The Effect of Social Security
Auxiliary Spouse and Survivor Benefits
on the Household Retirement Decision

David M. K. Knapp
University of Michigan

Appendix

A Tax Rates

I assume households jointly file their tax return if both individuals are alive, otherwise the household files as a single. The household pays federal, state, and payroll taxes on income from both household members. Income includes earnings from each individual's job, pension income, Social Security income, and asset returns (both defined contribution and savings).

In 1992, 50% of Social Security benefit income was taxed for jointly filing household with incomes over $32,000, and single filers with income over $25,000. In December 1993, the 50% threshold was kept in place, but a second bracket, 85% of Social Security benefit income, was added for households with incomes over $44,000 for joint filers and $34,000 for single, unmarried filers. In my analysis, I assume that the 1993 rules hold for every year. For example, if single John received $10000 in Social Security benefits and earned an additional $25,000 for part-time work, then $0.5\,(32000 - 25000) + 0.85(35000 - 32000) = \$6,050$ of his Social Securiy benefit would be taxable as income. However, John will never have more than 85% of his Social Security benefits taxed, implying that if he earned $\$50,000$ for his part time work, then only $0.85(\$10,000) = \$8,500$ would be taxable. Note that these rules and levels have not changes since 1993, and therefore are not indexed for inflation.

I use the IRS tax rules from 1992 and reported state tax rates in NBER's TAXSIM calculator.[31] I weight state tax rates by the U.S. Census's projections of population in each state in July 1992 for ages 50 and greater.[32] I assume that all individuals are not self-employed for tax purposes, meaning that he or she only pays half of the payroll tax. In table 7, I report the tax rates for married, jointly filing households and single households. For joint households, the 3rd tax bracket (ending at $55,500) represents the maximum Social Security contribution level. In table 7b, I assume that only one individual earns a total of $55,500. If both individuals are working, then the third through fifth tax brackets could change depending on each individual's earning levels. Notice that the addition of the fifth tax bracket between the two tables is due to the correspondence between the top income tax bracket and the maximum Social Security contribution level in 1992.

[31] See http://users.nber.org/~taxsim/ for more details

[32] See State Characteristics, 1990 to 1999 Annual Time Series of State Population Estimates by Age and Sex, 5-Year Age Groups by Sex at http://www.census.gov/popest/data/historical/1990s/state.html

Table 7: Taxes

(a) Single

Pre-Tax Income (1992 $)	Marginal Tax Rates			
	Federal	State	Payroll	Combined
0-3,600	0.00%	0.00%	7.65%	7.65%
3,601-25,050	15.00%	4.56%	7.65%	27.21%
25,051-55,500	27.86%	5.03%	7.65%	40.54%
55,500 +	30.68%	5.37%	1.45%	37.50%

(b) Married - Filing Jointly

Pre-Tax Income (1992 $)	Marginal Tax Rates			
	Federal	State	Payroll	Combined
0-6,000	0.00%	0.00%	7.65%	7.65%
6,001-41,800	15.00%	4.51%	7.65%	27.15%
41,801-55,500	28.00%	4.79%	7.65%	40.44%
55,501-92,500	27.83%	5.05%	1.45%	34.33%
92,500 +	31.34%	5.21%	1.45%	38.00%

Note: For each household member above the age of 65, the income threshold increases by $900 for single households and $700 for married housedholds.

B Recursive Methods

Recursion is commonly used in structural models, but the typical design of a decision tree taught in standard game theory can be difficult or impossible to reproduce due to finite computational time. Often this can arise when decisions are continuous (such as how much to consume or save), when the number of periods covered are large, or when choices in each period require historical variables. While this list is not extensive, it does represent all the challenges faced in a life-cycle model of labor force participation and benefit claiming. There exist significant computational tradeoffs that must be considered when developing a structural model of this variety, and these can only be understood if the reader first has an appreciation for how the backward recursion is actually conducted and approximated.

First, it is currently impossible to come close to calculating an entire decision tree. Instead, it is approximated at each decision period by a discretized set of the state variables. In the model by French and Jones (2011), this is done with 9 state variables: (1) benefit application decision, (2) preference type, (3) whether or not there is a cost of reentering the labor force for that period (i.e. un-retire), (4) health insurance transition, (5) health status, (6) health care cost transition, (7) Social Security AIME level, (8) wage change, (9) asset level. Given their discretization, this implies $2 \times 3 \times 2 \times 3 \times 2 \times 3 \times 16 \times 5 \times 32 = 552,960$ state combinations. The calculation of decision rules through backward recursion is based (in theory) on the history of choices an agent has made up until the current period's decision node, but due to the continuous nature of state variables, such as assets, and the long history required for other state variables, such as Social Security's Average

Indexed Monthly Earnings (AIME) measure, it becomes impossible to permit state variables to depend on prior decisions.

The calculation of AIME provides an excellent example of the challenges presented by backward recursion when future choices depend upon histories in addition to current states. AIME is calculated using the best 35 years of earnings. Even if we coarsely discretize potential earnings into 5 levels, and assume everyone started their significant earning at age 25, then just for the calculation of AIME at age 60, we would have $5^{35} \approx 2.91 \times 10^{24}$ possibles wage combinations required to calculate the AIME at age 60 (think about how bad it gets at age 61!). Instead, French and Jones (2011) take the AIME in the period as given, abstracting from the history that led to its level. Since what is relevant for the current decision is both the current AIME and the AIME for continuing to work, the lack of wage history requires the modeler to approximate AIME if the agent continues to work. In French and Jones' work, they use estimated replacement wages for the population based on age. For example, for an agent continuing to work at age 62, they assume that the current wage replaces 58.9% of one year's wages relative to the individual's AIME:

$$AIME_{t+1} = (1 + CPI) \cdot AIME_t - \frac{1}{35} \{0, W_t N_t - (0.589) \cdot (1 + CPI) \cdot AIME_t\}$$

In this setup, 58.9% is meant to approximate the ratio of the lowest earnings year to AIME. As the population gets older, the ratio approaches 100%, such that the AIME does not grow through replacement of the lowest earnings. At younger ages (before 55), they assume that entire years are replaced (i.e. that the ratio is 0%).

This setup presents two major challenges to a life-cycle model of labor supply and benefit claiming. First, it smooths out accruals in retirement programs (both Social Security and defined benefit pensions), possibly reducing or eliminating the incentive to delay claiming for some individuals. Moreover, since French and Jones' tie pension benefits to the AIME, it becomes less clear how to separate the actual effects of Social Security from pensions on labor supply outcomes. Second, since this setup only approximates AIME in the next period, it cannot account for any possible notches in benefit calculations that might exist from delayed claiming beyond a year. For example, an agent in good health who is replacing a zero earnings year in the AIME calculation for each additional year of work (e.g. a woman who took a decade off from the labor force) will not only have a significant incentive to delay claiming at 62, but may face a much larger incentive to delay claiming beyond a single year due to her likelihood of survival relative to Social Security's delayed benefit adjustments.

In this paper, I take an approach similar Gustman and Steinmeier (2005), where, in order to capture the AIME and pension calculations, I take the wage paths for a worker as given, allowing AIME and pension benefits to be calculated directly. This requires calculating the decision rules for each individual in sample, and thus requires a simplification in the number of states to achieve computational feasibility. Moreover, it eliminates the feasibility of a modeler incorporating wage uncertainty into the model for fear of quickly increasing the computational burden. Choosing a fixed wages allows my model to reflect the institutional details of Social Security and individual pension plans, as well as be able to appropriately account for individuals' unique earnings histories.

Since I estimate each household's decision rules separately, I use the husband and wife's earnings history at baseline to determine the rate at which lowest earnings are replaced.

C Numerical Methods

The recursive formulation of a household's value function is given by:

$$\begin{aligned} V_t\left(X_t\right) \quad = \max_{C_{h,t},N_{h,t},B_{h,t}} \quad & \Big\{U\left(C_{h,t},L_{h,t}\right)+\delta_\tau\left(1-s_{t+1}^h\right)\left(1-s_{t+1}^w\right)b(A_{h,t+1}) \\ & +\delta_\tau\left(1-s_{t+1}^h\right)s_{t+1}^w\mathbb{E}\left[V_{t+1}\left(X_{t+1}\mid X_t,t,C_{h,t},B_{h,t},N_{h,t},\text{ wife survives}\right)\right] \\ & +\delta_\tau s_{t+1}^h\left(1-s_{t+1}^w\right)\mathbb{E}\left[V_{t+1}\left(X_{t+1}\mid X_t,t,C_{h,t},B_{h,t},N_{h,t},\text{ husband survives}\right)\right] \\ & +\delta_\tau s_{t+1}^h s_{t+1}^w\mathbb{E}\left[V_{t+1}\left(X_{t+1}\mid X_t,t,C_{h,t},B_{h,t},N_{h,t},\text{ both survive}\right)\right]\Big\} \end{aligned} \tag{C.1}$$

subject to a non-negative borrowing constraint and the consumption floor. The solution to the recursive formulation requires solving for each household's consumption, labor force participation, and benefit claiming choices at every age at and after baseline (1992), collectively referred to as the decision rules. These decision rules are calculated numerically, using the model detailed in the paper because no closed form solution exists. No closed form solution exists for several reasons, including that future state variables depend upon the history of those variables, and that there are several discontinuities in the budget set arising from taxation, pensions, and Social Security Benefits.

The recursive formulation above is solved using value function iteration beginning in period T, assumed to be age 110, and solved backward to the first period. The vector of possible states is discretized into 13 state variables: (1) the husband's stochastic preference for leisure, ε_H , (2) the wife's stochastic preference for leisure, ε_H, (3) household marital status (this is important for widowhood), (4) household health insurance status, (5) husband health status, (6) wife's health status, (7) husband's Social Security PIA, (8) wife's Social Security PIA, (9) husband's pension level, (10) wife's pension level, (11) household asset level. Given the discretization, this implies $3\times3\times3\times3\times2\times2\times2\times2\times2\times2\times8=41,472$ state combinations solved for each of the 948 households in the sample. The time T decision rule is found by assuming that everyone knows they will die at period T, such that $V_T=U(C_T,0)+\delta\cdot b(A_{T+1})$. For each set of state variables, X_T, we calculate the optimal consumption (and hence savings) decision for period T. This yields the value function at time T, which can be used in calculating V_{T-1} to find the decision rules at period $T-1$ according to the Bellman equation in (C.1). This process is repeated from period $T-1$ back to period 0, which in this model corresponds to the male household member's age at baseline (i.e. 1992).[33]

The value function is evaluated at each state combination and linear interpolation is used for continuous variables (i.e. assets, AIME, pension benefit, ε_H, and ε_W). Discretization is finer at

[33]The appendix on *Recursive Methods* goes into greater detail about how I handle approximating next periods assets given our discretization method.

lower levels of assets since I would expect greater responsiveness at lower levels to changes in asset accumulation. In my initial estimate process I keep the number of states for Social Security and pension benefits small (2 states each), but these states reflects the individual's worst and best possible benefits based on his or her own earnings history. In robustness checks, I will investigate whether the results are sensitive to this rough discretization.

Each period, the household chooses the level of consumption, labor supply, and benefit application that maximizes their discounted lifetime utility. Consumption is a continuous choice in the model, however, implying that for each state combination the household must determine the optimal level of consumption. Given the discrete nature of the other choice variables, there is no reason to expect the value function to be globally concave with respect to consumption. I discretize the consumption space into 36 choice states and allow the household to solve for V_T based on each choice state, from which the household will choose the level of consumption that maximizes its discounted lifetime utility. When the problem is solved again for period $T-1$, the agent will test only a local range of consumption choices. As the backward induction process continues, the range of consumption states tested will depend upon the male household members' age, with a larger range being used during periods of critical life choices (e.g. age 65 when respondent reaches normal retirement age for Social Security). If a value on the boundary of the consumption range is chosen, then the range is expanded by three choice states in the direction of increasing utility until a local optimum is found.

Once the decision rules are calculated, the rules are then used to generate simulated household histories. 200 random outcomes of health, medical expenses, mortality, and unobserved individual heterogeneity ($\varepsilon_{i,t}$) are generated per household. Using each household's period 0 state vector, the household's decisions in period 0 are determined from the appropriate decision rules. When the state vector does not precisely lie on the discretized grid of state combinations, I use linear interpolation to approximate the household's decisions. Combining the states and decisions from period 0, I use the budget constraint and asset accumulation conditions from the model, in addition to the health, mortality and medical expense shocks after period 0, and the appropriate Social Security and pension rules for the household, to calculate the state vector in period 1. This process is repeated, creating a life-cycle history for the household. In generating the shocks, the actual (not discretized) value of the shock is used. In generated the pension and Social Security benefit levels, the actual earnings history and pension rules are used in the calculation.

In order to reduce the computational burden, individuals are assumed to claim their benefits and cease work by age 70.

D Moment Conditions and Method of Simulated Moments

D.1 Moments Conditions

This section is a more detailed description of §5.2 in the main text. For expositional clarity, I reproduce the moments cases as they are in the main text and describe the technical details of the moments that are matched.

I divide any moments using household assets into three quantiles to capture the dispersion of assets in the data. I match the following moment conditions for ages 58-69 ($T = \{58, 59, ...69\}$) for a total of $34T = 408$ moments.

1. Mean assets by quantile and men's age, for the lowest two quantiles ($2T$ moments)

   I divide any moments using household assets into three quantiles to capture the dispersion of assets in the data. The υ_j percentage of households (h) with assets below $Q_{\upsilon_j}(A_{ht}, t)$ is defined as

   $$Pr\left(A_{ht} \leq Q_{\upsilon_j}(A_{ht}, t) \mid t\right) = \upsilon_j$$

   where the quantile index is denoted by j. Put another way, $Q_{\upsilon_j}(A_{ht}, t)$ is the υ_jth age-condition asset quantile. The model analog to $Q_{\upsilon_j}(A_{ht}, t)$ is $\hat{Q}_{\upsilon_j}(t; \theta_0, \chi_0)$ from the simulated asset distribution. Note that $t = age_i$ is individual i's age, where here it is assumed to be the male's age ($i = H$). Let $\bar{A}_j(t, \theta_0; \chi_0)$ represent the model's prediction of the mean asset level observed in asset quantile j at age t. The implied conditional moment then becomes

   $$\mathbb{E}\left[A_{ht} \mid t, Q_{\upsilon_{j-1}}(A_{ht}, t) \leq A_{ht} \leq Q_{\upsilon_j}(A_{ht}, t)\right] = \bar{A}_j(t, \theta_0; \chi_0) \ .$$

   This can then be converted into an unconditional moment that can be estimated from the simulation results by rearranging the previous equation and plugging in for the model analogs:

   $$\mathbb{E}\left[A_{ht} - \bar{A}_j(t, \theta_0; \chi_0) \mid t\right] \times \mathbf{1}\left\{\hat{Q}_{\upsilon_{j-1}}(t; \theta_0, \chi_0) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}(t; \theta_0, \chi_0)\right\} = 0 \ . \tag{D.1}$$

2. Share of a preference type's household population within each asset quantile by age (lowest two quantiles only) for men ($10T$ moments)

   Let $\bar{h}_j(\tau, t; \theta_0, \chi_0)$ represent the model's prediction of share of households, h, where the husband is $t = age_i$ years old in asset quantile interval j with preference type τ. If the model is true then:

   $$\mathbb{E}\left[h \mid Q_{\upsilon_{j-1}}(A_{ht}, t) \leq A_{ht} \leq Q_{\upsilon_j}(A_{ht}, t), t, \tau\right] = \bar{h}_j(\tau, t; \theta_0, \chi_0) \ .$$

   Empirically when estimating the moment vector, $m(LC_i, \theta_0; \chi_0)$ (see next section), I convert

this relationship into an unconditional moment equation:

$$\mathbb{E}\left[h - \bar{h}_j\left(\tau, t; \theta_0, \chi_0\right) \mid t, \tau\right] \times \mathbf{1}\left\{\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right)\right\} = 0 \tag{D.2}$$

for asset quantiles $j \in \{1, 2\}$.[34] I exclude the share of the third asset quantile as the shares are constrained to add to one, and so it is identified by the other two moments.

3. Percent participating in the labor force by preference type, age, and sex ($10T$ moments)

   Recall that each household, h, is comprised of two members of each gender $i \in \{H, W\}$ at baseline. $LFPR_{hit}$ represents i's labor force participation at $t = age_i$. I match the following unconditional moment for men and women by age:

$$\mathbb{E}\left[LFPR_{hit} - \bar{LFPR}_i\left(t, \tau; \theta_0, \chi_0\right) \mid t, \tau\right] = 0 \ , \tag{D.3}$$

   where $\bar{LFPR}_i\left(t, \tau; \theta_0, \chi_0\right)$ is the model's prediction of average labor force participation for each gender with household preference type τ.

4. Percent working full-time, conditional on working, by preference type and sex (excluding first preference type which does not work in the first period - $8T$ moments)

   Similar to case (3), each household, h, is comprised of two members of each gender $i \in \{H, W\}$ at baseline. FT_{hit} represents i's labor force status conditional on participation at $t = age_i$. If $\bar{FT}_i\left(t, \tau; \theta_0, \chi_0\right)$ represents the model's prediction of individuals working full-time conditional on participation for preference type τ at age t, then the implied conditional moment condition becomes:

$$\mathbb{E}\left[FT_{hit} \mid LFPR_{hit} = 1, t, \tau\right] = \bar{FT}_i\left(t, \tau; \theta_0, \chi_0\right).$$

   I then convert this relationship to an unconditional moment condition:

$$\mathbb{E}\left[FT_{hit} - \bar{FT}_i\left(t, \tau; \theta_0, \chi_0\right) \mid t, \tau\right] \times \mathbf{1}\left\{LFPR_{hit} = 1\right\} = 0 \ , \tag{D.4}$$

   which is used as $8T$ of the moment conditions. Note that I exclude the type where both individuals are out of the labor force at baseline, $\tau = 0$, because the moment condition may be empty for certain ages.

5. Labor force participation by individual health status, age, and sex ($4T$ moments)

   As in case (3), I match labor force participation moments conditional on health status, $health \in \{good, bad\}$, and sex. Therefore the moment condition is:

$$\mathbb{E}\left[LFPR_{hit} - \bar{LFPR}_i\left(health, t, \tau; \theta_0, \chi_0\right) \mid t, \tau, health_{it} = health\right] = 0 \ .$$

[34] I define $\hat{Q}_{\upsilon_0}\left(age_i; \theta_0, \chi_0\right) = -\infty$ and $\hat{Q}_3\left(age_i; \theta_0, \chi_0\right) = +\infty$

I then convert this relationship to an unconditional moment condition:

$$\mathbb{E}\left[LFPR_{hit} - L\bar{F}PR_i\left(health, t, \tau; \theta_0, \chi_0\right) \mid t, \tau\right] \times \mathbf{1}\left\{health_{it} = health\right\} = 0 \ . \tag{D.5}$$

D.2 Method of Simulated Moments

Using the moment conditions discussed in the previous section, I use 408 moment conditions to over-identify the 48 preference parameters, denoted by θ. Let $m(\bullet)$ represents the moment condition based on observed life-cycle histories LC_i for individual i in household h, and let θ_0 represent the true value of the preference parameters θ, from the data generating process, χ_0. Note that the life cycle histories, LC_i, comprises all observables, including endogenous outcomes, exogenous or potentially endogenous state variables, X_t, and instrumental variables. Given the vector of moment conditions such that

$$\mathbb{E}\left[m\left(LC_i, \theta_0; \chi_0\right)\right] = 0 \ ,$$

then the generalized method of moments (GMM) estimator, $\hat{\theta}_{gmm}$ minimizes:

$$Q_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{N} m\left(LC_i, \theta_0; \chi_0\right)\right]' \mathbf{W}_n \left[\frac{1}{n}\sum_{i=1}^{N} m\left(LC_i, \theta_0; \chi_0\right)\right] \ ,$$

where $\mathbf{W}_n$ is the symmetric positive definite weighting matrix that does not depend on θ. Now if there is no closed-form solution for $m\left(LC_i, \theta; \chi_0\right)$ such that:

$$m\left(LC_i, \theta; \chi_0\right) = \int k\left(LC_i, u_i, \theta; \chi_0\right) g(u_i) du_i$$

then $m\left(LC_i, \theta; \chi_0\right)$ can be replaced by $\hat{m}\left(LC_i, u_{is}, \theta; \chi_0\right)$, an unbiased simulator, and u_i denotes s draws from the marginal density $g(u_i)$. The method of simulated moments (MSM) estimator $\hat{\theta}_{msm}$ instead minimizes:

$$\hat{Q}_n(\theta) = \left[\frac{1}{n}\sum_{i=1}^{n} \hat{m}\left(LC_i, u_{is}, \theta; \chi_0\right)\right]' \hat{\mathbf{W}}_n \left[\frac{1}{n}\sum_{i=1}^{n} \hat{m}\left(LC_i, u_{is}, \theta; \chi_0\right)\right] \tag{D.6}$$

where $\hat{m}\left(LC_i, u_{is}, \theta; \chi_0\right)$ is defined by the moment conditions in (D.1)-(D.5) above, and $\hat{\mathbf{W}}_n$ is the optimal weighting matrix from the simulated data. Following Gourieroux and Monfort (1996), as $n \to \infty$ and for a fixed number of simulations s, $\hat{\theta}_{msm}$ is both consistent and asymptotically normally distributed:

$$\sqrt{n}\left(\hat{\theta}_{msm} - \theta_0\right) \xrightarrow[n\to\infty]{d} N\left(0, \hat{\Upsilon}\right),$$

where:

$$\hat{\Upsilon} = \left(D'WD\right)^{-1} D'WSWD\left(D'WD\right)^{-1} \tag{D.7}$$

such that $D = {}^{\partial m(\cdot)}/_{\partial \theta'} \mid_{\theta=\theta_0}$ and $W = plim_{n\to\infty}\hat{W}$, which is estimated by:

$$\hat{W} \quad = \quad \left\{ \tilde{V}\left(m\left(LC_i, \theta; \chi_0\right)\right) + \frac{1}{s}\tilde{V}\left(\hat{m}\left(LC_i, u_{is}, \theta; \chi_0\right)\right) \right\}^{-1}$$

where $\tilde{V}(\cdot)$ is the estimated variance with respect to a larger simulation sample and S is the variance-covariance matrix of the simulated sample. Thus the first term represents the moment condition from the data with respect to the larger simulated sample, and the second term represents the moment condition with respect to the smaller simulation sample from which the estimates are selected. Note that the optimal choice of W, corresponds to $W = S^{-1}$, simplifying the asymptotic variance-covariance matrix to

$$\hat{\Upsilon} = \left(D'\hat{W}D\right)^{-1}$$

In practice, I use only the diagonal terms of $\tilde{V}\left(m\left(LC_i, \theta; \chi_0\right)\right)$ when calculating $\hat{W}$ in order to minimize (D.6). This is to ensure invertibility (non-singularity) and because S may be biased in small samples. When I calculate the standard errors of the preference parameter vector $\hat{\theta}_{msm}$ and test the moment conditions (i.e. over-identified restrictions of the model) against the zero restrictions implied by the model, I use equation (D.7) as the approximate variance-covariance matrix, $\hat{\Upsilon}$.

When calculating, $D = {}^{\partial m(\cdot)}/_{\partial \theta'} \mid_{\theta=\theta_0}$, most calculations are done by taking the straightforward numerical derivative using a two-sided approach with a 1 percent variation in the underlying parameter. However, the first two moment conditions, since they are based on asset quantiles, require additional simplification. Recall that equation (D.1) was written as

$$\mathbb{E}\left[A_{ht} - \bar{A}_j(t, \theta_0; \chi_0) \mid t\right] \times \mathbf{1}\left\{\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right)\right\} = 0 \ .$$

This equation can be rewritten as

$$\int_{Q_{\upsilon_{j-1}}(A_{ht},t)}^{Q_{\upsilon_j}(A_{ht},t)} \left\{\mathbb{E}\left[A_{ht} \mid t\right] - \bar{A}_j(t, \theta_0; \chi_0)\right\} \times f\left(A_{ht} \mid t\right) dA_{ht} = 0.$$

Applying Liebnitz's rule, the first-order condition becomes,

$$\begin{aligned} D \quad = \quad & -Pr\left[\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right) \mid t\right] \times \frac{\partial \bar{A}_j(t, \theta_0; \chi_0)}{\partial \theta'} \\ & + \left\{\mathbb{E}\left[\hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right) \mid t\right] - \bar{A}_j(t, \theta_0; \chi_0)\right\} \times f\left(\hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right) \mid t\right) \times \frac{\partial \hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right)}{\partial \theta'} \\ & - \left\{\mathbb{E}\left[\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \mid t\right] - \bar{A}_j(t, \theta_0; \chi_0)\right\} \times f\left(\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \mid t\right) \times \frac{\partial \hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right)}{\partial \theta'}. \end{aligned}$$

Similarly, recall equation (D.2):

$$\mathbb{E}\left[h - \bar{h}_j\left(\tau, t; \theta_0, \chi_0\right) \mid t, \tau\right] \times \mathbf{1}\left\{\hat{Q}_{\upsilon_{j-1}}\left(t; \theta_0, \chi_0\right) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}\left(t; \theta_0, \chi_0\right)\right\} = 0$$

It can be rewritten as

$$\int_{Q_{\upsilon_{j-1}}(A_{ht},t)}^{Q_{\upsilon_j}(A_{ht},t)} \left\{\mathbb{E}\left[h \mid A_{ht}, t, \tau\right] - \bar{h}_j\left(\tau, t; \theta_0, \chi_0\right)\right\} \times f\left(A_{ht} \mid t\right) dA_{ht} = 0,$$

where the first order condition becomes,

$$\begin{aligned} D \;=\; & -Pr\left[\hat{Q}_{\upsilon_{j-1}}\left(t;\theta_0,\chi_0\right) \leq A_{ht} \leq \hat{Q}_{\upsilon_j}\left(t;\theta_0,\chi_0\right) \mid t,\tau\right] \times \frac{\partial \bar{h}_j\left(\tau,t;\theta_0,\chi_0\right)}{\partial \theta'} \\ & + \left\{\mathbb{E}\left[h \mid \hat{Q}_{\upsilon_j}\left(t;\theta_0,\chi_0\right), t, \tau\right] - \bar{h}_j\left(\tau,t;\theta_0,\chi_0\right)\right\} \times f\left(\hat{Q}_{\upsilon_j}\left(t;\theta_0,\chi_0\right) \mid t,\tau\right) \times \frac{\partial \hat{Q}_{\upsilon_j}\left(t;\theta_0,\chi_0\right)}{\partial \theta'} \\ & - \left\{\mathbb{E}\left[h \mid \hat{Q}_{\upsilon_{j-1}}\left(t;\theta_0,\chi_0\right), t, \tau\right] - \bar{h}_j\left(\tau,t;\theta_0,\chi_0\right)\right\} \times f\left(\hat{Q}_{\upsilon_{j-1}}\left(t;\theta_0,\chi_0\right) \mid t,\tau\right) \times \frac{\partial \hat{Q}_{\upsilon_{j-1}}\left(t;\theta_0,\chi_0\right)}{\partial \theta'}. \end{aligned}$$

E Data and Sample Selection

This appendix provides greater detail on the data used in estimating the model described in §3.

E.1 Data

I use the original cohort of the Health and Retirement Study (HRS), which was born between 1931 and 1941, and has 12,652 respondents and 7,704 households in the main analysis. However when calculating the transition probabilities for health and mortality, as well as medical expenses, I also use the Asset and Health Dynamics among the Oldest Old (AHEAD) cohort from the HRS, which consists of non-institutionalized individuals born before 1923.

I use the RAND HRS cross-wave supplement (version L) as the initial data set. I then import Social Security earnings history from a separate file where I have calculated, conditional on the assumptions specified in §F, each individual's AIME and PIA as of 1992 as well as each individual's defined benefit levels for every possible age of retirement between 1992 and 2010. Using the combined data set, I use the RAND tenure variable to determine the number of jobs, including baseline job, that are observed between 1992 and 2010.

I define an individual to have retiree health insurance if they report having health insurance coverage that persists after retirement or have access to VA or CHAMPUS benefits (retired or active duty U.S. military benefits). An individual who has health insurance but does not meet these criteria is considered to have tied health insurance. If an individual has medicaid, private health insurance, or another type of means-tested health insurance, I treat them as having no health insurance, since these individuals are more likely to resemble to pool of individuals with no health insurance. I create a household health insurance variable by assuming that if one individual is eligible for retiree health insurance then everyone is. If no one in the household has retiree health insurance, but at least one individual has tied health insurance, then the household acts as if it has tied health insurance. Finally, no member of the household has health insurance, then the household is treated as having no health insurance.

Since the HRS is conducted at two year intervals, I use the reported labor force status in the RAND HRS supplement for labor force participation in years that correspond to survey waves, and then use information regarding last job and data on Social Security earnings history to fill in labor force participation between survey waves. To be participating in the labor force, an individual must report being employed full-time, employed part-time, unemployed, or partially retired. Additionally, the individual must work more than 300 hours per year. If an individual continues in the same job, then I assume that the hours in non-survey years are the same as the previous survey year. I use information on when a person ended his or her last job to deduce between-wave labor force participation and job changes. I only use Social Security data, when an individual has changed jobs and cannot use the surrounding waves' information regarding the employment (i.e. the inter-wave job was very short, or the individual was not surveyed in adjacent waves). When I do use the Social Security data, I assume the individual is participating in the labor force if they have a positive earnings history.

I consider an individual to be working full-time if he or she reports working full-time and if he or she reports working in excess of 1600 hours per year. An individual is considered working part-time if he or she reports being employed part-time or reports being partially retired with between 300 and 1600 annual hours of work. If I am relying on Social Security reports to determine the individual's work status, then I assume that 4 quarters of coverage corresponds to full-time work and between 1 and 3 quarters corresponds to part-time work. Using Social Security's earnings records is an imperfect measure since the burden for reaching 4 quarters is low, but this is rarely used since most people's work histories can be achieved based on respondent's reports of when they stopped working at his or her last job.

As described in the section on health, individuals provide a self-reported health status to the interviewer on a scale of *excellent, very good, good, fair,* and *poor*. I reduce these self-reports to a binary measure $good \in \{excellent,\ very\ good,\ good\}$ or $bad \in \{fair,\ poor\}$.

I use the RAND HRS measures of household assets. To create my measure of household assets, I sum the value of the household's primary residence, and the net value of other real estate, businesses, vehicles, stocks, mutual funds, other investment trusts, checking accounts, certificates of deposits, savings accounts, government savings bonds, treasury bills, bonds, bond funds, and any other reported savings, and subtract debt from the household primary residence's mortgage, any other debts based on the primary residence, and and remaining non-residence based debt.

E.2 Sample Selection

The original HRS sample has 7,704 households, which includes 5,813 households with at least one male. Of the male households, 1 is eliminated because the birth year of the respondent is unknown [5,812], 968 are not married [4,844] and 260 are eliminated for missing spousal information in the first wave [4,584]. I keep households that (1) are married in wave 1 and not missing spousal information [4,584], (2) are not missing information on their labor force participation in 1992 [4,575], (3) have never applied for Social Security disability benefits [3,300], (4) are without missing pension [2,628]

or Social Security information [2,197], (5) have a spousal age difference of less than 10 years [1,943], (6) are not missing information on either household member's baseline earnings [1,899], and, for computational tractability, (7) households with no more than one defined benefit pension [1,729]. Additionally, I drop annual observations if employment or health status of either household member is not reported, and if health insurance status cannot be determined when the household is less than age 65 (Medicare age) [1,728].

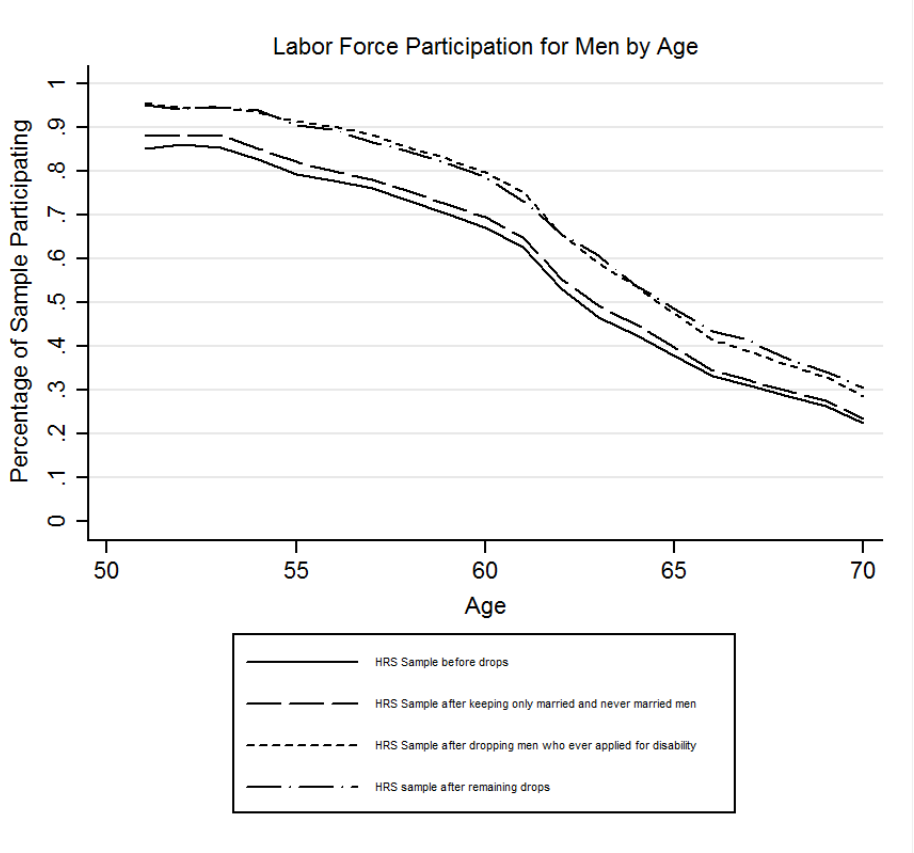


Figure 10: Sample Selection by Labor Force Participation of Men

After this sample selection, I am left with 1,728 married households. I use the Social Security Administrative data for earnings and labor force participation histories and respondent reports for periods not covered by the Social Security data. Doing so yields an average of 14.95 annual observations per household (out of a maximum possible of 20), providing a long history of observations. Figure 10 shows how my sample selection effects the average rate of male labor force participation. The omission of divorced, separated, and widowed households increases labor force participation slightly, but eliminating those household that ever apply for Social Security disability benefits increases labor force participation at all ages by approximately 10%. This result is not surprising since individuals who credibly apply for disability will likely have a reduced ability to participate

Table 8: Sample Statistics from the selected HRS Sample.

		Men	Women
Age	Mean	57.81	54.89
	Median	57.75	55.08
	Standard Dev.	4.72	4.69
Earnings	Mean	$31,697.86	$12,614.91
	Median	$24,400	$8,320
	Standard Dev.	$35,786	$18,689.36
AIME	Mean	$2,187.74	$703.8
	Median	$2,321.5	$489
	Standard Dev.	$971.85	$692.8
Predicted Annual Pension Benefit	Mean	$14,829.28	$6,696.37
	Median	$7,195.43	$2,611.16
	Standard Dev.	$27,029.28	$12,061.7
% with Current Pension Benefit		25.78	24.34
% Working		78.36	59.43
% Working Full-time		85.16	62.41
% in Bad Health		10.76	9.84
% White		89.93	90.1
Average Years of Education		12.65	12.48

		Household
Assets	Mean	$324,744.2
	Median	$163,272.5
	Standard Dev.	$564,908.1
% with Retiree Health Insurance		64.06
% with Tied Health Insurance		18.34
% with No Health Insurance		17.54
Preference Type (Work, Spousal)	Out	12.04
	Low, Low	23.26
	High, Low	20.54
	Low, High	21.59
	High, High	22.45
Fraction of Women Eligible for Spousal Benefit	Overall	57.18
	1st Asset Quantile	56.77
	2nd Asset Quantile	58.68
	3rd Asset Quantile	56.08
Number of Households		1,728

Note: Sample consists of only those households with one member between the ages of 51 and 61 in 1992. Individual income is conditional on participating in the labor force in 1992. Predicted Annual Pension Benefit is defined benefit pensions that are vested and is conditional on having a pension. The percentage with current pension is conditional on participating in the labor force in 1992. The percentage working full-time is conditional on participating in the labor force in 1992.

in the labor force.

Table 8 provides sample statistics for the entire sample, while table 1 in the main text provide the sample statistics for the estimation sample. Finally, table 9 provides the sample statistics for the model validation sample.

F Pensions

F.1 Defined Benefit Plans

DB plans provide a guaranteed payment to an employee who is vested. An employee typically becomes vested after 5 or 10 years of service, at which point they will be eligible for a pension benefit based on years of service. Many pension plans define the workers annual benefit ($db_{i,t}$) as:

$$db_{i,t} = (\text{Years of Service}) \times (\text{PoFS}) \times (\text{AFS})$$

Table 9: Sample Statistics from the portion of the HRS Sample used for Model Validation.

		Men	Women
Age	Mean	54.78	51.75
	Median	54.17	52.08
	Standard Dev.	3.62	3.29
Earnings	Mean	$35,131.12	$13,734.66
	Median	$28,000	$9,800
	Standard Dev.	$37,067.81	$16,785.95
AIME	Mean	$2,355.99	$733.6
	Median	$2,504	$524
	Standard Dev.	$1,015.35	$714.67
Predicted Annual Pension Benefit	Mean	$12,133	$6,631.55
	Median	$5,267	$2,225
	Standard Dev.	$17,493.59	$12,610.25
% with Current Pension Benefit		27.74	23.19
% Working		88.06	66.59
% Working Full-time		88.56	63.65
% in Bad Health		8.98	9.53
% White		90.69	91.24
Average Years of Education		12.81	12.55

		Household
Assets	Mean	$305,626.9
	Median	$141,600
	Standard Dev.	$532,144.8
% with Retiree Health Insurance		59.36
% with Tied Health Insurance		22.45
% with No Health Insurance		18.18
Preference Type (Work, Spousal)	Out	5.48
	Low, Low	24.42
	High, Low	24.21
	Low, High	22.02
	High, High	23.88
Fraction of Women Eligible for Spousal Benefit	Overall	57.5
	1st Asset Quantile	59.34
	2nd Asset Quantile	57.57
	3rd Asset Quantile	55.59
Number of Households		913

Note: Sample consists of only those households with one member born between 1937 and 1941. Individual income is conditional on participating in the labor force in 1992. Predicted Annual Pension Benefit is defined benefit pensions that are vested and is conditional on having a pension. The percentage with current pension is conditional on participating in the labor force in 1992. The percentage working full-time is conditional on participating in the labor force in 1992.

where PoFS is the percent of final salary, usually between 1.5% and 2.5%, and AFS is the average final salary, usually the best three or five years of service. PoFS may follow a bend point system based on years of service (e.g. 2.2% for the first 20 years of service and 2% thereafter). Note that to accommodate more gradual retirement, most plans take the best average annual salaries over a worker's lifetime. Depending on the plan, these best years may be required to be consecutive. Most plans offer an early retirement option, usually at ages 50, 55, 58, 60, or 62 assuming the employee is vested. Individuals taking early retirement may have their annual benefit reduced, but this reduction can vary widely by plan. For example, the California State Teacher Retirement System reduces monthly benefits by 50% of the PoFS for each month before age 60 and then keeps it at this rate for the same number of months after age 60.[35] Alternatively, Michigan's teacher pension system permanently reduces monthly benefits by an annualized amount of 6% for each year before age 60.

Once an individual reaches the full-retirement age, usually age 60, 62 or 65, some plans may offer delayed retirement benefits, such as a higher PoFS, but many offer no benefit beyond increased years of service increments. Other employers may offer a longevity bonus to a monthly benefit (e.g. an extra $300 for employees with at least 3 years of service). Re-employment at the same place of employment after claiming a pension plan is discouraged by most plans through benefit reduction or elimination. Alternatively, some employers in an effort to retain older workers have implemented deferred retirement option plans (often called DROPs) which permit a worker to claim his or her benefit, but this benefit is placed in an interest bearing account payable upon retirement.

Those employees who are not vested can receive a refund with interest on the amount that they personally contributed to the pension plan. Some non-vested plans allow for some payment of the employer portion if the employee has greater than 5 years of service.

In the model presented in §3, if an individual has too few years of service to qualify for an annual benefit, then the vested benefit level is treated as a lump sum benefit when the individual leaves the baseline job. Additionally, if the individual's plan is like the California plan above (lower benefits early, higher benefits later) or has a Social Security topper (higher benefits before age 62 and lower benefits after) then the long term rate (i.e. what the benefit level is 10 years after claiming) is treated as the monthly benefit level and the individual has to pay a lump sum payment at claiming that makes up for the difference. This is done for computational feasibility.

Some individuals have access to multiple defined benefit pension plans. I assume that the individual cannot claim until the early eligibility date of his or her largest DB plan. If smaller plans are not yet eligible, then I still assume that the individual receives the same annual benefit that they would receive in the long-term, but he or she makes lump sum payment upon claiming to cover those additional benefits.

Employees eligible for a benefit can generally elect to have a survivor benefit that is 0-100% of their benefit amount, where benefits are reduced according to the actuarially fair rate of adjustment

[35] For example, consider Jane who is eligible for a $2000 monthly benefit if she retires at her full retirement age in June. If Jane claims in April and she will receive a $1000 monthly benefit from April to August of that year and then she will receive $2000 per month thereafter.

(i.e. the pension provider will consider the possible survivor's gender and relative age). Most plans include a survivor option should the employee die prior to retirement that pays a fixed benefit at death (similar to life insurance) and may pay a monthly benefit that makes an assumption about what the employee would have done if he or she had survived and chosen a plan. For example, in the California State Teacher Retirement System, the survivor receives a benefit based on a 50% beneficiary option, so that the survivor would be eligible to receive 50% of the employee's benefit, which would be reduced based on the survivor's gender and relative claim age.

DB plans work much like Social Security, often providing disability insurance to the employee and life insurance benefits to spouses and children. Due to data limitations, and similar to what I do for Social Security, I ignore these benefits here as most couples in the HRS do not have children living with them. For survivor benefits, I will assume that individuals have claimed a 0% beneficiary option to simplify the analysis. Otherwise, benefits from DB plans will be defined as in the summary plan descriptions provided to HRS.[36]

F.2 Defined Contribution Plans

DC plans do not provide a guaranteed payment to an employee upon retirement. Many plans require an employee to be vested, usually 3 to 5 years, before he or she is eligible to retain employer contributions. Employers generally match employee contributions, up to some maximum level, such as a \$1 match for every \$1 contributed or \$1 match for every \$2 contributed. Most of these plans are administered by private entities, and provide the employee with a wide range of investment possibilities. These plans generally do not act as a form of insurance for the employee, so employees have to separately subscribe to disability or life insurance plans. Any surviving beneficiaries receive access to the DC plan's account balance.

Taxation of defined contribution plans is based on the taxable amount, which is generally considered the amount that has not previously been taxed. In most cases, this is comprised of the deferred wage, taxed based on your income tax bracket in the year the individual receives the annuity and any gains in those contributions over the lifetime. There are only two major notches in a household budget constraint based on DC plans: the 10% tax penalty for withdrawals before age $59\frac{1}{2}$ and the required withdrawals from investment retirement account in the year after an individual turns $70\frac{1}{2}$. I do not expect these constraints to be binding for the majority of the HRS sample. One of the major reasons an agent would want to withdraw from a DC plan before $59\frac{1}{2}$ would be due to a medical expense shock, which would be exempted under Internal Revenue Service rules.

I will treat defined contribution plans as additional post-income tax assets, therefore these plans will be subject only to a personal income tax on any growth and I will assume that they grow at the rate of return, r. This is a strong assumption because standard 401(k) contributions by the worker and all contributions made by an employer are generally are not taxed as income until disbursement.

[36] This assumption can be quite strong since it is possible that the different household types, described in §5.1.4, could vary in their choice of benefit plan. However, this reflects an income effect, and does not induce further notches in the budget constraint, so I view this as a reasonable assumption to make in light of the computational difficulties that this would otherwise entail.

I omit this detail for computational feasibility because a more accurate model of these assets would require both knowledge on HRS's part of which assets are and are not pre-tax (which HRS does not know) and an additional state variable in the estimation of the model that tracks pre-tax assets.

F.2.1 Defined Contribution Imputations

The Health and Retirement Study (HRS) has released imputations for DC plan wealth through the sixth wave (2002). To fill in DC plan wealth for 2002 to 2010, I impute the DC plan wealth following a procedure similar to how RAND imputes income and wealth levels, and compare my imputations with the earlier imputations done by the HRS staff for the overlapping years (2000,2002 or waves 5 and 6).[37]

The HRS collects information for up to 4 pensions each interview wave. If the respondent reports having either a DC or a combination plan and is missing information on the plan's balance, I impute a balance amount. Individual's who did not know their balance amount were asked a series of unfolding brackets to help approximate the balance (i.e. if you do not know your pension, is it greater or less than $20,000). Unlike RAND's imputation procedure, I do not impute ownership. Conditional on reported ownership of a DC plan, I impute the bracket if none is given, and then conditional on bracket, I impute an account balance.

I impute brackets for individuals who report a DC or combination plan, but do not provide a complete range. I begin by estimating an ordered logit model of the DC balance bracket on the sample of individuals who report complete brackets but do not report a balance. The covariates include dummies for if there is greater than 50% chance of leaving a bequest of more than $10,000, greater than 50% chance of leaving a bequest of more than $100,000, high school diploma or higher, college degree or higher, whether the respondent self reports excellent or very good health, whether the respondent self reports poor or fair health, whether the respondent works in a professional occupation, self-employed, married, spouse-age missing, and non-white, as well as continuous measures of tenure, own-age, own-age squared, spouse-age, and spouse-age squared. All covariates with the exception of the bequest arguments are interacted with the individual's gender. Second, I use the fitted model to predict the probability of being in each of the five brackets, and then use these probabilities to generate a cumulative distribution. Third, I draw a random number from a uniform distribution, and compare the random number to the cumulative distribution in order to assign each of the individuals with missing bracket information to a bracket.

Finally, I impute account balances for all individuals who report a DC or combination plan, but do not provide a specific balance amount. I begin by estimating a standard regression on log account balances using the same covariates as used in the bracket imputation in addition to the dummies for each individual's respective balance level. Second, I use the fitted model to predict DC account balances for all individuals who report a DC or combination plan. Using a modified hot-deck approach, I sort the data by the imputed account balances and then assign account balances

[37] See *Imputations for Pension-Related Variables, Final, Version 1.0 (June 2005)* by the Health and Retirement Study for a description of the HRS's imputation process for waves 1-6.

to missing observations based on a weighted average of the nearest-neighbors.

This imputation procedure produced similar results to the HRS imputation procedure used for the first six waves. Waves 5 and 6 were estimated for both samples. In the 6th wave, the imputation procedure produced 558 additional balances, bringing the total observed to 1,569. The mean [standard deviation] of log account balances before imputation was 10.19 [1.84] and after this imputation procedure it was 9.87 [2.00]. HRS's imputation procedure produced a mean [standard deviation] of 9.79 [1.99]. In the 5th wave, the imputation procedure produced 524 additional balances, bringing the total observed to 1,882. The mean [standard deviation] of log account balances before imputation was 10.05 [1.90] and after this imputation procedure it was 9.84 [1.91]. HRS's imputation procedure produced a mean [standard deviation] of 10.07 [1.89]. Since both imputation methods produce similar results, I use the HRS's imputations for the first 4 waves (1992-1998), and the aforementioned imputation method for the remaining waves (2000-2010).

F.2.2 Tax Treatment of IRAs and tax-deferred accounts

At baseline, 1992, I observe the respondent's report of how much money the household has saved in defined contribution plans, such as 401(k), 403(b), and IRA accounts. Standard accounts like these are usually composed of pre-tax earnings, meaning the individual has not paid income tax on this money. Therefore, when the money is disbursed, it would be taxed as unearned income (i.e. it will not be subject to payroll taxes, but it will be subject to income tax). The model, as currently estimated, does not include a distinction between pre- and post-tax savings because of the computational burden associated with estimating these separately. Since defined contribution plans are treated as post-tax savings in the model, I must make an assumption about how much of the account at baseline should be reduced to reflect the future payout of income taxes.

I assume that the money is disbursed based on the expected joint life-expectancy from the 1994 IRS joint life expectancy table.[38] I account for the couple's estimated Social Security benefits and defined benefit if both claimed their Social Security at age 62 and DB pension benefits at age of first eligibility, and assume the DC disbursements start at age 70. Put simply, an individual will draw on his or her DC account starting at age 70, and will take the minimal disbursements required by the IRS. Therefore, the respective amount that they will be taxed will be based on their annual income comprised of Social Security benefits, DB benefits, and DC disbursements. I account for the limited amount of Social Security income that is taxable. I set the tax attributable to the DC disbursement assuming that DC disbursements are the last dollars taxed. I then sum tax payments from the DC plan across years and subtract this from the total DC account at baseline. The account is then assumed to be comprised of post-tax dollars.[39]

[38] The 1994 IRS publication 590 was the earliest I could locate. Age 62 corresponds to the mean age people plan to begin collecting benefits.

[39] Note that this procedure is ad hoc: While I account for age differences within the couple, I do not account for the individual decision of when it is disbursed and whether the couple continues to work. Since the only source of income for these individuals is via annuity payments from Social Security, DB plans and minimal DC plan disbursments, and not based on earnings from work, the taxed amount should be lower than expected.

G Earnings Profiles

The model described in §3 assumes that each individual can choose between employment in her baseline full-time job (FT-B), a non-baseline full-time job (FT-NB), a part-time non-baseline job (PT), and no job. Earnings in these employment states are assumed to be non-stochastic and known to the individual, similar to Blau and Gilleskie (2006). However, unlike Blau and Gilleskie (2006), I allow earnings to change with age in all possible employment states. This is done to reflect diminished employment prospects with age and the fewer hours worked after age 58 among those participating in the labor force. In this section, I will first specify how baseline full-time earnings are determined, and then consider how non-baseline full-time and part-time earnings are determined.

I define the baseline job as the full-time job an individual currently holds at baseline (1992). If an individual leaves there FT-B job for any other state, then he or she cannot return to the FT-B job. Annual earnings for FT-B jobs are determined from individual self-reports in the Health and Retirement Study, and grow at a constant rate, consistent with the HRS pension calculator. The HRS pension calculator uses information collected from employer reported "summary plan descriptions" in combination with the worker's reported annual earnings and user-specified assumptions regarding nominal wage growth, inflation, and real interest rates to predict the worker's annual benefit levels by respective quit dates. Consequently, the earnings model must reflect the same assumptions used in the pension calculator to ensure that the correct benefit levels are predicted. The assumptions used in the pension calculator are a real interest rate of 4%, inflation of 2%, and nominal wage growth of 0%. This is consistent with the realized negative real wage growth rate of approximately 2%, following baseline, among individuals with pension plans in the sample specified in §4.

The situation is more complicated for non-baseline earnings. Approximately 58.9% of men and 35.0% of women are in a FT-B job at baseline. From the men (women) who have a FT-B job at baseline, 17.6% (22.8%) will transition to a PT job from the FT-B job, and 15.6% (15.5%) will transition to a FT-NB job from the FT-B job. Of the men (women) transitioning from FT-B to a FT-NB job, 31.6% (35.9%) will receive earnings increases after the move. Median annual earnings for men at FT-NB jobs rise until about age 57 and then decline, as seen in figure 11. This is despite median annual hours falling prior to age 57 and then remaining relatively constant for FT-NB jobs (as in figure 12). Alternatively, the story for women in FT-NB jobs is that annual earnings decline after 54 and then becomes noisy for ages 60+, despite annual hours remaining largely unchanged. Finally, part-time earnings decline as hours decline for both sexes.

Non-baseline jobs represent an alternative employment option for individuals at baseline and each subsequent period (up until the maximum working age of 70). Therefore, it is important to assign a feasible wage that a worker might believe is available to her outside of her baseline job (if she is working), or if she was to return to the workforce (if she was not working at baseline).

I estimate individual log earnings profiles (separately by sex and employment status), $\ln w_{it}$, for jobs that begin after baseline - the first sampling wave of the HRS in 1992. Baseline in my full sample corresponds to an average age of 57.6 for men and 54.6 for women. These jobs represent

Figure 11: Median Earnings - Non-baseline jobs

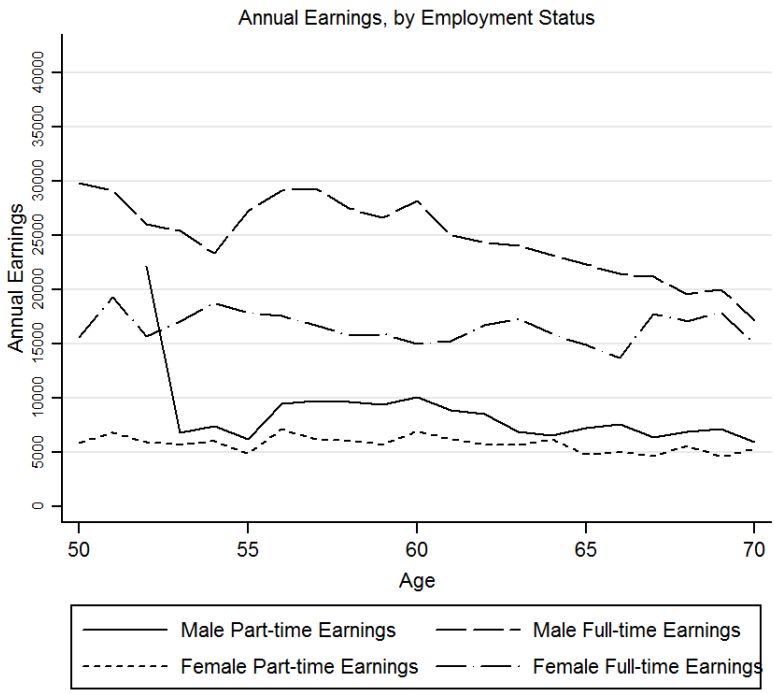
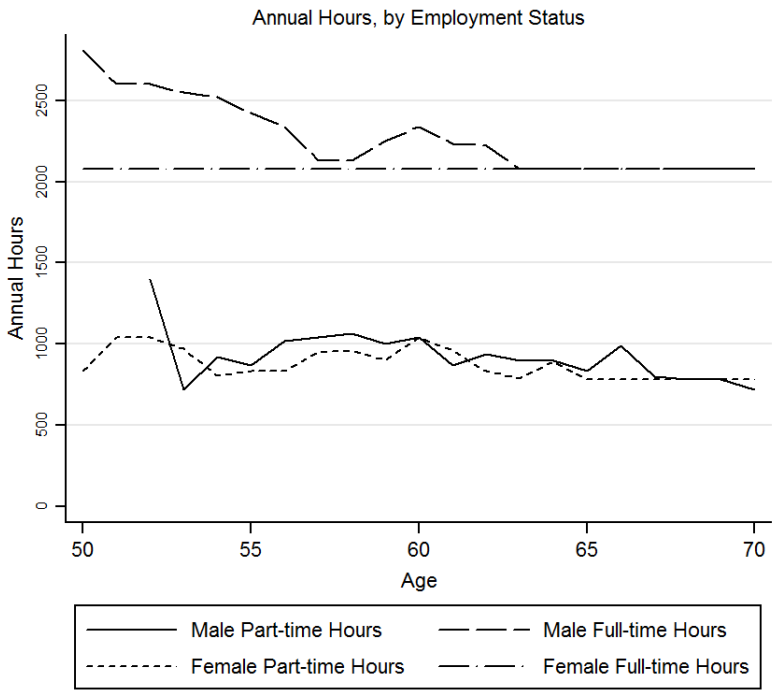


Figure 12: Median Hours - Non-baseline jobs



alternatives to the individual's baseline job, which most individuals have held for a long time. The independent variables, x_{it}, include a quartic in age and a quadratic in tenure (tenure is only included for FT-NB jobs). At this late age, I model wages as being primarily determined by an individual i's time invariant ability, c_i^j, if he or she is working in job $j \in$ {FT-NB, PT}:

$$\ln w_{it}^j \quad = \quad x_{it}^{'}\beta^j + c_i^j + \varepsilon_{it}^j \ , \tag{G.1}$$

where ε_{it}^j is a the model error term such that $\mathbb{E}\left[\varepsilon_{it}^j | j, x_{i1}...x_{iT_i}, c_i^j\right] = 0$, and T_i corresponds the last observed period for individual i . The model can then be used to predict the time invariant fixed effect, $\hat{c}_i^j = \bar{\ln w}_i^j - \bar{x}_i\hat{\beta}$ where $\bar{\ln w}_i = \sum_t(\ln w_{i,t}/T_i)$

When (G.1) is estimated, a value of $\hat{c}_i^j$ can be calculated for all individuals with at least two periods where non-baseline jobs are observed. The β^j terms in equation (G.1) are identified by variation within individuals over time.

Some individuals will not have a predicted fixed effect, $\hat{c}_i^j$. Specifically, individuals who (i) never work in another job after quitting his or her baseline job, and (ii) individuals who never work. In order to predict a fixed effect for these individuals, I regress

$$\hat{c}_i^j = \theta_0^j + \theta_1^j educ_i + \theta_2^j AIME1992_i + \theta_3^j EarningsBaseline_i + \eta_i^j \tag{G.2}$$

on the same individuals used in estimating equation (G.1), and then use (G.2) to predict $\hat{c}_i^j$ for those missing individuals due to (i). I do the same thing for individuals who never work, but exclude baseline earnings.

Predicted earnings profiles for individual i at each age t in job j are made by substituting the respective values into equation (G.1). Predicted profiles for the mean worker are included in figure 13.

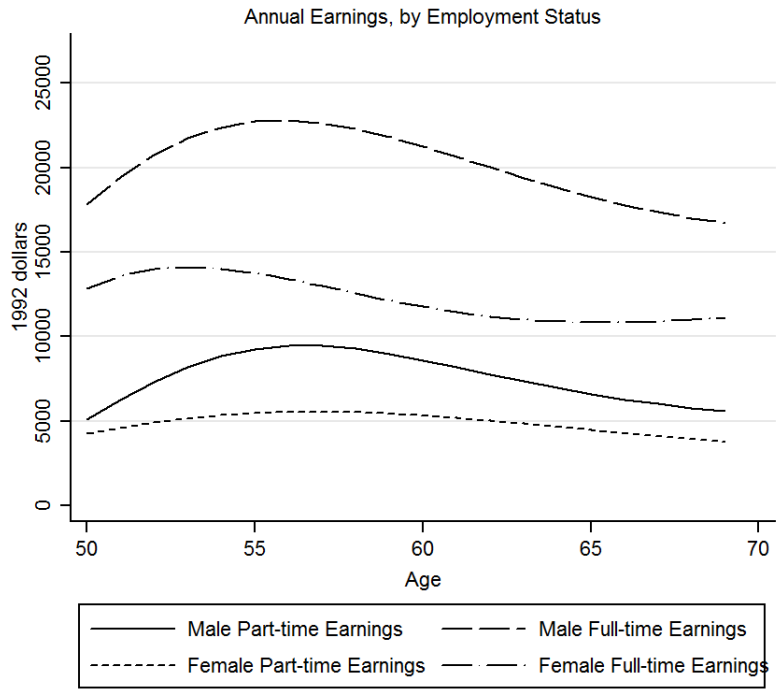
I do not estimate a combined model in (G.1), because this model specifically prevents the change in the quality of the match, $\Delta\varepsilon_{it}$, from being correlated with change in employment status, which rules out most types of endogenous job search. This is particularly problematic in my setting, where I observe workers occasionally getting higher wages on part-time jobs relative to full-time jobs. In fact, since 23% of individuals who have both a PT and FT-NB jobs after baseline have higher part-time wages, it is very likely that changes in observed employment status may be driven by positive shocks during job search.

H Health and Mortality Transitions

Health and mortality transitions are estimated using logit model based on a cubic in age, and lagged health status.

Figure 14 shows the 1 year transition probabilities from good to bad health and bad to bad health, for men and women. Men are more likely to move into and stay in bad health (relative to women) as they age. The probability of the average man (woman) remaining in bad health steadily

Figure 13: Annual Earnings by Employment Status



increases from around 72% (72%) at age 50 to 98% (96%) at age 100. Likewise, the probability of the average man (woman) transitioning from good health to bad health increases from 7% (7%) at age 50, to 50% (40%) at age 100.

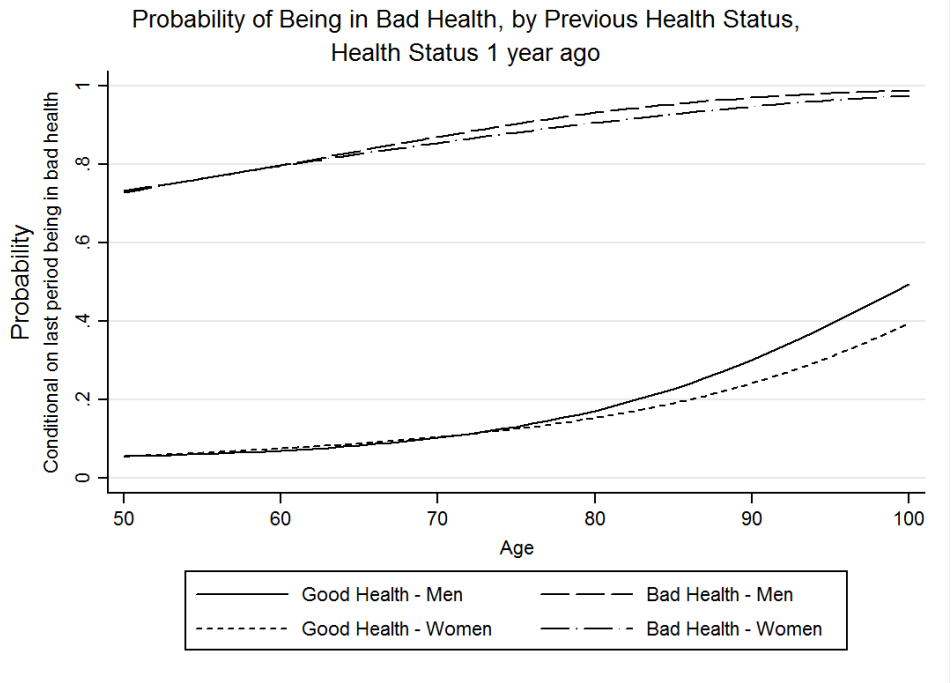
Figure 15a shows the probability of death for men conditional on health status. As a point of reference, I include information from the Social Security actuarial tables for the 1933 birth cohort. The figure indicates that at younger ages, my model under-predicts the conditional population mortality rate, which is to be expected since the sample used is going to be more likely to have worked and includes younger cohorts. At older ages, the model over-predicts the mortality rate, which is also expected since members of the AHEAD cohort, comprised of birth cohorts before 1924, are used identify mortality rates at these ages. Additionally, figure 15b shows the comparable result for women.

I Medical Expense Distribution

Each period, the household faces a medical expense shock based on its health status. As discussed in §5.1.3, I use a transitory shock from a distribution that is based on the the original HRS sample.

The HRS collects data on self-reported out of pocket medical expenditure ($M_{i,t}$) , which is imputed by the Labor and Population Program at the RAND Institute on Aging. In estimating the medical expense distribution, I include members from the Asset and Health Dynamics among the Oldest Old (AHEAD) cohort from the HRS. This sample consists of individuals born in 1923 or

Figure 14: Probability of being in bad health
Conditional on previous health status



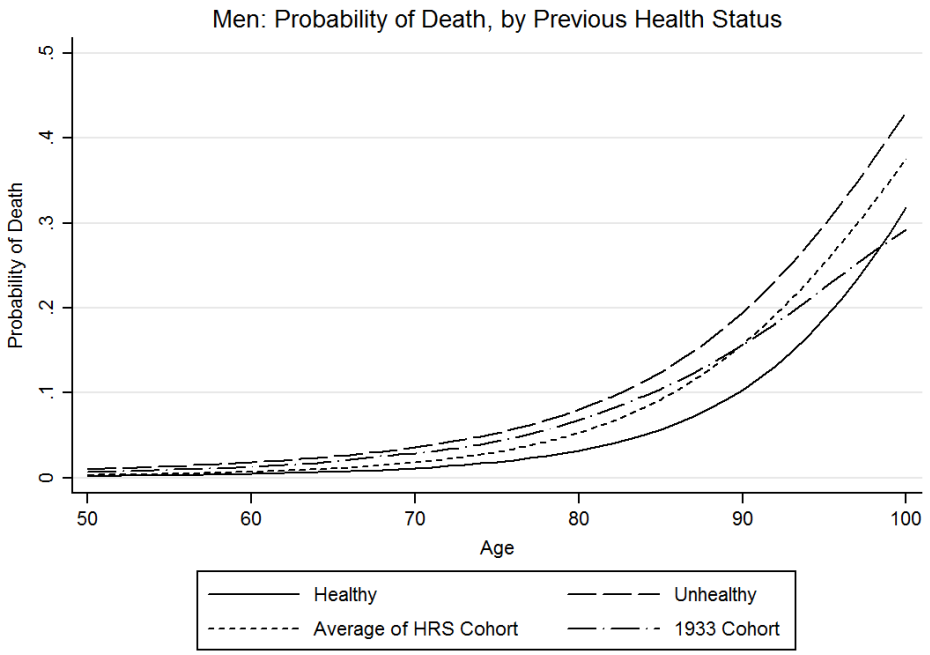
before. The combined sample is used to identify the distribution of medical expenses into old age.

I estimate the distribution of medical expense separately for ages above and below age 65, by regressing the logarithm of out-of-pocket medical expenses on a quadratic in age conditional on health insurance, labor force participation, and health status, which represent states of the structural model. Age 65 is chosen as a break-point since most individuals qualify for Medicare at this age and it becomes the primary insurer of the population above 65. As a result, the expense distribution can be expected to differ across groups on either side of age 65.

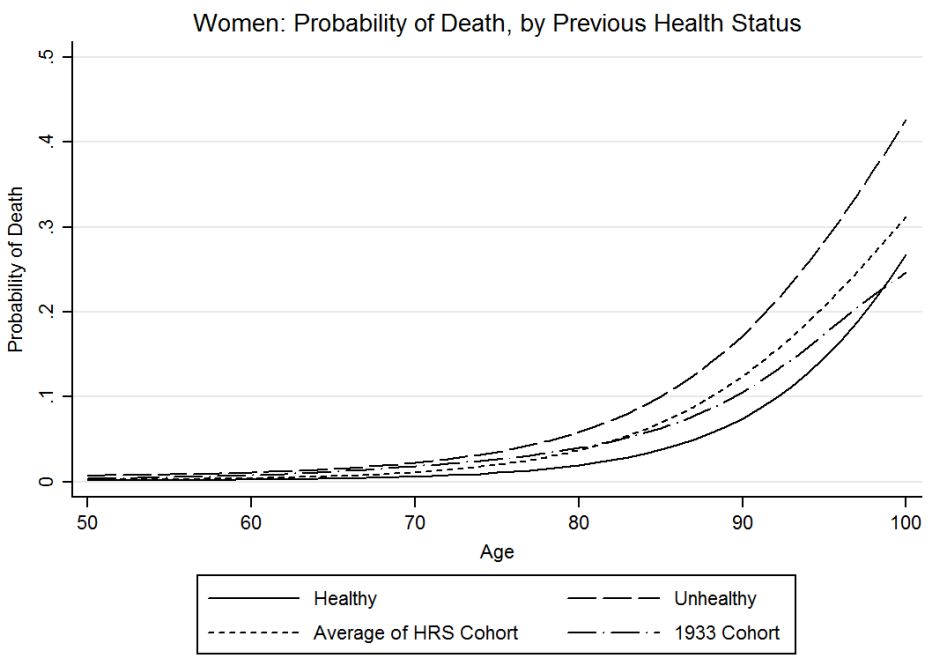
Previous work has used estimates of total medical expenses, and has generally used another data source for total medical expenditure because it is not observed by the HRS. I compare the distribution of $M_{i,t}$ to total medical expenditure found by Blau and Gilleskie (2006), who use an external survey - the 1987 National Medical Expenditure Survey. I observe that my medical expenditure estimates are generally lower at every level, particularly they are much lower at higher levels of medical expenditures. This is to be expected since they were attempting to estimate total medical expense, and health insurance limits catastrophic medical expenses.

Past literature that has included medical expense uncertainty has usually been focused on how health insurance alters retirement behavior. Due to computational limitations, I am unable to include a persistent process for medical expenses. Persistence in medical expenditures does exist indirectly through persistence in health status. I expect that this will lead to underestimating the household's lifetime medical expense risk.

Figure 15: Probability of Death by Sex
Conditional on previous health status



(A) Men



(B) Women

J Preference Types

As described in §5.1.4, households can vary based on characteristics that will be reflected in their preference for consumption versus leisure, but are not otherwise captured by the typical state variables. For this reason I include a preference index, as in Keane and Wolpin (1997), van der Klaauw and Wolpin (2008), and French and Jones (2011), to capture heterogeneity in preferences for consumption, own-leisure, spousal leisure, time, and household decision-making.

I construct my preference index by regressing each individual's labor force participation on a quartic in age, household health status, assets, earnings, health insurance status, the individual's AIME, defined benefit flow (if eligible), marital status, and a full set of interactions of these terms. Furthermore, I include in this regression three variables pertaining to the individual's preference for work:

1. Even if I didn't need the money, I would probably keep on working. (Agree or disagree)
2. When you think about the time when you and your husband or wife will retire, are you looking forward to it, are you uneasy about it, or what?
3. On a scale of 1 to 10, how much do you enjoy your job?

and, I include four more variables the pertain to the individual's preference for his or her spouse:

1. Generally speaking, would you say that the time you spend together with your husband or wife is extremely enjoyable, very enjoyable, somewhat enjoyable, or not too enjoyable?
2. When it comes to making major family decisions, who has the final say – you or your husband or wife?
3. Some couples like to spend their free time doing things together, while others like to do different things in their free time. What about you and your husband or wife? (together, separate, or sometimes together and sometimes separate)
4. I am going to read you a list of things that some people say are good about retirement. For each one, please tell me if, for you, they are very important, moderately important, somewhat important, or not important at all. Having more time with husband or wife.

For each of the above questions, I create a binary variable for each, either lumping answers such as agree and strongly agree together, or partitioning it by the median answer. I estimate the above regression separately for men and women. For each individual, the work preference index is the sum of the work preference coefficients multiplied by their respective independent variables, and similarly for the spouse preference index. The household's work or spouse preference index is simply the equally weighted sum for each household member's respective preference indices. The household preference indices are then converted into binary measures by partitioning them at each measures' median.

I observe that the work preference index is positively correlated with marriage, earnings, assets, AIME, defined-benefit pension flows, and negatively correlated with health. The spouse preference index is positively correlated with assets and health, but negatively correlated with earnings and AIME. An "out' preference index is created for households who were not asked the work questions in the first period because they were not working. As noted in table 1, the initial distribution consists of 17.4% of the "out" preference type, and then a relatively even distribution between the four other preference types.

K Additional Figures depicting Moment Matching

Figure 16: Asset Quantiles (by thirds) by Male Age

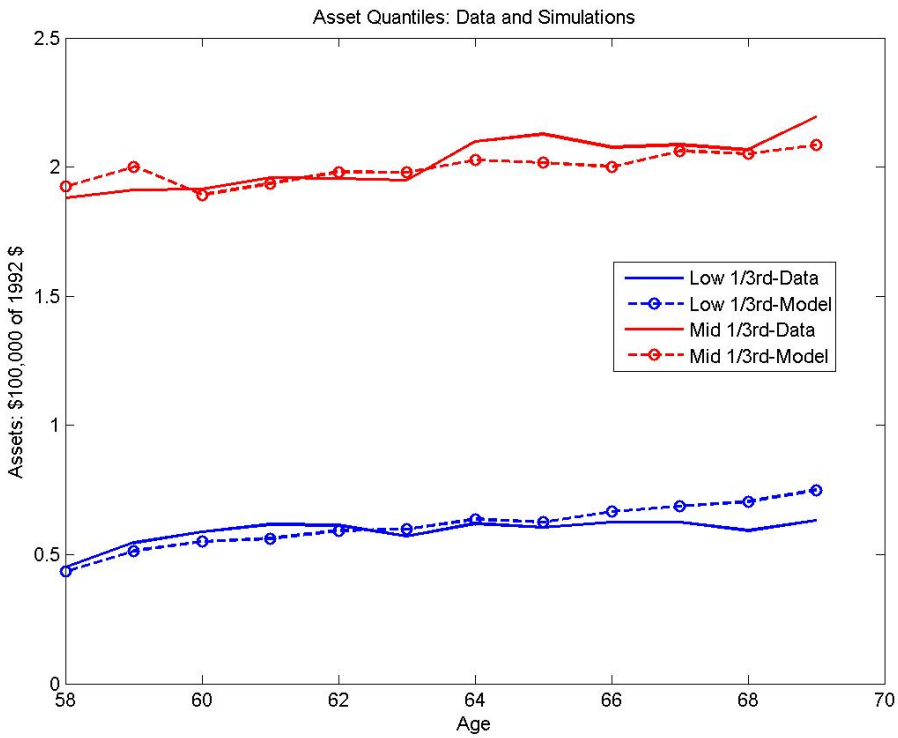
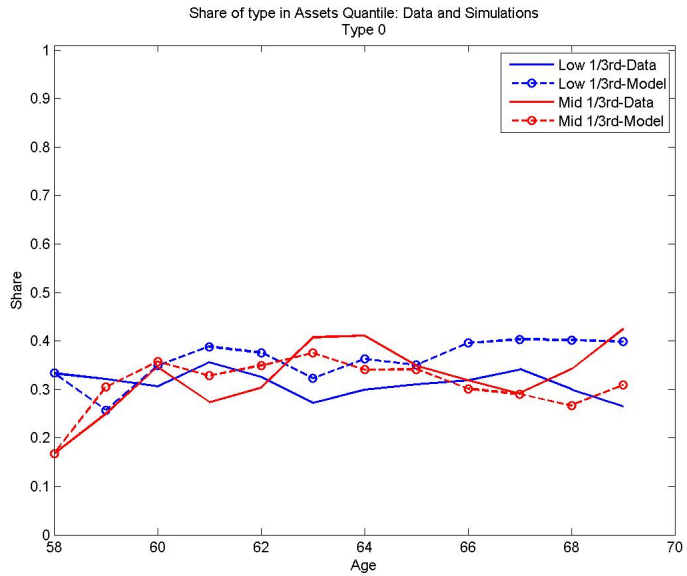
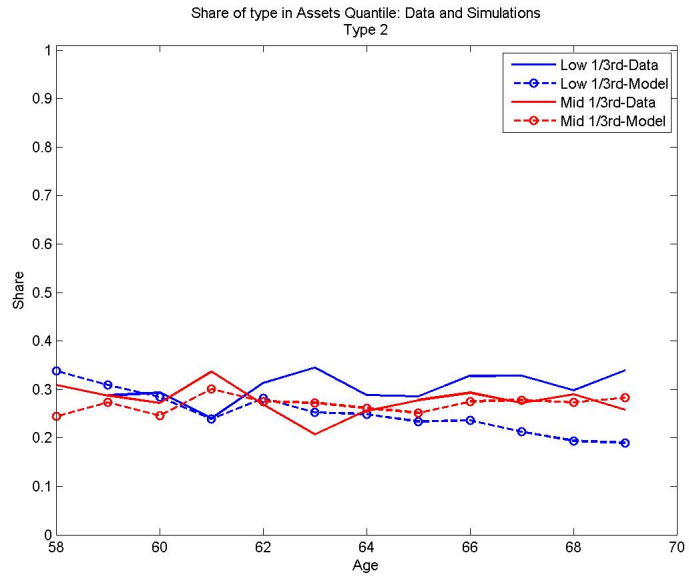


Figure 17: Asset Quantile Shares by Preference Type

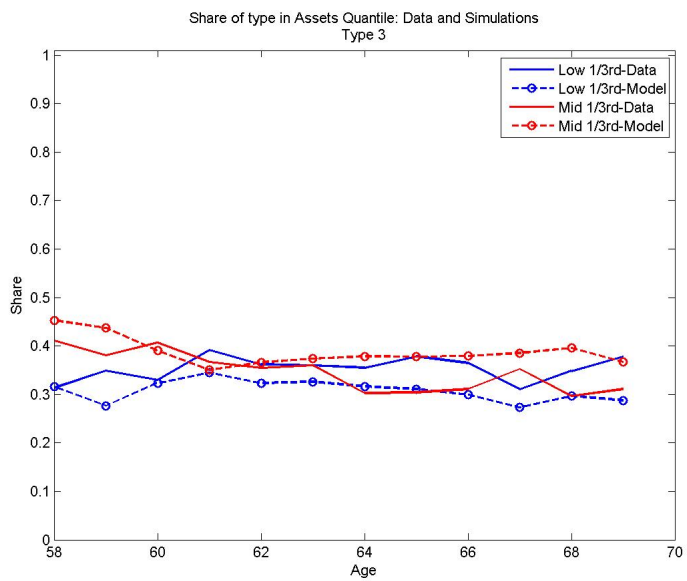
(a) Type 0
the Out Type



(b) Type 2
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(c) Type 3
High Preference for Own Leisure, High Preference for Spousal Leisure



(d) Type 4
Low Preference for Own Leisure, High Preference for Spousal Leisure

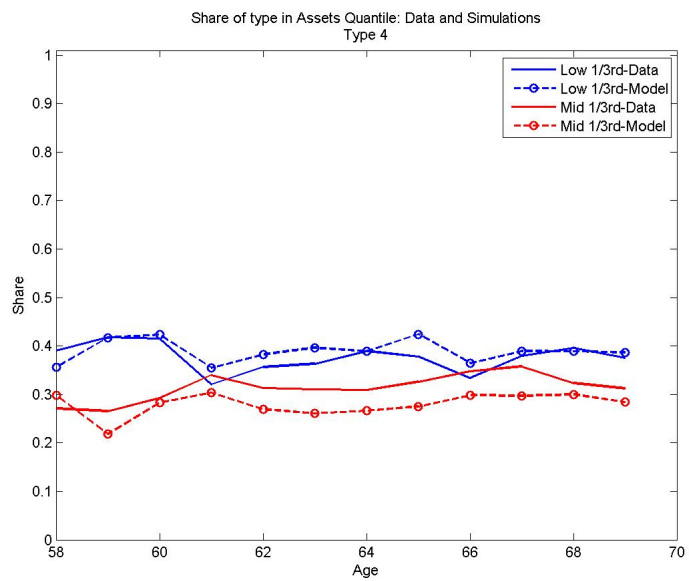
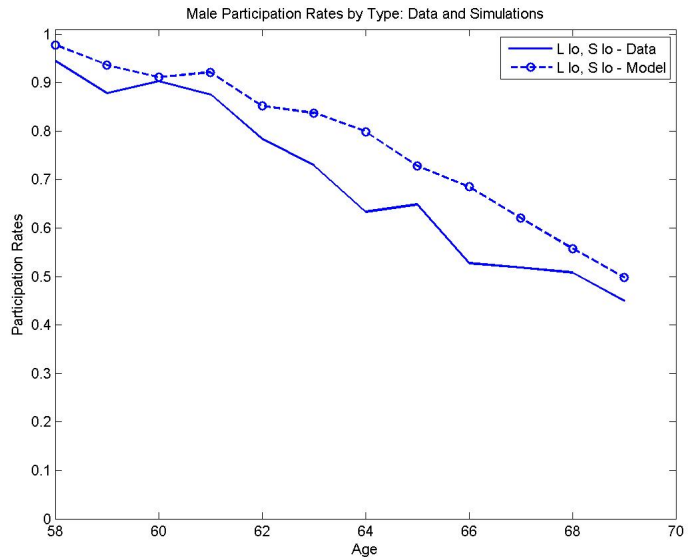
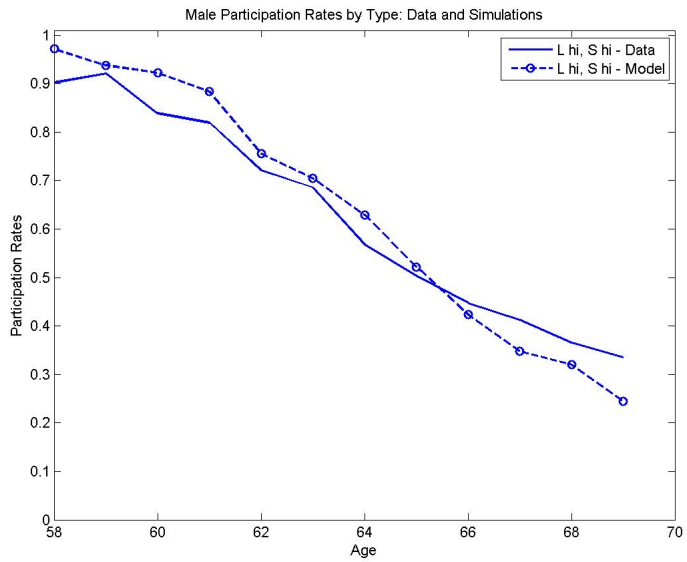


Figure 18: Men Labor Force Participation by Preference Type

(a) Type 2
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(b) Type 3
High Preference for Own Leisure, High Preference for Spousal Leisure



(c) Type 4
Low Preference for Own Leisure, High Preference for Spousal Leisure

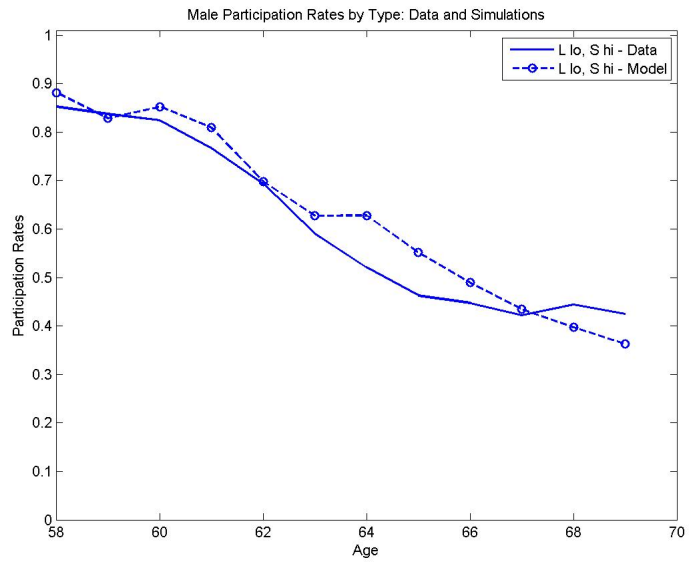
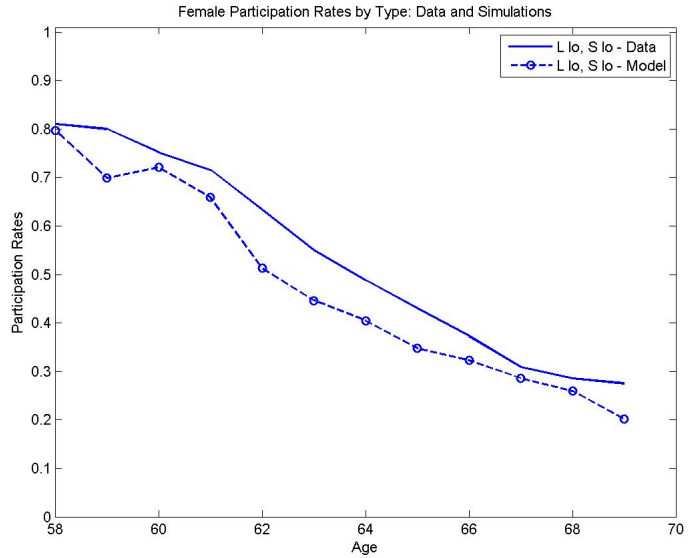
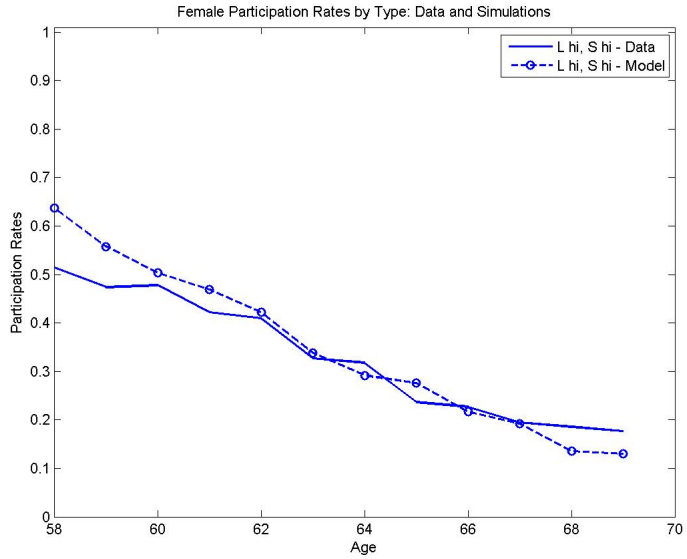


Figure 19: Women Labor Force Participation by Preference Type

(a) Type 2
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(b) Type 3
High Preference for Own Leisure, High Preference for Spousal Leisure



(c) Type 4
Low Preference for Own Leisure, High Preference for Spousal Leisure

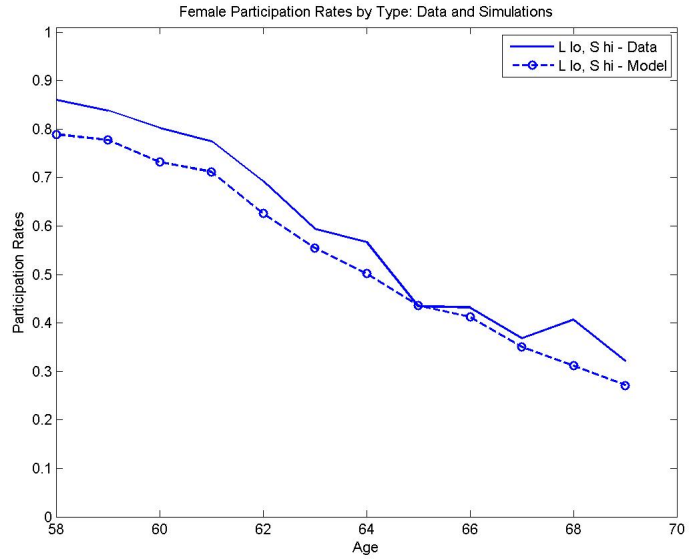
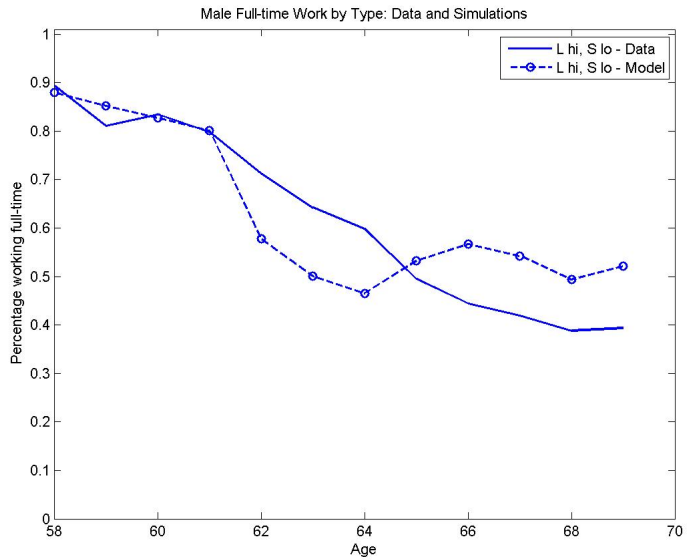
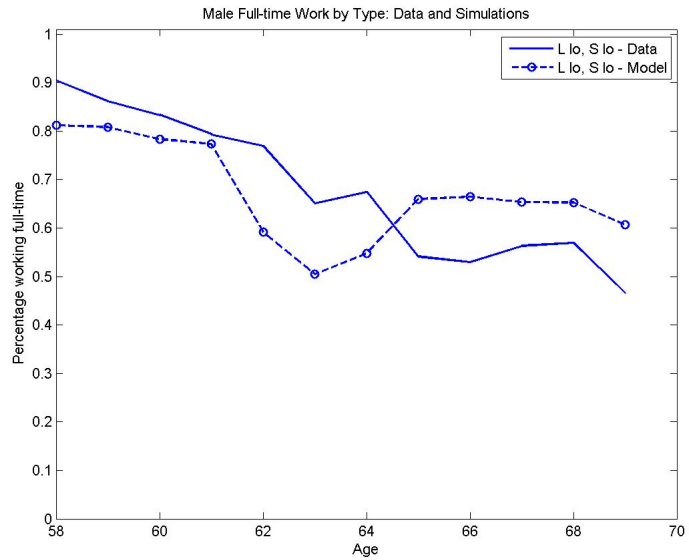


Figure 20: Men Full-time work by Preference Type

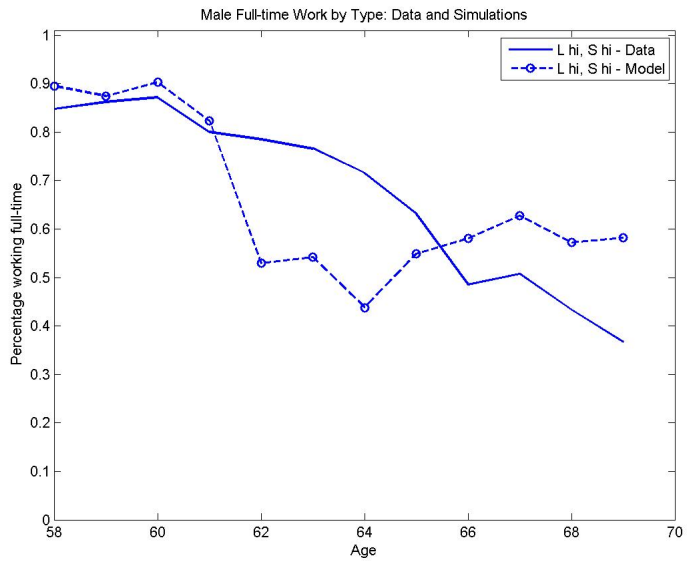
(a) Type 1
High Preference for Own Leisure, Low Preference for Spousal Leisure



(b) Type 2
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(c) Type 3
High Preference for Own Leisure, High Preference for Spousal Leisure



(d) Type 4
Low Preference for Own Leisure, High Preference for Spousal Leisure

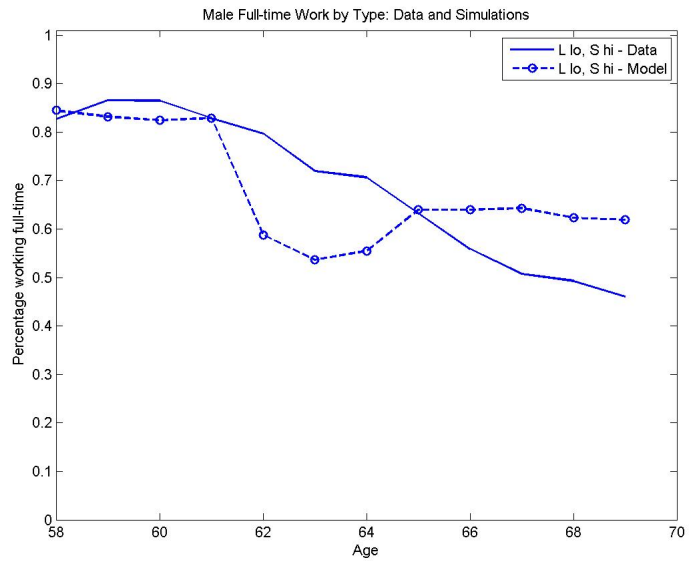
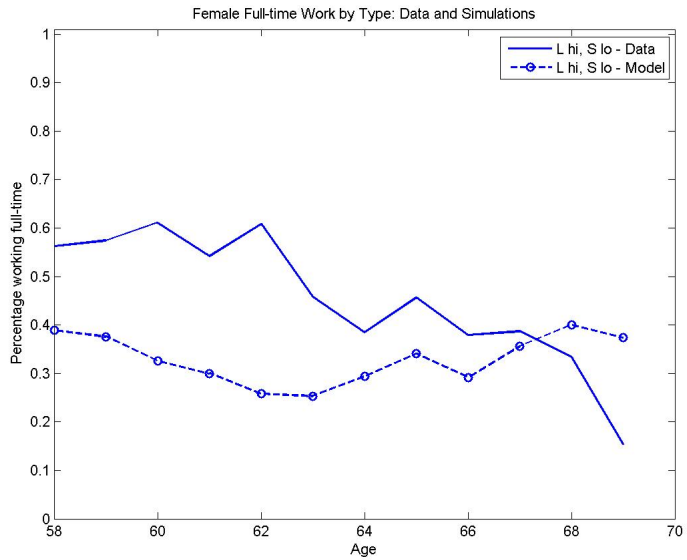
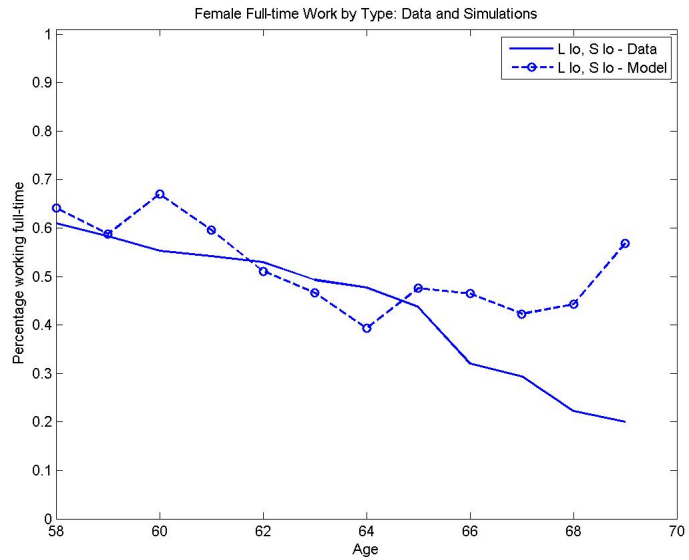


Figure 21: Women Full-time work by Preference Type

(a) Type 1
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(b) Type 2
Low Preference for Own Leisure, Low Preference for Spousal Leisure



(c) Type 3
High Preference for Own Leisure, High Preference for Spousal Leisure

(d) Type 4
Low Preference for Own Leisure, High Preference for Spousal Leisure

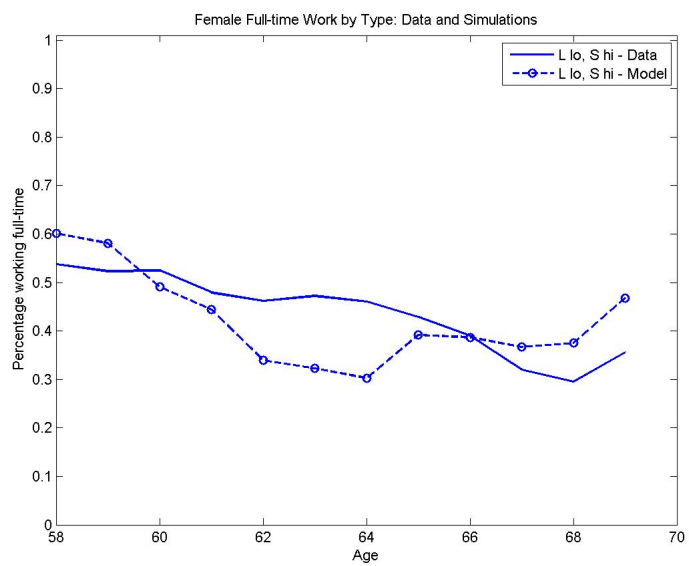
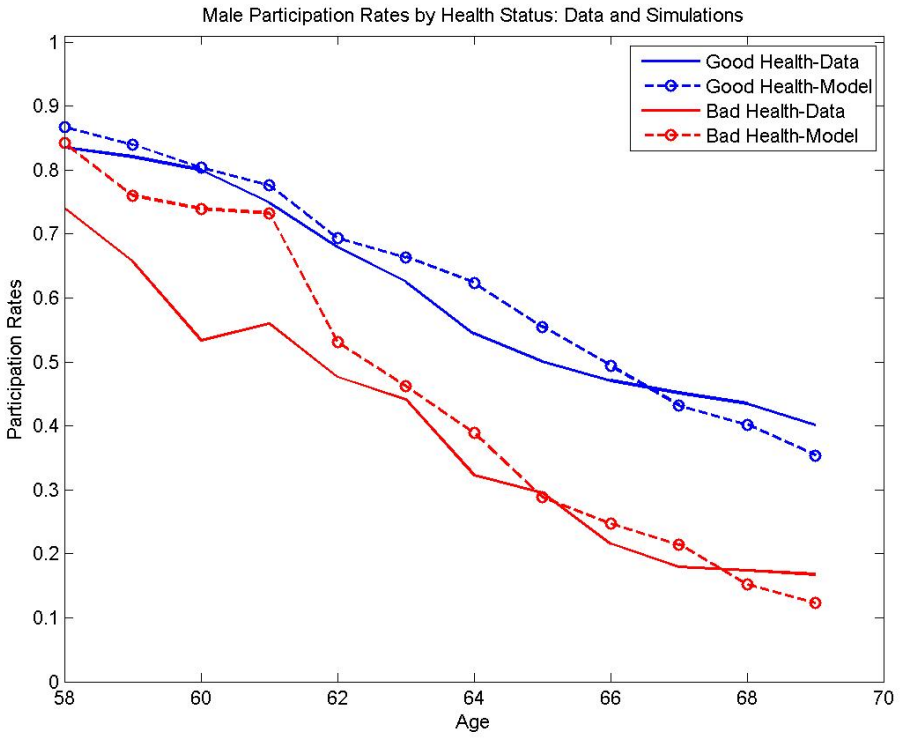


Figure 22: Participation Rate by Health Status

(a) Men



(b) Women