delta q^{m+1}(\delta|D_n, S_n) \log(f_t(S_{nt+1}|D_{nt}\ ,\ S_{nt}, \delta; \theta_f)$$

We should repeat these steps until reaching the relative convergence. Following Arcidiacono and Miller (2011) and using CI feature, we have:

$$\begin{aligned} \theta^{m+1} = \underset{\theta}{\operatorname{argmax}} \sum_n \sum_t \sum_\delta \{ & q^{m+1}(\delta|D_n, S_n)\ (\log(p_t(D_{nt}|S_{nt}, \delta; \theta_u)) \\ & + \ \log(f_t(S_{nt+1}\ |D_{nt}, S_{nt}, \delta; \theta_f))\ \ )\} \end{aligned}$$

Transition Probabilities

Table 6: Social Security Benefit Amount if claimed at 62

	Social Security Benefit Amount in 1000 Dollars			
	.1	3	6	9
Married	1.231*	1.044	1.445***	1.572***
	(0.12)	(0.10)	(0.15)	(0.17)
Lag SSB	0.870***	1.471***	1.717***	1.823***
	(0.03)	(0.05)	(0.06)	(0.07)
Income	0.923***	0.988**	1.047***	1.070***
	(0.00)	(0.00)	(0.00)	(0.00)
Gender	5.657***	0.304***	0.040***	0.013***
	(1.02)	(0.05)	(0.01)	(0.00)
Tenure	0.969	1.038*	1.057***	1.065***
	(0.02)	(0.02)	(0.02)	(0.02)
Experience	0.956***	1.149***	1.223***	1.248***
	(0.01)	(0.01)	(0.01)	(0.01)
N	65,775			

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 7: Social Security Benefit Amount if claimed at Full Retirement Age

	Social Security Benefit Amount in 1000 Dollars			
	.1	3	6	9
Married	0.738	0.422***	0.425***	0.479***
	(0.12)	(0.07)	(0.07)	(0.08)
Lag SSB	0.553***	0.980	1.262***	1.363***
	(0.02)	(0.04)	(0.05)	(0.05)
Income	0.895***	0.917***	0.998	1.053***
	(0.01)	(0.00)	(0.00)	(0.01)
Gender	40.246***	1.813**	0.419***	0.073***
	(9.32)	(0.37)	(0.09)	(0.01)
Tenure	1.015	1.016	1.042*	1.068***
	(0.02)	(0.02)	(0.02)	(0.02)
Experience	0.787***	1.031**	1.110***	1.146***
	(0.01)	(0.01)	(0.01)	(0.01)
N	65,748			

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 8: Social Security Benefit Amount if claimed at 70

	Social Security Benefit Amount in 1000 Dollars			
	.1	3	6	9
Married	0.671*	0.527***	0.568***	0.703*
	(0.11)	(0.08)	(0.09)	(0.11)
Lag SSB	0.526***	0.734***	1.137**	1.371***
	(0.02)	(0.03)	(0.05)	(0.06)
Income	0.895***	0.880***	0.936***	1.017***
	(0.01)	(0.00)	(0.00)	(0.00)
Gender	38.882***	3.823***	0.734	0.171***
	(11.41)	(0.91)	(0.17)	(0.04)
Tenure	1.034	0.955*	1.023	1.047**
	(0.02)	(0.02)	(0.02)	(0.02)
Experience	0.782***	0.973**	1.073***	1.116***
	(0.01)	(0.01)	(0.01)	(0.01)
N	64,298			

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 9: Health Insurance Premiums and Income

	(1) Private Insurance	(2) Employer Insurance	(3) Medicare Insurance	(4) Income
Married	0.712***	1.475***	0.607***	-0.293*
	(0.02)	(0.02)	(0.02)	(0.12)
Gender	-0.623***	-0.178***	0.026	-3.147***
	(0.02)	(0.02)	(0.02)	(0.13)
Age	-0.315***	-0.267***	0.008	-1.438***
	(0.01)	(0.01)	(0.01)	(0.07)
Age^2	0.002***	0.001***	-0.000***	0.009***
	(0.00)	(0.00)	(0.00)	(0.00)
Education	-0.123***	-0.074***	-0.097***	-1.029***
	(0.01)	(0.01)	(0.01)	(0.04)
Education^2	0.006***	0.005***	0.004***	0.097***
	(0.00)	(0.00)	(0.00)	(0.00)
Health Conditions	-0.497***	-0.021	0.410***	
	(0.04)	(0.04)	(0.03)	
$\text{Health Conditions}^2$	0.085***	-0.004	-0.008	
	(0.01)	(0.01)	(0.01)	
CES-D	-0.039**	0.160***	0.067***	-0.666***
	(0.01)	(0.02)	(0.01)	(0.10)
CESD^2	-0.007*	-0.019***	-0.003	0.079***
	(0.00)	(0.00)	(0.00)	(0.02)
Doctor Visit	0.335***	0.127**	0.121**	
	(0.04)	(0.04)	(0.04)	
Doctor Visits^2	-0.053***	-0.019*	-0.006	
	(0.01)	(0.01)	(0.01)	
Hospital	0.365***	-0.081*	0.347***	0.124
	(0.03)	(0.03)	(0.03)	(0.21)
Hospital^2	-0.018	0.035*	-0.041**	0.005
	(0.01)	(0.01)	(0.01)	(0.08)
Smoke	-0.033	0.006	-0.071**	
	(0.03)	(0.03)	(0.02)	
Full Retirement Age	-0.481***	-0.231***	-0.248***	
	(0.03)	(0.03)	(0.03)	
Tenure				0.567***
				(0.05)
Tenure^2				-0.001
				(0.00)
Experience^2				0.000
				(0.00)
Phisical Health				-1.636***
				(0.17)
Physical Health ^{2}				0.213***
				(0.03)
N	73,689	74,629	74,066	82,225

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 10: Probability of Receiving Social Security Benefits

	(1) SS_Receipient	(2) IV Shock	(3) Married
Married	0.209***		
	(0.02)		
CES-D	-0.044***		-0.272***
	(0.01)		(0.01)
Age	-0.001		-0.077***
	(0.00)		(0.00)
Phisical Health	-0.182***		-0.010
	(0.01)		(0.02)
Leisure	1.341***		
	(0.03)		
Income	-0.023***		-0.001
	(0.00)		(0.00)
Hospital	-0.029*		-0.045
	(0.01)		(0.03)
Doctor Visit	0.010*		0.017*
	(0.00)		(0.01)
Health Conditions	0.199***		
	(0.01)		
L.Assets	0.000***		
	(0.00)		
Female	-0.177***		
	(0.02)		
Education	0.016***		
	(0.00)		
L.Involuntary Layoff	-0.254***		
	(0.05)		
Consumption	-0.026		
	(0.02)		
L.Married		-0.364***	6.160***
		(0.05)	(0.05)
Bad Economic Environment		-0.125	
		(0.08)	
L.Tenure		-0.006	
		(0.00)	
L.Age		-0.081***	
		(0.00)	
L.Income		0.003***	
		(0.00)	
L.Health Conditions			0.017
			(0.02)
Private Insurance			0.817***
			(0.07)
Employer Insurance			-0.161*
			(0.06)
Medicare Insurance			0.234***
			(0.07)
Exercise			0.018
			(0.04)
Smoke			-0.494***
			(0.05)
N	65,859	65,865	65,865

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 11: Total Medical Expenditure

	Total Medical Expenditure					
	5	10	20	30	40	50
Married	1.118***	1.151***	1.199***	1.141**	1.166**	1.053
	(0.03)	(0.03)	(0.05)	(0.05)	(0.07)	(0.04)
Age	1.009***	1.009***	1.041***	1.021***	1.010*	1.019***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Phisical Health	0.758***	1.120***	0.836***	0.904***	1.207***	1.237***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)
Hospital	0.036***	0.234***	4.599***	5.040***	4.831***	6.837***
	(0.00)	(0.01)	(0.15)	(0.16)	(0.19)	(0.21)
Doctor Visit	1.060***	1.529***	1.018	1.268***	1.308***	1.305***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)
Health Conditions	1.434***	1.779***	1.190***	1.730***	1.933***	1.480***
	(0.01)	(0.02)	(0.02)	(0.04)	(0.07)	(0.03)
Female	0.902***	0.728***	1.053	0.770***	0.621***	0.695***
	(0.02)	(0.02)	(0.05)	(0.04)	(0.04)	(0.03)
Income	1.003***	1.002***	1.000	1.001	1.001	0.997*
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Education	1.031***	1.028***	0.988*	1.033***	1.049***	1.039***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)	(0.00)
Private Insurance	0.969	0.892**	0.898	1.075	0.910	1.122
	(0.04)	(0.04)	(0.08)	(0.09)	(0.10)	(0.09)
Employer Insurance	1.035	0.941	0.923	1.064	0.902	0.909
	(0.04)	(0.04)	(0.07)	(0.08)	(0.10)	(0.07)
Medicare Insurance	0.850***	0.631***	0.682***	0.819*	0.619***	0.697***
	(0.03)	(0.03)	(0.05)	(0.06)	(0.06)	(0.05)
N	65,862					

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 12: Mental Health

	CES-D		
	1	2	6
L.Married	0.994	0.957	1.065
	(0.02)	(0.02)	(0.05)
L.CES-D	1.849***	2.552***	3.480***
	(0.02)	(0.03)	(0.05)
L.Age	0.989***	0.994***	0.977***
	(0.00)	(0.00)	(0.00)
L.Phisical Health	1.158***	1.415***	1.737***
	(0.02)	(0.02)	(0.05)
L.Leisure	1.015	1.028	1.142
	(0.04)	(0.04)	(0.08)
L.Income	0.997***	0.997***	0.996**
	(0.00)	(0.00)	(0.00)
L.Hospital	0.968	1.014	1.053
	(0.02)	(0.02)	(0.03)
L.Doctor Visit	1.000	1.022***	1.020
	(0.01)	(0.01)	(0.01)
L.Health Conditions	1.052***	1.063***	1.159***
	(0.01)	(0.01)	(0.02)
L.Assets	1.000	1.000***	0.999***
	(0.00)	(0.00)	(0.00)
Female	0.944*	0.962	0.993
	(0.02)	(0.02)	(0.05)
Education	0.997	0.980***	0.968***
	(0.00)	(0.00)	(0.00)
L.Involuntary Layoff	1.148*	1.131	1.446**
	(0.07)	(0.07)	(0.17)
L.Consumption	0.871***	0.869***	0.888*
	(0.02)	(0.02)	(0.04)
L.Smoke	1.056	1.239***	1.696***
	(0.03)	(0.04)	(0.09)
L.Exercise	1.042	1.019	0.852***
	(0.02)	(0.02)	(0.04)
N	65,859		

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 13: Physical Health

	Physical Health (1: Excellent , 6: Death)				
	Very_Good	Good	Fair	Poor	6
L.Married	1.091***	1.014	0.870**	1.008	0.712***
	(0.03)	(0.03)	(0.04)	(0.13)	(0.04)
L.Phisical Health	3.232***	9.301***	29.641***	56.134***	9.008***
	(0.06)	(0.20)	(1.00)	(4.83)	(0.29)
Age	1.010***	1.028***	1.040***	1.059***	1.112***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Leisure	1.098*	1.325***	2.137***	3.405**	2.301***
	(0.04)	(0.06)	(0.20)	(1.45)	(0.29)
CES-D	1.212***	1.430***	1.670***	1.533***	1.394***
	(0.01)	(0.02)	(0.02)	(0.05)	(0.02)
Income	0.998**	0.998*	0.996	0.988	1.000
	(0.00)	(0.00)	(0.00)	(0.01)	(0.00)
Hospital	1.279***	1.591***	1.984***	2.621***	2.382***
	(0.03)	(0.04)	(0.07)	(0.15)	(0.08)
Doctor Visit	1.045***	1.104***	1.141***	1.101**	1.122***
	(0.01)	(0.01)	(0.01)	(0.04)	(0.01)
Health Conditions	1.250***	1.463***	1.588***	1.440***	1.481***
	(0.01)	(0.02)	(0.04)	(0.12)	(0.04)
L.Assets	1.000***	0.999***	0.999***	0.999***	0.999***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Female	1.145***	1.386***	1.393***	1.248	0.715***
	(0.03)	(0.04)	(0.07)	(0.18)	(0.03)
Education	0.982***	0.974***	0.956***	0.965**	0.967***
	(0.00)	(0.00)	(0.00)	(0.01)	(0.01)
Exercise	0.735***	0.522***	0.353***	0.178***	0.298***
	(0.02)	(0.02)	(0.02)	(0.04)	(0.02)
Smoke	1.068	1.175***	1.165*	1.278	2.130***
	(0.04)	(0.05)	(0.08)	(0.26)	(0.15)
N	65,859				

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 14: Number of Health Conditions

	Number of Health Conditions	
	1	3
Married	0.972	0.931
	(0.05)	(0.06)
L.Health Conditions	1.211e+10***	2.846e+11
	(4.51e+08)	(.)
Private Insurance	1.076	1.015
	(0.08)	(0.10)
Employer Insurance	0.968	0.862
	(0.07)	(0.08)
Medicare Insurance	1.334**	1.305*
	(0.12)	(0.14)
L.Phisical Health	1.175***	1.255***
	(0.04)	(0.05)
CES-D	1.094***	1.214***
	(0.02)	(0.03)
Age	1.002	1.004
	(0.00)	(0.01)
Income	1.000	0.997
	(0.00)	(0.00)
Hospital	2.110***	4.274***
	(0.12)	(0.27)
Doctor Visit	1.237***	1.412***
	(0.01)	(0.02)
Exercise	0.867**	0.771***
	(0.04)	(0.05)
Smoke	1.110	1.171
	(0.07)	(0.10)
N	65,865	

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 15: Number of Nights in Hospital

	Number of Nights in Hospital	
	1	3
L.Married	0.988	0.886*
	(0.02)	(0.04)
Private Insurance	1.157***	1.029
	(0.05)	(0.12)
Employer Insurance	1.155**	0.993
	(0.05)	(0.15)
Medicare Insurance	1.165***	1.228*
	(0.05)	(0.12)
L.Phisical Health	1.228***	1.497***
	(0.02)	(0.04)
L.Health Conditions	1.129***	1.301***
	(0.01)	(0.04)
CES-D	1.048***	1.126***
	(0.01)	(0.02)
Age	1.022***	1.023***
	(0.00)	(0.00)
Leisure	1.307***	2.433***
	(0.06)	(0.36)
Income	1.001	1.004
	(0.00)	(0.00)
Exercise	0.879***	0.698***
	(0.02)	(0.04)
Smoke	0.957	0.838*
	(0.03)	(0.07)
Education	1.001	0.996
	(0.00)	(0.01)
L.Hospital	1.974***	3.221***
	(0.03)	(0.08)
L.Doctor Visit	1.053***	1.083***
	(0.01)	(0.01)
N	65,865	

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Table 16: Number of Doctor Visits

	Number of Doctor Visits	
	1	5
L.Married	1.422***	1.433***
	(0.05)	(0.05)
Private Insurance	2.184***	2.795***
	(0.14)	(0.19)
Employer Insurance	2.275***	2.827***
	(0.14)	(0.19)
Medicare Insurance	1.845***	2.401***
	(0.11)	(0.15)
L.Phisical Health	0.914***	1.019
	(0.02)	(0.02)
L.Health Conditions	1.195***	1.456***
	(0.02)	(0.02)
CES-D	0.962**	0.983
	(0.01)	(0.01)
Age	1.002	0.999
	(0.00)	(0.00)
Leisure	1.423***	1.708***
	(0.09)	(0.11)
Income	1.011***	1.011***
	(0.00)	(0.00)
Exercise	1.054	0.985
	(0.04)	(0.04)
Smoke	0.734***	0.588***
	(0.03)	(0.03)
Education	1.054***	1.083***
	(0.00)	(0.00)
Hospital	1.775***	2.852***
	(0.10)	(0.15)
L.Doctor Visit	1.317***	1.840***
	(0.02)	(0.02)
N	65,865	

Standard errors in parentheses
$*p < 0.05, **p < 0.01, ***p < 0.001$.

Analysis of the Moments

To evaluate how well the model simulation matches the data I compare the first moments of the outcomes conditional on age between the simulated results based on the current policy (Medicare Eligibility age at 65) and data from HRS. The results of this comparison is shown in the Table 17 and Figures 19 through 21. Overall results show that the model is replicating the truth (data) closely and is reliable to be used for the out of sample prediction. In what follows, I discuss all the variables in five different groups: Labor supply, Monetary variables, Health, Health investment, and health insurance.

Labor Supply Moments

Labor supply outcome, measured by annual hours of work, closely replicates the data moment. When an individual works a full-time job, he supplies 1260 hours annually. While both simulated results and data show a high volume of job exits between age 60 and 70, simulation indicates a faster reduction in labor supply than the data before age 65 and slower reduction after age 65. The difference is rooted in both intensive as well as extensive margins. In another word, faster quits from full-time work and accepting a part-time job and higher retirement rate before age 65 in the simulated data in comparison to the real data. This trend is reversed for age 65 to 70. The trends are shown in Figure 17.

Monetary Variables Moments

Consumption, Assets (Wealth), Annual Medical Expenditure and social security benefits are considered as monetary variables. Simulated consumption is shown in the top-left graph in Figure 18 as well as column 5 of Table 17 and replicates the data closely and shows a slow increase in consumption as people are getting older. However, there is an anomaly in consumption at age 50 that shows high consumption that is not in accordance with the lifetime trends we can observe after age 50. Note that the data variation is high for before age 60 that reflects the smaller sample size for this age group.
Wealth shows a steep positive slope as people are getting older and increases from 50,000 $ at age 50 to about 180,000 at age 84 in the real data. Simulated data replicates this increase. This is shown in the top-right graph in Figure 18. Next monetary variable is Annual Medical expenditure, shown in bottom-left graph of Figure 18 and column 9 of Table 17. Both the real data and simulated results show a jump in medical expenditure at age 52 and 54 and slow increase after age 54. However, real data shows faster growth than simulated data between 60 to 68. Social security benefit is traced using two variables: Probability of receiving social security benefit and the amount of the social security benefit. While probability of receiving

social security rises rapidly from 44 % at age 62 (Early retirement Age) to more than 90 % at age 70 in both real and simulated data, the growth is slower in simulated data and reaches higher value of 96 % in comparison to 93 % in the real data. The amount of annual social security benefit is higher in simulated data in almost all of the age groups and is more persistent. Results are shown in columns 16 and 17 of Table 17 and bottom-right graph in Figure 18.

Health Moments

Physical and mental health status, as well as the number of health conditions and the probability of dying, are analyzed under the same title of health variables. Real and simulated data moments for the physical health status are shown in the top-left graph of the Figure 19 and column 10 of table 17. The overall trend of the reduction in health capital is grasped by the simulation closely (Higher values of physical health status show worse health). However, health status in simulated data diverges from the real data moment after age 80 and shows slightly better health status.

Mental health as shown by the probability of depression in the bottom-right graph of the Figure 19 and as the CES-D measure in column 12 of the table 17, shows improvements in the mental health until age 70 and deterioration after that. Like the physical health, mental health is replicated closely by the simulated data and the only divergence is after age 80 where simulated data predict better mental health than the real data. While the anomaly of the bad mental health can be seen beginning at age 50 due to the small sample size, mental health is persistently low until age 65 and deterioration begins after age 68. All of the changes in trends are captured closely by the simulation.

Number of health conditions is can vary between 0 and 8, is represented in top-right graph in figure 19 and in column 11 of the table 17. Simulated and real data show no statistically significant difference in 95 % confidence. In particular, simulation replicates the hump in the data between age 54 and 58 and follows the normal pattern of increase in the number of health conditions after age 60.

The last variable in the group of the health-related variables is the probability of dying that is shown in the bottom-left graph of the figure 19 and column 13 of the table 17. While the increasing probability trend is replicated closely by simulated data, simulation overpredicts the data for age groups of 50-54 and 80-84.

Health Investment Moments

Health investment refers to four variables: Probability of exercise (vigorous activity), the probability of smoke, the number of nights spent in the hospital and the number of doctor visits. The first two variables are modeled as decisions and appear in the utility function, while the number of nights in the hospital and the number of doctor visits are modeled as belief. The simulated and real data moments for the first health investment variable (probability of exercise) are provided in column 15 of the table 17 and the top-left graph of the figure 20. Data shows that this probability is close to one for the 50 years old people in the sample, and noting the definition of the variable which is having vigorous activity is reasonable since comparingly younger people are pursuing more active lifestyle than elderly people. The probability drops rapidly to about .5 between age 50 to 56, and then the reduction in the probability of exercise continues at a slower pace until it reaches .2 at age 84. The model does a good job in capturing the tends. However, the model underpredicts the probability of vigorous activity before age 54 and after age 78. For the probability of smoke, the model does a better job comparing to the probability of the vigorous activity and replicates the data moments closely. Especially, it captures the rapid reduction in probability between age 50 and 58. The results for this comparison is shown in column 14 of the table 17 and top-right graph in figure 20.

The number of nights in the hospital is the next variable. bottom-left graph of the figure 20 and column 8 of the table 17 show the data and simulated moments for this variable. The difference between the two moments is not statistically significant at 95 % confidence for any of the age groups, and the increase in utilization is captured by the simulation. The number of nights in the hospital is increasing from .22 for age 50 to .57 at age 84. The next measure of medical services utilization is the number of doctor visits that is shown in column 7 of the table 17 and the bottom-right graph in the figure 20. Like the number of nights in the hospital, the number of doctor visits is closely replicated by the simulated data. After a slight reduction in the number of doctor visits between age 50 to 54, it increases from 2.1 at age 54 about 3.5 at age 84.

Health Insurance Moments

This group of variables includes the probability of buying private health insurance and employer health insurance. Private health insurance refers to the health insurance that is bought directly by the individual and does not include the Medigap health insurance. The real and simulated moments of the probability of buying private health insurance is shown in the left graph in figure 21 as well as column 6 of the table 17. While the overall trends of the data are grasped by the simulation, it is underpredicted for age 58 to 64 and overpredicted for age 64 to 70. Employer health insurance is the next health insurance decision in the model and is shown by the right graph in figure 21. The model overpredicts the employer health insurance for age 58 to 64 and underpredicts it for age 64 to 68. It confirms that there is a complementary relationship between Private health insurance and employer health insurance and individuals are substituting one for another in the model.

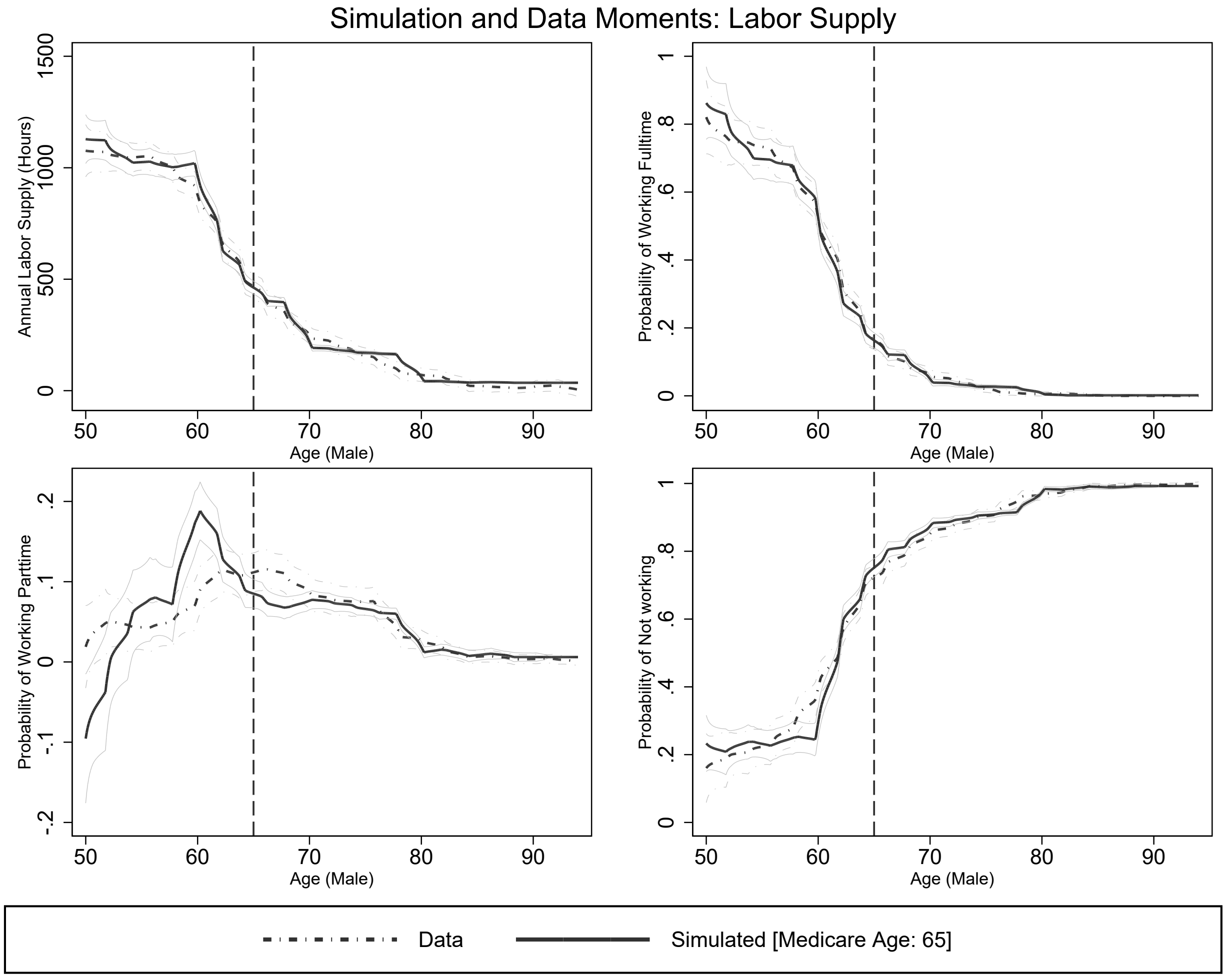


Figure 17: First moment comparison of the real data and simulated data under the current policy for the different measures of **Labor Supply**.

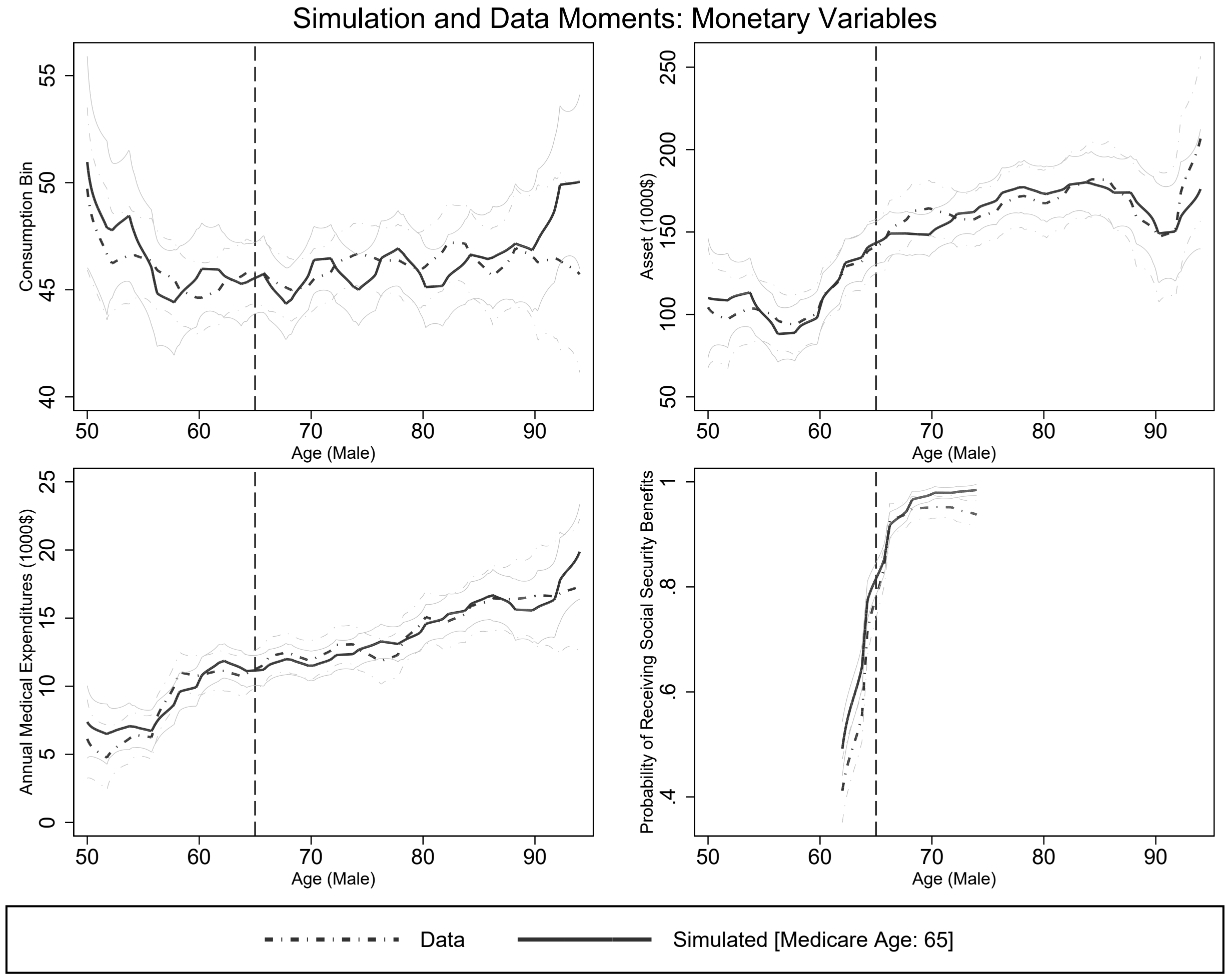


Figure 18: First moment comparison of the real data and simulated data under the current policy for the **Monetary Variables**.

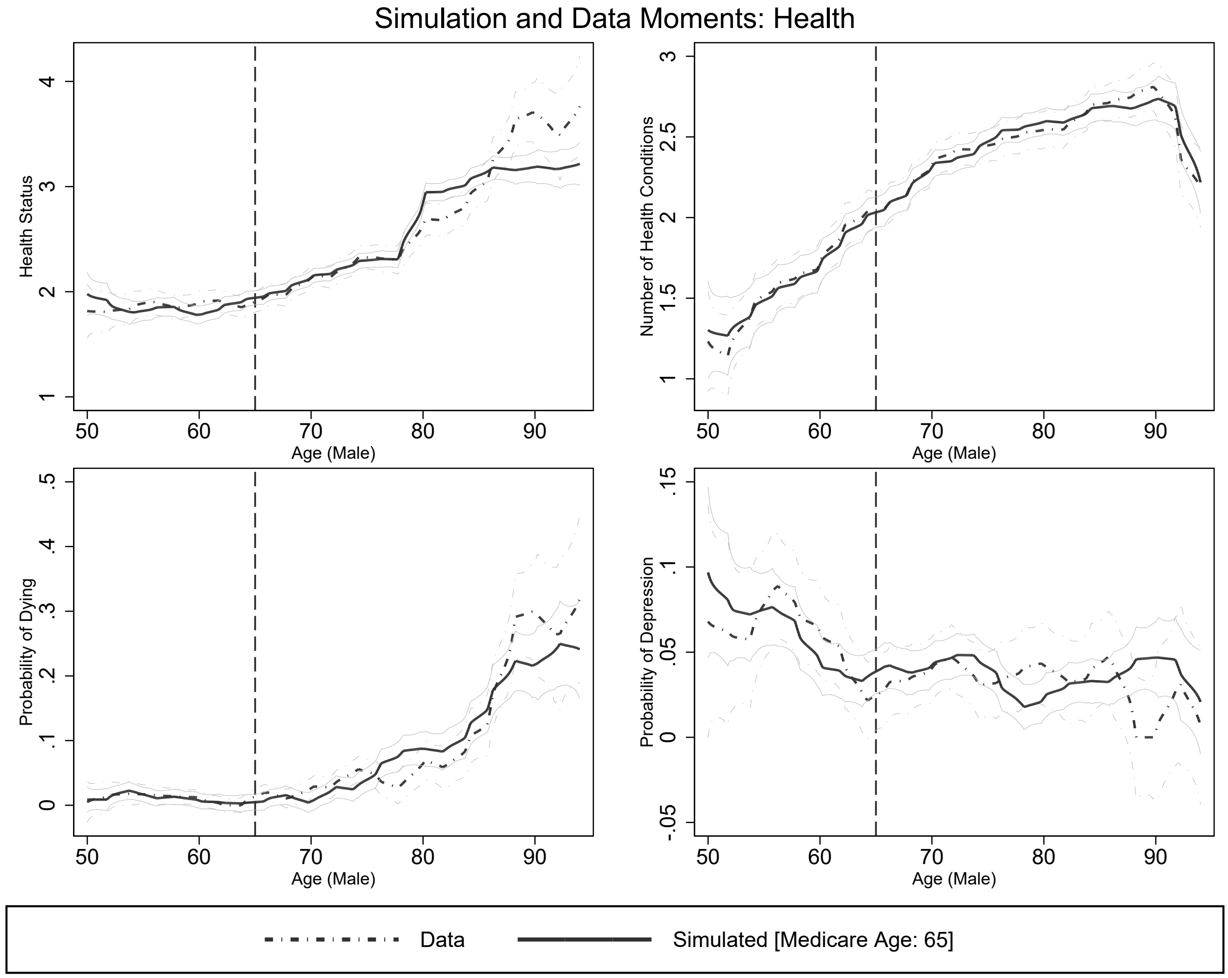


Figure 19: First moment comparison of the real data and simulated data under the current policy for the different measures of **Health**.

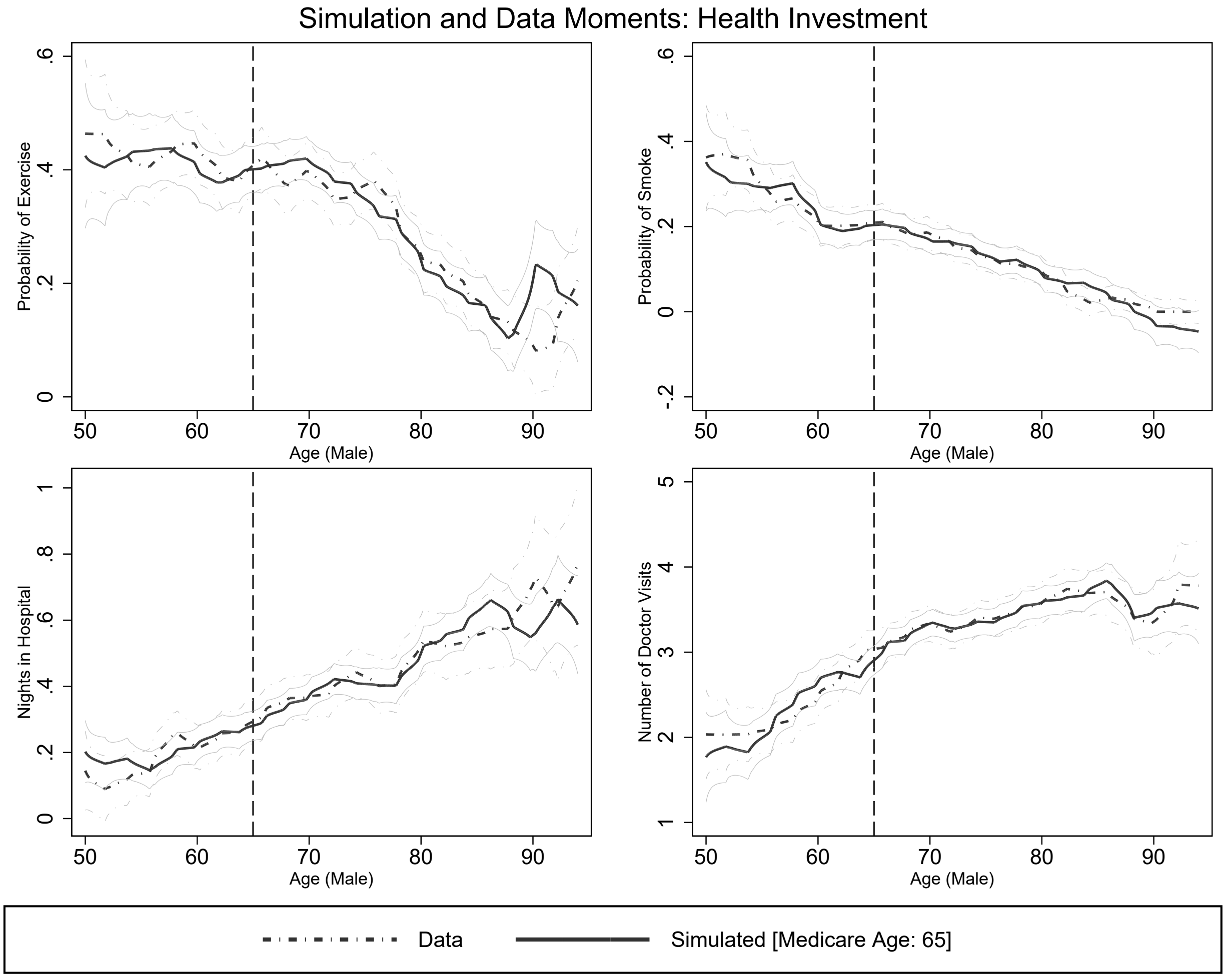


Figure 20: First moment comparison of the real data and simulated data under the current policy for the different types of the **Health Investments**.

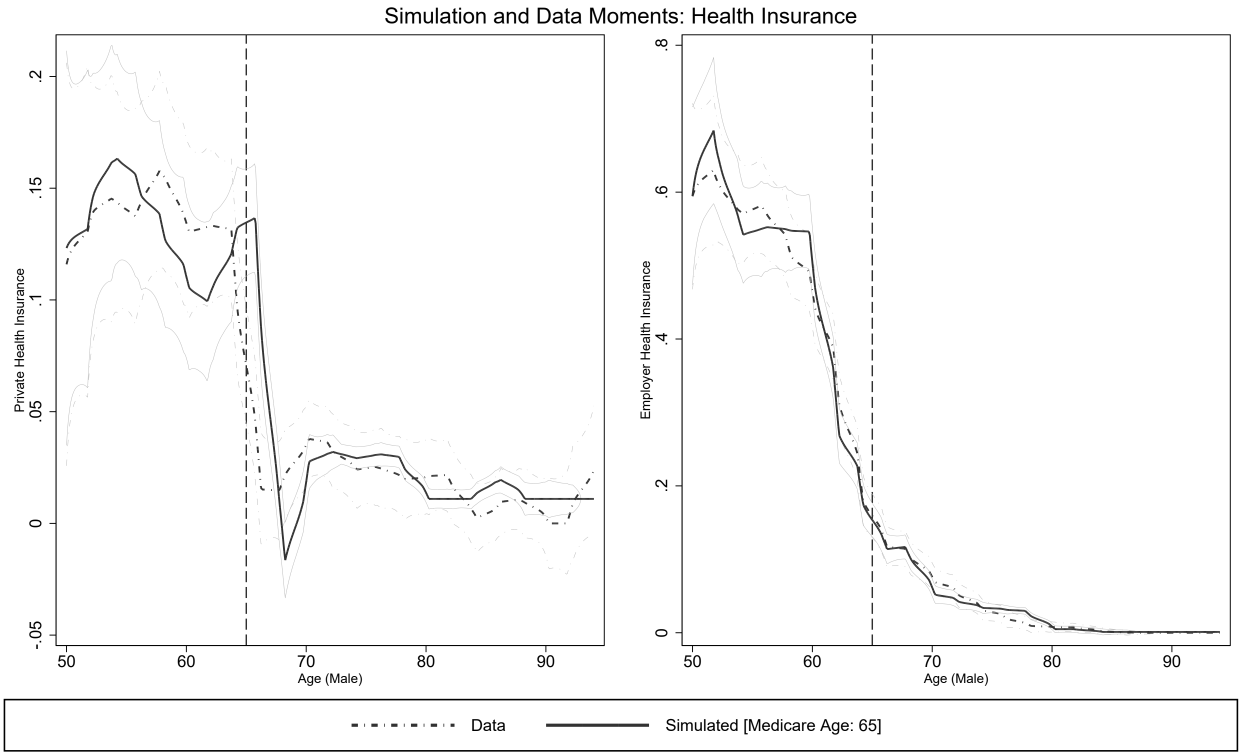
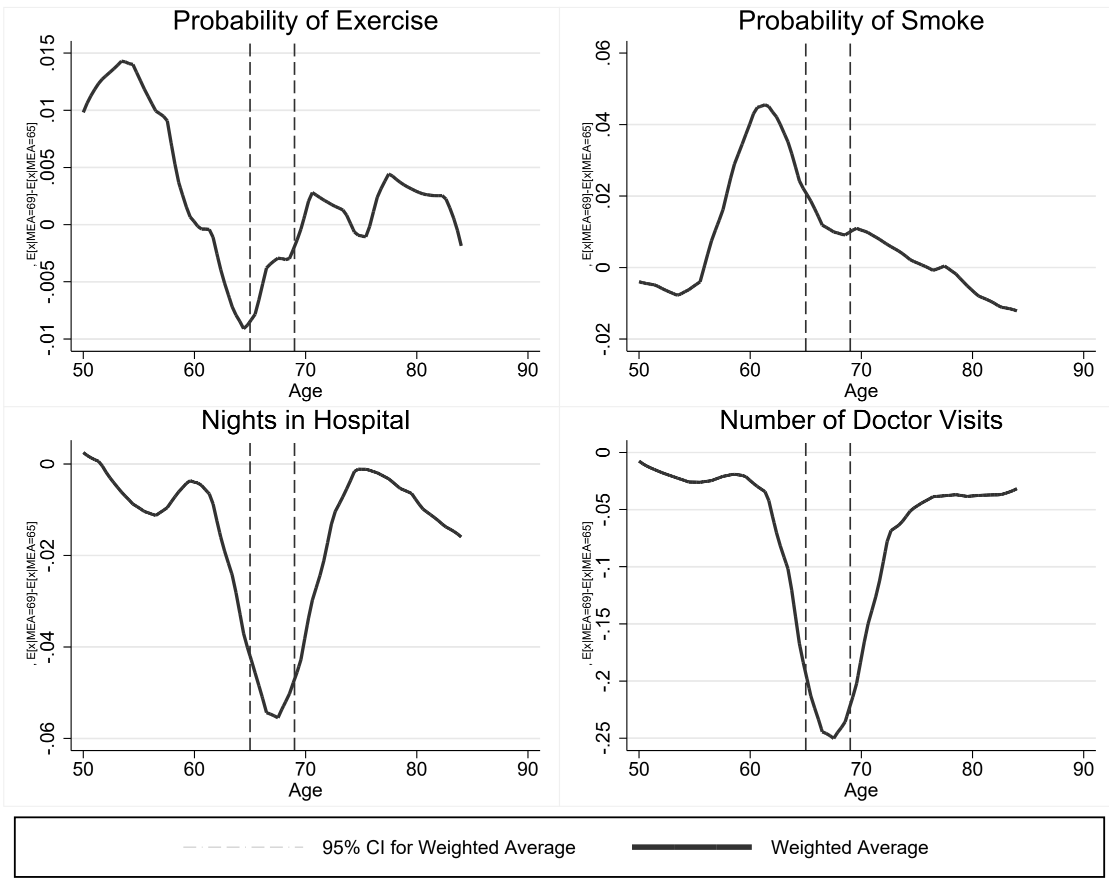


Figure 21: First moment comparison of the real data and simulated data under the current policy for the purchasing probability of different types of **Health Insurance**.

Marginal Effects for the Different sub-Samples

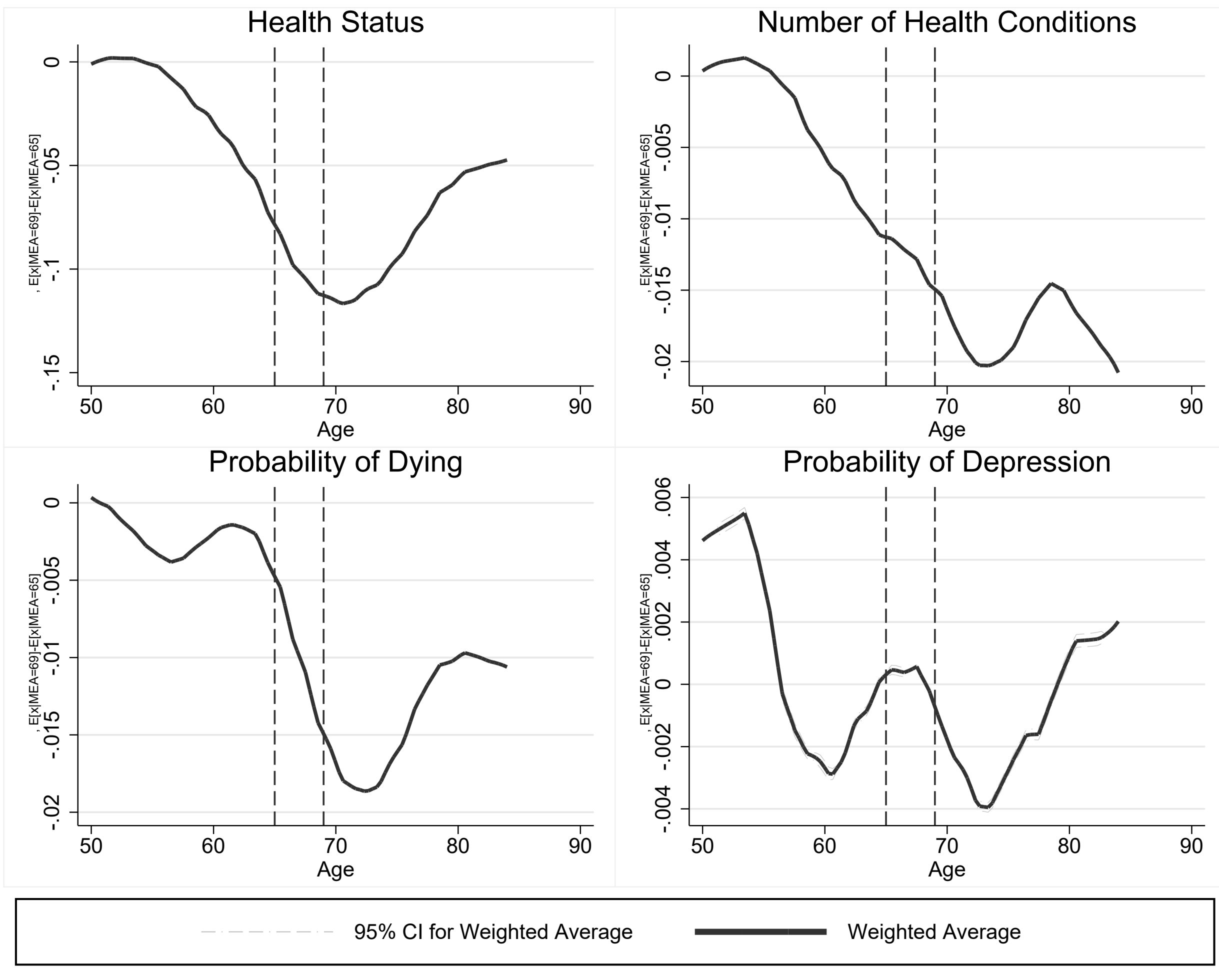
Medicare at 61: Simulation Results and Analysis

Figure 22: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different **Health Investments**



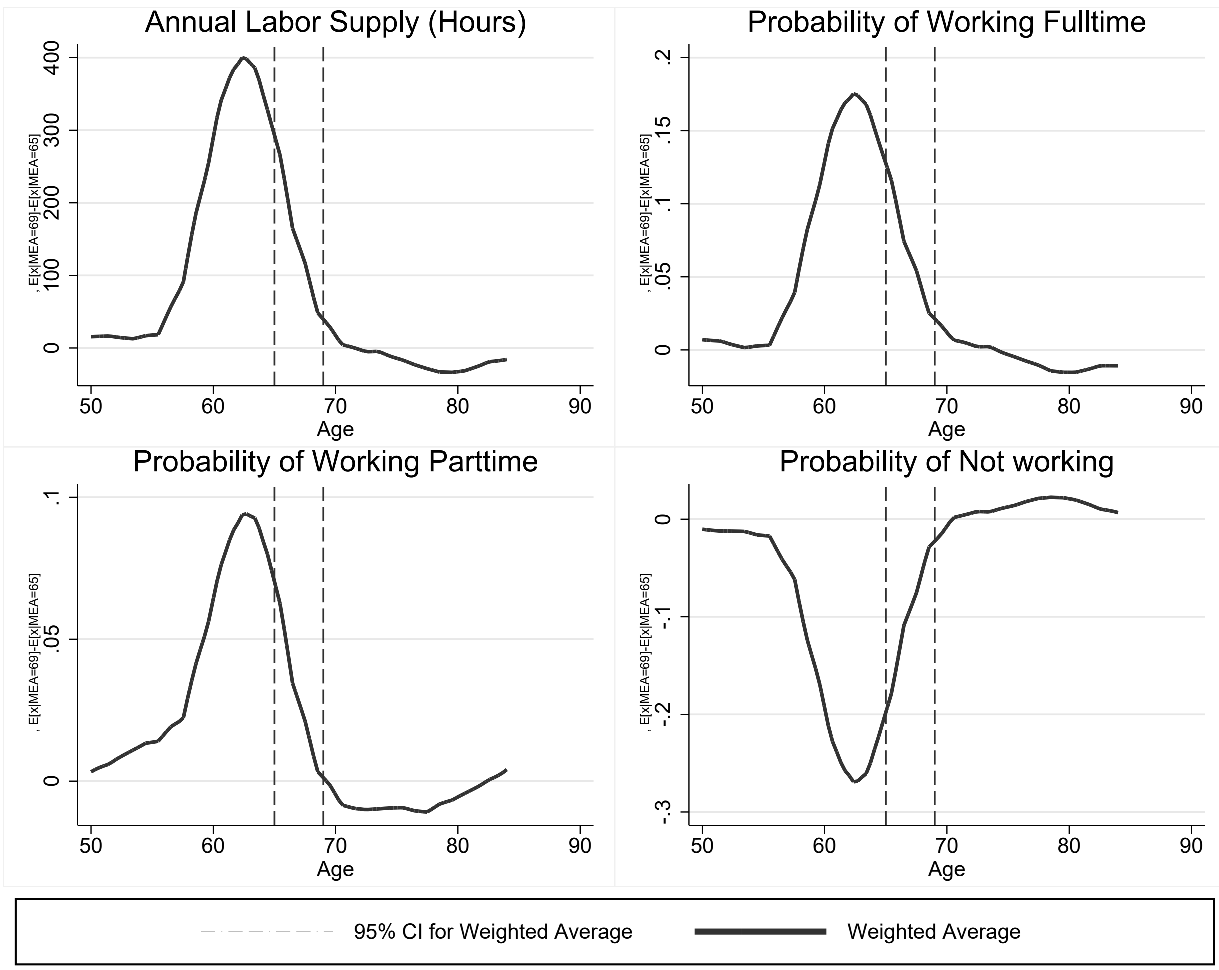
The effect of a change in policy ($E[x|MEA = 69] - E[x|MEA = 65]$) on health insurance for two different samples: Weighted average sample and sample of poor people in poor health. The confidence interval is calculated at the 95 % confidence level, where the source of variation is the estimated covariance matrix of the utility parameters. Simulation pool (sample) consists of the people who are 50 years old. The pool is weighted for different characteristics to match the characteristics, observed in the real data for 50 years old people. To produce the simulation pool for the poor people in poor health, assets, health status and the number of health conditions reset to the 10^{th}, 90^{th} and 90^{th} percentiles at age 50, respectively. This leads to 50,000 \$ of assets, fair health ($h^p = 3$) and three health conditions at age 50.

Figure 23: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on the different measures of **Health**



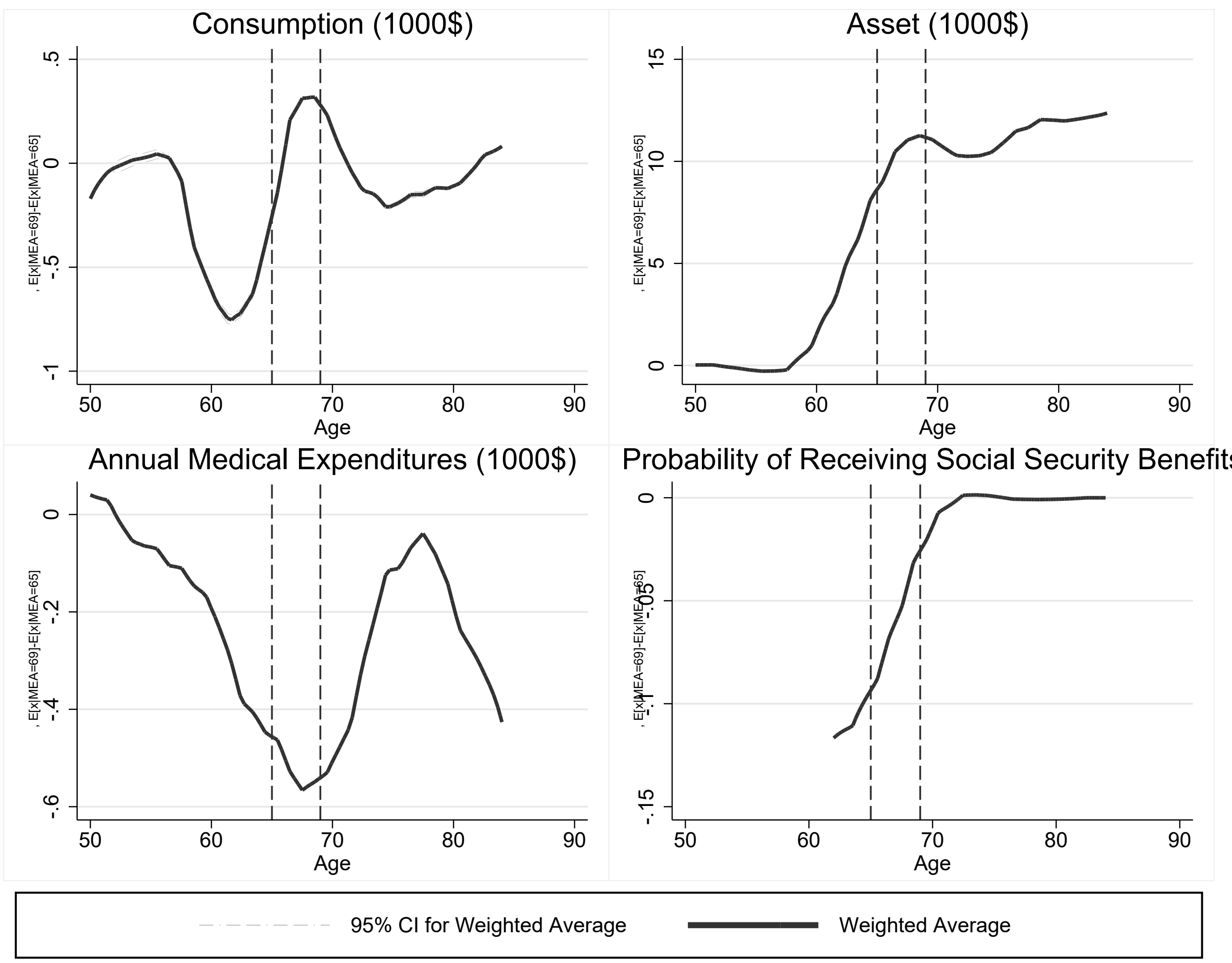
The effect of a change in policy ($E[x|MEA = 69] - E[x|MEA = 65]$) on health for two different samples: Weighted average sample and sample of poor people in poor health. The confidence interval is calculated at the 95 % confidence level, where the source of variation is the estimated covariance matrix of the utility parameters. Simulation pool (sample) consists of the people who are 50 years old. The pool is weighted for different characteristics to match the characteristics, observed in the real data for 50 years old people. To produce the simulation pool for the poor people in poor health, assets, health status and the number of health conditions reset to the 10^{th}, 90^{th} and 90^{th} percentiles at age 50, respectively. This leads to 50,000 $ of assets, fair health ($h^p = 3$) and three health conditions at age 50.

Figure 24: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on **Labor Supply**



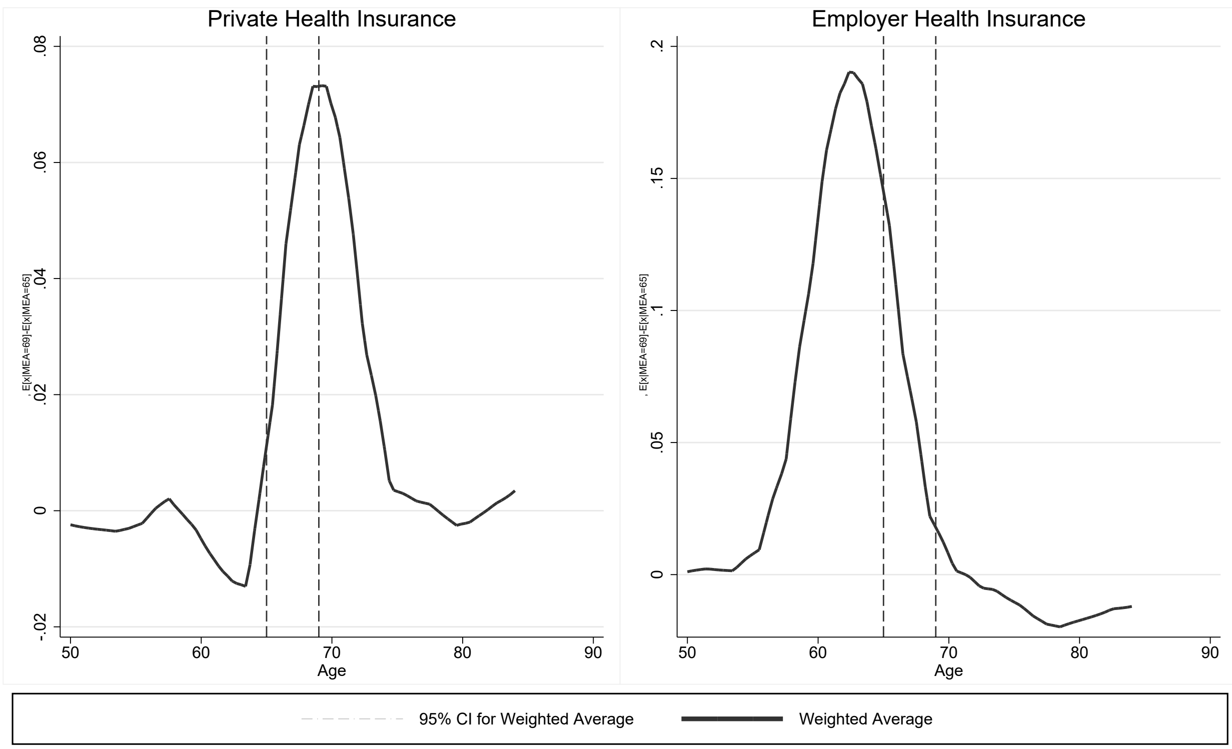
The effect of a change in policy ($E[x|MEA = 69] - E[x|MEA = 65]$) on labor supply for two different samples: Weighted average sample and sample of poor people in poor health. The confidence interval is calculated at the 95 % confidence level, where the source of variation is the estimated covariance matrix of the utility parameters. Simulation pool (sample) consists of the people who are 50 years old. The pool is weighted for different characteristics to match the characteristics, observed in the real data for 50 years old people. To produce the simulation pool for the poor people in poor health, assets, health status and the number of health conditions reset to the 10^{th}, 90^{th} and 90^{th} percentiles at age 50, respectively. This leads to 50,000 $ of assets, fair health ($h^p = 3$) and three health conditions at age 50.

Figure 25: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on the **Monetary Variables**



The effect of a change in policy ($E[x|MEA = 69] - E[x|MEA = 65]$) on the different monetary variables for two different samples: Weighted average sample and sample of poor people in poor health. The confidence interval is calculated at the 95 % confidence level, where the source of variation is the estimated covariance matrix of the utility parameters. Simulation pool (sample) consists of the people who are 50 years old. The pool is weighted for different characteristics to match the characteristics, observed in the real data for 50 years old people. To produce the simulation pool for the poor people in poor health, assets, health status and the number of health conditions reset to the 10^{th}, 90^{th} and 90^{th} percentiles at age 50, respectively. This leads to 50,000 $ of assets, fair health ($h^p = 3$) and three health conditions at age 50.

Figure 26: The average marginal effects of change in Medicare Eligibility Age from 65 to 69 on different types of the **Health Insurance**



The effect of a change in policy ($E[x|MEA = 69] - E[x|MEA = 65]$) on the different types of health insurance for two different samples: Weighted average sample and sample of poor people in poor health. The confidence interval is calculated at the 95 % confidence level, where the source of variation is the estimated covariance matrix of the utility parameters. Simulation pool (sample) consists of the people who are 50 years old. The pool is weighted for different characteristics to match the characteristics, observed in the real data for 50 years old people. To produce the simulation pool for the poor people in poor health, assets, health status and the number of health conditions reset to the 10^{th}, 90^{th} and 90^{th} percentiles at age 50, respectively. This leads to 50,000 \$ of assets, fair health ($h^p = 3$) and three health conditions at age 50.

Table 17: Data and simulated moments

		50-54	55-59	60-64	65-69	70-74	75-79	80-84	Total
Hours of Work	Data	1,055.250	1,018.577	704.477	373.215	211.625	113.174	53.904	402.659
	Simulate	1,068.712	1,010.089	716.864	399.711	185.142	165.123	40.615	335.526
		-(00.315)	(00.228)	-(00.413)	-(01.241)	(02.106)	-(04.629)	(01.740)	(06.971)
Leisure	Data	0.165	0.194	0.439	0.708	0.838	0.913	0.959	0.694
	Simulate	0.154	0.201	0.433	0.684	0.858	0.874	0.971	0.737
		(00.312)	-(00.226)	(00.270)	(01.407)	-(02.019)	(04.486)	-(01.983)	-(05.662)
Years of Tenure	Data	9.448	9.396	6.996	3.382	1.705	0.829	0.421	3.576
	Simulate	9.339	9.473	6.749	4.045	1.458	1.317	0.379	3.072
		(00.152)	-(00.121)	(00.645)	-(02.568)	(01.759)	-(04.047)	(00.506)	(04.499)
Assets	Data	102.709	94.784	122.201	154.738	160.417	169.275	173.639	147.624
	Simulate	112.026	88.521	124.072	149.128	156.912	175.342	176.490	151.948
		-(00.672)	(00.583)	-(00.231)	(00.629)	(00.493)	-(00.631)	-(00.331)	-(01.316)
Consumption	Data	47.155	45.708	45.145	45.306	46.029	46.365	46.647	45.943
	Simulate	48.784	44.614	45.610	44.807	45.843	46.694	45.636	45.966
		-(00.996)	(00.772)	-(00.488)	(00.533)	(00.252)	-(00.332)	(01.202)	-(00.067)
Private Health Insurance	Data	0.138	0.151	0.131	0.015	0.033	0.023	0.015	0.058
	Simulate	0.150	0.142	0.113	0.047	0.029	0.030	0.011	0.051
		-(00.364)	(00.311)	(01.006)	-(02.827)	(00.673)	-(01.103)	(00.920)	(01.676)
Doctor Visits	Data	2.034	2.180	2.732	3.150	3.324	3.446	3.666	3.102
	Simulate	1.826	2.321	2.709	3.126	3.331	3.429	3.631	3.176
		(01.090)	-(00.841)	(00.199)	(00.211)	-(00.081)	(00.158)	(00.357)	-(01.786)
Hospital	Data	0.118	0.219	0.239	0.344	0.400	0.403	0.535	0.373
	Simulate	0.182	0.173	0.250	0.321	0.401	0.402	0.546	0.392
		-(01.995)	(01.224)	-(00.404)	(00.658)	-(00.010)	(00.017)	-(00.314)	-(01.506)
Total Medical Expenditure	Data	5.911	8.903	10.820	12.332	12.560	12.058	15.124	12.170
	Simulate	7.019	8.070	11.186	11.783	11.947	13.200	15.038	12.555
		-(01.062)	(00.932)	-(00.538)	(00.700)	(01.020)	-(01.439)	(00.110)	-(01.351)
Health	Data	1.828	1.874	1.882	1.973	2.204	2.310	2.740	2.276
	Simulate	1.855	1.854	1.849	1.996	2.202	2.310	2.974	2.328
		-(00.318)	(00.289)	(00.718)	-(00.483)	(00.049)	(00.007)	-(04.358)	-(02.387)
Health Conditions	Data	1.315	1.608	1.888	2.120	2.393	2.496	2.592	2.199
	Simulate	1.351	1.577	1.865	2.118	2.365	2.543	2.617	2.258
		-(00.321)	(00.317)	(00.354)	(00.035)	(00.653)	-(00.920)	-(00.581)	-(02.647)
CES-D	Data	1.059	1.223	0.900	0.850	0.860	0.867	1.002	0.948
	Simulate	1.159	1.151	0.874	0.882	0.917	0.763	0.984	0.934
		-(00.734)	(00.618)	(00.388)	-(00.493)	-(01.045)	(01.600)	(00.321)	(00.584)
Probability of Dying	Data	0.015	0.014	0.004	0.015	0.039	0.032	0.074	0.050
	Simulate	0.018	0.012	0.004	0.013	0.017	0.070	0.094	0.062
		-(00.317)	(00.308)	-(00.077)	(00.233)	(02.612)	-(02.879)	-(01.460)	-(02.643)
Smoke	Data	0.360	0.263	0.203	0.195	0.162	0.113	0.059	0.162
	Simulate	0.311	0.298	0.198	0.200	0.160	0.120	0.074	0.149
		(01.066)	-(00.995)	(00.193)	-(00.241)	(00.166)	-(00.355)	-(00.971)	(01.731)
Exercise	Data	0.438	0.424	0.403	0.394	0.372	0.354	0.223	0.344
	Simulate	0.421	0.437	0.390	0.409	0.393	0.316	0.203	0.334
		(00.371)	-(00.317)	(00.507)	-(00.559)	-(01.037)	(01.430)	(00.883)	(00.993)
Probability of Receiving SSB	Data			0.342	0.933	0.946	0.916	0.616	0.584
	Simulate			0.424	0.931	0.981	0.981	0.887	0.765
				-(03.179)	(00.090)	-(04.534)	-(05.316)	-(13.372)	-(19.966)
SSB	Data			4.442	5.918	5.402	4.488	2.532	3.857
	Simulate			4.234	6.425	6.375	6.239	5.833	5.325
				(01.221)	-(04.579)	-(11.852)	-(15.806)	-(27.365)	-(24.827)
Probability of Being Married	Data	0.547	0.547	0.515	0.456	0.476	0.481	0.399	0.463
	Simulate	0.558	0.539	0.518	0.460	0.494	0.447	0.416	0.452
		-(0.237)	(0.196)	-(0.099)	-(0.146)	-(0.883)	(1.226)	-(0.752)	(1.201)

t-statistics in parentheses
Expected outcome is calculated for the weighted average of people at age 50. The probability of living each period affects the expected value of the outcome of interests.

Table 18: Average effects of change in the policy

		50-54	55-59	60-64	65-69	70-74	75-79	80-84
Hours of Work	MEA : 65	1,369.307	1,237.003	653.902	325.312	306.780	315.946	291.935
	MEA : 61	1,333.178	823.668	444.614	400.000	395.231	366.015	344.284
		-(02.73)	-(18.00)	-(10.79)	(02.73)	(04.04)	(01.67)	(02.09)
Assets	MEA : 65	65.745	65.745	70.589	70.589	106.309	106.309	118.449
	MEA : 61	66.238	66.238	65.054	65.054	101.269	101.269	112.295
		(00.14)	-(00.58)	-(01.63)	-(00.87)	-(01.01)	-(00.84)	-(00.77)
Consumption	MEA : 65	45.746	45.011	45.591	45.148	44.697	45.167	42.479
	MEA : 61	45.795	46.039	46.507	45.576	44.766	45.219	42.358
		(00.13)	(01.66)	(02.20)	(00.66)	(00.15)	(00.07)	-(00.19)
Doctor Visits	MEA : 65	2.468	2.815	2.961	3.411	3.562	3.674	3.414
	MEA : 61	2.439	2.810	3.175	3.516	3.623	3.677	3.411
		-(00.55)	-(00.06)	(04.08)	(01.46)	(01.12)	(00.04)	-(00.05)
Hospital	MEA : 65	0.148	0.121	0.167	0.258	0.295	0.407	0.392
	MEA : 61	0.146	0.130	0.202	0.265	0.291	0.392	0.388
		-(00.17)	(00.71)	(03.12)	(00.37)	-(00.24)	-(00.54)	-(00.20)
Total Medical Expenditure	MEA : 65	6.159	7.350	9.141	10.839	11.217	14.243	12.575
	MEA : 61	6.167	7.442	9.208	11.145	11.325	14.070	12.471
		(00.04)	(00.30)	(00.24)	(00.64)	(00.34)	-(00.29)	-(00.23)
Health	MEA : 65	1.745	1.797	2.012	2.337	2.808	3.394	4.023
	MEA : 61	1.749	1.815	2.083	2.410	2.852	3.437	4.082
		(00.16)	(00.46)	(02.01)	(01.28)	(00.84)	(00.58)	(00.96)
CES-D	MEA : 65	0.958	0.726	0.689	0.686	0.577	0.588	0.518
	MEA : 61	0.947	0.731	0.717	0.724	0.610	0.616	0.557
		-(00.31)	(00.13)	(01.05)	(00.84)	(01.16)	(00.66)	(01.13)
Life Expectancy	MEA : 65	1.00	1.00	0.99	0.99	0.98	0.98	0.93
	MEA : 61	1.00	1.00	0.99	0.99	0.98	0.98	0.93
		-(00.93)	-(00.38)	-(00.21)	-(00.02)	-(00.17)	-(00.08)	-(00.45)
Probability of Receiving SSB	MEA : 65	0.000	0.000	0.373	0.870	0.948	0.968	0.929
	MEA : 61	0.000	0.000	0.422	0.875	0.946	0.975	0.929
				(04.98)	(00.36)	-(00.31)	(00.93)	-(00.01)
SSB	MEA : 65	0.000	0.000	3.915	6.490	6.480	6.525	6.212
	MEA : 61	0.000	0.000	3.825	6.300	6.273	6.286	6.058
				-(02.61)	-(02.38)	-(02.88)	-(02.55)	-(01.61)
Probability of Being Married	MEA : 65	0.598	0.602	0.576	0.542	0.462	0.397	0.307
	MEA : 61	0.599	0.609	0.596	0.565	0.471	0.387	0.277
		(00.07)	(00.33)	(01.20)	(01.08)	(00.54)	-(00.41)	-(01.52)

t-statistics in parentheses
For each outcome, the first two rows are the outcome under the current (MEA: 65) or counterfactual policy (MEA: 61). Values in parentheses are t-statistics of the null hypothesis, $H_0 : E(X_{\text{MEA }:65}|Age) = E(X_{\text{MEA }:61}|Age)$. The effects under the current and the counterfactual policies are simulated by employing the pool of individuals at age 50 and weight for other characteristics to reflect the population of 50 years old in the sample. The pool size is 3000 and the simulation is continued for 40 years.

Table 19: Expected value of the outcome at age 50

Variable	MEA : 65	MEA : 61	Difference
Total Hours of Work	22,420.17	20,357.22	-2,062.96
	(208.42)	(158.74)	-(11.37)
Assets	97,807.99	92,864.77	-4,943.22
	(116.44)	(116.76)	-(5.51)
Total Consumption	1,506,692.17	1,500,493.91	-6,198.26
	(363.46)	(362.29)	-(1.06)
Total Doctor Visits	106.02	106.30	0.27
	(296.34)	(295.18)	(.54)
Total Hospital	8.47	8.55	0.07
	(150.52)	(145.73)	(.88)
Total Medical Expenditure	342,173.91	338,747.83	-3,426.09
	(223.78)	(216.94)	-(1.55)
Health	5.45	5.58	0.13
	(508.35)	(509.91)	(4.06)
CES-D	0.68	0.71	0.02
	(195.39)	(191.58)	(2.9)
Life Expectancy	33.01	32.59	-0.42
	(385.09)	(375.12)	-(3.4)
Number of Years Receiving SSB	18.84	18.81	-0.03
	(224.55)	(223.29)	-(.26)
Total SSB	137,342.61	131,269.57	-6,073.04
	(216.62)	(206.68)	-(6.76)
Number of years being married	16.17	16.17	0.00
	(151.08)	(155.52)	-(.59)

t-statistics in parentheses
Values in parantheses are t-statistics of the null hypothesis, $H_0 : E(X_{Data}|Age) = E(X_{Simulatied}|Age)$