

Labor Force Attachment Beyond Normal Retirement Age

Berk Yavuzoglu*

November 4, 2014

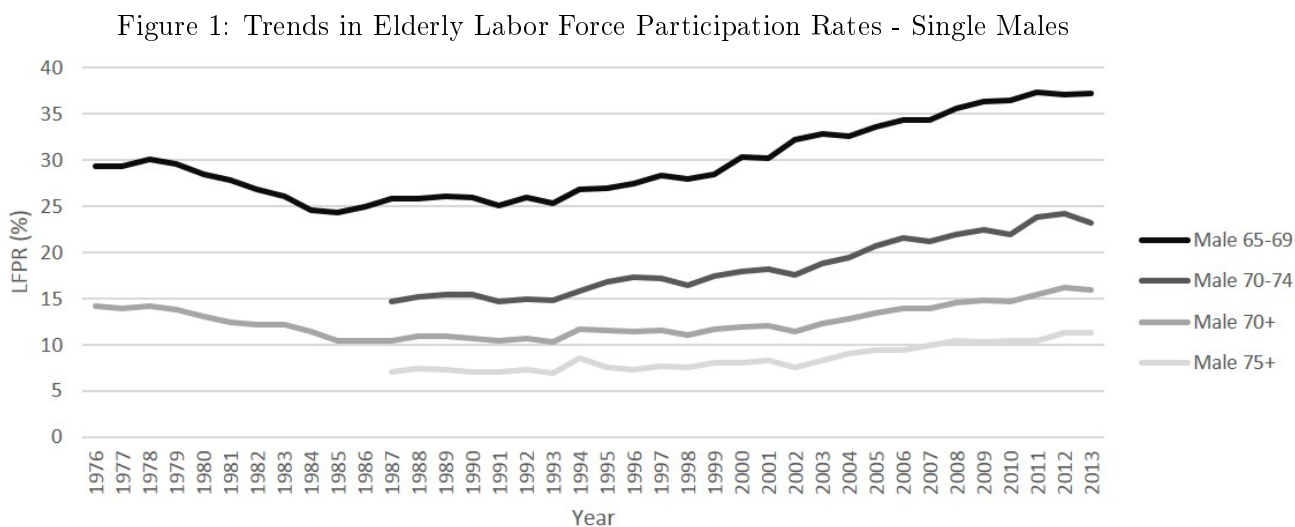
Abstract

The labor force participation decisions of the elderly in the U.S. deserve special attention due to their high participation rates beyond the normal retirement age, which is currently 66. Behavioral models studying the labor supply decisions of this age group are scant in the literature. This paper analyzes the joint determination of labor supply, consumption (savings) and the decision to apply for Social Security (SS) benefits of elderly single males. I use a dynamic programming formulation and restricted data from the Health and Retirement Study. In my study, I focus on the participation decision rather than the retirement decision because a significant portion of the elderly return to work after being non-participant for a while. I account for this through positive wage and health shocks. The estimated model helps explain the role of incentives provided by the SS system to the elderly. Counterfactual analyses show that the labor force participation decision is sensitive to changes in SS benefits and FICA tax amounts on the extensive margin, but the effects on the intensive margin are not substantial. While decreasing SS benefits by 20 percent increases the participation rate of the elderly aged 66 – 75 by 37 percent, decreasing FICA taxes by 50 percent causes the participation rate to increase by 6.6 percent for the age group 66 – 70. I further find that abolishing the year 2000 SS amendment was an important determinant of the recent increase in LFPR. Applying the earnings test on my sample decreases LFPR by 2.7 percentage points and mean hours worked by 115 hours at the age group 66-70.

*Department of Economics, University of Wisconsin-Madison, Madison, WI, 53705. E-mail: yavuzoglu@wisc.edu I am grateful to John Kennan, James Walker, Rasmus Lentz, Christopher Taber, Insan Tunali, Karl Scholz, Salvador Navarro, Eric French, Serafin Grundl and Francisco Franchetti for their helpful comments. I also thank National Bureau of Economic Research for providing me pre-doctoral fellowship through this paper. I obtained the access to the Restricted Health and Retirement Survey dataset through James Walker's NICHD grant.

1 Motivation

The labor force participation rate (LFPR) beyond normal retirement age is high in the U.S. In 2006, it was 26.2 percent for the age group 66 – 69, 20.5 percent for the age group 70 – 74, and 7.0 percent for 75+ for single males.¹ These levels have exhibited an upward trend since 1995 as shown in Figure 1.² This upward trend in the elderly participation behavior helps finance some of the fiscal burden of SS. Moreover, the U.S. population is growing older steadily, which reflects both aging of the baby boom generation and increased longevity. With the increasing stock of elderly population, it is essential to understand behavioral responses of these people to the changes in the SS system to come up with any policy analysis.



Source: Bureau of Labor Statistics

Mandatory retirement was a widespread practice in the U.S. labor market prior to the 1978 and 1986 amendments in the Age Discrimination in Employment Act.³ Since all the elderly can decide whether to work at any age after these amendments, the recent literature treats retirement as a decision. Though, it is not obvious what the term retirement stands for. It can either mean

¹These statistics are enormous compared to the developed European countries. See Table A.1 in the Appendix for a comparison. Moreover, male and female life expectancies at age 65 are higher in most of these countries than the corresponding U.S. levels. See Table A.2 in the Appendix.

²During that time, real value of the mean asset levels have been increasing as well except a temporary decrease in 2009. See Figure A.1 in the Appendix.

³Lazear (1979) argues that mandatory retirement can be seen as a life-cycle Pareto optimal contract solving the “agency problem” where workers are paid less than their value of marginal productivity when young and more when old.

starting collecting retirement benefits or quitting the labor force. Notice that retirement is not necessarily a permanent decision in the latter case since an elderly might return to work after being non-participant for a while. This makes the meaning of the term retirement vague. Hence, I focus on the participation decision of individuals beyond normal retirement age.

In this paper I analyze the labor supply, consumption and Social Security benefits application decision of elderly single males jointly, using a dynamic programming formulation. The aim of the paper is understanding the labor supply decisions of single males beyond normal retirement age, currently 66⁴, which is not well studied in the literature. I focus only on single males to avoid complexities arising from modeling the joint decision making by couples.⁵ The estimated model is useful to understand the incentives provided by the SS system to the elderly. As the counterfactual analyses, I decrease SS benefit amounts by 20 percent, and FICA tax amounts by 50 percent, which is equivalent to 3.825 percent increase in wages. I further provide an estimate of what the effect of the “earnings test” would be on my sample if it was not abolished by the year 2000 SS amendment. This quantifies the effect of the year 2000 SS amendment on the recent increase in the elderly participation rates provided in Figure 1.

The specification of the dynamic programming model in this paper extends French (2005). Unlike French (2005), I include 3 different health status categories⁶, health expenses, medicare, education levels and allow limited borrowing. French (2005) shows that the “earnings test”⁷ is the main reason for the non-participation decision of elderly people and solves the early retirement puzzle by incorporating pension benefits into his model. Rust and Phelan (1997) find that health care expenses and Medicare as well as SS rules are the important determinants of the retirement decision for financially constrained people. Recent work by Blau and Goodstein (2010), using an econometric model which is a linear approximation to the decision rule for employment, estimates that 25 to 50 percent of the recent increase in elderly LFPR is attributable to the SS rules, 16 to 18 percent to increase in education and another 15 to 18 percent to increase in LFPR of married

⁴It was 65 in 2002 and increased by 2 months each year until 2009.

⁵22.6 percent of males aged 58 – 95, the age group of interest in this paper, are single. This corresponds to 9.6 percent of the population. Only 13.5 percent of them are never married. Note that I omit cohabiting elderly males in my study.

⁶French (2005) has difficulty in matching labor force participation of unhealthy individuals due to the binary discretization of health status.

⁷See Section 6.3 for a discussion about “earnings test”.

women.⁸

Blau and Gilleskie (2008) investigate the effect of health insurance on retirement behavior. They find that changes in the access to the retiree health insurance plans provided by employers or Medicare have substantial effects on participation behavior for people with poor health, but only a modest effect for people with good health. French and Jones (2011) have a similar context to Blau and Gilleskie (2008), and they find that Medicare and employer provided health insurance, value of which is closely tied to the health care uncertainty, are important determinants of the retirement decision. Casanova (2010) approaches the retirement problem as a joint couple decision allowing for leisure complementarity and shared budget constraint in a dynamic programming framework.⁹ She shows that individual models of retirement decision cannot capture the incentives of couples. All the papers mentioned above focus on the retirement decision and utilize structural models except Blau and Goodstein (2010). Departing from the recent literature, Maestas (2010) models participation behavior and focuses on returning to work after being non-participant (she calls it unretirement) using a reduced form model. She finds that in between 1992 and 2002, 26 percent of the elderly unretire and 82 percent of this was anticipated.

I aim to contribute to the scant pool of structural papers looking at the labor supply decisions of people beyond normal retirement age. Since the elderly population is steadily increasing and the fiscal burden of SS is sizable, understanding behavioral responses of the elderly people to the changes in the SS system is essential to come up with any policy analysis. My paper aims to accomplish this by specifying a flexible model capturing most of the documented determinants of the elderly non-participation decision in the literature.

⁸Figure A.2 in the Appendix shows that even LFPRs of singles with high school or college diploma have an increasing pattern since 1995. Since we control for marital status and education, there should be another reason behind the recent increase in the elderly LFPR which Blau and Goodstein (2010) fail to explain. The reason could be the increase in the overall health status of the elderly.

⁹Casanova (2010) focuses on married people and models participation as a dichotomous decision including full-time work, part-time work and non-participation rather than a continuous hours worked decision. She further assumes that individuals start receiving Social Security benefits in the first period they choose not to participate in the labor force. Casanova (2010) does not account for changes in health status in her model.

2 Data and Preliminary Examination

Data

I use Health and Retirement Survey (HRS) data, which is a nationally representative panel data of adults in the U.S. aged 50+, conducted biannually and first fielded in 1992. It contains information on labor force participation, health, financial variables, family characteristics and a host of other topics. The results in this paper are obtained using a subsample of the HRS data comprising non-disabled single males aged 58 – 95 from 2002 to 2008. The working sample consists of 1,437 individuals with a total of 3,651 observations. Appendix A explains the steps I used to obtain the working sample from the raw data. I assume that attrition is missing completely at random (or ignorable).

Preliminary Examination

This section provides a multinomial logit analysis of the labor force participation decision of single men beyond normal retirement age. The aim is to provide basic information about the data before executing a structural labor force participation analysis of single elderly males. Since the normal retirement age has gradually increased from 65 in 2002 to 66 in 2008 with 2 month increments, and the HRS provides age data with 1 year increments, I consider 66 years as the cutoff age in this analysis. LFPR of single males aged 66 to 69 is 31.0 percent, aged 70 to 74 is 22.1 percent whereas the same statistic for single males aged 75+ is 7.8 percent in my sample. I observe only a few unemployed respondents in my sample and therefore I do not distinguish unemployment and out of the labor force states like Rust and Phelan (1997) .

Tables 1 and 2 provide summary statistics for the variables used in the multinomial logit analysis by labor force status for two age groups: 66 – 74 and 75+. I define part-time work as working less than 1,600 hours in a year.¹⁰ As seen from these Tables 1 and 2, people in the labor force are younger, more educated and healthier on average. Moreover, 93 percent of full-time workers and 98 percent of part-time workers in the age group 66 – 74 receive SS benefits.

¹⁰This assumption causes me to assign elderly people who are working full-time (more than 30 hours a week) for a while then quitting the labor force in a given year as part-time workers. However, this is the case for only 3.9 percent of the workers in my sample using the data about usual hours worked in a week in HRS. Since this is a small statistic and the unit of time is one year in my structural model, I stick to the part-time work definition given by the yearly hours worked.

Table 1: Sample Means (Standard Deviations) of Variables by Labor Force Participation Status for Single Males Aged 66-74

| Variable | Full sample | Full-Time Workers | Part-Time Workers | Out of Labor Force |
|---------------------------------|-----------------------|-------------------------|-------------------------|-----------------------|
| Age | 69.860 | 69.070 | 69.560 | 70.060 |
| High School Dropout (reference) | 0.308 | 0.243 | 0.190 | 0.340 |
| High School Graduate | 0.500 | 0.513 | 0.417 | 0.513 |
| University Graduate | 0.192 | 0.243 | 0.393 | 0.146 |
| “Fair” Health | 0.340 | 0.204 | 0.238 | 0.381 |
| Good Health (reference) | 0.314 | 0.303 | 0.321 | 0.315 |
| “Very Good” Health | 0.346 | 0.493 | 0.440 | 0.304 |
| Black | 0.217 | 0.289 | 0.226 | 0.203 |
| Medicare | 0.953 | 0.882 | 0.946 | 0.966 |
| # of Other Health Insurance | 0.727 (0.652) | 0.770 (0.581) | 0.744 (0.701) | 0.716 (0.654) |
| Health Expenses | 985.881 (2915.982) | 1,048.217 (1795.559) | 1,107.935 (2870.910) | 952.644 (3074.140) |
| Assets (in \$1,000) | 322.363 (727.838) | 370.297 (814.995) | 543.656 (1285.935) | 273.068 (536.773) |
| # of Children | 2.637 (2.236) | 2.803 (2.516) | 2.375 (1.746) | 2.659 (2.265) |
| Receiving Social Security | 0.953 | 0.934 | 0.982 | 0.951 |
| Sample size | 1222 | 152 | 168 | 902 |

To estimate a multinomial logit model of labor force status, consider the latent utility model:

$$y_{ij}^* = \theta'_{ij} z_i + \eta_{ij} \text{ for } j = 1, 2, 3. \quad (1)$$

where i denotes individuals, y_{ij}^* 's denote the unobserved utilities obtained from the choice of labor force participation status j , z_i is the vector of explanatory variables given in Tables 1 and 2, θ_{ij} 's are the corresponding vectors of unknown coefficients and η_{ij} 's are the random disturbances. Let $r = \max(y_1^*, y_2^*, y_3^*)$. Then, the labor status is given by

$$lfp = \left\{ \begin{array}{l} 1 = \text{full-time, if } r = y_1^*, \\ 2 = \text{part-time, if } r = y_2^*, \\ 3 = \text{out of labor force, if } r = y_3^*, \end{array} \right\} \quad (2)$$

Table 2: Sample Means (Standard Deviations) of Variables by Labor Force Participation Status for Single Males Aged 75+

| Variable | Full sample | Full-Time Workers | Part-Time Workers | Out of Labor Force |
|---------------------------------|-------------|----------------------|----------------------|-----------------------|
| Age | 82.616 | 78.651 | 80.059 | 82.873 |
| High School Dropout (reference) | 0.420 | 0.349 | 0.318 | 0.428 |
| High School Graduate | 0.434 | 0.465 | 0.412 | 0.434 |
| University Graduate | 0.146 | 0.186 | 0.271 | 0.138 |
| “Fair” Health | 0.415 | 0.256 | 0.212 | 0.431 |
| Good Health (reference) | 0.319 | 0.488 | 0.459 | 0.307 |
| “Very Good” Health | 0.266 | 0.256 | 0.329 | 0.262 |
| Black | 0.165 | 0.093 | 0.118 | 0.170 |
| Medicare | 0.969 | 0.977 | 0.965 | 0.969 |
| | 0.773 | 0.910 | 0.624 | 0.777 |
| # of Other Health Insurance | (0.576) | (0.526) | (0.597) | (0.575) |
| | 1,949.107 | 1,489.279 | 1,309.553 | 1,998.235 |
| Health Expenses | (7,433.505) | (2,783.727) | (1,912.462) | (7,713.378) |
| | 326.909 | | | 280.228 |
| Assets (in \$1,000) | (805.645) | 812.201 | 910.138 | (605.999) |
| | | (1,441.032) | (2,122.816) | |
| | 3.029 | | | 3.032 |
| # of Children | (2.386) | 2.721 | 3.129 | (2.411) |
| | | (2.004) | (2.109) | |
| Receiving Social Security | 0.968 | 0.977 | 0.976 | 0.967 |
| Sample size | 1,637 | 43 | 85 | 1,509 |

We assume that η_j 's satisfy the Independence of Irrelevant Alternatives (IIA) hypothesis, so they have type I extreme value distribution. McFadden (1974) proves that this specification corresponds to the Multinomial Logit model.

The choice probabilities are given by

$$\pi_j = \Pr(lfp = j \mid z) = \frac{\exp(\theta'_j z)}{\sum_{k=1}^3 \exp(\theta'_k z)}, \quad j = 1, 2, 3. \quad (3)$$

Since $\sum_{l=1}^3 \pi_l = 1$, we choose people who are out of the labor force as the reference group and set $\theta_3 = 0$. Then, we obtain consistent estimates for θ_j 's by maximizing the following likelihood function

$$L = \prod_{lfp=1} \pi_1 \prod_{lfp=2} \pi_2 \prod_{lfp=3} \pi_3. \quad (4)$$

Table 3: Multinomial Logit Estimates of Labor Force Status on Some Possible Determinants for Single Males Aged 66-74

| Variable | Full-Time | | Part-Time | |
|--------------------------------|-----------|-----------|-----------|-----------|
| | Coef. | Std. Err. | Coef. | Std. Err. |
| Age | -0.152*** | 0.036 | -0.073** | 0.034 |
| High School Graduate | 0.201 | 0.220 | 0.344 | 0.242 |
| University Graduate | 0.605** | 0.275 | 1.535*** | 0.282 |
| “Fair” Health | -0.524** | 0.250 | -0.415* | 0.229 |
| Very Good Health | 0.418** | 0.207 | 0.135 | 0.212 |
| Black | 0.661*** | 0.205 | 0.461** | 0.216 |
| Has Other Health Insurance | 0.350* | 0.194 | -0.097 | 0.184 |
| Health Expenses (in \$1000) | 0.021 | 0.021 | 0.022 | 0.022 |
| Assets (in \$100000) | 0.009 | 0.013 | 0.021** | 0.009 |
| Has Children | -0.032 | 0.204 | 0.248 | 0.210 |
| Receiving Social Security | -0.115 | 0.363 | 1.321** | 0.662 |
| Constant | 8.213*** | 2.479 | 1.298 | 2.586 |
| No. of observations | | | 1,231 | |
| Log-likelihood w/o covariates | | | -931.929 | |
| Log-likelihood with covariates | | | -865.430 | |

Robust standard errors are in parentheses.

* significant at 10%; ** significant at 5%; *** significant at 1%.

Good health is the reference group for health status. High school dropouts is the reference group for education.

The results of this estimation can be found in Table 3 for the age group 66–74. Noticed that the log odds of staying in the labor force decrease with age and increase with health status. Having high school dropouts as the reference category, being a university graduate increases the participation probability compared to staying out of the labor force whereas being a high school graduate does not have a significant effect on the participation decision.

The results also suggests that being black increases the participation probability. Having other health insurance is slightly significant for full-time work, but not for part-time work. There is a question in HRS inquiring about the primary health insurance plan of a subset of the respondents. In my sample, 15.3 percent of people in the age group 66 – 74 who responded to this question identified their primary insurance as different than Medicare. A further inspection by labor force status reveals that 51.9 percent of full-time workers, 6.0 percent of part-time workers and 9.7 percent of non-participants have a primary health insurance different than Medicare in that age group. Moreover, part-time participation probability increases with asset levels, but assets do not have a significant effect on full-time participation probability.

3 Model

I use a dynamic programming formulation. I have a three dimensional vector of control variables: consumption, hours worked in a year and a dummy variable indicating whether the individual applied for SS benefits. Consumption (c_t) and hours worked (h_t) are continuous variables obtained via splines after using discretizations.¹¹ b_t denotes the dummy variable indicating whether the individual applied for SS benefit or not.

I have a five dimensional vector of state variables: assets, wages, Principal Insurance Amount (PIA)¹², health status and education. I use 11 asset states denoted by A_t , 6 wage states denoted by w_t and 5 PIA states.¹³ There are 4 health status categories: “very good”, good, “fair” and dead¹⁴ denoted by hs_t and taking values 1, 2, 3 and 4, respectively. I have 3 education (ed_t) groups: no high school diploma ($ed < 12$ years of education), high school graduates ($12 \leq ed < 16$ years of education) and university graduates ($ed \geq 16$ years of education).¹⁵ I use projection method to accommodate continuous state space of assets, wages and PIA. I control for Medicare (m_t) in my model, and include SS benefits (ss_t) and Medicare premium (mp) in the budget constraint.

I model the problem as a discrete control process where subjects make decisions every year. Denote the control variables by d , state variables by x , and preference parameters by θ . The flow utility function, for each health status category $i \in \{very\ good\ health, good\ health\ and\ fair\ health\}$, is given by:

$$U(x_t, d_t, \theta) = \frac{1}{1-v} \left(c_t^{\theta_{C_i}} \hat{L}^{\theta_{L_i}} \right)^{1-v} \quad (5)$$

where

$$\hat{L} = L - (h_t + \theta_{P,f} + \theta_{P,good} I(\text{good health}) + \theta_{P,fair} I(\text{fair health}) + \theta_{PA} (age_t - 57)^\gamma) I(h_t > 0), \quad (6)$$

¹¹The initial discretization used for consumption is 3,000, 13,000, 23,000, 33,000, 53,000, 73,000, 93,000, 113,000, 143,000, 173,000 and 203,000. Moreover, the initial discretization used for hours worked is 0, 750, 1500, 2250, 3000 and 3750.

¹²See Section 4.3 for a discussion about the relationship between Average Indexed Monthly Earnings (AIME) and PIA.

¹³The initial asset states are given by $-15,000, 0, 15,000, 40,000, 80,000, 120,000, 200,000, 300,000, 500,000, 800,000$ and $1,300,000$. The initial wage states are given by 2, 8, 14, 20, 32 and 44. The initial PIA states are 0, 25th percentile, 50th percentile, 75th percentile and the maximum observed amount.

¹⁴HRS has 5 self-reported health status categories: excellent, very good, good, fair and poor. I combine the self-reported excellent and very good health status categories and call the new category as “very good”, and combine fair and poor health status categories and call the new category as “fair”.

¹⁵My sample is not big enough to conduct separate analyses by education groups.

The coefficient of relative risk aversion is given by v . For each health status category i , θ_{Ci} and θ_{Li} measure the consumption and leisure weights, respectively. $I(\cdot)$ is the indicator function. θ_{Pf} is the fixed cost of work, and $\theta_{P,good}$ and $\theta_{P,fair}$ are the additional participation costs depending on health status level, with $\theta_{P,very\ good}$ normalized to zero. $\theta_{PA}(age_t - 57)^\gamma$ measures the portion of the labor force participation cost explained by age.

Following De Nardi (2004), people who die value asset bequests according to the function

$$b(A_t) = \theta_B \frac{(A_t + K)^{\theta_{C2}(1-v)}}{1-v} \quad (7)$$

where K measures the curvature of the function. With $K > 0$, the disutility of leaving non-positive bequests in the amount of less than K dollars becomes finite. The curvature implicitly sets a borrowing constraint since the elderly face mortality uncertainty each period.

The constraints are the wage determination equation, health status determination equation, the health expenses determination equation and the asset accumulation equation.

I do not observe wages for more than half of the employed workers. I impute them for each cross-section separately, using the solution methodology for double selection problems provided by Tunali and Yavuzoglu (2012). This methodology relaxes the trivariate normality assumption among the error terms of the two selection equations and the regression equation by following the Edgeworth expansion approach of Lee (1982). The details can be found in Appendix B.

Log wages in the current period depend on age, education and PIA:¹⁶

$$\ln(w_t) = \varsigma_0 + \varsigma_1 age_t + \varsigma_2 \frac{age_t^2}{100} + \delta_{high} I(12 \leq ed_t < 16) + \delta_{uni} I(ed_t \geq 16) + \delta_{PIA} \frac{PIA}{100} + AR_t, \quad (8)$$

where

$$AR_t = \rho_{AR} AR_{t-1} + \eta_t, \quad \eta_t \propto N(0, \sigma_\eta^2). \quad (9)$$

According to the human capital theory, workers should be paid their marginal product which decreases over the time due to the decrease in health stock and human capital investment. The resulting wage process is approximated through Equations 8 and 9. PIA is included as a proxy for work experience.

¹⁶Having no high school diploma is the reference category for wage premium coefficients for high school and university graduates.

Health status next period (including being dead) depends on the current health status, age and education:¹⁷

$$\mu_{j,i,age_t,ed} = \Pr(hs_{t+1} = j | hs_t = i, age_t, ed). \quad (10)$$

Out of pocket health expenses depend on age, health status, medicare and asset levels:

$$\begin{aligned} \ln(he_t) &= \varphi_0 + \varphi_1 age_t + \frac{\varphi_2}{100} age_t^2 + \delta_{fair} I(fair\ health) + \delta_{good} I(good\ health) \\ &+ \delta_{medicare} m_t + \delta_{assets} \left(\frac{A_t}{100,000} \right) + \xi_t, \end{aligned} \quad (11)$$

where

$$\xi_t \propto N(0, \sigma_\xi^2). \quad (12)$$

The age dependency of out of pocket health expenses arises from the increasing hazard rates of serious illnesses with age. I assume everyone is entitled to Medicare at age 65, which causes a reduction in out-of-pocket health expenses. This, in turn, provides an incentive for the elderly to leave the labor force. I include asset levels in Equation 11 because of the positive correlation between wealth and the quality of care demanded. Moreover, poor people might be covered by Medicaid when confronted with high out-of-pocket health expenses.¹⁸

The asset accumulation equation is given by:

$$A_{t+1} = (1 + r)A_t + Y_1(w_t h_t, \tau_1) b_t ss_t - he_t - c_t - mp - Y_2(G_t, \tau_2), \quad (13)$$

where r is the interest rate, $Y_1(w_t h_t, \tau_1)$ is the level of post-FICA tax wage earnings, τ_2 is the tax structure regarding state and federal taxes and $Y_2(G_t, \tau_2)$ is the level of tax amount paid out of gross taxable earnings, G_t . It is generated via:

$$G_t = w_t h_t + Y_3(b_{t-1} ss_{t-1}, \tau_3) \quad (14)$$

where $Y_3(b_{t-1} ss_{t-1}, \tau_3)$ is the taxable portion of the SS benefits.

¹⁷See Section 4.1 for the functional form.

¹⁸The magnitude of the standard deviation of the out of pocket health expenses corresponds to the 95th percentile value. However, most of the elderly would never face those extreme expenses amounts or uncertainty (standard deviation) since it depends on their choices. For that reason, I use health expenses values up to 96th percentile to calculate data moments in my simulated method of moments estimates.

I assume that wage decrease and health deterioration are the main determinants of the non-participation decision of the elderly. However, non-participation is not a permanent decision so that an elderly might return to work after being a non-participant for a while. The data reveals that 5.6 percent of the non-participants aged 66 – 67 return to work within 2 years, 10.0 percent within 4 years and 10.3 percent within 6 years.¹⁹ I account for this through wage, health and health expenses shocks.

$$\begin{aligned}
V_t(x_t) = & \max_{c_t, b_t, h_t} [u_t(x_t, d_t, \theta) + \beta (\sum_{j=1}^3 \Pr(hs_{t+1} = j | hs_t, m_t, ed, t) \times \\
& \int \int V(x_{t+1}) dF(w_t | w_{t-1}, ed, PIA, t) dG(he_t | hs_t, m_t, t) \\
& + \Pr(hs_{t+1} = 4 | hs_t, m_t, ed, t) \times \\
& \int \int b(A_{t+1}) dF(w_t | w_{t-1}, ed, PIA, t) dG(he_t | hs_t, m_t, t))] \quad (15)
\end{aligned}$$

Equation 15 provides the Bellman equation where $F(\cdot|\cdot)$ and $G(\cdot|\cdot)$ denote the conditional distributions of next period wages and current period health expenses respectively, and β denotes the intertemporal discount factor. Each period, people transit into “very good”, good or “fair” health statuses, or they die. If they live, they get a continuation value dependent on their health status, and if they die, they receive bequest value. Both the continuation and bequest values next period depend on the health expenses and wage shocks this period, which I integrate over to obtain expected values. I assume that terminal age is 95 and solve the problem recursively. This assumption does not mean that everyone dies at 95, but people die with probability 1 at age 95. This is an innocuous assumption since the mortality rate is very high beyond 95 and simplifies the problem computationally. The optimal decision rule will be given by $\delta = (\delta_0, \delta_1, \dots, \delta_T)$ where $d_t = \delta_t(x_t)$ specifies optimal decision d_t as a function of the state variables x_t .

The model will be estimated in 2 steps. In the first step, I estimate some parameters and calibrate some others given by $\{r, L, mp, \Pr(hs_{t+1} | hs_t, age_t, m_t, ed), PIA, \tau_1, \tau_2$ and $\tau_3\}$. I assume rational expectations. Given the first stage estimation, I estimate the following parameters in the model using the simulated method of moments $\phi = \{v, \theta_{Ci}$'s, θ_{Li} 's, θ_{Pf} , $\theta_{P,fair}$, $\theta_{P,good}$, θ_{PA} , γ , v in

¹⁹2.3 percent of all the non-participants beyond normal retirement age return back to work within 2 years, 3.1 percent within 4 years and 4.7 percent within 6 years.

the flow utility function, ϱ , ς_0 , ς_1 , ς_2 , δ_{high} , δ_{uni} , $\delta_{PIA,\rho AR}$, σ_η^2 in the wage determination equation, φ_0 , φ_1 , φ_2 , $\delta_{medicare}$, δ_{fair} , δ_{good} in the health expenses determination equation, θ_B and K in the bequest function, β }.

4 First Stage Estimation

I set the yearly interest rate, r , equal to 0.04, the time endowment, L , equal to 5000, and yearly Medicare premium, mp , equal to \$1,062 for people subscribed to Medicare.²⁰

4.1 Health Transition Matrix

It is not viable to estimate Equation 10 non-parametrically since it involves a health transition matrix for each possible education, medicare and age combination.²¹ Consequently, I estimate a parametric model of transition rates via maximum likelihood utilizing the methodology of Robinson (2002):

$$p(j|i) = \Pr(hs_{t+1} = j | hs_t = i, age_t, ed) = \exp(a_{ij,ed} + b_{ij,ed}(age_t - 57) + c_{ij,ed} \frac{(age_t - 57)^2}{100})$$

for $i < 4$ and $i \neq j$. (16)

Table 4: Maximum Likelihood Estimates of the Health Status Determination Equation for Male High School Graduates

| $i \setminus j$ | $\hat{a}_{ij,ed=high-school}$ | | | |
|-------------------------|-------------------------------|-------------------|-------------------------------|-------------------|
| | “very good” | good | “fair” | dead |
| “very good” | — | -2.018 (0.052) | -3.772 (0.147) | -5.510 (0.225) |
| good | -1.469 (0.108) | — | -2.244 (0.064) | -4.697 (0.195) |
| “fair” | -3.013 (0.207) | -1.421 (0.124) | — | -3.855 (0.204) |
| | $\hat{b}_{ij,ed=high-school}$ | | $\hat{c}_{ij,ed=high-school}$ | |
| $i < j$ (recovery) | -0.045 (0.015) | | 0.092 (0.048) | |
| $j = 4$ (death) | 0.073 (0.016) | | 0.029 (0.036) | |
| $i > j$ (deterioration) | 0.001 (0.0009) | | 0.091 (0.013) | |

²⁰This corresponds to the 2006 Medicare Part B premium.

²¹This corresponds to $3 \times 2 \times 37 = 222$ different health transition matrices, and $222 \times 9 = 1,998$ parameters.

Note that $\Pr(hs_{t+1}|hs_t = 4, ed_t, age) = 0$ since $hs_t = 4$ corresponds to death. There is no restriction on $a_{ij,ed}$ values. The age adjustment parameters, $b_{ij,ed}$ and $c_{ij,ed}$, are restricted to 3 values: one for recovery ($i < j$), one for mortality ($j = 4$) and one for health deterioration ($i > j$). The parameters estimates for high school graduates can be found in Table 4.^{22,23} Notice that the higher the estimate is (in absolute value) the lower the probability.

To assess the performance of the estimation, I compare the implied 2 year transition rates in the model with the data at the first quartile, median and third quartile of the age distribution, provided in Table 5.²⁴ The model fit looks reasonable.

Table 5: Observed and Fitted Health Status Transition Matrices for Male High School Graduates

| $i \setminus j$ | Observed Probabilities | | | | Fitted Probabilities | | | |
|----------------------------------|----------------------------------|-------|--------|-------|----------------------|-------|--------|-------|
| | At the First Age Quartile (= 64) | | | | | | | |
| | “very good” | good | “fair” | dead | “very good” | good | “fair” | dead |
| “very good” | 70.8% | 23.0% | 3.8% | 2.4% | 71.2% | 22.0% | 5.4% | 1.6% |
| good | 24.2% | 56.1% | 17.5% | 2.2% | 26.8% | 53.1% | 16.6% | 3.3% |
| “fair” | 10.5% | 29.5% | 52.1% | 7.9% | 8.9% | 26.6% | 57.3% | 6.8% |
| At the Median Age (= 69) | | | | | | | | |
| “very good” | 64.5% | 25.5% | 9.1% | 1.0% | 68.0% | 23.9% | 6.1% | 2.4% |
| good | 26.1% | 54.2% | 15.8% | 3.9% | 23.2% | 53.6% | 18.4% | 4.8% |
| “fair” | 5.6% | 21.7% | 63.1% | 9.6% | 7.5% | 23.6% | 58.9% | 10.1% |
| At the Third Age Quartile (= 76) | | | | | | | | |
| “very good” | 63.1% | 25.8% | 6.6% | 4.5% | 60.6% | 28.0% | 7.9% | 4.4% |
| good | 20.4% | 48.5% | 22.3% | 8.8% | 19.9% | 50.1% | 22.0% | 8.8% |
| “fair” | 5.1% | 20.9% | 57.3% | 16.6% | 6.3% | 20.6% | 56.2% | 17.6% |

4.2 Taxes

FICA is a federal payroll tax imposed on workers. It has two components: Social Security tax and Medicare tax. During the period 1990-2010, the Social Security tax rate was 6.2 percent of an employee’s wages up to a threshold of earnings known as the Social Security Wage Base²⁵, and the Medicare tax rate was 1.45 percent of an employee’s wages without any cap. I use these values to set τ_1 .

²²For space considerations, I do not provide the estimates for high school dropouts and university graduates here.

²³HRS only provides mortality information up to the year 2006 currently. Therefore, I only use the information on health transition observations between years 2002 – 2004 and 2004 – 2006 to obtain these estimates. The implied biannual transition rates from the model are utilized to get the maximum likelihood estimates.

²⁴It corresponds to the ages 64, 69 and 76, respectively.

²⁵In the time period under study, Social Security Wage Base increased from \$84,900 to \$102,000. For simplicity, I fix the Social Security Wage Base at the year 2006 value, \$94,200, in my analysis.

The second portion of the tax structure, τ_2 , includes federal and state tax rates. I take the federal tax rates from the 2006 annual tax rate schedules accounting for standard deductions by age and personal exemptions which phase out after an income threshold. As state taxes, I use the 2006 Rhode Island tax rate schedule following French and Jones (2011).

The current regulation for federal income taxation of SS benefits is determined by The Deficit Reduction Act of 1993. For a single elderly individual, up to 50% of the SS benefits are subject to taxation if his combined income (the sum of adjusted gross income plus nontaxable interest plus one-half of SS benefits) is between \$25,000 and \$34,000. If his combined income is more than \$34,000, up to 85 percent of his SS benefits are taxable. I generate the precise taxable income using IRS Publication Number 915 to set τ_3 . In doing this I omit nontaxable interest since I do not have a measure of it.

4.3 Social Security Benefit Levels

Social Security benefit levels are calculated using Average Indexed Monthly Earnings (AIME), which is the average of 35 highest indexed earnings years.²⁶ Then, a formula is applied on AIME to compute Primary Insurance Amount (PIA) which gives the basis for SS benefit level.

I obtain the AIME levels for 67.5 percent of respondents exploiting their work history from the restricted data set using 2006 as index year. I observe the SS benefit amount of another 21.7 percent of the sample even though I cannot see their full work history. I generate AIME values for this subsample through an inverse function of the benefit levels.²⁷ I impute the AIME values for the rest of the sample. PIA is given by 90 percent of the first \$656 of AIME plus 32 percent of AIME over \$656 and through \$3,955, plus 15 percent of AIME over \$3,955.

I assume that AIME values are constant, so working another year does not have any effect on that value. For people having at least 35 years of work history, the incremental increase in AIME level is either zero (if the earnings in the extra year does not exceed 35th highest earning year) or close to zero. Moreover, at least 10 years of working history is required to be entitled to SS benefits. These two groups constitute 75 percent of the sample.

²⁶For AIME calculation, earnings levels in any year cannot exceed the maximum taxable earnings level of that year determined by the SSA. The index used for AIME is called the “national average wage index”.

²⁷In doing so, I increase SS benefit amount of early retirees by 25 percent which is equivalent to assuming that they retired 36 months earlier than their full retirement age. I index the benefit amounts according to the 2006 level. I also consider Medicare premiums deducted from SS benefit check while calculating AIME levels.

5 Results

5.1 Solution Methodology

I employ the simulated method of moments strategy where I match the following moments:

- By age, participation rate for the age group 60 – 85 and average hours worked for participants for the age group 60 – 72 to identify $\theta_{C,i}$, $\theta_{L,i}$ for each health status i , $\theta_{P,A}$, γ and v .
- For each health status, average of participation rates between ages 66 and 74 to identify $\theta_{P,f}$, $\theta_{P,good}$ and $\theta_{P,fair}$.
- By age, mean wage for the age group 60 – 75 to identify ς_0 , ς_1 and ς_2 .
- For each education level, average of mean wages of ages from 61 to 70 to identify δ_{high} and δ_{uni} .
- For three PIA intervals, average of mean wages between ages 62 and 67 to identify δ_{PIA} .
- Covariance of wages between ages 65 and 67 for participants to identify ρ_{AR} .
- Average of standard deviation of wages between ages 62 and 67 to identify σ_{η}^2 .
- By health status, average of mean out-of-pocket health expenses between ages 67 – 70 and 71 – 74. This helps me identify γ_0 , γ_1 , γ_2 , δ_{good} and δ_{fair} .
- Average out-of-pocket health expenses between ages 61 – 64 and 67 – 70 to identify $\delta_{medicare}$.²⁸
- Average out-of-pocket health expenses between ages 67 – 75 for people with assets levels 0 – 40,000, 40,000 – 200,000 and 200,000 – 1,000,000 to identify δ_{assets} .
- Average of standard deviation of out-of-pocket health expenses between ages 62 and 67 to identify σ_{ξ}^2 .

I assume that at the terminal age agents are non-participants and consume all of their assets. In solving the model, I calculate the expectations of value and bequest functions using the Gauss-Hermite quadratures of order 5 to account for the wage and health expenses shocks. The next step

²⁸The out-of pocket health expenses is highest as soon as people become eligible for Medicare and decreases later on. That can be seen as a temporary effect of being entitled to Medicare which is not carried through older ages. This is why I do not have any health expenses moment including ages 65 – 66.

is to randomly draw 1,000 observations from the data using the Mersenne Twister random number generator and simulate their behavior with interpolation/extrapolation. Subsequently, the distance between the simulated and the sample moments are computed. In doing this, I use the the inverse of the variance covariance matrix of the data moments as the weight matrix. This methodology provides consistent estimates. The variance covariance matrix of data moments is estimated via bootstrap using 1,000 replications. This process is repeated with different parameter vector choices using the Nelder-Mead algorithm. The solution is given by the parameters minimizing the distance between the simulated and the true data moments. The computational details can be found in Appendix C.

5.2 Parameter Estimates

The estimates are provided in Table 6. While consumption share parameters are positively associated with health status, the leisure share parameters are negatively correlated except good health. The participation cost increases as health deteriorates. The wage shows a decreasing pattern after age 49. Compared to people having no high school diploma, high school graduates earn 28 percent more while college graduates earn 45 percent more. The part of wages unexplained by the observables shows 85 percent persistency over a year.

Given the same asset level and age, the elderly with good health pay 14 percent more out-of-pocket health expenses than ones with “very good” health, whereas the elderly with “fair” health pay 39 percent more on average. Having Medicare decreases out-of-pocket health expenses by 5 percent. Given the same age level and health status, an increase of \$100,000 in asset levels are associated with a 8 percent increase in out-of health expenses on average.

The curvature estimate implies that the elderly can have unsecured debt up to \$5,320²⁹, which can be thought of maxing out credit cards rather than borrowing against SS benefits. Figure 2 provides the participation cost due to age.

²⁹The elderly can have more debts as long as they have corresponding assets for these debts like mortgage. The bequest function implies having an asset level less than $-\$5,320$ produces infinite disutility. The data suggest that some elderly do borrow small amount of money. 3.7 percent of the elderly have negative assets levels in the data whereas only 1.8 percent have asset levels less than $-\$5,320$.

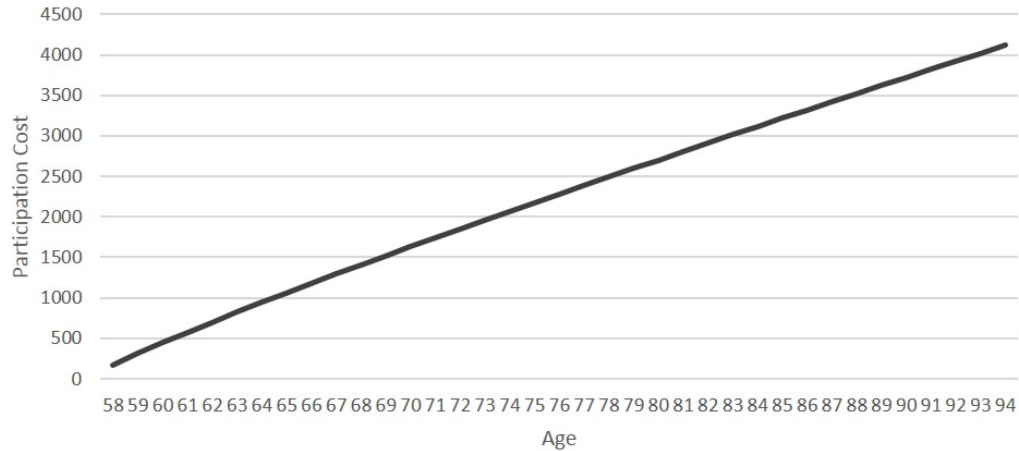
Table 6: The Estimates of the Preference Parameters

| Parameter | Explanation | Coef. | Parameter | Explanation | Coef. |
|------------------------------------|--|---------|--|--------------------------------|---------|
| <i>Flow Utility Parameters</i> | | | <i>Wage Equation Parameters</i> | | |
| $\theta_{C,verygood}$ | Cons. weight, "very good" health | 0.588 | ς_0 | Constant | 0.690 |
| $\theta_{C,good}$ | Cons. weight, good health | 0.530 | ς_1 | Age | 0.069 |
| $\theta_{C,fair}$ | Cons. weight, "fair" health | 0.495 | ς_2 | Age squared/100 | -0.070 |
| $\theta_{L,verygood}$ | Leisure weight, "very good" health | 0.421 | $\delta_{high\ school}$ | High school wage premium | 0.276 |
| $\theta_{L,good}$ | Leisure weight, good health | 0.566 | $\delta_{university}$ | University wage premium | 0.453 |
| $\theta_{L,fair}$ | Leisure weight, "fair" health | 0.530 | δ_{PIA} | PIA/100 (proxy for experience) | 0.005 |
| θ_{Pf} | Fixed cost of work (hours worked) | 102.3 | ρ_{AR} | AR term | 0.853 |
| $\theta_{P,good}$ | Add. part. cost - good health | 5.1 | σ_{η}^2 | Variance of the error | 0.045 |
| $\theta_{P,fair}$ | Add. part. cost -"fair" health | 761.2 | <i>Health Expenses Equation Parameters</i> | | |
| θ_{PA} | Participation Cost due Age - Shifter | 168.8 | γ_0 | Constant | 5.855 |
| γ | Participation Cost due Age - Convexity | 0.886 | γ_1 | Age | 0.0011 |
| v | Relative risk aversion | 4.850 | γ_2 | Age squared/100 | -0.0018 |
| <i>Bequest Function Parameters</i> | | | δ_{good} | Premium for good health | 0.135 |
| θ_B | Bequest Shifter | 0.00005 | δ_{fair} | Premium for "fair" health | 0.386 |
| K | Curvature | 5,320 | $\delta_{medicare}$ | Premium for Medicare | -0.051 |
| | | | δ_{assets} | Premium for Assets (\$100,000) | 0.078 |
| β | Intertemporal discount factor | 0.882 | σ_{ξ}^2 | Variance of the error | 0.179 |

Notes: No high school diploma is the reference category for wage premium parameters.

"Very good" health is the reference category for health expenses premium coefficients.

Figure 2: Participation Cost Explained by Age



5.3 Model Fit

Figures 3, 4 and 5 provides the model fit of participation rate, mean hours worked and mean wages for participants, respectively. Table 7 provides the model fit of the average of mean wages between ages 61 and 70 by education group. Table 8 shows the model fit of the average of mean wages between ages 62 and 67 by three PIA intervals. Table 9 provides the model fit of the average of participation rates between ages 66 and 74 by health status. Table 10 provides the model fit of the

average of mean health expenses between ages 67 – 70 and 71 – 74 by health status. Table 11 shows model fit of the average health expenses between ages 67 – 75 by asset levels. Table 12 provides the model fit of the remaining moments. The model fits the data well with reasonable estimates.

Figure 3: Model Fit - Participation Rate by Age

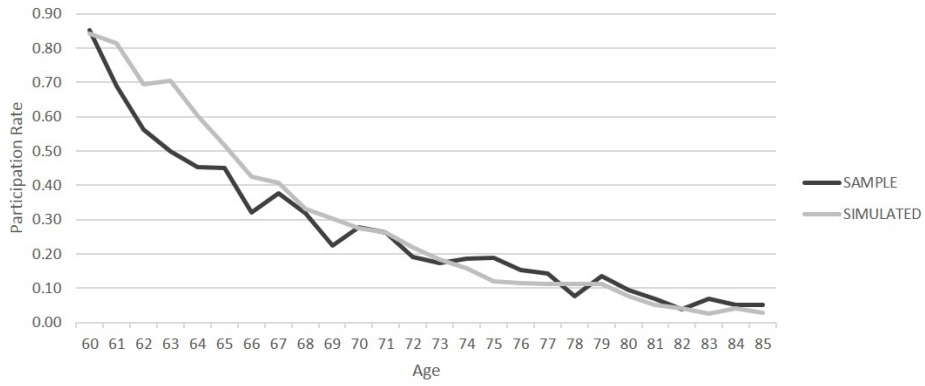


Figure 4: Model Fit - Mean Hours Worked for Participants by Age

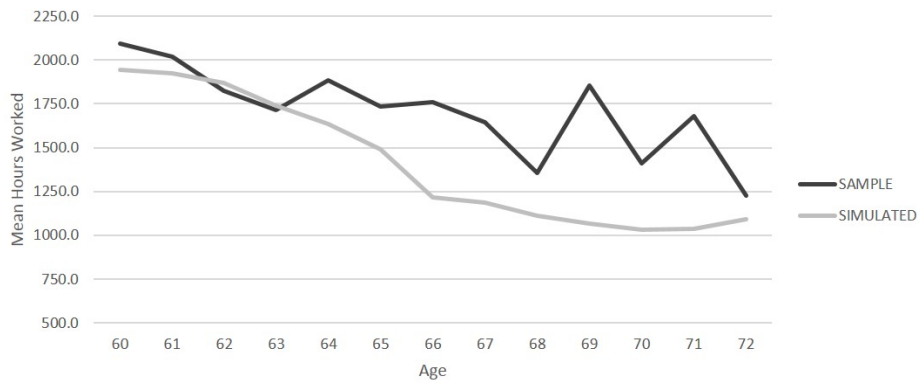


Figure 5: Model Fit - Mean Wages for Participants by Age

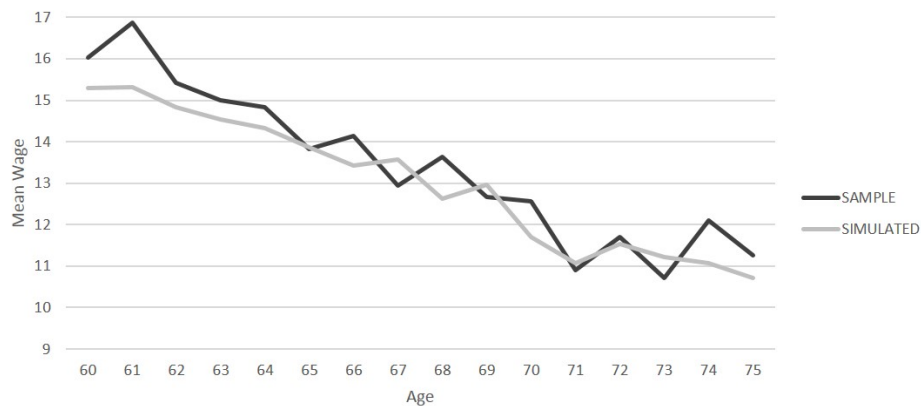


Table 7: Model Fit - Average of Mean Wages of Each Age Between 61 – 70 by Education

| Education Status | Data | Simulation |
|------------------------|-------|------------|
| No High School Diploma | 10.72 | 10.24 |
| High School Graduates | 13.86 | 14.33 |
| University Graduates | 17.28 | 16.73 |

Table 8: Model Fit - Average of Mean Wages Between Ages 62 and 67 by PIA

| PIA Level | Data | Simulation |
|---------------------|-------|------------|
| PIA < 1,000 | 12.04 | 12.21 |
| 1,000 < PIA < 1,500 | 14.59 | 13.75 |
| PIA > 1,500 | 15.58 | 15.62 |

Table 9: Model Fit- Average of Participation Rates Between Ages 66 and 74 by Health Status

| Health Status | Data | Simulation |
|---------------|-------|------------|
| “Very Good” | 0.345 | 0.328 |
| “Good” | 0.264 | 0.318 |
| “Fair” | 0.178 | 0.176 |

Table 10: Average of Mean Health Expenses Between Ages 67 – 70 and 71 – 74 by Health Status

| Health Status | Ages 67 – 70 | | Ages 71 – 74 | |
|---------------|--------------|------------|--------------|------------|
| | Data | Simulation | Data | Simulation |
| “Very Good” | 560 | 567 | 604 | 550 |
| “Good” | 625 | 566 | 640 | 663 |
| “Fair” | 662 | 626 | 672 | 683 |

Table 11: Average Health Expenses Between Ages 67 – 75 by Assets

| Assets | Data | Simulation |
|---------------------|------|------------|
| 0 – 40,000 | 452 | 487 |
| 40,000 – 200,000 | 639 | 495 |
| 200,000 – 1,000,000 | 730 | 608 |

Table 12: Model Fit - Rest

| | Data | Simulation |
|---|-------|------------|
| Covariance of Wages Between Ages 65 and 67 (For Participants in Both Periods) | 16.62 | 19.90 |
| Average of Standard Deviation of Wages Between Ages 62 and 67 | 5.02 | 5.43 |
| Average of Health Expenses Between Ages 61 – 64 | 707 | 676 |
| Average of Health Expenses Between Ages 67 – 70 | 614 | 583 |
| Average of Standard Deviation of Health Expenses Between Ages 62 and 67 | 1,043 | 976 |

6 Counterfactuals

6.1 Changing Social Security Benefit Amounts

In my first counterfactual, I decrease SS benefit amounts by 20 percent. This is mainly an income effect for the elderly with a small substitution effect arising from a possible change in the decision to start collecting retirement benefits. The participation decision is sensitive to Social Security benefits as seen in Figure 6. 20 percent decrease in SS benefits is associated with a 37 percent increase in LFPR of the age group 66 – 75. However, there is not a significant response in the intensive margin as presented in Figure 7.

Figure 6: Participation Rates Under 20% Decreased SS Benefit Level

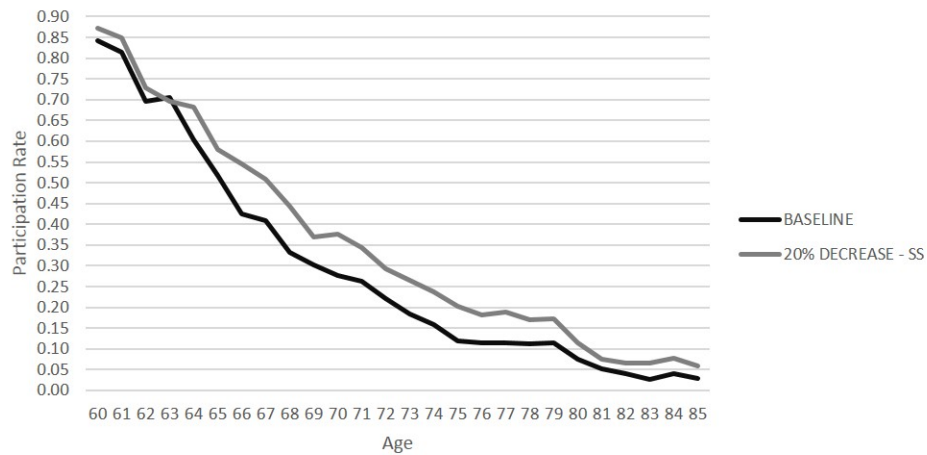
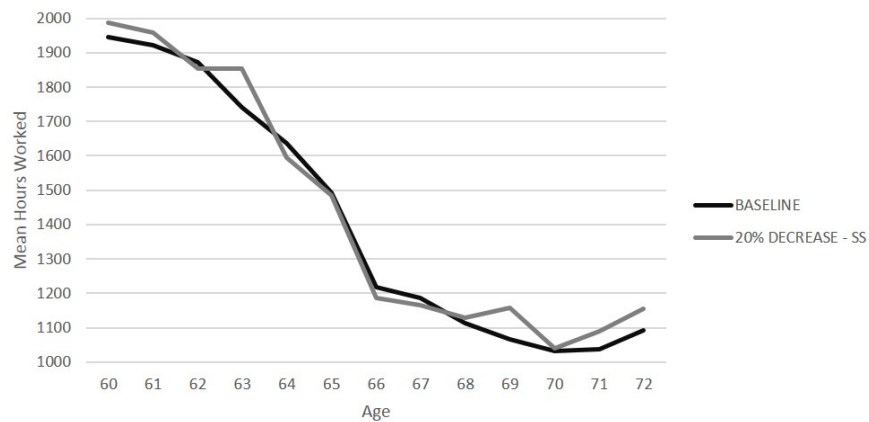


Figure 7: Mean Hours Worked Under Decreased SS Benefit Levels



6.2 Changing FICA tax amounts

In my second counterfactual analysis, I decrease the employee portion of FICA tax amounts by 50 for everyone starting at the age 58, the initial age in my dynamic programming set-up. This can be thought as a 3.825 percent increase in wages as well. This kind of analysis will have both income and substitution effects on the elderly. Figure 8 shows that such a policy change affects the extensive margin mainly beyond normal retirement age. The corresponding increase in the LFPR for people aged 66 – 70 is 6.6 percent, which corresponds to a labor supply elasticity of 1.73.

Figure 8: Participation Rates Under Decreased FICA Amounts for Everyone

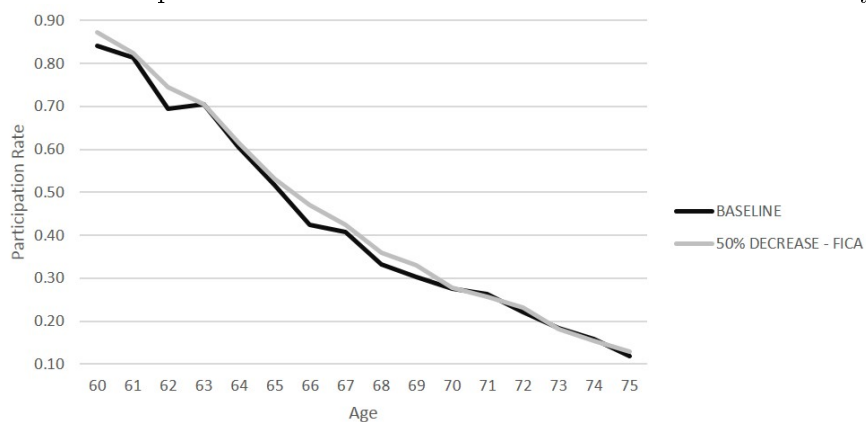
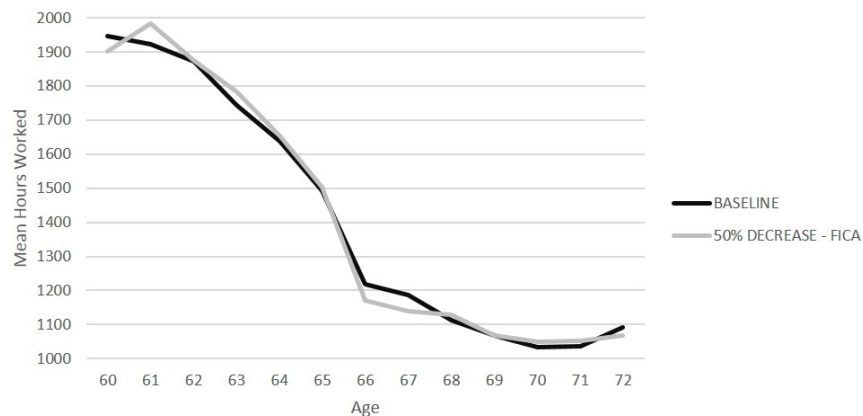


Figure 9 provides the labor supply response to decreased FICA levels on the intensive margin. The effects are not substantial. The elderly increase their hours supplied by an average of 25 hours on average between ages 61 – 65, but decrease it by an average of 50 hours at ages 66 and 67.

Figure 9: Mean Hours Worked Under Decreased FICA Amounts for Everyone



If FICA taxes are reduced only for people aged 70+, the response in the extensive margin is observed mainly between ages 70 and 73. The corresponding increase in LFPR is 3.8 percent on these age groups, which corresponds to unit labor supply elasticity. The effect at the intensive margin is not substantial.

Figure 10: Participation Rates Under Decreased FICA Amounts for People Aged 70+

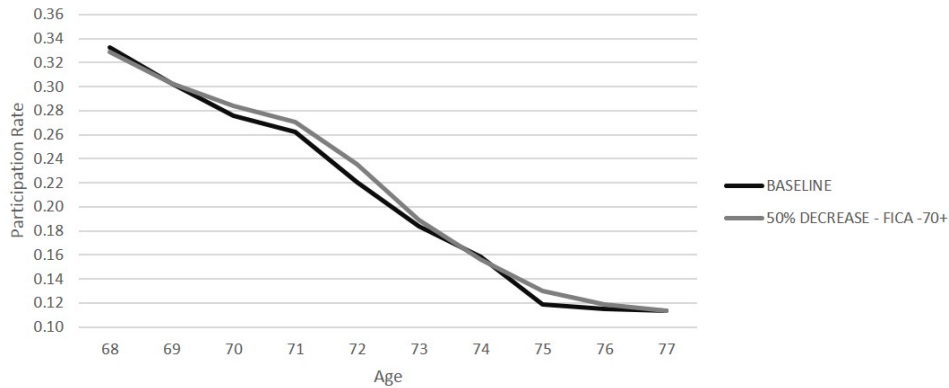
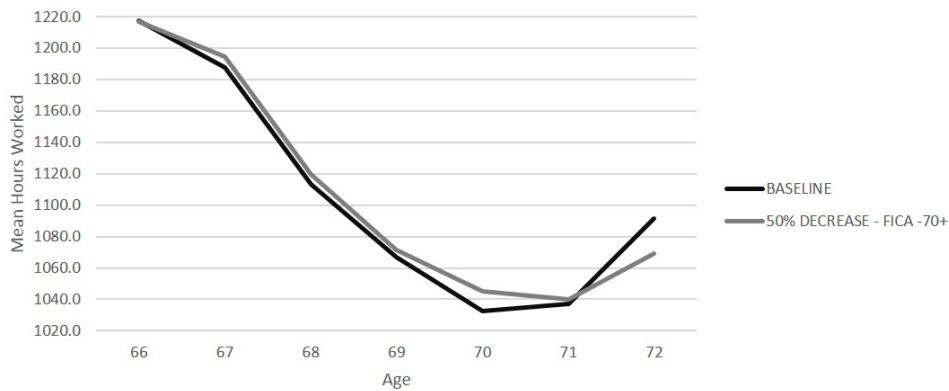


Figure 11: Mean Hours Worked Under Decreased FICA Amounts for People Aged 70+



6.3 The Effect of Year 2000 Social Security Amendments

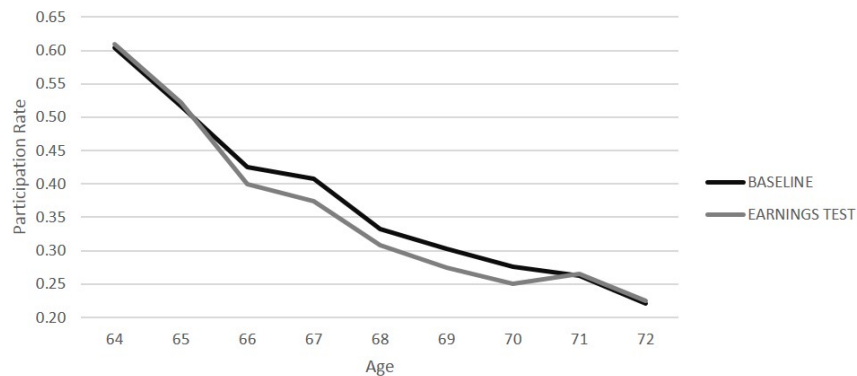
“Earning test” is a program deferring part (or all) of SS benefits of people whose earnings exceed a threshold level to later years by indexing the withheld amount with the delayed retirement credit. Until year 2000, it applied to the elderly until the age 70, and it currently applies only on the elderly who start collecting their SS benefits before normal retirement age. The annual delayed retirement credit was 3.0 percent in 1989 and was raised by 0.5 percentage point every two years since then until 2008. That corresponded to 5.5 percent delayed retirement credit right before the year 2000

SS amendment, which was actuarially unfair. It is 8 percent now and can be considered actuarially fair.³⁰ “Earnings test” withholds \$1 in benefits for every \$2 of earnings in excess of the lower exempt amount, and \$1 in benefits for every \$3 of earnings in excess of the higher exempt amount. The lower and higher exempt amount are determined by the Social Security Administration.

The time period studied in the paper is 2002 – 2008, right after the abolishment of the earnings test. It is possible to see the behavioral effects of the year 2000 SS amendment by applying the pre-2000 rules on my sample. I set the delayed retirement credit to 4.5 percent and use the 2006 values of lower and higher exempt amounts rather than 2000 values in this analysis.

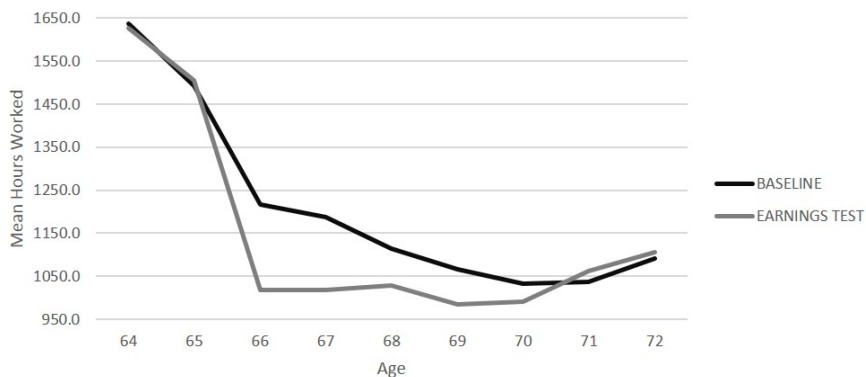
Figure 12 shows that LFPR of the elderly aged 66 – 70 decreases by 2.7 percentage points with the “earnings test”. The effect on the intensive margin is substantial and is shown in Figure 13. Notice mean hours worked decreases by 115 hours in the same age group. The mean earnings of participants at age 66 with the introduction of earnings test, \$13,680, gets very close to the lower exempt amount of earnings test , \$12,480. It can be concluded that the “earnings test” provided a significant labor supply disincentive.

Figure 12: The Effect of Earnings Test-Extensive Margin



³⁰ Assume that the yearly retirement benefits of a SS beneficiary is equal to \$10,000. The CDC report in 2009 indicates that the life expectancy at age 65 was around 19 years. Since the SS makes the yearly cost-of-living adjustment on the retirement benefits, we assume that the real value of the benefits stays the same. The current inflation rate is around 2 percent. If this beneficiary delays getting retirement for a year, he/she gets 10,800/1.02 in today’s value for 18 years on average, and if he does not delay the retirement, he/she gets \$10,000 for 19 years on average. Observe that $(10,800/1.02) * 18 \simeq 10,000 * 19$.

Figure 13: The Effect of The Earnings Test - Intensive Margin



7 Conclusion

This paper analyzes the joint determination of labor supply, consumption (savings) and the decision to apply for Social Security (SS) benefits of the elderly single males using a dynamic programming formulation and restricted data from the Health and Retirement Study (HRS). I first conduct a preliminary multinomial logit analysis to see the correlations in the data, then formulate a dynamic programming model enhancing the understanding of the elderly labor force decision. In doing so, I focus on the participation decision rather than retirement since a significant portion of the elderly return to work after being non-participants for a while.

The specification of my model is flexible in terms of capturing most of the documented determinants of the elderly non-participation decision in the literature. Using counter-factual analyses, I find that decreasing SS benefits by 20 percent increases the participation rate of the elderly single males aged 66 – 75 by 37 percent whereas decreasing FICA taxes by 50 percent causes the participation rate to increase by 6.6 percent for the age group 66 – 70. The response in the intensive margin for either case is not substantial. It is essential to understand the incentives provided by the SS system on the elderly labor supply decision since the U.S. population is steadily aging and the fiscal burden of SS is a big concern. These results suggest that the policy recommendations arising within the public debate to change the SS rules might have a marked effect on the participation decision of people beyond normal retirement age.

I further apply “earnings test” , which was abolished by the year 2000 SS amendment, on my sample via a counter-factual analysis to quantify the effect of the year 2000 SS amendment on the

recent increase in the elderly participation rates. I find that the abolishment of the “earnings test” increased the participation rate of the elderly single males aged 66 – 70 by 2.7 percentage points on the extensive margin and mean hours worked by 115 hours on the intensive margin. The effect on the extensive margin explains one-fourth of the recent increase in the elderly participation rates. Moreover, the decrease in the intensive margin brings mean earnings level very close to the lower exempt amount of “earnings test”. This finding suggests that prior to the year 2000 SS amendments, the elderly limited their hours supplied to avoid the implicit taxation imposed by the “earnings test” through an unfair delayed retirement credit. For a more thorough understanding, I plan to include part-time wage penalty, Supplemental Security Income and pension benefits in my model in the future.

References

- BLAU, D. M. AND D. B. GILLESKIE (2008): “The Role Of Retiree Health Insurance In The Employment Behavior Of Older Men,” *International Economic Review*, 49, 475–514.
- BLAU, D. M. AND R. M. GOODSTEIN (2010): “Can Social Security Explain Trends in Labor Force Participation of Older Men in the United States?” *Journal of Human Resources*, 45, 328–363.
- CASANOVA, M. (2010): “Happy Together: A Structural Model of Couples’ Joint Retirement Choices,” Working paper.
- DE NARDI, M. (2004): “Wealth Inequality and Intergenerational Links,” *The Review of Economic Studies*, 71, 743–768.
- FRENCH, E. (2005): “The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior,” *Review of Economic Studies*, 72, 395–427.
- FRENCH, E. AND J. B. JONES (2011): “The effects of Health Insurance and Self-Insurance on Retirement Behavior,” *Econometrica*, 79, 693–732.
- LAZEAR, E. P. (1979): “Why Is There Mandatory Retirement?” *The Journal of Political Economy*, 87, 1261–1284.
- LEE, L. F. (1982): “Some Approaches to the Correction of Selectivity Bias,” *Review of Economic Studies*, 49, 355–372.
- MAESTAS, N. (2010): “Back to Work: Expectations and Realizations of Work after Retirement,” *Journal of Human Resources*, 45, 718–748.
- MCFADDEN, D. L. (1974): “Conditional Logit Analysis of Qualitative Choice Behavior,” in *Frontiers in Econometrics*, ed. by P. Zarembka, Academic Press, New York, 105–142.
- ROBINSON, J. (2002): “A Long-Term-Care Status Transition Model,” Unpublished Paper.
- RUST, J. AND C. PHELAN (1997): “How Social Security and Medicare Affect Retirement Behavior In a World of Incomplete Markets,” *Econometrica*, 65, 781–831.

TUNALI, I. AND B. YAVUZOGLU (2012): “Edgeworth Expansion Based Correction of Selectivity Bias in Models of Double Selection,” Working Paper.

APPENDIX

A Data

HRS includes some confirmation questions for the health insurance section. While generating the health insurance data, I exploit these confirmation questions. I also use the tracker file released by HRS which accounts for misspecified cases of age and marital status. I define marital status as a dummy variable where the non-married class is composed of separated, divorced, widowed, never married and other categories. Health expenses are obtained by summing up out of pocket expenses for hospital, nursing home, outpatient surgery, doctor visit, dental, prescription drugs, in-home health care and special facility and other health service costs in the last 2 years. I exploit HRS Core Income and Wealth Imputations data for the missing asset values, which is consistent with the HRS and provided by the RAND Corporation. There are 5 different health status categories; excellent, very good, good, fair and bad. These are the self-reported measures of health. I use a dummy variable for each category. I also have dummy variables for blacks, Social Security benefit recipients and Medicare. Number of other health insurance includes private insurance, employment insurance and government insurance other than medicare. In defining labor force participation status, I first impute the hours worked and the weeks worked observations for 1.02 percent of the workers who report at least one of the hours worked or weeks worked. Then, using hours and weeks worked information, I assign workers as full-time employed if they work more than 1,600 hours and part-time employed otherwise. I further assign people who are listed as temporarily laid off with blank usual hours and weeks worked observations as non-participant.

Now, I explain the steps I used to obtain the working sample for the year 2006 from HRS data. The same procedure is followed for any other year. I use the respondent sample of HRS. Originally, it has 18,469 observations. I drop 1,703 disabled people as well as 62 observations who report their labor force status other than employed, unemployed and out of the labor force, 5 observations who do not know their labor force status, 2 observations who refuse to report it and 10 observations

who take a partial interview where this question is skipped. I drop 30 respondents who work in one job and refuse to report or do not know both how many hours in a week and weeks in a year he/she works as well as 5 respondents working in two jobs who do not know or refuse to report either hours worked or weeks worked in each job. The removal of these 35 respondents does not induce an important bias since they correspond to the 0.27 percent of the final sample. When I limit ages to 58 and above, I lose 3,283 observations. I exclude 17 respondents who do not know about their health status as well a respondent who does not know his marital status. I exclude 10 respondents who do not know if they are receiving Social Security, 10 respondents who refused to answer this question and a respondent with blank Social Security information. I exclude 15 more people who do not know if they are covered by Medicare and another respondent with blank Medicare information. I drop 6 observations with gender inconsistencies from the data. I drop 100 observations with blank years of education. I also drop 64 observations who do not know if they get medicare, 2 observations who refuses to answer this question and 2 respondents with blank Medicaid information. Moreover, I drop 16 observations who do not know if they get Champus, Champ-Va, Tri-Care or any other military health plan and 2 respondents who refuse to answer this question. Finally, I drop 38 respondents who do not know the number of private health insurance they have and 7 respondents who refuses to answer this question. Finally, I drop 34 outliers with asset levels more than \$20 million and 3 outliers with hourly wage levels lower than \$2 or higher than \$100.³¹ In the end, I am left with a sample size of 13,040 for the year 2006.

B Imputation of Wages

Wage estimates are obtained for each cross section separately using the solution provided by Tunali and Yavuzoglu (2012) for double selection problems, which relaxes the trivariate normality assumption among the disturbances of the two selection equations and the regression equation by following the Edgeworth expansion approach of Lee (1982). Instead, Tunali and Yavuzoglu (2012) do not impose any condition on the form of the distribution of the random disturbance in the regression (partially observed outcome) equation, but conveniently assume bivariate normality between the

³¹If these outliers have asset levels less than \$20 million or hourly wage levels between \$2 and \$100 in the other years, I keep them in the data for those years. In solving the model, I exploit the data to get the first stage estimates and to get the initial values in the solution of the DP model.

random disturbances of the two selection equations.

$$\text{Home - work utility : } U_0^* = \theta'_0 z + \nu_0, \quad (17)$$

$$\text{Part - time work utility : } U_1^* = \theta'_1 z + \nu_1, \quad (18)$$

$$\text{Full - time work utility : } U_2^* = \theta'_2 z + \nu_2. \quad (19)$$

Assume that home-work (or non-participation), part-time employment and full-time employment utilities can be expressed as follows where z is a vector of observed variables, θ_j 's are the corresponding vectors of unknown coefficients and ν_j 's are the random disturbances. Assuming that individuals choose the state with highest utility, their decisions can be captured using the utility differences:

$$y_1^* = U_1^* - U_0^* = (\theta'_1 - \theta'_0)z + (\nu_1 - \nu_0) = \beta'_1 z + \sigma_1 u_1, \quad (20)$$

$$y_2^* = U_2^* - U_1^* = (\theta'_2 - \theta'_1)z + (\nu_2 - \nu_1) = \beta'_2 z + \sigma_2 u_2. \quad (21)$$

Note that y_1^* can be expressed as the propensity to be part-time employed rather than being a non-participant and y_2^* as the incremental propensity to engage in full-time employment rather than part-time employment. Then, $y_1^* + y_2^*$ gives the propensity to engage in full-time employment over home-work. The three way classification observed in the sample arises as follows:

$$lfp = \left\{ \begin{array}{l} 1 = \text{full-time employment, if } y_2^* > 0 \text{ and } y_1^* + y_2^* > 0, \\ 2 = \text{part-time employment, if } y_1^* > 0 \text{ and } y_2^* < 0, \\ 3 = \text{home-work, if } y_1^* < 0 \text{ and } y_1^* + y_2^* < 0, \end{array} \right\} \quad (22)$$

In this case the support of (y_1^*, y_2^*) is broken down into three mutually exclusive regions, which respectively correspond to $lfp = 1, 2,$ and 3 . The classification in the sample is obtained via a pair from the triplet $\{y_1^*, y_2^*, y_1^* + y_2^*\}$. Normalizing the variances of y_1^* and $y_1^* + y_2^*$ to 1 has an implication for the variance of y_2^* ($\sigma_2^2 = -2\rho_{12}$ where ρ_{12} is the correlation between u_1 and u_2). This is why I may apply the normalization to one of $\sigma_{11} = \sigma_1^2$ and $\sigma_{22} = \sigma_2^2$, but must leave the other variance free to take on any positive value. In the analysis, I take $\sigma_{11} = 1$ and let σ_{22} be free.

In the first step, I rely on maximum likelihood estimation and obtain consistent estimates of β_1 , β_2 , ρ_{12} and σ_2 subject to $\sigma_1 = 1$. The likelihood function is given by

$$L = \prod_{lfp=1} P_1 \prod_{lfp=2} P_2 \prod_{lfp=3} P_3, \quad (23)$$

where $P_j = Pr(lfp = j)$ for $j = 1, 2, 3$.

The regression equation for this problem is a Mincer-type wage equation given below where X_3 is the set of explanatory variables³²:

$$\log(wage) = \beta_3' X_3 + \sigma_3 u_3. \quad (24)$$

The aim is to estimate β_3 for $lfp = 1, 2$. After forming the estimates of selectivity correction terms via first step estimates, I run a linear regression equation with 9 selectivity correction terms coming from the Edgeworth expansion in the second step. Note that robust correction obtained via Edgeworth expansion nests the conventional trivariate normality correction, and therefore both the conventional trivariate normality specification and the presence of the selectivity bias can be tested via this estimation. Details can be found in Tunali and Yavuzoglu (2012).

I present only the 2006 cross section results here to demonstrate the employed methodology. Table A.3 provides the results of the first step. Very low ρ_{12} value implies that unobserved characteristics affecting the decision of part-time employment over non-participation do not affect the decision of full-time employment compared to part-time employment. As expected, females have a lower participation probability. Blacks are more likely to be part-time employed compared to being a non-participant. This may be caused by the low asset levels of blacks. Along with the line of my expectations, participation profile is concave with respect to age.

Moreover, as years of education increases, people are more likely to work part-time rather than being out of the labor force or working full-time. Since people want to realize some return on their educational investments, they are more likely to be a participant. However, these people should have enough savings making them unlikely to work full-time. Receiving Social Security benefits decreases the full-time employment and part-time employment probabilities. Having Medicare decreases the

³²I use the explanatory variables provided in Tables 3 and ?? except race and number of children. See discussion in Section 4.

participation probability which is reasonable since one of the main concerns in the labor force participation decision is health insurance as documented by Rust and Phelan (1997). I do not have a good explanation regarding why number of other types of health insurance increases full-time employment probability over part-time employment. With good health as the reference category, there is positive correlation between health and participation probability. Being married decreases the participation probability which might be caused by that the high joint savings levels of couples.³³

Using the estimates of the first step, I provide least squares estimates of the $\log(\text{wage})$ equation for full-time and part-time employed separately in Table A.4. $\hat{\lambda}$'s denote the selectivity correction terms. The presence of selectivity bias can be tested by looking at the joint significance of all the selectivity correction terms. For both full-time and part time employment, the evidence is in favor of the non-random selection ($p - \text{value} \simeq 0.000$ for both cases).

Conventional trivariate normality specification uses only $\hat{\lambda}_1$ and $\hat{\lambda}_2$. The test for the joint significance of the remaining $\hat{\lambda}$'s provides evidence in favor of the robust selectivity correction for both full-time ($p - \text{value} = 0.031$) and part-time employed ($p - \text{value} \simeq 0.000$).

The evidence is in favor of the non-random selection for both full-time and part-time employed in all the years under study at 5 percent level of significance. Moreover, the evidence is in favor of the robust selectivity correction for part-time employed in 2002 and 2004, and full-time employed in 2004 and 2008, but in favor of the conventional trivariate normality specification for full-time employed in 2002 and part-time employed in 2008 at 5 percent level of significance.

An interesting finding is that the wage gap between blacks and whites disappears for elderly workers. Wages are concave with respect to age and increases with education as expected. While being female decreases full-time wages by 11.3 percent, it does not affect the part-time wages. With good health as the reference category, it can be concluded that wages are positively correlated with health.

Using these estimates, predicted wage values for workers without an observed wage rate are obtained.

³³Note that the even though both couples contribute to the savings level, their expenses do not double due to the shared consumption like housing or utilities

C Computational Issues in Solving DP Model

Since the value function depends on 2 continuous variables, namely assets and wage, I use a weighted average of the 4 grid points closest to the intermediary point to interpolate the value function if the intermediary point is in the set covered by the grid points. Otherwise, I extrapolate the value function using the closest point on this set and assuming linearity.

If the respondent is non-participant in the previous period, I generate his current wage by ignoring the autoregressive part of the wage equation since I do not observe any wage in the previous period, but include uncertainty. This is reasonable since the autoregressive part of the wage is the portion unexplained by the observables.

I interpolate or extrapolate the consumption and labor force participation decisions using weighted average of the relevant points in the grid to generate moments for simulated individuals, as I did in solving the model. Then, I generate wages and assets using wage determination and asset accumulation equations in the simpler model where I include randomly generated normally distributed wage shocks.

D Tables

Table A.1: LFPRs of Different Age Groups along with Retirement Ages in Different Countries

| Country | Early Retirement Age | Normal Retirement Age | LFPR, 50-54 | LFPR, 55-59 | LFPR, 60-64 | LFPR, 65-69 | LFPR, 70-74 | LFPR, 75+ |
|-------------|----------------------|-----------------------|-------------|-------------|-------------|-------------|-------------|-----------|
| Austria | 62 (57) | 65 (60) | 81.2% | 55.2% | 15.8% | 7.1% | 3.0% | 1.3% |
| Belgium | 60 | 65 (64) | 71.3% | 44.8% | 16.0% | 4.5% | n/a | n/a |
| Denmark | 60 | 65 | 87.3% | 83.2% | 42.1% | 13.1% | n/a | n/a |
| Finland | 62 | 65 | 86.2% | 72.9% | 38.7% | 7.6% | 3.9% | n/a |
| France | none | 60 | 84.1% | 58.1% | 15.1% | 2.8% | 1.2% | 0.3% |
| Germany | 63 | 65 | 85.0% | 73.9% | 33.3% | 6.7% | 3.0% | 1.0% |
| Greece | 60 (55) | 62 (57) | 70.3% | 53.5% | 32.7% | 9.8% | n/a | n/a |
| Ireland | none | 65 | 73.9% | 62.7% | 44.8% | 17.2% | 7.8% | 3.4% |
| Italy | 57 | 65 (60) | 71.2% | 45.1% | 19.2% | 7.5% | 2.9% | 0.9% |
| Netherlands | none | 65 | 79.5% | 63.9% | 26.9% | 8.2% | n/a | n/a |
| Norway | none | 67 | 84.6% | 77.4% | 57.3% | 20.6% | 6.0% | n/a |
| Spain | 60 | 65 | 71.3% | 57.5% | 34.6% | 5.3% | 1.6% | 0.4% |
| Sweden | 61 | 65 | 88.0% | 83.0% | 62.5% | 13.2% | 6.8% | n/a |
| UK | none | 65 (60) | 82.6% | 71.2% | 44.3% | 16.3% | 6.0% | 1.6% |
| USA | 62 | 65.5 | 78.3% | 69.9% | 48.4% | 29.5% | 17.8% | 6.1% |

Notes: Parentheses indicate the eligibility age for women when different. Columns 2-3 are obtained from “Social Security Programs throughout the World: Europe, 2006” by U.S. Social Security Administration. Columns 4 – 9 are obtained from 2006 Health and Retirement Survey for the U.S. and 2006 OECD database for the rest of the countries. Note that the 2006 OECD database includes agricultural workers. Labor force participation rates of elderly people in countries with high agricultural production, like Ireland, can be naturally high since the definition of agricultural work is vague and scope of it is very broad. This further reinforces the discrepancy in elderly LFPRs among the U.S. and the developed European countries. Note that the LFPR for the age group 66 – 69 in the U.S. is 26.2 percent (accounting for the normal retirement age, 65.5 years, in 2006).

Table A.2: Female and Male Life Expectancy at Age 65 in Various Countries

| Country | Life Expectancy at Age 65 | |
|-------------|---------------------------|--------|
| | Male | Female |
| Austria | 82.3 | 85.7 |
| Belgium | 82 | 85.6 |
| Denmark | 81.2 | 84.2 |
| Finland | 82.0 | 86.3 |
| France | 83.2 | 87.7 |
| Germany | 82.2 | 85.5 |
| Greece | 82.5 | 84.4 |
| Ireland | 81.8 | 85.3 |
| Italy | 82.9 | 86.8 |
| Netherlands | 81.9 | 85.4 |
| Norway | 82.7 | 85.8 |
| Spain | 82.9 | 87.0 |
| Sweden | 82.7 | 85.9 |
| UK | 82.5 | 85.2 |
| USA | 82.0 | 84.7 |

Notes: The statistics are obtained from Centers for Disease Control and Prevention (CDC) for the U.S. and United Nations Economic Commission for Europe (UNECE) Statistical Database for the rest of the countries for the calendar year 2006.

Table A.3: Maximum Likelihood Estimates of Reduced Form Participation Equations (Normalized Version)

| Variable | First Selection | | Second Selection | |
|-----------------------------------|-----------------|----------------|------------------|-----------|
| | Coef. | Std. Err. | Coef. | Std. Err. |
| Black | 0.085* | 0.045 | -0.057 | 0.060 |
| Married | -0.097*** | 0.036 | -0.216 | 0.134 |
| Age | 0.085* | 0.048 | 0.161 | 0.116 |
| Age squared/100 | -0.098*** | 0.032 | -0.148 | 0.104 |
| Female | -0.270*** | 0.047 | -0.499 | 0.321 |
| Years of Education | 0.042*** | 0.006 | -0.024*** | 0.009 |
| Receive SS | -0.282*** | 0.109 | -0.958* | 0.551 |
| Medicare | -0.238*** | 0.055 | -0.071 | 0.118 |
| # of Other Health Insurance | -0.006 | 0.027 | 0.126* | 0.072 |
| Poor Health | -0.675*** | 0.084 | -0.321 | 0.344 |
| Fair Health | -0.200*** | 0.044 | -0.188 | 0.151 |
| Very Good Health | 0.137*** | 0.038 | -0.051 | 0.052 |
| Excellent Health | 0.253*** | 0.051 | 0.027 | 0.080 |
| Constant | -1.805 | 1.779 | -3.096 | 3.063 |
| σ_{11} | | 1 [normalized] | | |
| σ_{22} | | 0.896 (1.070) | | |
| ρ_{12} | | 0.049 (0.289) | | |
| No. of observations | | 13040 | | |
| Log-likelihood without covariates | | -13474.844 | | |
| Log-likelihood with covariates | | -7880.2924 | | |

Robust standard errors are reported.

* significant at 10%; ** significant at 5%; *** significant at 1%.

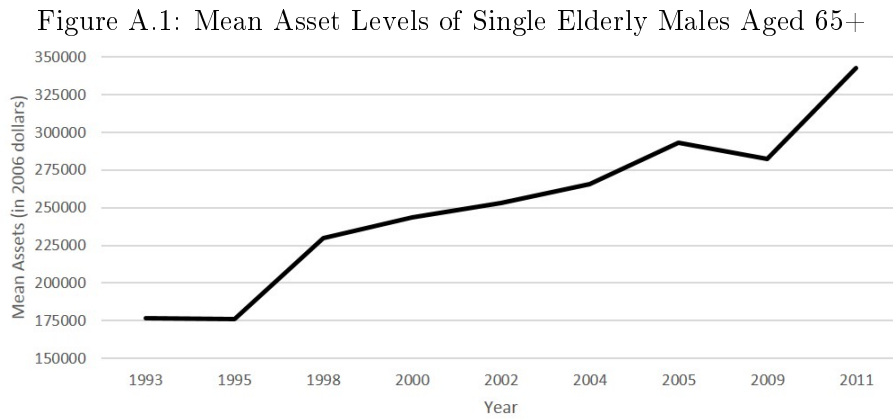
Table A.4: Least Squares Estimates of the Wage Equations

| Variable | Full-Time Employed | | Part-Time Employed | |
|---------------------|--------------------|-----------|--------------------|-----------|
| | Coef. | Std. Err. | Coef. | Std. Err. |
| Black | 0.036 | 0.048 | -0.019 | 0.052 |
| Age | 0.185** | 0.083 | 0.252** | 0.100 |
| Age squared/100 | -0.150** | 0.063 | -0.195*** | 0.070 |
| Female | -0.113*** | 0.043 | -0.068 | 0.055 |
| Years of Education | 0.093*** | 0.014 | 0.090*** | 0.014 |
| Poor Health | -0.621*** | 0.173 | -0.408* | 0.218 |
| Fair Health | -0.130** | 0.052 | -0.049 | 0.064 |
| Very Good Health | 0.180*** | 0.053 | 0.106** | 0.054 |
| Excellent Health | 0.187** | 0.081 | 0.254*** | 0.075 |
| $\hat{\lambda}_1$ | -164.019*** | 54.843 | 0.052 | 0.123 |
| $\hat{\lambda}_2$ | 135.448*** | 45.377 | -0.052 | 0.786 |
| $\hat{\lambda}_3$ | 13.081* | 6.441 | -0.033 | 0.091 |
| $\hat{\lambda}_4$ | -28.254** | 11.967 | -0.383 | 0.520 |
| $\hat{\lambda}_5$ | -6.821 | 5.309 | -0.027 | 0.469 |
| $\hat{\lambda}_6$ | 2.776** | 1.311 | 0.006 | 0.020 |
| $\hat{\lambda}_7$ | -8.949** | 3.556 | -0.446 | 0.353 |
| $\hat{\lambda}_8$ | 8.613** | 3.389 | -0.571*** | 0.199 |
| $\hat{\lambda}_9$ | 1.861** | 0.790 | 1.014* | 0.525 |
| Constant | -3.024 | 2.785 | -6.635** | 3.343 |
| No. of observations | | 788 | | 738 |
| R^2 | | 0.236 | | 0.199 |

Robust standard errors are reported.

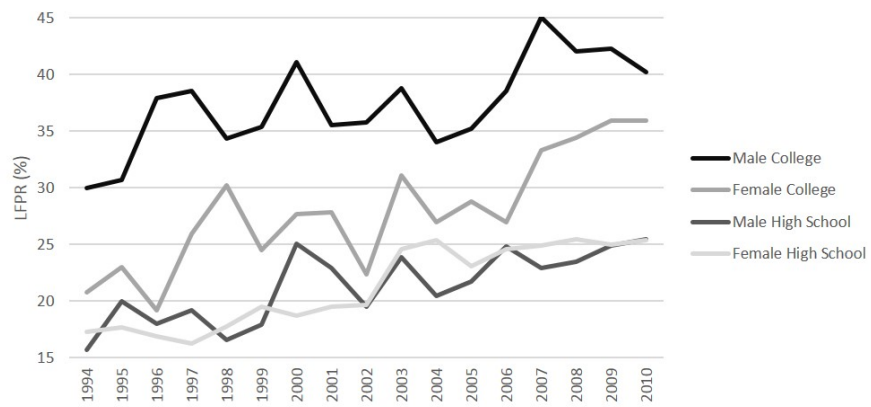
* significant at 10%; ** significant at 5%; *** significant at 1%.

E Figures



Source: Obtained using Wealth and Asset Ownership Data from U.S. Census Bureau.

Figure A.2: Trends in Labor Force Participation for Single Elderly Aged 65 – 74 by Gender and Education Level



Source: Obtained using March CPS Data