The Effects of Health Insurance and Self-Insurance on Retirement Behavior

Eric French
Federal Reserve Bank of Chicago

John Bailey Jones*
University at Albany, SUNY

June 10, 2010

Abstract

This paper provides an empirical analysis of the effects of employer-provided health insurance, Medicare, and Social Security on retirement behavior. Using data from the Health and Retirement Study, we estimate a dynamic programming model of retirement that accounts for both saving and uncertain medical expenses. Our results suggest that Medicare is important for understanding retirement behavior, and that uncertainty and saving are both important for understanding the labor supply responses to Medicare. Half the value placed by a typical worker on his employer-provided health insurance is the value of reduced medical expense risk. Raising the Medicare eligibility age from 65 to 67 leads individuals to work an additional 0.074 years over ages 60-69. In comparison, eliminating two years worth of Social Security benefits increases years of work by 0.076 years.

*Comments welcome at `efrench@frbchi.org` and `jbjones@albany.edu`. We thank Joe Altonji, Peter Arcidiacono, Gadi Barlevy, David Blau, John Bound, Chris Carroll, Mariacristina De Nardi, Tim Erikson, Hanming Fang, Donna Gilleskie, Lars Hansen, John Kennan, Spencer Krane, Hamp Lankford, Guy Laroque, John Rust, Dan Sullivan, Chris Taber, the editors and referees, students of Econ 751 at Wisconsin, and participants at numerous seminars for helpful comments. We received advice on the HRS pension data from Gary Englehardt and Tom Steinmeier, and excellent research assistance from Kate Anderson, Olesya Baker, Diwakar Choubey, Phil Doctor, Ken Housinger, Kirti Kamboj, Tina Lam, Kenley Peltzer, and Santadarshan Sadhu. The research reported herein was supported by the Center for Retirement Research at Boston College (CRR) and the Michigan Retirement Research Center (MRRC) pursuant to grants from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium. The opinions and conclusions are solely those of the authors, and should not be construed as representing the opinions or policy of the SSA or any agency of the Federal Government, the CRR, the MRRC, or the Federal Reserve System. Recent versions of the paper can be obtained at `http://www.albany.edu/~jbjones/papers.htm`.

1 Introduction

One of the largest social programs for the rapidly growing elderly population is Medicare. In 2008, Medicare had 44.1 million beneficiaries and $481 billion of expenditures, making it only slightly smaller than Social Security.[1]

Prior to receiving Medicare at age 65, many individuals receive health insurance only if they continue to work. This work incentive disappears at age 65, when Medicare provides health insurance to almost everyone. An important question, therefore, is whether Medicare significantly affects the labor supply of the elderly. This question is crucial when considering Medicare reforms; the fiscal effects of such reforms depend on how labor supply responds. However, there is relatively little research on the labor supply responses to Medicare.

This paper provides an empirical analysis of the effect of employer-provided health insurance and Medicare in determining retirement behavior. Using data from the Health and Retirement Study, we estimate a dynamic programming model of retirement that accounts for both saving and uncertain medical expenses. Our results suggest that Medicare is important for understanding retirement behavior, because it insures against medical expense shocks that can exhaust a household's savings.

Our work builds upon, and in part reconciles, several earlier studies. Assuming that individuals value health insurance at the cost paid by employers, Lumsdaine et al. (1994) and Gustman and Steinmeier (1994) find that health insurance has a small effect on retirement behavior. One possible reason for their results is that they find that the average employer contribution to health insurance is modest, and declines by only a small amount after age 65. If workers are risk-averse, however, and if health insurance allows them to smooth consumption when facing volatile medical expenses, they could value employer-provided health insurance well beyond the cost paid by employers. Medicare's age-65 work disincentive thus comes not only from the reduction in average medical costs paid by those without employer-provided health insurance, but also from the reduction in the volatility of those costs.

[1] Figures taken from 2009 Medicare Annual Report (The Boards of Trustees of the Hospital Insurance and Supplementary Medical Insurance Trust Funds, 2009).

Addressing this point, Rust and Phelan (1997) and Blau and Gilleskie (2006, 2008) estimate dynamic programming models that account explicitly for risk aversion and uncertainty about out-of-pocket medical expenses. Their estimated labor supply responses to health insurance are larger than those found in studies that omit medical expense risk. Rust and Phelan and Blau and Gilleskie, however, assume that an individual's consumption equals his income net of out-of-pocket medical expenses. In other words, they ignore an individual's ability to smooth consumption through saving. If individuals can self-insure against medical expense shocks by saving, prohibiting saving will overstate the consumption volatility caused by medical cost volatility. It is therefore likely that Rust and Phelan and Blau and Gilleskie overstate the value of health insurance, and thus the effect of health insurance on retirement.

In this paper we construct a life-cycle model of labor supply that not only accounts for medical expense uncertainty and health insurance, but also has a saving decision. Moreover, we include the coverage provided by means-tested social insurance to account for the fact that Medicaid provides a substitute for other forms of health insurance. To our knowledge, ours is the first study of its kind. While van der Klaauw and Wolpin (2008) and Casanova (2010) also estimate retirement models that account for both savings and uncertain medical expenses, they do not focus on the role of health insurance, and thus use much simpler models of medical expenses.

Almost everyone becomes eligible for Medicare at age 65. However, the Social Security system and pensions also provide retirement incentives at age 65. This makes it difficult to determine whether the high job exit rates observed at age 65 are due to Medicare, Social Security, or pensions. One way we address this problem is to exploit variation in employer-provided health insurance. Some individuals receive employer-provided health insurance only while they work, so that their coverage is *tied* to their job. Other individuals have *retiree* coverage, and receive employer-provided health insurance even if they retire. If workers value access to health insurance, those with *retiree* coverage should be more willing to retire before age 65. Our data show that individuals with *retiree* coverage tend to retire about a half year earlier than individuals with *tied* coverage. This suggests that employer-provided health

insurance is a determinant of retirement.

One problem with using employer-provided health insurance to identify Medicare's effect on retirement is that individuals may choose to work for a firm because of its post-retirement benefits. The fact that early retirement is common for individuals with *retiree* coverage may not reflect the effect of health insurance on retirement. Instead, individuals with preferences for early retirement may be self-selecting into jobs that provide *retiree* coverage. To address this issue, we measure self-selection into jobs with different health insurance plans. We allow the value of leisure and the time discount factor to vary across individuals. Modelling preference heterogeneity with the approach used by Keane and Wolpin (1997), we find that individuals with strong preferences for leisure are more likely to work for firms that provide *retiree* health insurance. However, self-selection does not affect our main results.

Estimating the model by the Method of Simulated Moments, we find that the model fits the data well with reasonable parameter values. Next, we simulate the labor supply response to changing some of the Medicare and Social Security retirement program rules. Raising the Medicare eligibility age from 65 to 67 would increase years worked by 0.074 years. Eliminating two years worth of Social Security benefits would increase years worked by 0.076 years. Thus, even after allowing for both saving and self-selection into health insurance plans, the effect of Medicare on labor supply is as large as the effect of Social Security. One reason why we find that Medicare is important is that we find that medical expense risk is important. Even when we allow individuals to save, they value the consumption smoothing benefits of health insurance. We find that about half the value a typical worker places on his employer-provided health insurance comes from these benefits.

The rest of paper proceeds as follows. Section 2 develops our dynamic programming model of retirement behavior. Section 3 describes how we estimate the model using the Method of Simulated Moments. Section 4 describes the HRS data that we use in our analysis. Section 5 presents life cycle profiles drawn from these data. Section 6 contains preference parameter estimates for the structural model, and an assessment of the model's performance, both within and outside of the estimation sample. In Section 7, we conduct several policy experiments.

In Section 8 we consider a few robustness checks. Section 9 concludes.

2 The Model

In order to capture the richness of retirement incentives, our model is very complex and has many parameters. Appendix A provides definitions for all the variables used in the main text.

2.1 Preferences and Demographics

Consider a household head seeking to maximize his expected discounted (where the subjective discount factor is β) lifetime utility at age t, $t = 59, 60, ..., 95$. Each period that he lives, the individual derives utility from consumption, C_t, and hours of leisure, L_t. The within-period utility function is of the form

$$U(C_t, L_t) = \frac{1}{1-\nu}\left(C_t^{\gamma} L_t^{1-\gamma}\right)^{1-\nu}. \tag{1}$$

We allow both β and γ to vary across individuals. Individuals with higher values of β are more patient, while individuals with higher values of γ place less weight on leisure.

The quantity of leisure is

$$L_t = L - N_t - \phi_{Pt} P_t - \phi_{RE} RE_t - \phi_H H_t, \tag{2}$$

where L is the individual's total annual time endowment. Participation in the labor force is denoted by P_t, a 0-1 indicator equal to one when hours worked, N_t, are positive. The fixed cost of work, ϕ_{Pt}, is treated as a loss of leisure. Including fixed costs helps us capture the empirical regularity that annual hours of work are clustered around 2000 hours and 0 hours (Cogan, 1981). Following a number of studies,[2] we allow preferences for leisure, in our case the value of ϕ_{Pt}, to increase linearly with age. Workers that leave the labor force

[2]Examples include Rust and Phelan (1997), Blau and Gilleskie (2006, 2008), Gustman and Steinmeier (2005), Rust et al. (2003), and van der Klaauw and Wolpin (2008).

can re-enter; re-entry is denoted by the 0-1 indicator $RE_t = 1\{P_t = 1 \text{ and } P_{t-1} = 0\}$, and individuals re-entering the labor market incur the cost ϕ_{RE}. The quantity of leisure also depends on an individual's health status through the 0-1 indicator $H_t = 1\{health_t = bad\}$, which equals one when his health is bad.

Workers alive at age t survive to age $t+1$ with probability s_{t+1}. Following De Nardi (2004), workers that die value bequests of assets, A_t, according to the function $b(A_t)$:

$$b(A_t) = \theta_B \frac{\left(A_t + \kappa\right)^{(1-\nu)\gamma}}{1-\nu}. \tag{3}$$

The survival probability s_t, along with the transition probabilities for the health variable H_t, depend on age and previous health status.

2.2 Budget Constraints

The individual holds three forms of wealth: assets (including housing); pensions; and Social Security. He has several sources of income: asset income, rA_t, where r denotes the constant pre-tax interest rate; labor income, $W_t N_t$, where W_t denotes wages; spousal income, ys_t; pension benefits, pb_t; Social Security benefits, ss_t; and government transfers, tr_t. The asset accumulation equation is

$$A_{t+1} = A_t + Y_t + ss_t + tr_t - M_t - C_t. \tag{4}$$

M_t denotes medical expenses. Post-tax income, $Y_t = Y(rA_t + W_t N_t + ys_t + pb_t, \tau)$, is a function of taxable income and the vector τ, described in Appendix B, that captures the tax structure.

Individuals face the borrowing constraint

$$A_t + Y_t + ss_t + tr_t - C_t \geq 0. \tag{5}$$

Because it is illegal to borrow against future Social Security benefits and difficult to borrow

against many forms of future pension benefits, individuals with low non-pension, non-Social Security wealth may not be able to finance their retirement before their Social Security benefits become available at age 62 (Kahn, 1988; Rust and Phelan, 1997; Gustman and Steinmeier, 2005).[3]

Following Hubbard et al. (1994, 1995), government transfers provide a consumption floor:

$$tr_t = max\{0, C_{min} - (A_t + Y_t + ss_t)\}. \tag{6}$$

Equation (6) implies that government transfers bridge the gap between an individual's "liquid resources" (the quantity in the inner parentheses) and the consumption floor. Treating C_{min} as a sustenance level, we further require that $C_t \geq C_{min}$. Our treatment of government transfers implies that individuals will always consume at least C_{min}, even if their out-of-pocket medical expenses exceed their financial resources.

2.3 Medical Expenses, Health Insurance, and Medicare

We define M_t as the sum of all out-of-pocket medical expenses, including insurance premia and expenses covered by the consumption floor. We assume that an individual's medical expenses depend upon five components. First, medical expenses depend on the individual's employer-provided health insurance, I_t. Second, they depend on whether the person is working, P_t, because workers who leave their job often pay a larger fraction of their insurance premiums. Third, they depend on the individual's self-reported health status, H_t. Fourth, medical expenses depend on age. At age 65, individuals become eligible for Medicare, which is a close substitute for employer-provided coverage.[4] Offsetting this, as people age their health declines (in a way not captured by H_t), raising medical expenses. Finally, medical

[3]We assume time-t medical expenses are realized after time-t labor decisions have been made. We view this as preferable to the alternative assumption that the time-t medical expense shocks are fully known when workers decide whether to hold on to their employer-provided health insurance. Given the borrowing constraint and timing of medical expenses, an individual with extremely high medical expenses this year could have negative net worth next year. Because many people in our data have unresolved medical expenses, medical expense debt seems reasonable.

[4]Individuals who have paid into the Medicare system for at least 10 years become eligible at age 65. A more detailed description of the Medicare eligibility rules is available at `http://www.medicare.gov/`.

expenses depend on the person-specific component ψ_t, yielding:

$$\ln M_t = m(H_t, I_t, t, P_t) + \sigma(H_t, I_t, t, P_t) \times \psi_t. \tag{7}$$

Note that health insurance affects both the expectation of medical expenses, through $m(.)$ and the variance, through $\sigma(.)$

Even after controlling for health status, French and Jones (2004a) find that medical expenses are very volatile and persistent. Thus we model the person-specific component of medical expenses, ψ_t, as

$$\psi_t = \zeta_t + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2), \tag{8}$$

$$\zeta_t = \rho_m \zeta_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2), \tag{9}$$

where ξ_t and ϵ_t are serially and mutually independent. ξ_t is the transitory component, while ζ_t is the persistent component, with autocorrelation ρ_m.

We assume that medical expenditures are exogenous. It is not clear *ex ante* whether this causes us to understate or overstate the importance of health insurance. On the one hand, individuals with health insurance receive better care. Our model does not capture this benefit, and in this respect understates the value of health insurance. Conversely, treating medical expenses as exogenous ignores the ability of workers to offset medical shocks by adjusting their expenditures on medical care. This leads us to overstate the consumption risk facing uninsured workers, and thus the value of health insurance. Evidence from other structural analyses suggests that our assumption of exogeneity leads us to overstate the effect of health insurance on retirement.[5]

[5]To our knowledge, Blau and Gilleskie (2008) is the only estimated, structural retirement study to have endogenous medical expenditures. Although Blau and Gilleskie (2008) do not discuss how their results would change if medical expenses were treated as exogenous, they find that even with several mechanisms (such as prescription drug benefits) omitted, health insurance has "a modest impact on employment behavior among older males". De Nardi, French and Jones (2010) study the saving behavior of retirees. They find that the effects of reducing means-tested social insurance are smaller when medical care is endogenous, rather than exogenous. They also find, however, that even when medical expenditures are a choice variable, they are a major reason why the elderly save.

Differences in labor supply behavior across health insurance categories are an integral part of identifying our model. We assume that there are three mutually exclusive categories of health insurance coverage. The first is *retiree* coverage, where workers keep their health insurance even after leaving their jobs. The second category is *tied* health insurance, where workers receive employer-provided coverage as long as they continue to work. If a worker with *tied* health insurance leaves his job, he can keep his health insurance coverage for that year. This is meant to proxy for the fact that most firms must provide "COBRA" health insurance to workers after they leave their job. After one year of *tied* coverage and not working, the individual's insurance ceases.[6] The third category consists of individuals whose potential employers provide no health insurance at all, or *none*. Workers move between these insurance categories according to

$$I_t = \begin{cases} retiree & \text{if} \quad I_{t-1} = retiree \\ tied & \text{if} \quad I_{t-1} = tied \quad \text{and} \quad N_{t-1} > 0 \\ none & \text{if} \quad I_{t-1} = none \quad \text{or} \quad (I_{t-1} = tied \quad \text{and} \quad N_{t-1} = 0) \end{cases} . \tag{10}$$

2.4 Wages and Spousal Income

We assume that the logarithm of wages at time t, $\ln W_t$, is a function of health status (H_t), age (t), hours worked (N_t) and an autoregressive component, ω_t:

$$\ln W_t = W(H_t, t) + \alpha \ln N_t + \omega_t. \tag{11}$$

The inclusion of hours, N_t, in the wage determination equation captures the empirical regularity that, all else equal, part-time workers earn relatively lower wages than full time workers. The autoregressive component ω_t has the correlation coefficient ρ_W and the

[6] Although there is some variability across states as to how long individuals are eligible for employer-provided health insurance coverage, by Federal law most individuals are covered for 18 months (Gruber and Madrian, 1995). Given a model period of one year, we approximate the 18-month period as one year. We do not model the option to take up COBRA, assuming that the take-up rate is 100%. Although the actual take-up rate is around $\frac{2}{3}$ (Gruber and Madrian, 1996), we simulated the model assuming that the rate was 0%, so that individuals transitioned from *tied* to *none* as soon as they stopped working, and found very similar labor supply patterns. Thus assuming a 100% take-up rate does not seem to drive our results.

normally-distributed innovation η_t:

$$\omega_t = \rho_W \omega_{t-1} + \eta_t, \quad \eta_t \sim N(0, \sigma_\eta^2). \tag{12}$$

Because spousal income can serve as insurance against medical shocks, we include it in the model. In the interest of computational simplicity, we assume that spousal income is a deterministic function of an individual's age and health status:

$$ys_t = ys(H_t, t). \tag{13}$$

2.5 Social Security and Pensions

Because pensions and Social Security generate potentially important retirement incentives, we model the two programs in detail.

Individuals receive no Social Security benefits until they apply. Individuals can first apply for benefits at age 62. Upon applying the individual receives benefits until death. The individual's Social Security benefits depend on his Average Indexed Monthly Earnings ($AIME$), which is roughly his average income during his 35 highest earnings years in the labor market.

The Social Security System provides three major retirement incentives.[7] First, while income earned by workers with less than 35 years of earnings automatically increases their $AIME$, income earned by workers with more than 35 years of earnings increases their $AIME$ only if it exceeds earnings in some previous year of work. Because Social Security benefits increase in $AIME$, this causes work incentives to drop after 35 years in the labor market. We describe the computation of $AIME$ in more detail in Appendix D.

Second, the age at which the individual applies for Social Security affects the level of benefits. For every year before age 65 the individual applies for benefits, benefits are reduced

[7] A description of the Social Security rules can be found in recent editions of the *Green Book* (Committee on Ways and Means). Some of the rules, such as the benefit adjustment formula, depend on an individual's year of birth. Because we fit our model to a group of individuals that on average were born in 1933, we use the benefit formula for that birth year.

by 6.67% of the age-65 level. This is roughly actuarially fair. But for every year after age 65 that benefit application is delayed, benefits rise by 5.5% up until age 70. This is less than actuarially fair, and encourages people to apply for benefits by age 65.

Third, the Social Security Earnings Test taxes labor income of beneficiaries at a high rate. For individuals aged 62-64, each dollar of labor income above the "test" threshold of $9,120 leads to a 1/2 dollar decrease in Social Security benefits, until all benefits have been taxed away. For individuals aged 65-69 before 2000, each dollar of labor income above a threshold of $14,500 leads to a 1/3 dollar decrease in Social Security benefits, until all benefits have been taxed away. Although benefits taxed away by the earnings test are credited to future benefits, after age 64 the crediting rate is less than actuarially fair, so that the Social Security Earnings Test effectively taxes the labor income of beneficiaries aged 65-69.[8] When combined with the aforementioned incentives to draw Social Security benefits by age 65, the Earnings Test discourages work after age 65. In 2000, the Social Security Earnings Test was abolished for those 65 and older. Because those born in 1933 (the average birth year in our sample) turned 67 in 2000, we assume that the earnings test was repealed at age 67. These incentives are incorporated in the calculation of ss_t, which is defined to be net of the earnings test.

Pension benefits, pb_t, are a function of the worker's age and pension wealth. Pension wealth (the present value of pension benefits) in turn depends on pension accruals. We assume that pension accruals are a function of a worker's age, labor income, and health insurance type, using a formula estimated from confidential HRS pension data. The data show that pension accrual rates differ greatly across health insurance categories; accounting for these differences is essential in isolating the effects of employer-provided health insurance. When finding an individual's decision rules, we assume further that the individual's existing pension wealth is a function of his Social Security wealth, age, and health insurance type. Details of our pension model are described in Section 4.3 and Appendix C.

[8]The credit rates are based on the benefit adjustment formula. If a year's worth of benefits are taxed away between ages 62 and 64, benefits in the future are increased by 6.67%. If a year's worth of benefits are taxed away between ages 65 and 66, benefits in the future are increased by 5.5%.

2.6 Recursive Formulation

In addition to choosing hours and consumption, eligible individuals decide whether to apply for Social Security benefits; let the indicator variable $B_t \in \{0,1\}$ equal one if an individual has applied. In recursive form, the individual's problem can be written as

$$V_t(X_t) = \max_{C_t,N_t,B_t} \left\{ \frac{1}{1-\nu}\left(C_t^{\gamma}(L - N_t - \phi_{Pt}P_t - \phi_{RE}RE_t - \phi_H H_t)^{1-\gamma}\right)^{1-\nu} + \beta(1-s_{t+1})b(A_{t+1}) \right.$$
$$\left. + \beta s_{t+1} \int V_{t+1}(X_{t+1}) dF(X_{t+1}|X_t, t, C_t, N_t, B_t) \right\}, \tag{14}$$

subject to equations (5) and (6). The vector $X_t = (A_t, B_{t-1}, H_t, AIME_t, I_t, P_{t-1}, \omega_t, \zeta_{t-1})$ contains the individual's state variables, while the function $F(\cdot|\cdot)$ gives the conditional distribution of these state variables, using equations (4) and (7) - (13).[9] The solution to the individual's problem consists of the consumption rules, work rules, and benefit application rules that solve equation (14). These decision rules are found numerically using value function iteration. Appendix E describes our numerical methodology.

3 Estimation

To estimate the model, we adopt a two-step strategy, similar to the one used by Gourinchas and Parker (2002) and French (2005). In the first step we estimate or calibrate parameters that can be cleanly identified identified without explicitly using our model. For example, we estimate mortality rates and health transitions straight from demographic data. In the second step, we estimate the preference parameters of the model, as well as the consumption floor, using the method of simulated moments (MSM).

[9] Spousal income and pension benefits (see Appendix C) depend only on the other state variables and are thus not state variables themselves.

3.1 Moment Conditions

The objective of MSM estimation is to find the preference vector that yields simulated life-cycle decision profiles that "best match" (as measured by a GMM criterion function) the profiles from the data. The moment conditions that comprise our estimator are:

1. Because an individual's ability to self-insure against medical expense shocks depends upon his asset level, we match 1/3rd and 2/3rd asset quantiles by age. We match these quantiles in each of T periods (ages), for a total of $2T$ moment conditions.

2. We match job exit rates by age for each health insurance category. With three health insurance categories (*none*, *retiree* and *tied*), this generates $3T$ moment conditions.

3. Because the value a worker places on employer-provided health insurance may depend on his wealth, we match labor force participation conditional on the combination of asset quantile and health insurance status. With 2 quantiles (generating 3 quantile-conditional means) and 3 health insurance types, this generates $9T$ moment conditions.

4. To help identify preference heterogeneity, we utilize a series of questions in the HRS that ask workers about their preferences for work. We combine the answers to these questions into a time-invariant index, $pref \in \{high, low, out\}$, which is described in greater detail in Section 4.4. Matching participation conditional on each value of this index generates another $3T$ moment conditions.

5. Finally, we match hours of work and participation conditional on our binary health indicator. This generates $4T$ moment conditions.

Combined, the five preceding items result in $21T$ moment conditions. Appendix F provides a detailed description of the moment conditions, the mechanics of our MSM estimator, the asymptotic distribution of our parameter estimates, and our choice of weighting matrix.

3.2 Initial Conditions and Preference Heterogeneity

A key part of our estimation strategy is to compare the behavior of individuals with different forms of employer-provided health insurance. If access to health insurance is an important factor in the retirement decision, we should find that individuals with *tied* coverage retire later than those with *retiree* coverage. In making such a comparison, however, we must account for the possibility that individuals with different health insurance options differ systematically along other dimensions as well. For example, individuals with retiree coverage tend to have higher wages and more generous pensions.

We control for this "initial conditions" problem in three ways. First, the initial distribution of simulated individuals is drawn directly from the data. Because households with retiree coverage are more likely to be wealthy in the data, households with retiree coverage are more likely to be wealthy in our initial distribution. Similarly, in our initial distribution households with high levels of education are more likely to have high values of the persistent wage shock ω_t.

Second, we model carefully the way in which pension and Social Security accrual varies across individuals and groups.

Finally, we control for unobservable differences across health insurance groups by introducing permanent preference heterogeneity, using the approach introduced by Heckman and Singer (1984) and adapted by (among others) Keane and Wolpin (1997) and van der Klaauw and Wolpin (2008). Each individual is assumed to belong to one of a finite number of preference "types", with the probability of belonging to a particular type a logistic function of the individual's initial state vector: his age, wealth, initial wages, health status, health insurance type, medical expenditures, and preference index.[10] We estimate the type probability parameters jointly with the preference parameters and the consumption floor.

[10]These discrete type-based differences are the only preference heterogeneity in our model. For this reason many individuals in the data make decisions different from what the model would predict. Our MSM procedure circumvents this problem by using moment conditions that average across many individuals. One way to reconcile model predictions with individual observations is to introduce measurement error. In earlier drafts of this paper (French and Jones, 2004b) we considered this possibility by estimating a specification where we allowed for measurement error in assets. Adding measurement error, however, had little effect on either the preference parameter estimates or policy experiments, and we dropped this case.

In our framework, correlations between preferences and health insurance emerge because people with different preferences systematically select jobs with different types of health insurance coverage. Workers in our data set are first observed in their fifties; by this age, all else equal, jobs that provide generous post-retirement health insurance are more likely to be held by workers that wish to retire early. One way to measure this self-selection is to structurally model the choice of health insurance at younger ages, and use the predictions of that model to infer the correlation between preferences and health insurance in the first wave of the HRS. Because such an approach is computationally expensive, we instead model the correlation between preferences and health insurance in the initial conditions.

3.3 Wage Selection

We estimate a selection-adjusted wage profile using the procedure developed in French (2005). First, we estimate a fixed effects wage profile from HRS data, using the wages observed for individuals who are working. The fixed-effects estimator is identified using wage growth for workers. If wage growth rates for workers and non-workers are the same, composition bias problems—the question of whether high wage individuals drop out of the labor market later than low wage individuals—are not a problem. However, if individuals leave the market because of a wage drop, such as from job loss, then wage growth rates for workers will be greater than wage growth for non-workers. This selection problem will bias estimated wage growth upward.

We control for selection bias by finding the wage profile that, when fed into our model, generates the same fixed effects profile as the HRS data. Because the simulated fixed effect profiles are computed using only the wages of those simulated agents that work, the profiles should be biased upwards for the same reasons they are in the data. We find this bias-adjusted wage profile using the iterative procedure described in French (2005).

4 Data and Calibrations

4.1 HRS Data

We estimate the model using data from the Health and Retirement Survey (HRS). The HRS is a sample of non-institutionalized individuals, aged 51-61 in 1992, and their spouses. With the exception of assets and medical expenses, which are measured at the household level, our data are for male household heads. The HRS surveys individuals every two years, so that we have 8 waves of data covering the period 1992-2006. The HRS also asks respondents retrospective questions about their work history that allow us to infer whether the individual worked in non-survey years. Details of this, as well as variable definitions, selection criteria, and a description of the initial joint distribution, are in Appendix G.

As noted above, the Social Security rules depend on an individual's year of birth. To ensure that workers in our sample face a similar set of Social Security retirement rules, we fit our model to the data for the cohort of individuals aged 57-61 in 1992. However, when estimating the stochastic processes that individuals face we use the full sample, plus Assets and Health Dynamics of the Oldest Old (AHEAD) data, which provides information on these processes at older ages. With the exception of wages, we do not adjust the data for cohort effects. Because our subsample of the HRS covers a fairly narrow age range, this omission should not generate much bias.

4.2 Health Insurance and Medical Expenses

We assign individuals to one of three mutually exclusive health insurance groups: *retiree*, *tied*, and *none*, as described in Section 2. Because of small sample problems, the *none* group includes those with private health insurance as well as those with no insurance at all. Both face high medical expenses because they lack employer-provided coverage. Private health insurance is a poor substitute for employer-provided coverage, as high administrative costs and adverse selection problems can result in prohibitively expensive premiums. Moreover, private insurance is much less likely to cover pre-existing medical conditions. Because the

model includes a consumption floor to capture the insurance provided by Medicaid, the *none* group also includes those who receive health care through Medicaid. We assign those who have health insurance provided by their spouse to the *retiree* group, along with those who report that they could keep their health insurance if they left their jobs. Both of these groups have health insurance that is not tied to their job. We assign individuals who would lose their employer-provided health insurance after leaving their job to the *tied* group. Appendix H shows our estimated (health insurance-conditional) job exit rate profiles are robust to alternative coding decisions.

The HRS has data on self-reported medical expenses. Medical expenses are the sum of insurance premia paid by households, drug costs, and out-of-pocket costs for hospital, nursing home care, doctor visits, dental visits, and outpatient care. Because our model explicitly accounts for government transfers, the appropriate measure of medical expenses includes expenses paid for by government transfers. Unfortunately, we observe only the medical expenses paid by households, not those paid by Medicaid. Therefore, we impute Medicaid payments for households that received Medicaid benefits, as described in Appendix G.

We fit these data to the medical expense model described in Section 2. Because of small sample problems, we allow the mean, $m(.)$, and standard deviation, $\sigma(.)$, to depend only on the individual's Medicare eligibility, health insurance type, health status, labor force participation and age. Following the procedure described in French and Jones (2004a), $m(.)$ and $\sigma(.)$ are set so that the model replicates the mean and 95th percentile of the cross-sectional distribution of medical expenses in each of these categories. Details are in Appendix I.

Table 1 presents summary statistics, conditional on health status. Table 1 shows that for healthy individuals who are 64 years old, and thus not receiving Medicare, average annual medical expenses are \$3,360 for workers with *tied* coverage and \$6,010 for those with *none*, a difference of \$2,650. With the onset of Medicare at age 65, the difference shrinks to \$1,030.[11]

[11]The pre-Medicare cost differences are roughly comparable to EBRI's (1999) estimate that employers on average contribute \$3,288 per year to their employees' health insurance. They are larger than Gustman and Steinmeier's (1994) estimate that employers contribute about \$2,500 per year before age 65 (1977 NMES data, adjusted to 1998 dollars with the medical component of the CPI).

Thus, the value of having employer provided health insurance coverage largely vanishes at age 65.

	Retiree - Working	Retiree - Not Working	Tied - Working	Tied - Not Working	None
Age = 64, without Medicare, Good Health					
Mean	\$3,160	\$3,880	\$3,360	\$5,410	\$6,010
Standard Deviation	\$5,460	\$7,510	\$5,040	\$10,820	\$15,830
99.5th Percentile	\$32,700	\$44,300	\$30,600	\$63,500	\$86,900
Age = 65, with Medicare, Good Health					
Mean	\$3,320	\$3,680	\$3,830	\$4,230	\$4,860
Standard Deviation	\$4,740	\$5,590	\$5,920	\$9,140	\$7,080
99.5th Percentile	\$28,800	\$33,900	\$35,800	\$52,800	\$43,000
Age = 64, without Medicare, Bad Health					
Mean	\$3,930	\$4,830	\$4,170	\$6,730	\$7,470
Standard Deviation	\$6,940	\$9,530	\$6,420	\$13,740	\$20,060
99.5th Percentile	\$41,500	\$56,100	\$38,900	\$80,400	\$109,500
Age = 65, with Medicare, Bad Health					
Mean	\$4,130	\$4,580	\$4,760	\$5,260	\$6,040
Standard Deviation	\$6,030	\$7,120	\$7,530	\$11,590	\$9,020
99.5th Percentile	\$36,600	\$43,000	\$45,500	\$66,700	\$54,700

Table 1: Medical Expenses, by Medicare and Health Insurance Status

As Rust and Phelan (1997) emphasize, it is not just differences in mean medical expenses that determine the value of health insurance, but also differences in variance and skewness. If health insurance reduces medical expense volatility, risk-averse individuals may value health insurance at well beyond the cost paid by employers. To give a sense of the volatility, Table 1 also presents the standard deviation and 99.5th percentile of the medical expense distributions. Table 1 shows that for healthy individuals who are 64 years old, annual medical expenses have a standard deviation of \$5,040 for workers with *tied* coverage and \$15,830 for those with *none*, a difference of \$10,790. With the onset of Medicare at age 65, the difference shrinks to \$1,160. Therefore, Medicare not only reduces average medical expenses for those without employer-provided health insurance. It reduces medical expense volatility as well.

The parameters for the idiosyncratic process ψ_t, $(\sigma_\xi^2, \sigma_\epsilon^2, \rho_m)$, are taken from French and Jones (2004a, "fitted" specification). Table 2 presents the parameters, which have been normalized so that the overall variance, σ_ψ^2, is one. Table 2 reveals that at any point in time,

Parameter	Variable	Estimate (Standard Errors)
ρ_m	autocorrelation of persistent component	0.925 (0.003)
σ^2_ϵ	innovation variance of persistent component	0.04811 (0.008)
σ^2_ξ	innovation variance of transitory component	0.6668 (0.014)

Table 2: Variance and Persistence of Innovations to Medical Expenses

the transitory component generates almost 67% of the cross-sectional variance in medical expenses. The results in French and Jones reveal, however, that most of the variance in cumulative *lifetime* medical expenses is generated by innovations to the persistent component. For this reason, the cross sectional distribution of medical expenses reported in Table 1 understates the lifetime risk of medical expenses. Given the autocorrelation coefficient ρ_m of 0.925, this is not surprising.

4.3 Pension Accrual

Appendix C describes how we use confidential HRS pension data to construct the accrual rate formula. Figure 1 shows the average pension accrual rates generated by this formula when we simulate the model.

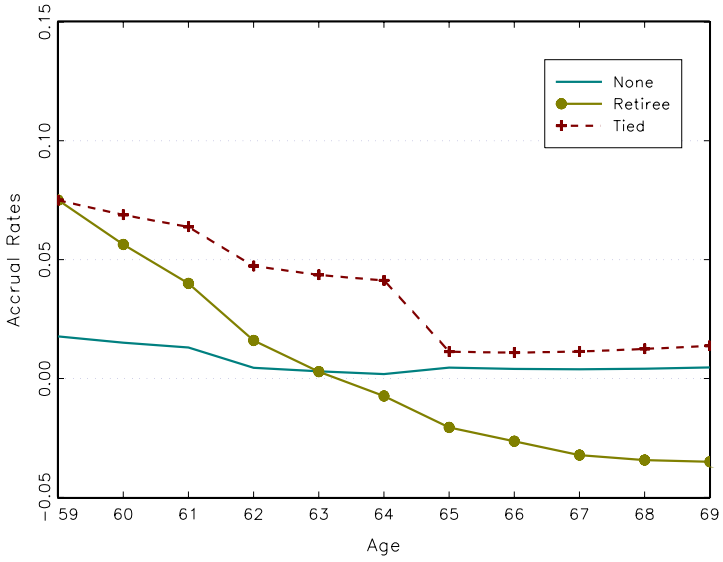

Figure 1: Average Pension Accrual Rates, by Age and Health Insurance Coverage

Figure 1 reveals that workers with *retiree* coverage face the sharpest drops in pension

accrual after age 60.[12] While *retiree* coverage in and of itself provides an incentive for early retirement, the pension plans associated with *retiree* coverage also provide the strongest incentives for early retirement. Failing to capture this link will lead the econometrician to overstate the effect of *retiree* coverage on retirement.

4.4 Preference Index

In order to better measure preference heterogeneity in the population (and how it is correlated with health insurance), we estimate a person's "willingness" to work using three questions from the first (1992) wave of the HRS. The first question asks the respondent the extent to which he agrees with the statement, "Even if I didn't need the money, I would probably keep on working." The second question asks the respondent, "When you think about the time when you will retire, are you looking forward to it, are you uneasy about it, or what?" The third question asks, "How much do you enjoy your job?"

To combine these three questions into a single index, we regress wave 5-7 (survey year 2000-2004) participation on the response to the three questions along with polynomials and interactions of all the state variables in the model: age, health status, wages, wealth, and AIME, medical expenses, and health insurance type. Multiplying the numerical responses to the three questions by their respective estimated coefficients and summing yields an index. We then discretize the index into three values: *high*, for the top 50% of the index for those working in wave 1; *low*, for the bottom 50% of the index for those working in wave 1; and *out* for those not working in wave 1. Appendix J provides additional details on the construction of the index. Figure 6 below shows that the index has great predictive power: at age 65, participation rates are 56% for those with an index of *high*, 39% for those with an index of *low*, and 12% for those with an index of *out*.

[12] Because Figure 1 is based on our estimation sample, it does not show accrual rates for earlier ages. Estimates that include the validation sample show, however, that those with retiree coverage have the highest pension accrual rates in their early and middle 50s.

4.5 Wages

Recall from equation (11) that $\ln W_t = \alpha \ln(N_t) + W(H_t, t) + \omega_t$. Following Aaronson and French (2004), we set $\alpha = 0.415$, which implies that a 50% drop in work hours leads to a 25% drop in the offered hourly wage. This is in the middle of the range of estimates of the effect of hours worked on the offered hourly wage.

We estimate $W(H_t, t)$ using the methodology described in section 3.3.

The parameters for the idiosyncratic process ω_t, (σ_η^2, ρ_W) are estimated by French (2005). The results indicate that the autocorrelation coefficient ρ_W is 0.977; wages are almost a random walk. The estimate of the innovation variance σ_η^2 is 0.0141; one standard deviation of an innovation in the wage is 12% of wages.

4.6 Remaining Calibrations

We set the interest rate r equal to 0.03. Spousal income depends upon an age polynomial and health status. Health status and mortality both depend on previous health status interacted with an age polynomial.

5 Data Profiles and Initial Conditions

5.1 Data Profiles

Figure 2 presents some of the labor market behavior we want our model to explain. The top panel of Figure 2 shows empirical job exit rates by health insurance type. Recall that Medicare should provide the largest labor market incentives for workers that have *tied* health insurance. If these people place a high value on employer-provided health insurance, they should either work until age 65, when they are eligible for Medicare, or they should work until age 63.5 and use COBRA coverage as a bridge to Medicare. The job exit profiles provide some evidence that those with tied coverage do tend to work until age 65. While the age-65 job exit rate is similar for those whose health insurance type is *tied* (20%), *retiree* (17%), or *none* (18%), those with *retiree* coverage have higher exit rates at 62 (22%) than those with

tied (14%) or *none* (18%).[13] At almost every age other than 65, those with *retiree* coverage have higher job exit rates than those with *tied* or no coverage. These differences across health insurance groups, while large, are smaller than the differences in the empirical exit profiles reported in Rust and Phelan (1997).

The low job exit rates before age 65 and the relatively high job exit rates at age 65 for those with *tied* coverage suggests that some people with *tied* coverage are working until age 65, when they become eligible for Medicare. On the other hand, job exit rates for those with *tied* coverage are lower than those with *retiree* coverage for every age other than 65, and are not much higher at age 65. This suggests that differences in health insurance coverage may not be the only reason for the differences in job exit rates.

The bottom panel of Figure 2 presents observed labor force participation rates. In comparing participation rates across health insurance categories, it is useful to keep in mind the transitions implied by equation (10): retiring workers in the *tied* insurance category transition into the *none* category. Because of this, the labor force participation rates for those with *tied* insurance are calculated for a group of individuals that were all working in the previous period. It is therefore not surprising that the *tied* category has the highest participation rates. Conversely, it is not surprising that the *none* category has the lowest participation rates, given that category includes *tied* workers who retire.

[13]The differences across groups are statistically different at 62, but not at 65. Furthermore, F-tests reject the hypothesis that the three groups have identical exit rates at all ages at the 5% level.

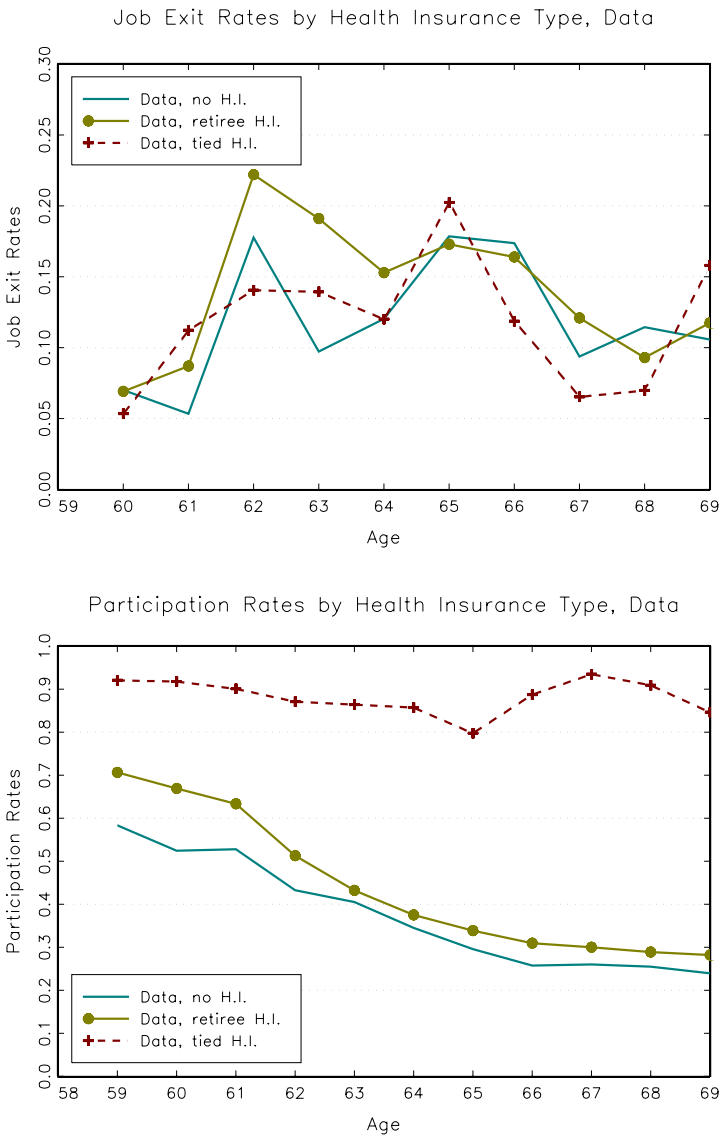

Figure 2: Job Exit and Participation Rates, Data

5.2 Initial Conditions

Each artificial individual in our model begins its simulated life with the year-1992 state vector of an individual, aged 57-61 in 1992, observed in the data. Table 3 summarizes this initial distribution, the construction of which is described in Appendix G. Table 3 shows that individuals with *retiree* coverage tend to have the most asset and pension wealth, while individuals in the *none* category have the least. The median individual in the *none* category has no pension wealth at all. Individuals in the *none* category are also more likely to be in bad health, and not surprisingly, less likely to be working. In contrast, individuals with *tied* coverage have high values of the preference index, suggesting that their delayed retirement reflects differences in preferences as well as in incentives.

	Retiree	*Tied*	*None*
Age			
Mean	58.7	58.6	58.7
Standard deviation	1.5	1.5	1.5
$AIME$ (in thousands of 1998 dollars)			
Mean	24.9	24.9	16.0
Median	27.2	26.9	16.2
Standard deviation	9.1	8.6	9.2
Assets (in thousands of 1998 dollars)			
Mean	231	205	203
Median	147	118	52
Standard deviation	248	251	307
Pension Wealth (in thousands of 1998 dollars)			
Mean	129	80	17
Median	62	17	0
Standard deviation	180	212	102
Wage (in 1998 dollars)			
Mean	17.4	17.6	12.0
Median	14.7	14.6	8.6
Standard deviation	13.4	12.4	11.2
Preference index			
Fraction *out*	0.27	0.04	0.48
Fraction *low*	0.42	0.44	0.19
Fraction *high*	0.32	0.52	0.33
Fraction in bad health	0.20	0.13	0.41
Fraction working	0.73	0.96	0.52
Number of observations	1,022	225	455

Table 3: Summary Statistics for the Initial Distribution

6 Baseline Results

6.1 Preference Parameter Estimates

The goal of our MSM estimation procedure is to match the life cycle profiles for assets, hours and participation found in the HRS data. In order to use these profiles to identify preferences, we make several identifying assumptions, the most important being that preferences vary with age in two specific ways; (1) through changes in health status; and (2) through the linear time trend in the fixed cost ϕ_{Pt}. Therefore, age can be thought of as an "exclusion restriction", which changes the incentives for work and savings in ways that can not be captured with changes in preferences.

Table 4 presents preference parameter estimates. The first 3 rows of Table 4 show the parameters that vary across the preference types. We assume that there are three types of individuals, and that the types differ in the utility weight on consumption, γ, and their time discount factor, β. Individuals with high values of γ have stronger preferences for work. Individuals with high values of β are more patient and thus more willing to defer consumption and leisure. Table 4 reveals significant differences in γ and β across preference types, which are discussed in some detail in Section 6.2.

Table 4 also shows the fraction of workers belonging to each preference type. Averaging over the three types reveals that the average value of β, the discount factor, implied by our model is 0.913, which is slightly lower than most estimates. The discount factor is identified by the intertemporal substitution of consumption and leisure, as embodied in the asset and labor supply profiles.

Another key parameter is ν, the coefficient of relative risk aversion for the consumption-leisure composite. A more familiar measure of risk aversion is the coefficient of relative risk aversion for consumption. Assuming that labor supply is fixed, it can be approximated as $-\frac{(\partial^2 U/\partial C^2)C}{\partial U/\partial C} = -(\gamma(1-\nu)-1)$. The weighted average value of the coefficient is 5.0. This value falls within the range of estimates found in recent studies by Cagetti (2003) and French (2005), but it is larger than the values of 1.1, 1.8, and 1.0 reported by Rust and Phelan

(1997), Blau and Gilleskie (2006), and Blau and Gilleskie (2008) respectively, in their studies of retirement.

Parameters that vary across individuals

	Type 0	Type 1	Type 2
γ: consumption weight	0.412	0.649	0.967
	(0.045)	(0.007)	(0.203)
β: time discount factor	0.945	0.859	1.124
	(0.074)	(0.013)	(0.328)
Fraction of individuals	0.267	0.615	0.118

Parameters that are common to all individuals

ν: coefficient of relative risk aversion, utility	7.49 (0.311)	θ_B: bequest weight†	0.0223 (0.0012)
κ: bequest shifter, in thousands	444 (28.4)	c_{min}: consumption floor	4,380 (167)
L: leisure endowment, in hours	4,060 (44)	ϕ_H: hours of leisure lost, bad health	506 (20.9)
ϕ_{P0}: fixed cost of work at age 60, in hours	826 (20.0)	ϕ_{P1}: fixed cost of work: age trend, in hours	54.7 (2.58)
ϕ_{RE}: hours of leisure lost when re-entering labor market	94.0 (8.63)		

χ^2 statistic = 775; Degrees of freedom = 171

Method of simulated moments estimates.
Diagonal weighting matrix used in calculations. See Appendix F for details.
Standard errors in parentheses.
†Parameter expressed as marginal propensity to consume out of final-period wealth.
Parameters estimated jointly with type probability prediction equation. See Appendix K for estimated coefficients of the type probability prediction equation.

Table 4: Estimated Structural Parameters

The risk coefficient ν and the consumption floor C_{min} are identified in large part by the asset quantiles, which reflect precautionary motives. The bottom quantile in particular depends on the interaction of precautionary motives and the consumption floor. If the consumption floor is sufficiently low, the risk of a catastrophic medical expense shock, which over a lifetime could equal over $100,000 (see French and Jones (2004a)), will generate strong precautionary incentives. Conversely, as emphasized by Hubbard, Skinner and Zeldes (1995), a high consumption floor discourages saving among the poor, since the consumption floor

effectively imposes a 100% tax on the saving of those with high medical expenses and low income and assets.

Our estimated consumption floor of $4,380 is similar to other estimates of social insurance transfers for the indigent. For example, when we use Hubbard, Skinner and Zeldes's (1994, Appendix A) procedures and more recent data, we find that the average benefit available to a childless household with no members aged 65 or older was $3,500. A value of $3,500 understates the benefits available to individuals over age 65; in 1998 the Federal SSI benefit for elderly (65+) couples was nearly $9,000 (Committee on Ways and Means, 2000, p. 229). On the other hand, about half of eligible households do not collect SSI benefits (Elder and Powers, 2006, Table 2), possibly because transactions or "stigma" costs outweigh the value of public assistance. Low take-up rates, along with the costs that probably underly them, suggest that the effective consumption floor need not equal statutory benefits.

The bequest parameters θ_B and κ are identified largely from the top asset quantile. It follows from equation (3) that when the shift parameter κ is large, the marginal utility of bequests will be lower than the marginal utility of consumption unless the individual is rich. In other words, the bequest motive mainly affects the saving of the rich; for more on this point, see De Nardi (2004). Our estimate of θ_B implies that the marginal propensity to consume out of wealth in the final period of life (which is a nonlinear function of θ_B, β, γ, ν and κ) is 1 for low income individuals and 0.022 for high-income individuals.

Turning to labor supply, we find that individuals in our sample are willing to intertemporally substitute their work hours. In particular, simulating the effects of a 2% wage change reveals that the wage elasticity of average hours is 0.486 at age 60. This relatively high labor supply elasticity arises because the fixed cost of work generates volatility on the participation margin. The participation elasticity is 0.353 at age 60, implying that wage changes cause relatively small hours changes for workers. For example, the Frisch labor supply elasticity of a type-1 individual working 2,000 hours per year at age 60 is approximated as $-\frac{L-N_t-\phi_{P0}}{N_t} \times \frac{1}{(1-\gamma)(1-\nu)-1} = 0.19$.

The fixed cost of work at age 60, ϕ_{P0}, is 826 hours per year, and increases by $\phi_{P1} = 55$

hours per year. The fixed cost of work is identified by the life cycle profile of hours worked by workers. Average hours of work (available upon request) do not drop below 1,000 hours per year (or 20 hours per week, 50 weeks per year) even though labor force participation rates decline to near zero. In the absence of a fixed cost of work, one would expect hours worked to parallel the decline in labor force participation. (See Rogerson and Wallenius, 2009.) The time endowment L is identified by the combination of the participation and hours profiles. The time cost of bad health, ϕ_H, is identified by noting that unhealthy individuals work fewer hours than healthy individuals, even after conditioning on wages. The re-entry cost, ϕ_{RE}, of 94 hours, is identified by exit rates. In the absence of a re-entry cost, workers are more willing to "churn" in and out of the labor force, raising exit rates.

6.2 Preference Heterogeneity and Health Insurance

Table 4 shows considerable heterogeneity in preferences. To understand these differences, Table 5 shows simulated summary statistics for each of the preference types. Table 5 reveals that Type-0 individuals have the lowest value of γ, i.e., they place the highest value on leisure. 92% of Type-0 individuals were *out* of the labor force in wave 1. Type-2 individuals, in contrast, have the highest value of γ. 84% of Type-2 individuals have a preference index of *high*, meaning that they were working in wave 1 and self-reported having a low preference for leisure. Type-1 individuals fall in the middle, valuing leisure less than Type-0 individuals, but more than Type-2 individuals. 54% of Type-1 individuals have a preference index value of *low*.

Including preference heterogeneity allows us to control for the possibility that workers with different preferences select jobs with different health insurance packages. Table 5 suggests that some self-selection is occurring, as it reveals while 14% of workers with *tied* coverage are Type-2 agents, who have the lowest disutility of work, only 5% are Type-0 agents, who have the highest disutility. In contrast, 11% of workers with *retiree* coverage are Type-2 agents, and 27% are Type-0 agents. This suggests that workers with *tied* coverage might be more willing to retire later than those with *retiree* coverage because they have a lower disutility

of work. However, Section 6.4 shows that accounting for this correlation has little impact on the estimated effect of health insurance on retirement.

	Type 0	Type 1	Type 2
Key preference parameters			
γ^*	0.412	0.649	0.967
β^*	0.945	0.859	1.124
Means by preference type			
Assets ($1,000s)	150	215	405
Pension Wealth ($1,000s)	92	97	74
Wages ($/hour)	11.3	19.0	11.1
Probability of health insurance type, given preference type			
Health insurance = *none*	0.371	0.222	0.261
Health insurance = *retiree*	0.607	0.603	0.581
Health insurance = *tied*	0.023	0.175	0.158
Probability of preference index value, given preference type			
Preference Index = *out*	0.922	0.068	0.034
Preference Index = *low*	0.039	0.539	0.131
Preference Index = *high*	0.039	0.392	0.835
Fraction of individuals	0.267	0.615	0.118
*Values of β and γ are from Table 4.			

Table 5: Mean Values by Preference Type, Simulations

6.3 Simulated Profiles

The bottom of Table 4 displays the overidentification test statistic. Even though the model is formally rejected, the life cycle profiles generated by the model match up well with the life cycle profiles found in the data.

Figure 3 shows the 1/3rd and 2/3rd asset quantiles at each age for the HRS sample and for the model simulations. For example, at age 64 about one third of the men in our sample live in households with less than $80,000 in assets, and about one third live in households with over $270,000 of assets. Figure 3 shows that the model fits both asset quantiles well. The model is able to fit the lower quantile in large part because of the consumption floor of $4,350; the predicted 1/3rd quantile rises when the consumption floor is lowered.

The three panels in the left hand column of Figure 4 show that the model is able to replicate the two key features of how labor force participation varies with age and health

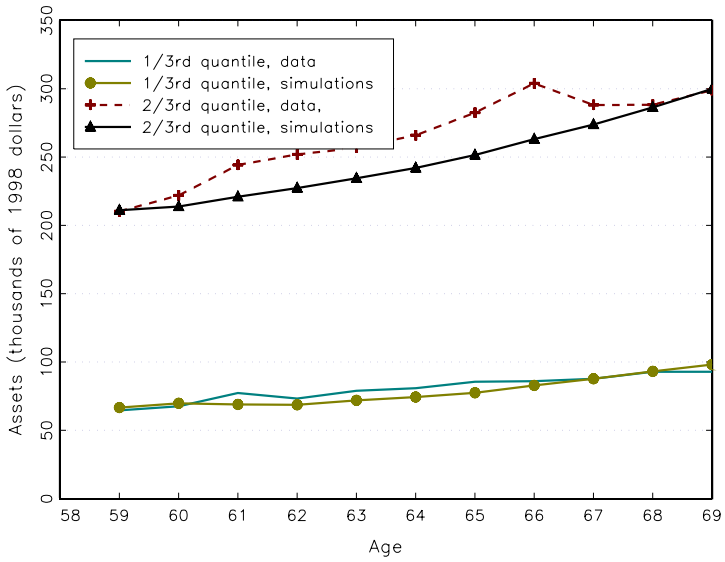

Figure 3: Asset Quantiles, Data and Simulations

insurance. The first key feature is that participation declines with age, and the declines are especially sharp between ages 62 and 65. The model underpredicts the decline in participation at age 65 (a 4.9 percentage point decline in the data versus a 3.5 percentage point decline predicted by the model), but comes closer at age 62 (a 10.6 percentage point decline in the data versus a 10.9 percentage point decline predicted by the model).

The second key feature is that there are large differences in participation and job exit rates across health insurance types. The model does a good job of replicating observed differences in participation rates. For example, the model matches the low participation levels of the uninsured. Turning to the lower left panel of Figure 5, the data show that the group with the lowest participation rates are the uninsured with low assets. The model is able to replicate this fact because of the consumption floor. Without a high consumption floor, the risk of catastrophic medical expenses, in combination with risk aversion, would cause the uninsured to remain in the labor force and accumulate a buffer stock of assets.

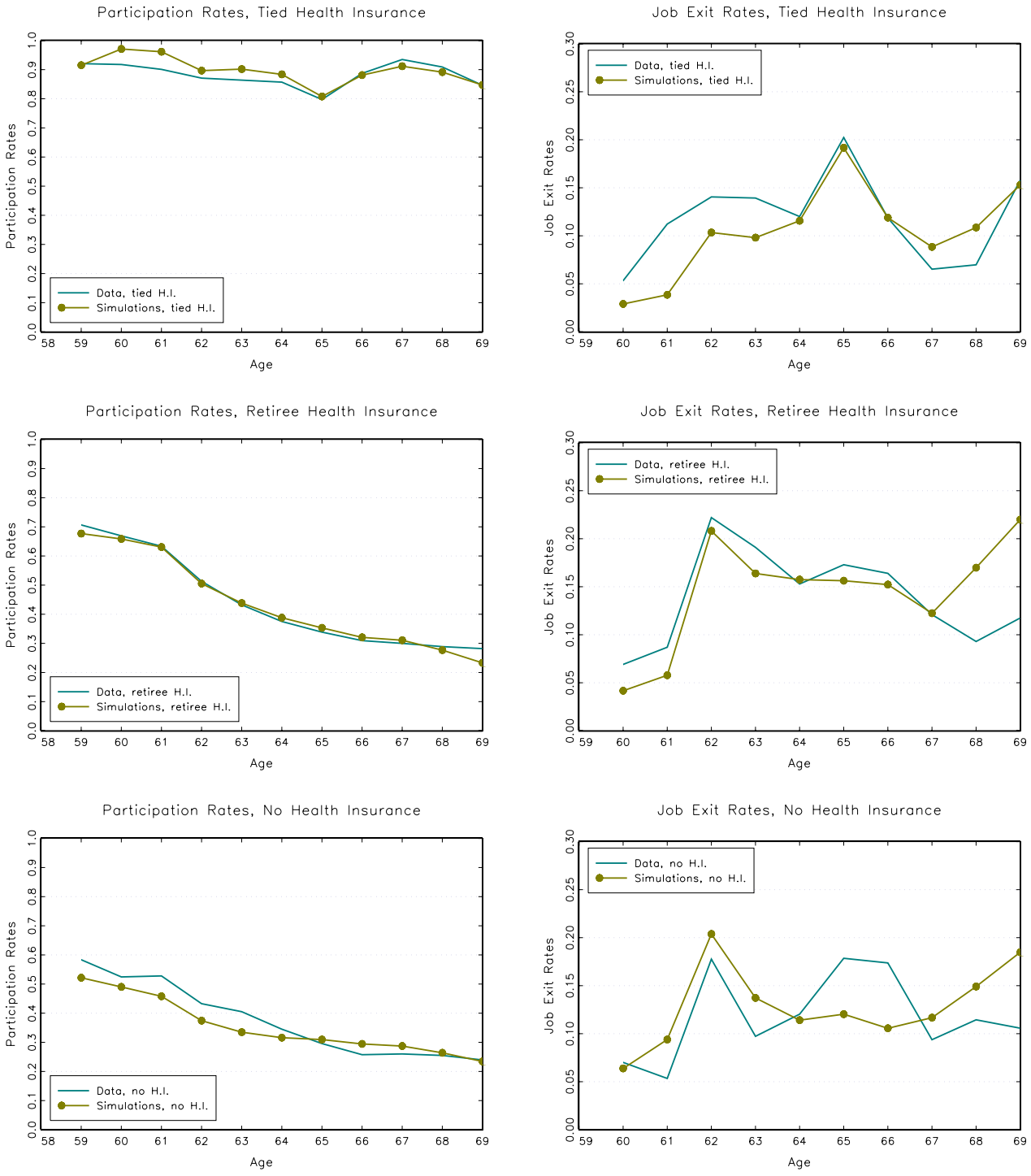

Figure 4: Participation and Job Exit Rates, Data and Simulations

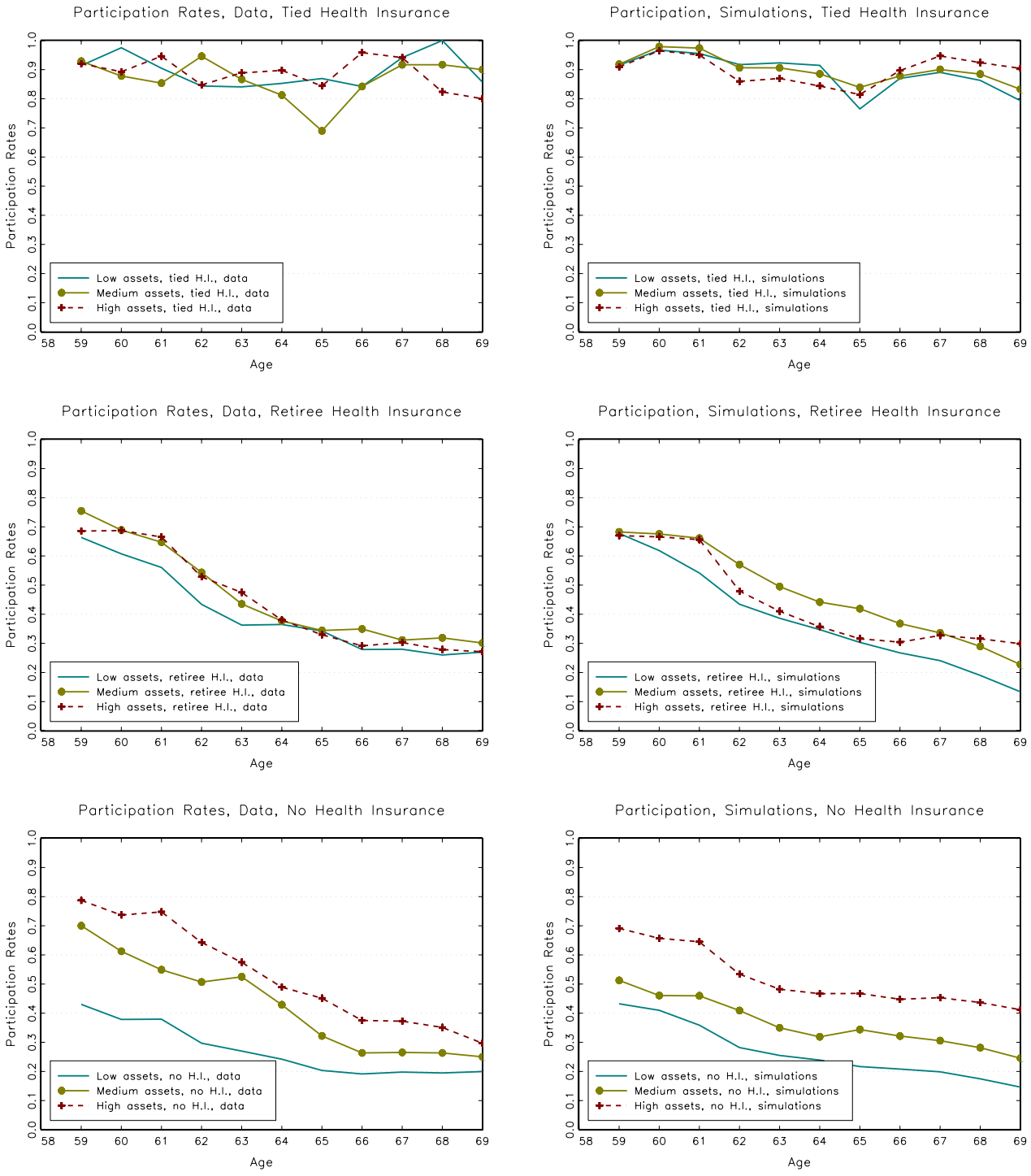

Figure 5: Labor Force Participation Rates by Asset Grouping, Data and Simulations

The panels in the right hand column of Figure 4 compare observed and simulated job exit rates for each health insurance type. The model does a good job of fitting the exit rates of workers with *retiree* or *tied* coverage. For example, the model captures the high age-62 job exit rates for those with *retiree* coverage and the high age-65 job exit rates for those with *tied* coverage. However, it fails to capture the high exit rates at age 65 for workers with no health insurance.

Figure 6 shows how participation differs across the three values of the discretized preference index constructed from HRS attitudinal questions. Recall that an index value of *out* implies that the individual was not working in 1992. Not surprisingly, participation for this group is always low. Individuals with positive values of the preference index differ primarily in the rate at which they leave the labor force. Although *low*-index individuals initially work as much as *high*-index individuals, they leave the labor force more quickly. As noted in our discussion of the preference parameters, the model replicates these differences by allowing the taste for leisure (γ) and the discount rate (β) to vary across preference types. When we do not allow for preference heterogeneity, the model is unable to replicate the patterns observed in Figure 6. This highlights the importance of the preference index in identifying preference heterogeneity.

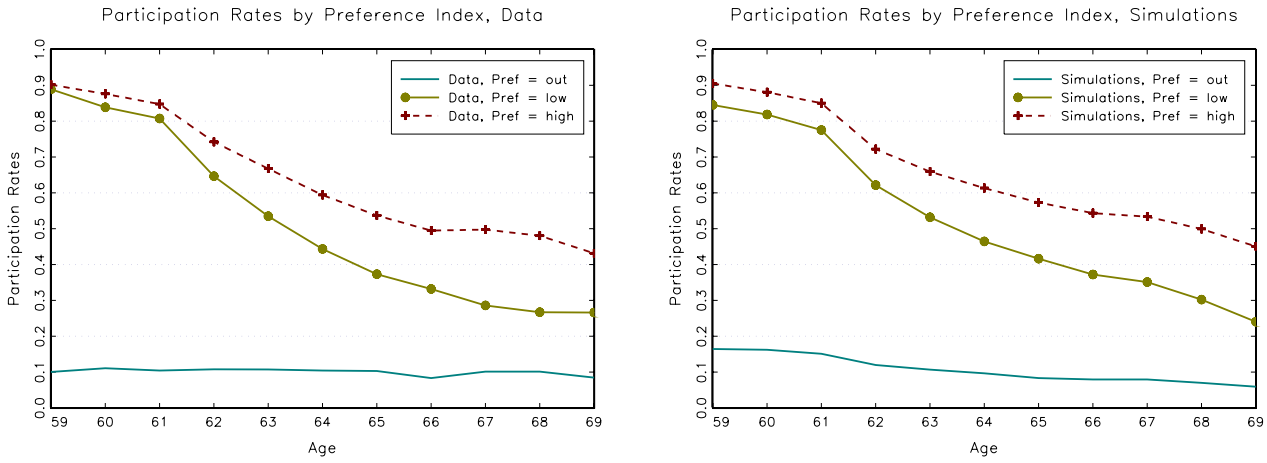

Figure 6: Labor Force Participation Rates by Preference Index, Data and Simulations

6.4 The Effects of Employer-Provided Health Insurance

The labor supply patterns in Figures 2 and 4 show that those with *retiree* coverage retire earlier than those with *tied* coverage. However, the profiles do not identify the effects of health insurance on retirement, for three reasons. First, as shown in Table 3, those with *retiree* coverage have greater pension wealth than other groups. Second, as shown in Figure 1, pension plans for workers with *retiree* coverage provide stronger incentives for early retirement than the pension plans held by other groups. Third, as shown in Table 5, preferences for leisure vary by health insurance type. In short, retirement incentives differ across health insurance categories for reasons unrelated to health insurance incentives.

To isolate the effects of employer-provided health insurance on labor supply, we conduct some additional simulations. We give everyone the pension accrual rates of *tied* workers so that pension incentives are identical across health insurance types. We then simulate the model twice, assuming first that all workers have *retiree* health insurance coverage at age 59, then *tied* coverage at age 59. Across the two simulations, households face different medical expense distributions, but in all other dimensions the distribution of incentives and preferences is identical.

This exercise reveals that if all workers had *retiree* coverage rather than *tied* coverage the job exit rate at age 62 would be 8.4 percentage points higher. In contrast, the raw difference in model-predicted exit rates at age 62 is 10.5 percentage points. (The raw difference in the data is 8.2 percentage points.) The high age-62 exit rates of those with *retiree* coverage are thus partly due to more generous pensions and stronger preferences for leisure. Even after controlling for these factors, however, health insurance is still an important determinant of retirement.

The effects of health insurance can also be measured by comparing participation rates. We find that the labor force participation rate for ages 60-69 would be 5.1 percentage points lower if everyone had *retiree*, rather than *tied*, coverage at age 59. Furthermore, moving everyone from *retiree* to *tied* coverage increases the average retirement age (defined as the oldest age at which the individual works plus one) by 0.34 years.

In comparison, Blau and Gilleskie's (2001) reduced-form estimates imply that having *retiree* coverage, rather than *tied* coverage, increases the job exit rate 7.5 percentage points at age 61. Blau and Gilleskie also find that accounting for selection into health insurance plans modestly increases the estimated effect of health insurance on exit rates. Other reduced form findings in the literature are qualitatively similar to Blau and Gilleskie. For example, Madrian (1994) finds that retiree coverage reduces the retirement age by 0.4-1.2 years, depending on the specification and the data employed. Karoly and Rogowski (1994), who attempt to account for selection into health insurance plans, find that retiree coverage increases the job exit rate 8 percentage points over a $2\frac{1}{2}$ year period. Our estimates, therefore, lie within the lower bound of the range established by previous reduced form studies, giving us confidence that the model can be used for policy analysis.

Structural studies that omit medical expense risk find smaller health insurance effects than we do. For example, Gustman and Steinmeier (1994) find that retiree coverage reduces years in the labor force by 0.1 years. Lumsdaine et al. (1994) find even smaller effects. Structural studies that include medical expense risk but omit self-insurance find bigger effects. Our estimated effects are larger than Blau and Gilleskie's (2006, 2008), who find that retiree coverage reduces average labor force participation 1.7 and 1.6 percentage points, respectively, but are smaller than the effects found by Rust and Phelan (1997).[14]

6.5 Model Validation

Following several recent studies (e.g., Keane and Wolpin, 2007), we perform an out-of-sample validation exercise. Recall that we estimate the model on a cohort of individuals aged 57-61 in 1992. We test our model by considering the HRS cohort aged 51-55 in 1992; we refer to this group as our validation cohort. These individuals faced different Social Security incentives than did the estimation cohort. The validation cohort did not face the

[14]Blau and Gilleskie (2006) consider the retirement decision of couples, and allow husbands and wives to retire at different dates. Blau and Gilleskie (2008) allow workers to choose their medical expenses. Because these modifications provide additional mechanisms for smoothing consumption over medical expense shocks, they could reduce the effect of employer-provided health insurance.

Social Security earnings test after age 65, had a later full retirement age, and faced a benefit adjustment formula that more strongly encouraged delayed retirement. In addition to facing different Social Security rules, the validation cohort possessed different endowments of wages, wealth, and employer benefits. A useful test of our model, therefore, is to see if it can predict the behavior of the validation cohort.

	Data			Model		
	1933	1939	Difference†	1933	1939	Difference*
Age	(1)	(2)	(3)	(4)	(5)	(6)
60	0.657	0.692	0.035	0.650	0.706	0.056
61	0.636	0.642	0.006	0.622	0.677	0.055
62	0.530	0.545	0.014	0.513	0.570	0.057
63	0.467	0.508	0.041	0.456	0.490	0.035
64	0.408	0.471	0.063	0.413	0.449	0.037
65	0.358	0.424	0.066	0.378	0.459	0.082
66	0.326	0.382	0.057	0.350	0.430	0.080
67	0.314	0.374	0.060	0.339	0.386	0.047
Total, 60-67	3.696	4.037	0.341	3.721	4.168	0.447
	† Column (2) − Column (1)			*Column (5) − Column (4)		

Table 6: Participation Rates by Birth Year Cohort

Columns (1)-(3) of Table 6 show the participation rates observed in the data for each cohort, and the difference. The data suggest that the change in the Social Security rules coincides with increased labor force participation, especially at later ages. By way of comparison, Song and Manchester (2007), examining Social Security administrative data, find that between 1996 and 2003, participation rates increased by 3, 4 and 6 percentage points for workers turning 62-64, 65, and 66-69, respectively. These differences are similar to the differences between the 1933 and 1939 cohorts in our data, as shown in column 3.

Columns (4)-(6) of Table 6 show the participation rates predicted by the model. The simulations for the validation cohort use the initial distribution and Social Security rules for the validation cohort, but use the parameter values estimated on the older estimation cohort.[15] Comparing Columns (3) and (6) shows that the model-predicted increase in labor

[15]We do not adjust for business cycle conditions. Because the validation cohort starts at age 53, 6 years before the estimation cohort, the validation exercise requires its own wage selection adjustment and pension prediction equation. Using the baseline preference estimates, we construct these inputs in the same way

supply (0.45 years), resembles the increase observed in the data (0.35 years).

7 Policy Experiments

The preceding sections showed that the model fits the data well, given plausible preference parameters. In this section, we use the model to predict how changing the Social Security and Medicare rules would affect retirement behavior. The results of these experiments are summarized in Table 7.

Age	SS = 65 MC = 65 (1)	SS = 67† MC = 65 (2)	SS = 65 MC = 67 (3)	SS = 67† MC = 67 (4)	Data (5)
60	0.650	0.651	0.651	0.652	0.657
61	0.622	0.625	0.623	0.626	0.636
62	0.513	0.526	0.516	0.530	0.530
63	0.456	0.469	0.460	0.472	0.467
64	0.413	0.426	0.422	0.433	0.407
65	0.378	0.386	0.407	0.415	0.358
66	0.350	0.358	0.374	0.381	0.326
67	0.339	0.346	0.341	0.347	0.314
68	0.307	0.311	0.307	0.312	0.304
69	0.264	0.270	0.264	0.270	0.283
Total 60-69	4.292	4.368	4.366	4.438	4.283

SS = Social Security normal retirement age
MC = Medicare eligibility age
† Benefits reduced by two years, as described in text

Table 7: Effects of Changing the Social Security Retirement and Medicare Eligibility Ages

The first column of Table 7 shows model-predicted labor market participation at ages 60 through 69 under the 1998 Social Security rules. Under the 1998 rules, the average person works a total of 4.29 years over this 10-year period. The fifth column of Table 7 shows that this is close to the total of 4.28 years observed in the data.

The Social Security rules are slowly evolving over time. If current plans continue, by 2030 the normal Social Security retirement age, the date at which workers can receive "full

we construct their baseline counterparts. In addition, we adjust the intercept terms in the type prediction equations so that the validation cohort generates the same distribution of preference types as the estimation sample.

benefits", will have risen from 65 to 67. Raising the normal retirement age to 67 effectively eliminates two years of Social Security benefits. Column (2) shows the effect of this change.[16] The wealth effect of lower benefits leads years of work to increase by 0.076 years, to 4.37 years.[17]

The third column of Table 7 shows participation when the Medicare eligibility age is increased to 67.[18] Over a 10-year period, total years of work increase by 0.074 years, so that the average probability of employment increases by 0.74 percentage points per year. This amount is larger than the changes found by Blau and Gilleskie (2006), whose simulations show that increasing the Medicare age increases the average probability of employment by 0.1 percentage points, but is smaller than the effects suggested by Rust and Phelan's (1997) analysis.

The fourth column shows the combined effect of cutting Social Security benefits and raising the Medicare eligibility age. The joint effect is an increase of 0.146 years, 0.072 years more than that generated by cutting Social Security benefits in isolation. In summary, the model predicts that raising the Medicare eligibility age will have almost the same effect on retirement behavior as the benefit reductions associated with a higher Social Security retirement age. Medicare has an even bigger effect on those with *tied* coverage at age 59.[19] Simulations reveal that for those with *tied* coverage, eliminating two years of Social Security benefits increases years in the labor force by 0.12 years, whereas shifting forward the Medicare

[16]Under the 2030 rules, an individual claiming benefits at age 65 would receive an annual benefit 13.3% smaller than the benefit he would have received under the 1998 rules (holding $AIME$ constant). We thus implement the two-year reduction in benefits by reducing annual benefits by 13.3% at every age.

[17]In addition to reducing annual benefits, the intended 2030 rules would impose two other changes. First, the rate at which benefits increase for delaying retirement past the normal age would increase from 5.5% to 8.0%. This change, like the reduction in annual benefits, should encourage work. However, raising the normal retirement age implies that the relevant earnings test for ages 65-66 would become the stricter, early-retirement test. This change should discourage work. We find that when we switch from the 1998 to the 2030 rules, the effects of the three changes cancel out, so that total hours over ages 60-69 are essentially unchanged.

[18]By shifting forward the Medicare eligibility age to 67, we increase from 65 to 67 the age at which medical expenses can follow the "with Medicare" distribution shown in Table 1.

[19]Only 13% of the workers in our sample had *tied* coverage at age 59. In contrast, Kaiser/HRET (2006) estimated that about 50% of large firms offered *tied* coverage in the mid-1990s. We might understate the share with *tied* coverage because, as shown in the Kaiser/HRET study, the fraction of workers with *tied* (instead of *retiree*) coverage grew rapidly in the 1990s, and our health insurance measure is based on wave-1 data collected in 1992. In fact, the HRS data indicate that later waves had a higher proportion of individuals with *tied* coverage than wave 1. We may also be understating the share with *tied* coverage because of changes in the wording of the HRS questionnaire; see Appendix H for details.

eligibility age to 67 would increase years in the labor force by 0.28 years.

To understand better the incentives generated by Medicare, we compute the value Type-1 individuals place on employer-provided health insurance, by finding the increase in assets that would make an uninsured Type-1 individual as well off as a person with retiree coverage. In particular, we find the compensating variation $\lambda_t = \lambda(A_t, B_t, H_t, AIME_t, \omega_t, \zeta_{t-1}, t)$, where

$$V_t(A_t, B_t, H_t, AIME_t, \omega_t, \zeta_{t-1}, retiree) = V_t(A_t + \lambda_t, B_t, H_t, AIME_t, \omega_t, \zeta_{t-1}, none).$$

Table 8 shows the compensating variation $\lambda(A_t, 0, good, \$32000, 0, 0, 60)$ at several different asset (A_t) levels.[20] The first column of Table 8 shows the valuations found under the baseline specification. One of the most striking features is that the value of employer-provided health insurance is fairly constant through much of the wealth distribution. Even though richer individuals can better self-insure, they also receive less protection from the government-provided consumption floor. These effects more or less cancel each other out over the asset range of -$5,700 to $147,000. However, individuals with asset levels of $600,000 place less value on retiree coverage, because they can better self-insure against medical expense shocks.

Part of the value of retiree coverage comes from a reduction in average medical expenses—because retiree coverage is subsidized—and part comes from a reduction in the volatility of medical expenses—because it is insurance. In order to separate the former from the latter, we eliminate medical expense uncertainty, by setting the variance shifter $\sigma(H_t, I_t, t, B_t, P_t)$ to zero, and recompute λ_t, using the same state variables and mean medical expenses as before. Without medical expense uncertainty, λ_t is approximately $11,000. Comparing the two values of λ_t shows that for the typical worker (with $147,000 of assets) about half of the value of health insurance comes from the reduction of average medical expenses, and half comes from the reduction of medical expense volatility.

The first two columns of Table 8 measure the lifetime value of health insurance as an asset

[20] In making these calculations, we remove health-insurance-specific differences in pensions, as described in section 6.4. It is also worth noting that for the values of H_t and ζ_{t-1} considered here, the conditional differences in expected medical expenses are smaller than the unconditional differences shown in Table 1.

	Compensating Assets		Compensating Annuity	
	With Uncertainty	Without Uncertainty	With Uncertainty	Without Uncertainty
Asset Levels	(1)	(2)	(3)	(4)
Baseline Case				
-$5,700	$20,400	$10,700	$4,630	$2,530
$51,600	$19,200	$10,900	$4,110	$2,700
$147,200	$21,400	$10,600	$4,180	$2,540
$600,000	$16,700	$11,900	$2,970	$2,360
No-Saving Cases				
(a) -$6,000	$112,000	$8,960	$11,220	$2,160
(b) -$6,000	$21,860	$6,862	$3,884	$2,170

Compensating variation between *retiree* and *none* coverages for agents with type-1 preferences.
Calculations described in text.
No-Saving case (a) uses benchmark preference parameter values; case (b) uses parameter values estimated for no-saving specification.

Table 8: Value of Employer-Provided Health Insurance

increment that can be consumed immediately. An alternative approach is to express the value of health insurance as an illiquid annuity comparable to Social Security benefits. Columns (3) and (4) show this "compensating annuity".[21] When the value of health insurance is expressed as an annuity, the fraction of its value attributable to reduced medical expense volatility falls from one-half to about 40 percent. In most other respects, however, the asset and annuity valuations of health insurance have similar implications.

To sum, allowing for medical expense uncertainty greatly increases the value of health insurance. It is therefore unsurprising that we find larger effects of health insurance on retirement than do Gustman and Steinmeier (1994) and Lumsdaine et al. (1994), who assume that workers value health insurance at its actuarial cost.

[21] To do this, we first find compensating $AIME$, $\widehat{\lambda}_t$, where

$$V_t(A_t, B_t, H_t, AIME_t, \omega_t, \zeta_{t-1}, retiree) = V_t(A_t, B_t, H_t, AIME_t + \widehat{\lambda}_t, \omega_t, \zeta_{t-1}, none).$$

This change in $AIME$ in turn allows us to calculate the change in expected pension and Social Security benefits that the individual would receive at age 65, the sum of which can be viewed as a compensating annuity. Because these benefits depend on decisions made after age 60, the calculation is only approximate.

8 Alternative Specifications

To consider whether our findings are sensitive to our modelling assumptions, we re-estimate the model under three alternate specifications.[22] Table 9 shows model-predicted participation rates under the different specifications, along with the data. The parameter estimates behind these simulations are shown in Appendix K. Column (1) of Table 9 presents our baseline case. Column (2) presents the case where individuals are not allowed to save. Column (3) presents the case with no preference heterogeneity. Column (4) presents the case where we remove the subjective preference index from the type prediction equations and the GMM criterion function. Column (5) presents the data. In general, the different specifications match the data profile equally well.

Age	Baseline (1)	No Saving (2)	Homogeneous Preferences (3)	No Preference Index (4)	Data (5)
60	0.650	0.648	0.621	0.653	0.657
61	0.622	0.632	0.595	0.625	0.636
62	0.513	0.513	0.517	0.516	0.530
63	0.456	0.457	0.453	0.459	0.467
64	0.413	0.429	0.409	0.417	0.407
65	0.378	0.380	0.365	0.381	0.358
66	0.350	0.334	0.351	0.357	0.326
67	0.339	0.327	0.345	0.346	0.314
68	0.307	0.308	0.319	0.314	0.304
69	0.264	0.282	0.286	0.273	0.283
Total 60-69	4.292	4.309	4.260	4.340	4.283

Table 9: Model Predicted Participation, by Age: Alternative Specifications

Table 10 shows how total years of work over ages 60-69 are affected by changes in Social Security and Medicare under each of the alternative specifications. In all specifications, decreasing the Social Security benefits and raising the Medicare eligibility age increase years of work by similar amounts.

[22] In earlier drafts of this paper (French and Jones, 2004b, 2007), we also estimated a specification where housing wealth is illiquid. Although parameter estimates and model fit for this case were somewhat different than our baseline results, the policy simulations were similar.

Rule Specification	Baseline (1)	No Saving (2)	Homogeneous Preferences (3)	No Preference Index (4)
Baseline: SS = 65, MC = 65	4.292	4.309	4.260	4.340
SS = 67: Lower benefits†	4.368	4.399	4.335	4.411
SS = 65, MC = 67	4.366	4.384	4.322	4.417
SS = 67† and MC = 67	4.438	4.456	4.395	4.482
SS = Social Security normal retirement age				
MC = Medicare eligibility age				
† Benefits reduced by two years, as described in text				

Table 10: Effects of Changing the Social Security Retirement and Medicare Eligibility Ages, Ages 60-69, Alternative Specifications

8.1 No Saving

We have argued that the ability to self-insure through saving significantly affects the value of employer-provided health insurance. One test of this hypothesis is to modify the model so that individuals cannot save, and examine how labor market decisions change. In particular, we require workers to consume their income net of medical expenses, as in Rust and Phelan (1997) and Blau and Gilleskie (2006, 2008).

The second column of Table 9 contains the labor supply profile generated by the no-saving specification. Comparing this profile to the baseline case in column (1) shows that, in addition to its obvious failings with respect to asset holdings, the no-saving case matches the labor supply data no better than the baseline case.[23]

Table 8 displays two sets of compensating values for the no-saving case. Case (a), which uses the parameter values from the benchmark case, shows that eliminating the ability to save greatly increases the value of retiree coverage: when assets are -$6,000, the compensating annuity increases from $4,600 in the baseline case (with savings) to $11,200 in no-savings case (a). When there is no medical expense uncertainty, the comparable figures are $2,530 in the baseline case and $2,160 in the no-savings case. Thus, the ability to self-insure through

[23]Because the baseline and no-savings cases are estimated with different moments, their overidentification statistics are not comparable. However, inserting the decision profiles generated by the baseline model into the moment conditions used to estimate the no-savings case produces an overidentification statistic of 349, while the no-saving specification produces an overidentification statistic of 398.

saving significantly reduces the value of employer-provided health insurance. Case (b) shows that using the parameter values estimated for the no-saving specification, which include a lower value of the risk parameter ν, also lowers the value of insurance.

Simulating the responses to policy changes, we find that raising the Medicare eligibility age to 67 leads to an additional 0.075 years of work, an amount almost identical to that of the baseline specification.

8.2 No Preference Heterogeneity

To assess the importance of preference heterogeneity, we estimate and simulate a model where individuals have identical preferences (conditional on age and health status). Comparing columns (1), (3) and (5) of Table 9 shows that the model without preference heterogeneity matches aggregate participation rates as well as the baseline model. However, the no-heterogeneity specification does much less well in replicating the way in which participation varies across the asset distribution, and, not surprisingly, does not replicate the way in which participation varies across our discretized preference index.

When preferences are homogeneous the simulated response to delaying the Medicare eligibility age, 0.062 years, is similar to the response in the baseline specification. This is consistent with our analysis in Section 6.4, where not accounting for preference heterogeneity and insurance self-selection appeared to only modestly change the estimated effects of health insurance on retirement.

8.3 No Preference Index

In the baseline specification, we use the preference index (described in Section 4.4) to predict preference type, and the GMM criterion function includes participation rates for each value of the index. Because labor force participation differs sharply across the index in ways not predicted by the model's other state variables, we interpret the index as a measure of otherwise unobserved preferences toward work. It is possible, however, that using the preference index causes us to overstate the correlation between health insurance and tastes for

leisure. For example, Table 3 shows that employed individuals with *retiree* coverage are more likely to have a preference index that is *low* than employed individuals with *tied* coverage. This means that workers with *retiree* coverage are more likely to report looking forward to retirement, and thus more likely to be assigned a higher desire for leisure. But workers with *retiree* coverage may be more likely to report looking forward to retirement simply because they would have health insurance and other financial resources during retirement. As a robustness test, we remove the preference index, and the preference index-related moment conditions, and re-estimate the model.

	Type 0	Type 1	Type 2
Key preference parameters			
γ^*	0.405	0.647	0.986
β^*	0.962	0.858	1.143
Means by preference type			
Assets ($1,000s)	115	231	376
Pension Wealth ($1,000s)	60	108	85
Wages ($/hour)	11.0	18.4	13.5
Probability of health insurance type, given preference type			
Health insurance = *none*	0.392	0.193	0.394
Health insurance = *retiree*	0.560	0.633	0.518
Health insurance = *tied*	0.047	0.174	0.089
Probability of preference index value, given preference type			
Preference Index = *out*	0.523	0.216	0.224
Preference Index = *low*	0.247	0.399	0.363
Preference Index = *high*	0.230	0.385	0.413
Fraction with preference type	0.246	0.635	0.119

Table 11: Mean Values by Preference Type, Alternative Specification

Table 11 contains summary statistics for the preference groups generated by this alternative specification. Comparing Table 11 to the baseline results contained in Table 5 reveals that eliminating the preference index from the type prediction equations changes only modestly the parameter estimates and the distribution of insurance coverage across the three preference types. The model without the preference index provides less evidence of self-selection: when the preference index is removed the fraction of high preference for work, type-2 individuals with *tied* coverage falls from 15.8% to 8.9%.

Table 10 shows that excluding the preference index only slightly changes the estimated

effect of Medicare and Social Security on labor supply. Given that self-selection has only a small effect on our results when we include the preference index, it should come as no surprise that self-selection has only a small effect when we exclude the index.

9 Conclusion

Prior to age 65, many individuals receive health insurance only if they continue to work. At age 65, however, Medicare provides health insurance to almost everyone. Therefore, a potentially important work incentive disappears at age 65. To see if Medicare benefits have a large effect on retirement behavior, we construct a retirement model that includes health insurance, uncertain medical costs, a savings decision, a non-negativity constraint on assets and a government-provided consumption floor.

Using data from the Health and Retirement Study, we estimate the structural parameters of our model. The model fits the data well, with reasonable preference parameters. In addition, the model does a satisfactory job of predicting the behavior of individuals who, by belonging to a younger cohort, faced different Social Security rules than the individuals upon which the model was estimated.

We find that health care uncertainty significantly affects the value of employer-provided health insurance. Our calculations suggest that about half of the value workers place on employer-provided health insurance comes from its ability to reduce medical expense risk. Furthermore, we find evidence that individuals with higher tastes for leisure are more likely to choose employers that provide health insurance to early retirees. Nevertheless, we find that Medicare is important for understanding retirement, especially for workers whose health insurance is tied to their job. For example, the effects of raising the Medicare eligibility age to 67 are just as large as the effects of reducing Social Security benefits.

References

[1] **Aaronson, D., and E. French,** "The Effect of Part-Time Work on Wages: Evidence from the Social Security Rules," *Journal of Labor Economics,* 2004, 22(2), 329-352.

[2] **Altonji, J., and L. Segal,** "Small Sample Bias in GMM Estimation of Covariance Structures," *Journal of Business and Economic Statistics*, 1996, 14(3), 353-366.

[3] **Blau, D.,** "Labor Force Dynamics of Older Men," *Econometrica*, 1994, 62(1), 117-156.

[4] **Blau, D. and D. Gilleskie,** "Retiree Health Insurance and the Labor Force Behavior of Older Men in the 1990's," *Review of Economics and Statistics*, 2001, 83(1), 64-80.

[5] **Blau, D. and D. Gilleskie,** "Health Insurance and Retirement of Married Couples," *Journal of Applied Econometrics*, 2006, 21(7), 935-953.

[6] **Blau, D. and D. Gilleskie,** "The Role of Retiree Health Insurance in the Employment Behavior of Older Men," *International Economic Review*, 2008, 49(2), 475-514.

[7] **The Boards of Trustees of the Hospital Insurance and Supplementary Medical Insurance Trust Funds,** *2009 Annual Report of the Boards of Trustees of the Hospital Insurance and Supplementary Medical Insurance Trust Funds*, 2009.

[8] **Bound, J., T. Stinebrickner, and T. Waidmann,** "Health, Economic Resources and the Work of Older Americans," *Journal of Econometrics*, 156(1): 106-129.

[9] **Buchinsky, M.,** "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *Journal of Human Resources*, 1998, 33, 88-126.

[10] **Cagetti, M.,** "Wealth Accumulation Over the Life Cycle and Precautionary Savings," *Journal of Business and Economic Statistics,* 2003, 21(3), 339-353.

[11] **Casanova, M.,** "Happy Together: A Structural Model of Couples' Joint Retirement Decisions," mimeo, 2010.

[12] **Chamberlain, G.,** "Comment: Sequential Moment Restrictions in Panel Data," *Journal of Business & Economic Statistics*, 1992, 10(1), 20-26.

[13] **Chernozhukov, V., and C. Hansen,** "An IV Model of Quantile Treatment Effects," MIT Working Paper 02-06, 2002.

[14] **Cogan, J.,** "Fixed Costs and Labor Supply," *Econometrica*, 1981, 49(4), 945-963.

[15] **Committee On Ways And Means, U.S. House Of Representatives,** *Green Book*, United States Government Printing Office, various years.

[16] **David, M., R. Little, M. Samuhel, and R. Triest,** "Alternative Methods for CPS Income Imputation," *Journal of the American Statistical Association*, 1986, 81(393), 29-41.

[17] **De Nardi, C.,** "Wealth Distribution, Intergenerational Links and Estate Taxation," *Review of Economic Studies,* 2004, 71(3), 743-768.

[18] **De Nardi, C., E. French, and J. Jones** "Why do the Elderly Save? The Role of Medical Expenses," *Journal of Political Economy,* 2010, 118(1), 39-75.

[19] **Duffie, D. and K. Singleton,** "Simulated Moments Estimation of Markov Models of Asset Prices," *Econometrica*, 1993, 61(4), 929-952.

[20] **Elder, T. and E. Powers,** "The Incredible Shrinking Program: Trends in SSI Participation of the Aged," *Research on Aging* 2006, 28(3), 341-358.

[21] **Employee Benefit Research Institute,** *EBRI Health Benefits Databook*, EBRI-ERF, 1999.

[22] **Epple, D. and H. Seig,** "Estimating Equilibrium Models of Local Jurisdictions," *Journal of Political Economy* 1999, 107(4), 645-681.

[23] **French, E.,** "The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior," *Review of Economic Studies,* 2005, 72(2), 395-427.

[24] **French, E., and J. Jones,** "On the Distribution and Dynamics of Health Care Costs," *Journal of Applied Econometrics,* 2004a, 19(4), 705-721.

[25] **French, E., and J. Jones,** "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," Center for Retirement Research Working Paper 2004-12, 2004b.

[26] **French, E., and J. Jones,** "The Effects of Health Insurance and Self-Insurance on Retirement Behavior," Michigan Retirement Research Center Working paper 2007-170, 2007.

[27] **Gourieroux, C., and A. Monfort,** *Simulation-Based Econometric Methods*, Oxford University Press, 1997.

[28] **Gourinchas, P. and Parker, J.,** "Consumption Over the Life Cycle," *Econometrica*, 2002, 70(1), 47-89.

[29] **Gruber, J., and B. Madrian,** "Health Insurance Availibility and the Retirement Decision," *American Economic Review*, 1995, 85(4), 938-948.

[30] **Gruber, J., and B. Madrian,** "Health Insurance and Early Retirement: Evidence from the Availability of Continuation Coverage," in D.A. Wise, ed., *Advances in the Economics of Aging* 1996, University of Chicago Press, 115-143.

[31] **Gustman, A., and T. Steinmeier,** "Employer-Provided Health Insurance and Retirement Behavior," *Industrial and Labor Relations Review* 1994, 48(1), 124-140.

[32] **Gustman, A., and T. Steinmeier,** "The Social Security Early Entitlement Age in a Structural Model of Retirement and Wealth," *Journal of Public Economics*, 2005, 89, 441-463.

[33] **Gustman, A., O. Mitchell, A. Samwick and T. Steinmeier,** "Evaluating Pension Entitlements," in O. Mitchell, P. Hammond, and A, Rappaport (eds.) *Forecasting Retirement Needs and Retirement Wealth*, 2000, University of Chicago Press, 309-326.

[34] **Heckman, J., and B. Singer,** "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration Data," *Econometrica*, 1984, 52(2), 271-320.

[35] **Hubbard, R., J. Skinner, and S. Zeldes,** "The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving," *Carnegie-Rochester Series on Public Policy*, 1994, 40, 59-125.

[36] **Hubbard, R., J. Skinner, and S. Zeldes,** "Precautionary Saving and Social Insurance," *Journal of Political Economy*, 1995, 103(2), 360-399.

[37] **Judd, K.,** *Numerical Methods in Economics*, MIT Press, 1998.

[38] **Kaiser/HRET,** *The 2006 Kaiser/HRET Employer Health Benefit Survey.* http://www.kff.org/insurance/7527/upload/7527.pdf, 2006.

[39] **Kahn, J.,** "Social Security, Liquidity, and Early Retirement," *Journal of Public Economics*, 1988, 35, 97-117.

[40] **Karoly, L., and J. Rogowski,** "The Effect of Access to Post-Retirement Health Insurance on the Decision to Retire Early," *Industrial and Labor Relations Review* 1994, 48(1), 103-123.

[41] **Keane, M., and K. Wolpin,** "The Career Decisions of Young Men," *Journal of Political Economy*, 1997, 105(3), 473-522.

[42] **Keane, M., and K. Wolpin,** "Exploring the Usefulness of a Non-Random Holdout Sample for Model Validation: Welfare Effects on Female Behavior," *International Economic Review*, 2007, 48(4), 1351-1378.

[43] **Little, R.,** "Missing Data Adjustments in Large Surveys," *Journal of Business and Economic Statistics,* 1988, 6(3), 287-301.

[44] **Lumsdaine, R., J. Stock, and D. Wise,** "Pension Plan Provisions and Retirement: Men, Women, Medicare and Models," in D. Wise (ed.) *Studies in the Economics of Aging*, 1994.

[45] **MaCurdy, T.,** "An Empirical Model of Labor Supply in a Life-Cycle Setting," *Journal of Political Economy*, 1981, 89(6), 1059-1085.

[46] **Madrian, B.,** "The Effect of Health Insurance on Retirement," *Brookings Papers on Economic Activity,* 1994, 181-252.

[47] **Manski, C.,** *Analog Estimation Methods in Econometrics*, Chapman and Hall, 1988.

[48] **Newey, W.,** "Generalized Method of Moments Specification Testing," *Journal of Econometrics*, 1985, 29(3), 229-256.

[49] **Newey, W. and D. McFadden,** "Large Sample Estimation abd Hypothesis Testing" in R. Engle and D. McFadden (eds.) *Handbook of Econometrics*, Vol. 4., Elsevier, 1994.

[50] **Pakes,A., and D. Pollard,** "Simulation and the Aysmptotics of Optimization Estimators," *Econometrica*, 1989, 57(5), 1027-1057.

[51] **Palumbo, M.,** "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," *Review of Economic Studies*, 1999, 66(2), 395-421.

[52] **Pischke, J-S.,** "Measurement Error and Earnings Dynamics: Some Estimates From the PSID Validation Study," *Journal of Business & Economics Statistics*, 1995, 13(3), 305-314.

[53] **Powell, J.,** "Estimation of Semiparametric Models" in R. Engle and D. McFadden (eds.) *Handbook of Econometrics*, Vol. 4., Elsevier, 1994.

[54] **Rogerson, R., and J. Wallenius,** "Retirement in a Life Cycle Model of Labor Supply with Home Production" Michigan Retirement Research Center Working Paper, 2009-205.

[55] **Rust, J. and C. Phelan,** "How Social Security and Medicare Affect Retirement Behavior in a World of Incomplete Markets," *Econometrica*, 1997, 65(4), 781-831.

[56] **Rust, J., Buchinsky, M., and H. Benitez-Silva,** "An Empirical Model of Social Insurance at the End of the Life Cycle," mimeo, 2003.

[57] **Song, J., and J. Manchester,** "New Evidence on Earnings and Benefit Claims Following Changes in the Retirement Earnings Test in 2000," *Journal of Public Economics*, 2007, 91(3), 669-700.

[58] **Tauchen, G.,** "Finite State Markov-chain Approximations to Univariate and Vector Autoregressions." *Economics Letters*, 1986, 20, 177-181.

[59] **United States Social Security Administration,** *Social Security Bulletin: Annual Statistical Supplement*, United States Government Printing Office, selected years.

[60] **van der Klaauw, W., and K. Wolpin,** "Social Security, Pensions and the Savings and Retirement Behavior of Households," *Journal of Econometrics*, 2008, 145(1-2), 21-42.

Appendix A: Cast of Characters

Preference Parameters		*Health-related Parameters*	
γ	consumption weight	H_t	health status
β	time discount factor	M_t	out-of-pocket medical expenses
ν	coefficient of RRA, utility	I_t	health insurance type
θ_B	bequest weight	$m(\cdot)$	mean shifter, logged medical expenses
κ	bequest shifter	$\sigma(\cdot)$	volatility shifter, logged medical expenses
C_{min}	consumption floor	ψ_t	idiosyncratic medical expense shock
L	leisure endowment	ζ_t	persistent medical expense shock
ϕ_H	leisure cost of bad health	ϵ_t	innovation, persistent shockk
ϕ_{Pt}	fixed cost of work	ρ_m	autocorrelation, persistent shock
ϕ_{P0}	fixed cost, intercept	σ^2_ϵ	innovation variance, persistent shock
ϕ_{P1}	fixed cost, time trend	ξ_t	transitory medical expense shock
ϕ_{RE}	re-entry cost	σ^2_ξ	variance, transitory shock
Decision Variables		*Wage-related Parameters*	
C_t	consumption	W_t	hourly wage
N_t	hours of work	$W(\cdot)$	mean shifter, logged wages
L_t	leisure	α	coefficient on hours, logged wages
P_t	participation	ω_t	idiosyncratic wage shock
A_t	assets	ρ_W	autocorrelation, wage shock
B_t	Social Security application	η_t	innovation, wage shock
Financial Variables		σ^2_η	innovation variance, wage shock
$Y(\cdot)$	after-tax income	*Miscellaneous*	
τ	tax parameter vector	s_t	survival probability
r	real interest rate	$pref$	discrete preference index
ys_t	spousal income	X_t	state vector, worker's problem
$ys(\cdot)$	mean shifter, spousal income	$\lambda(\cdot)$	compensating variation
ss_t	Social Security income	T	number of years in GMM criterion
$AIME_t$	Social Security wealth		
pb_t	pension benefits		

Table 12: Variable Definitions, Main Text

Appendix B: Taxes

Individuals pay federal, state, and payroll taxes on income. We compute federal taxes on income net of state income taxes using the Federal Income Tax tables for "Head of Household" in 1998. We use the standard deduction, and thus do not allow individuals to defer medical expenses as an itemized deduction. We also use income taxes for the fairly representative state of Rhode Island (27.5% of the Federal Income Tax level). Payroll taxes are 7.65% up to a maximum of $68,400, and are 1.45% thereafter. Adding up the three taxes generates the following level of post-tax income as a function of labor and asset income:

Pre-tax Income (Y)	Post-Tax Income	Marginal Tax Rate
0-6250	0.9235Y	0.0765
6250-40200	5771.88 + 0.7384(Y-6250)	0.2616
40200-68400	30840.56 + 0.5881(Y-40200)	0.4119
68400-93950	47424.98 + 0.6501(Y-68400)	0.3499
93950-148250	64035.03 + 0.6166(Y-93950)	0.3834
148250-284700	97515.41 + 0.5640(Y-148250)	0.4360
284700+	174474.21 + 0.5239(Y-284700)	0.4761

Table 13: After Tax Income

Appendix C: Pensions

Although the HRS pension data allow us to estimate pension wealth with a high degree of precision, Bellman's curse of dimensionality prevents us from including in our dynamic programming model the full range of pension heterogeneity found in the data. Thus we thus use the pension data to construct a simpler model of pensions. The fundamental equation behind our model of pensions is the accumulation equation for pension wealth, pw_t:

$$pw_{t+1} = \begin{cases} (1/s_{t+1})[(1+r)pw_t + pacc_t - pb_t] & \text{if living at } t+1 \\ 0 & \text{otherwise} \end{cases}, \tag{15}$$

where $pacc_t$ is pension accrual and pb_t is pension benefits. Two features of this equation bear noting. First, a pension is worthless once an individual dies. Dividing through by the survival probability s_{t+1} ensures that the expected value of pensions $E(pw_{t+1}|pw_t, pacc_t, pb_t)$ equals $(1+r)pw_t + pacc_t - pb_t$, the actuarially fair amount. Second, since pension accrual and pension interest are not directly taxed, the appropriate rate of return on pension wealth is the pre-tax one. Pension benefits, on the other hand, are included in the income used to calculate an individual's income tax liability.

Simulating equation (15) requires us to know pension benefits and pension accrual. We calculate pension benefits by assuming that at age t, the pension benefit is

$$pb_t = pf_t \times pb_t^{\max}, \tag{16}$$

where $pb_t^{\max}$ is the benefit received by individuals actually receiving pensions (given the

earnings history observed at time t) and pf_t the probability that a person with a pension is currently drawing pension benefits. We estimate pf_t as the fraction of respondents who are covered by a pension that receive pension benefits at each age; the fraction increases fairly smoothly, except for a 23-percentage-point jump at age 62. To find the annuity $pb_t^{\max}$ given pension wealth at time t (and assuming no further pension accruals so that $pacc_k = 0$ for $k = t, t+1, ..., T$), note first that recursively substituting equation (15) and imposing $pw_{T+1} = 0$ reveals that pension wealth is equal to the present discounted value of future pension benefits:

$$pw_t = \frac{1}{1+r} \sum_{k=t}^{T} \frac{S(k,t)}{(1+r)^{k-t}} pf_k pb_k^{\max}, \tag{17}$$

where $S(k,t) = (1/s_t) \prod_{j=t}^{k} s_j$ gives the probability of surviving to age k, conditional on having survived to time t. If we assume further that the maximum pension benefit is constant from time t forward, so that $pb_k^{\max} = pb_t^{\max}$, $k = t, t+1, ..., T$, this equation reduces to

$$pw_t = \Gamma_t pb_t^{\max}, \tag{18}$$

$$\Gamma_t \equiv \frac{1}{1+r} \sum_{k=t}^{T} \frac{S(k,t)}{(1+r)^{k-t}} pf_k. \tag{19}$$

Using equations (16) and (18), pension benefits are thus given by

$$pb_t = pf_t \Gamma_t^{-1} pw_t. \tag{20}$$

Next, we assume pension accrual is given by

$$pacc_t = \alpha_0(I_t, W_t N_t, t) \times W_t N_t, \tag{21}$$

where $\alpha_0(.)$ is the pension accrual rate as a function of health insurance type, labor income, and age. We estimate $\alpha_0(.)$ in two steps, estimating separately each component of:

$$\alpha_0 = E(pacc_t | W_t N_t, I_t, t, pen_t = 1) \Pr(pen_t = 1 | I_t, W_t N_t) \tag{22}$$

where $pacc_t$ is the accrual rate for those with a pension, and pen_t is a 0-1 indicator equal to 1 if the individual has a pension.

We estimate the first component, $E(pacc_t|W_tN_t, I_t, t, pen_t = 1)$, from restricted HRS pension data. To generate a pension accrual rate for each individual, we combine the pension data with the HRS pension calculator to estimate the pension wealth that each individual would have if he left his job at different ages. The increase in pension wealth gained by working one more year is the accrual. Assuming that pension benefits are 0 as long as the worker continues working, it follows from equation (15) that

$$pacc_t = s_{t+1}pw_{t+1} - (1+r)pw_t. \tag{23}$$

The HRS pension data have a high degree of employer- and worker-level detail, allowing us to estimate pension accrual accurately. With accruals in hand, we then estimate $E(pacc_t|W_tN_t, I_t, t, pen_t = 1)$ by regressing accrual rates on a fourth-order age polynomial, indicators for age greater than 62 or 65, log income, log income interacted with the age variables, health insurance indicators, and health insurance indicators interacted with the age variables, using the subset of workers that have a pension on their current job.

Figure 7 shows estimated pension accrual, by health insurance type and earnings. It shows that those with retiree coverage have the sharpest declines in pension accrual after age 60. It also shows that once health insurance and the probability of having a pension plan are accounted for, the effect of income on pension accrual is relatively small. Our estimated age (but not health insurance) pension accrual rates line up closely with Gustman et al. (1998), who also use the restricted firm-based HRS pension data.

In the second step, we estimate the probability of having a pension, $\Pr(pen_t = 1|I_t, W_tN_t, t)$, using unrestricted self-reported data from individuals who are working and are ages 51-55. The function $\Pr(pen_t = 1|I_t, W_tN_t, t)$ is estimated as a logistic function of log income, health insurance indicators, and interactions between log income and health insurance.

Table 14 shows the probability of having different types of pensions, conditional on health

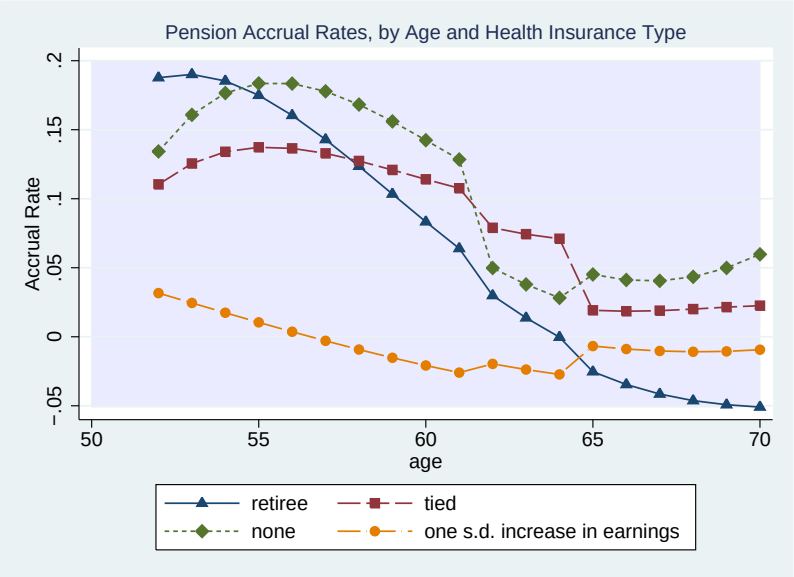

Figure 7: Pension Accrual Rates for Individuals with Pensions, by Age, Health Insurance Coverage and Earnings

insurance. The table shows that only 8% of men with no health insurance have a pension, but 64% of men with tied coverage and 74% of men with retiree insurance have a pension. Furthermore, it shows that those with retiree coverage are also the most likely to have defined benefit (DB) pension plans, which provide the strongest retirement incentives after age 62.

	Probability of Pension Type		
Variable	No Insurance	Retiree Insurance	Tied Insurance
Defined Benefit	.026	.412	.260
Defined Contribution	.050	.172	.270
Both DB and DC	.006	.160	.106
Total	.082	.744	.636
Number of Observations	343	955	369

Table 14: Probability of having a pension on the current job, by health insurance type, working men, age 51-55

Combining the restricted data with the HRS pension calculator also yields initial pension balances as of 1992. Mean pension wealth in our estimation sample is $93,300. Disaggregating by health insurance type, those with retiree coverage have $129,200, those with tied coverage have $80,000, and those with none have $17,300. With these starting values, we simulate pension wealth in our dynamic programming model with equation (15), using equation (21)

to estimate pension accrual, and using equation (20) to estimate pension benefits. Using these equations, it is straightforward to track and record the pension balances of each simulated individual.

But even though it is straightforward to use equation (15) when computing pension wealth in the simulations, it is too computationally burdensome to include pension wealth as a separate state variable when computing the decision rules. Our approach is to impute pension wealth as a function of age and AIME. In particular, we impute a worker's annual pension benefits as a function of his Social Security benefits:

$$\begin{aligned}\widehat{pb}_t(PIA_t, I_{t-1}, t) \;=\; & \sum_{k\in\{retiree,tied,none\}} [\gamma_{0,k,0} + \gamma_{0,k,1}t + \gamma_{0,k,2}t^2] \cdot 1\{I_{t-1} = k\} \\ & + \gamma_3 PIA_t + \; [\gamma_{4,0} + \gamma_{4,1}t + \gamma_{4,2}t^2] \cdot \max\{0, PIA_t - 9,999.6\} \\ & + [\gamma_{5,0} + \gamma_{5,1}t + \gamma_{5,2}t^2] \cdot \max\{0, PIA_t - 14,359.9\}, \qquad (24)\end{aligned}$$

where PIA_t is the Social Security benefit the worker would get if he were drawing benefits at time t; as shown in Appendix D below, PIA is a monotonic function of AIME. Using equations (18) and (24) yields imputed pension wealth, $\widehat{pw}_t = \Gamma_t \widehat{pb}_t$. Equation (24) is estimated with regressions on simulated data generated by the model. Since these simulated data depend on the γ's—$\widehat{pw}_t$ affects the decision rules used in the simulations—the γ's solve a fixed-point problem. Fortunately, estimates of the γ's converge after a few iterations.

This imputation process raises two complications. The first is that we use a different pension wealth imputation formula when calculating decision rules than we do in the simulations. If an individual's time-t pension wealth is $\widehat{pw}_t$, his time-$t+1$ pension wealth (if living) should be

$$\widehat{\widehat{pw}}_{t+1} = (1/s_{t+1})[(1+r)\widehat{pw}_t + pacc_t - pb_t].$$

This quantity, however, might differ from the pension wealth that would be imputed using PIA_{t+1}, $\widehat{pw}_{t+1} = \Gamma_{t+1}\widehat{pb}_{t+1}$ where $\widehat{pb}_{t+1}$ is defined in equation (24). To correct for this, we increase non-pension wealth, A_{t+1}, by $s_{t+1}(1-\tau_t)(\widehat{\widehat{pw}}_{t+1} - \widehat{pw}_{t+1})$. The first term in this

expression reflects the fact that while non-pension assets can be bequeathed, pension wealth cannot. The second term, $1-\tau_t$, reflects the fact that pension wealth is a pre-tax quantity—pension benefits are more or less wholly taxable—while non-pension wealth is post-tax—taxes are levied only on interest income.

A second problem is that while an individual's Social Security application decision affects his annual Social Security benefits, it should not affect his pension benefits. (Recall that we reduce PIA if an individual draws benefits before age 65.) The pension imputation procedure we use, however, would imply that it does. We counter this problem by recalculating PIA when the individual begins drawing Social Security benefits. In particular, suppose that a decision to accelerate or defer application changes PIA_t to $rem_t PIA_t$. Our approach is to use equation (24) find a value PIA_t^* such that

$$(1-\tau_t)\widehat{pb}_t(PIA_t^*) + PIA_t^* = (1-\tau_t)\widehat{pb}_t(PIA_t) + rem_t PIA_t,$$

so that the change in the sum of PIA and imputed after-tax pension income equals just the change in PIA, i.e., $(1-rem_t)PIA_t$.

Appendix D: Computation of AIME

We model several key aspects of Social Security benefits. First, Social Security benefits are based on the individual's 35 highest earnings years, relative to average wages in the economy during those years. The average earnings over these 35 highest earnings years are called Average Indexed Monthly Earnings, or AIME. It immediately follows that working an additional year increases the AIME of an individual with less than 35 years of work. If an individual has already worked 35 years, he can still increase his AIME by working an additional year, but only if his current earnings are higher than the lowest earnings embedded in his current AIME. To account for real wage growth, earnings in earlier years are inflated by the growth rate of average earnings in the overall economy. For the period 1992-1999, average real wage growth, g, was 0.016 (Committee on Ways and Means, 2000, p. 923). This indexing stops at the year the worker turns 60, however, and earnings accrued after age 60

are not rescaled.[24] Furthermore, AIME is capped. In 1998, the base year for the analysis, the maximum AIME level was $68,400.

Precisely modelling these mechanics would require us to keep track of a worker's entire earnings history, which is computationally infeasible. As an approximation, we assume that (for workers beneath the maximum) annualized AIME is given by

$$\begin{aligned} AIME_{t+1} = &(1 + g \times 1\{t \leq 60\})AIME_t \\ &+ \frac{1}{35}\max\left\{0,\ W_t N_t - \alpha_t(1 + g \times 1\{t \leq 60\})AIME_t\right\}, \end{aligned} \tag{25}$$

where the parameter α_t approximates the ratio of the lowest earnings year to $AIME$. We assume that 20% of the workers enter the labor force each year between ages 21 and 25, so that $\alpha_t = 0$ for workers aged 55 and younger. For workers aged 60 and older, earnings update $AIME_t$ only if current earnings replace the lowest year of earnings, so we estimate α_t by simulating wage (not earnings) histories with the model developed in French (2005), calculating the sequence $\left\{1\{\text{time-}t \text{ earnings do not increase } AIME_t\}\right\}_{t \geq 60}$ for each simulated wage history, and estimating α_t as the average of this indicator at each age. Linear interpolation yields α_{56} through α_{59}.

AIME is converted into a Primary Insurance Amount (PIA) using the formula

$$PIA_t = \begin{cases} 0.9 \times AIME_t & \text{if } AIME_t < \$5,724 \\ \$5,151.6 + 0.32 \times (AIME_t - 5,724) & \text{if } \$5,724 \leq AIME_t < \$34,500 \\ \$14,359.9 + 0.15 \times (AIME_t - 34,500) & \text{if } AIME_t \geq \$34,500 \end{cases}. \tag{26}$$

Social Security benefits ss_t depend both upon the age at which the individual first receives Social Security benefits and the Primary Insurance Amount. For example, pre-Earnings Test benefits for a Social Security beneficiary will be equal to PIA if the individual first receives benefits at age 65. For every year before age 65 the individual first draws benefits, benefits are reduced by 6.67% and for every year (up until age 70) that benefit receipt is delayed,

[24] After age 62, nominal benefits increase at the rate of inflation.

benefits increase by 5.0%. The effects of early or late application can be modelled as changes in AIME rather than changes in PIA, eliminating the need to include age at application as a state variable. For example, if an individual begins drawing benefits at age 62, his adjusted AIME must result in a PIA that is only 80% of the PIA he would have received had he first drawn benefits at age 65. Using equation (26), this is easy to find.

Appendix E: Numerical Methods

Because the model has no closed form solution, the decision rules it generates must be found numerically. We find the decision rules using value function iteration, starting at time T and working backwards to time 1. We find the time-T decisions by maximizing equation (14) at each value of X_T, with $V_{T+1} = b(A_{T+1})$. This yields decision rules for time T and the value function V_T. We next find the decision rules at time $T-1$ by solving equation (14), having solved for V_T already. Continuing this backwards induction yields decision rules for times $T-2, T-3, ..., 1$.

The value function is directly computed at a finite number of points within a grid, $\{X_i\}_{i=1}^{I}$;[25] We use linear interpolation within the grid (i.e., we take a weighted average of the value functions of the surrounding gridpoints) and linear extrapolation outside of the grid to evaluate the value function at points that we do not directly compute. Because changes in assets and AIME are likely to cause larger behavioral responses at low levels of assets and AIME, the grid is more finely discretized in this region.

At time t, wages, medical expenses and assets at time $t+1$ will be random variables. To capture uncertainty over the persistent components of medical expenses and wages, we convert ζ_t and ω_{t+1} into discrete Markov chains, following the approach of Tauchen (1986); using discretization rather than quadrature greatly reduces the number of times one has to interpolate when calculating $E_t(V(X_{t+1}))$. We integrate the value function with respect to

[25] In practice, the grid consists of: 32 asset states, $A_h \in [-\$55{,}000, \$1{,}200{,}000]$; 5 wage residual states, $\omega_i \in [-0.99, 0.99]$; 16 AIME states, $AIME_j \in [\$4{,}000, \$68{,}400]$; 3 states for the persistent component of medical expenses, ζ_k, over a normalized (unit variance) interval of $[-1.5, 1.5]$. There are also two application states, two health states, and two states for participation in the previous period. This requires solving the value function at $61,440$ different points for ages 62-69, when the individual is eligible to apply for benefits, at $31,260$ points before age 62 (when application is not an option) or at ages 70-71 (when we impose application), and at $15,360$ points after age 71 (when we impose retirement as well).

the transitory component of medical expenses, ξ_t, using 5-node Gauss-Hermite quadrature (see Judd, 1999).

Because of the fixed time cost of work and the discrete benefit application decision, the value function need not be globally concave. This means that we cannot find a worker's optimal consumption and hours with fast hill climbing algorithms. Our approach is to discretize the consumption and labor supply decision space and to search over this grid. Experimenting with the fineness of the grids suggested that the grids we used produced reasonable approximations.[26] In particular, increasing the number of grid points seemed to have a small effect on the computed decision rules.

We then use the decision rules to generate simulated time series. Given the realized state vector X_{i0}, individual i's realized decisions at time 0 are found by evaluating the time-0 decision functions at X_{i0}. Using the transition functions given by equations (4) through (13), we combine X_{i0}, the time-0 decisions, and the individual i's time-1 shocks to get the time-1 state vector, X_{i1}. Continuing this forward induction yields a life cycle history for individual i. When X_{it} does not lie exactly on the state grid, we use interpolation or extrapolation to calculate the decision rules. This is true for ζ_t and ω_t as well. While these processes are approximated as finite Markov chains when the decision rules are found, the simulated sequences of ζ_t and ω_t are generated from continuous processes. This makes the simulated life cycle profiles less sensitive to the discretization of ζ_t and ω_t than when ζ_t and ω_t are drawn from Markov chains.

Finally, to reduce the computational burden, we assume that all workers apply for Social Security benefits by age 70, and retire by age 72: for $t \geq 70$, $B_t = 1$; and for $t \geq 72$, $N_t = 0$.

[26] The consumption grid has 100 points, and the hours grid is broken into 500-hour intervals. When this grid is used, the consumption search at a value of the state vector X for time t is centered around the consumption gridpoint that was optimal for the same value of X at time $t+1$. (Recall that we solve the model backwards in time.) If the search yields a maximizing value near the edge of the search grid, the grid is reoriented and the search continued. We begin our search for optimal hours at the level of hours that sets the marginal rate of substitution between consumption and leisure equal to the wage. We then try 6 different hours choices in the neighborhood of the initial hours guess. Because of the fixed cost of work, we also evaluate the value function at $N_t = 0$, searching around the consumption choice that was optimal when $H_{t+1} = 0$. Once these values are found, we perform a quick, "second-pass" search in a neighborhood around them.

Appendix F: Moment Conditions, Estimation Mechanics, and the Asymptotic Distribution of Parameter Estimates

Following Gourinchas and Parker (2002) and French (2005), we estimate the parameters of the model in two steps. In the first step we estimate or calibrate parameters that can be cleanly identified without explicitly using our model. For example, we estimate mortality rates and health transitions from demographic data. As a matter of notation, we call this set of parameters χ. In the second step, we estimate the vector of "preference" parameters, $\theta = \big(\gamma_0, \gamma_1, \gamma_2, \beta_0, \beta_1, \beta_2, \nu, L, \phi_{P0}, \phi_{P1}, \phi_{RE}, \theta_B, \kappa, C_{min}, \text{preference type prediction coefficients}\big)$, using the method of simulated moments (MSM).

We assume that the "true" preference vector θ_0 lies in the interior of the compact set $\Theta \subset \mathbb{R}^{39}$. Our estimate, $\hat{\theta}$, is the value of θ that minimizes the (weighted) distance between the estimated life cycle profiles for assets, hours, and participation found in the data and the simulated profiles generated by the model. We match $21T$ moment conditions. They are, for each age $t \in \{1, ..., T\}$: two asset quantiles (forming $2T$ moment conditions), labor force participation rates conditional on asset quantile and health insurance type ($9T$), labor market exit rates for each health insurance type ($3T$), labor force participation rates conditional on the preference indicator described in the main text ($3T$), and labor force participation rates and mean hours worked conditional upon health status ($4T$).

Consider first the asset quantiles. As stated in the main text, let $j \in \{1, 2, ..., J\}$ index asset quantiles, where J is the total number of asset quantiles. Assuming that the age-conditional distribution of assets is continuous, the π_j-th age-conditional asset quantile, $Q_{\pi_j}(A_{it}, t)$, is defined as

$$\Pr\left(A_{it} \leq Q_{\pi_j}(A_{it}, t)|t\right) = \pi_j.$$

In other words, the fraction of age-t individuals with less than Q_{π_j} in assets is π_j. As is well known (see, e.g., Manski, 1988, Powell, 1994 or Buchinsky, 1998; or the review in Chernozhukov and Hansen, 2002), the preceding equation can be rewritten as a moment condition

by using the indicator function:

$$E\left(1\{A_{it} \leq Q_{\pi_j}(A_{it}, t)\}|t\right) = \pi_j. \tag{27}$$

The model analog to $Q_{\pi_j}(A_{it}, t)$ is $g_{\pi_j}(t; \theta_0, \chi_0)$, the jth quantile of the simulated asset distribution. If the model is true then the data quantile in equation (27) can be replaced by the model quantile, and equation (27) can be rewritten as:

$$E\left(1\{A_{it} \leq g_{\pi_j}(t; \theta_0, \chi_0)\} - \pi_j|t\right) = 0, \quad j \in \{1, 2, ..., J\},\ t \in \{1, ..., T\}. \tag{28}$$

Since $J = 2$, equation (28) generates $2T$ moment conditions.

Equation (28) is a departure from the usual practice of minimizing a sum of weighted absolute errors in quantile estimation. The quantile restrictions just described, however, are part of a larger set of moment conditions, which means that we can no longer estimate θ by minimizing weighted absolute errors. Our approach to handling multiple quantiles is similar to the minimum distance framework used by Epple and Seig (1999).[27]

The next set of moment conditions uses the quantile-conditional means of labor force participation. Let $\overline{P}_j(I, t; \theta_0, \chi_0)$ denote the model's prediction of labor force participation given asset quantile interval j, health insurance type I, and age t. If the model is true, $\overline{P}_j(I, t; \theta_0, \chi_0)$ should equal the conditional participation rates found in the data:

$$\overline{P}_j(I, t; \theta_0, \chi_0) = E[P_{it} \mid I, t, g_{\pi_{j-1}}(t; \theta_0, \chi_0) \leq A_{it} \leq g_{\pi_j}(t; \theta_0, \chi_0)], \tag{29}$$

with $\pi_0 = 0$ and $\pi_{J+1} = 1$. Using indicator function notation, we can convert this conditional

[27] Buchinsky (1998) shows that one could include the first-order conditions from multiple absolute value minimization problems in the moment set. However, his approach involves finding the gradient of $g_{\pi_j}(t; \theta, \chi)$ at each step of the minimization search.

moment equation into an unconditional one (e.g., Chamberlain, 1992):

$$E([P_{it} - \overline{P}_j(I,t;\theta_0,\chi_0)] \times 1\{I_{it} = I\} \\ \times 1\{g_{\pi_{j-1}}(t;\theta_0,\chi_0) \leq A_{it} \leq g_{\pi_j}(t;\theta_0,\chi_0)\} \mid t) = 0, \tag{30}$$

for $j \in \{1,2,...,J+1\}$, $I \in \{none, retiree, tied\}$, $t \in \{1,...,T\}$. Note that $g_{\pi_0}(t) \equiv -\infty$ and $g_{\pi_{J+1}}(t) \equiv \infty$. With 2 quantiles (generating 3 quantile-conditional means) and 3 health insurance types, equation (29) generates $9T$ moment conditions.

As described in Appendix J, we use HRS attitudinal questions to construct the preference index $pref \in \{high, low, out\}$. Considering how participation varies across this index leads to the following moment condition:

$$E\left(P_{it} - \overline{P}(pref,t;\theta_0,\chi_0) \mid pref_i = pref, t\right) = 0, \tag{31}$$

for $t \in \{1,...,T\}$, $pref \in \{0,1,2\}$. Equation (31) yields $3T$ moment conditions, which are converted into unconditional moment equations with indicator functions.

We also match exit rates for each health insurance category. Let $\overline{EX}(I,t;\theta_0,\chi_0)$ denote the fraction of time-$t-1$ workers predicted to leave the labor market at time t. The associated moment condition is

$$E\left([1 - P_{it}] - \overline{EX}(I,t;\theta_0,\chi_0 \mid I_{i,60} = I, P_{i,t-1} = 1, t\right) = 0, \tag{32}$$

for $I \in \{none, retiree, tied\}$, $t \in \{1,...,T\}$. Equation (32) generates $3T$ moment conditions, which are converted into unconditional moments as well.[28]

Finally, consider health-conditional hours and participation. Let $\overline{\ln N}(H,t;\theta_0,\chi_0)$ and $\overline{P}(H,t;\theta_0,\chi_0)$ denote the conditional expectation functions for hours (when working) and

[28] Because exit rates apply only to those working in the previous period, they normally do not contain the same information as participation rates. However, this is not the case for workers with *tied* coverage, as a worker stays in the *tied* category only as long as he continues to work. To remove this redundancy, the exit rates in equation (32) are conditioned on the individual's age-60 health insurance coverage, while the participation rates in equation (29) are conditioned on the individual's current coverage.

participation generated by the model for workers with health status H; let $\ln N_{it}$ and P_{it} denote measured hours and participation. The moment conditions are

$$E\left(\ln N_{it} - \overline{\ln N}(H,t;\theta_0,\chi_0) \mid P_{it} > 0, H_{it} = H, t\right) = 0, \tag{33}$$

$$E\left(P_{it} - \overline{P}(H,t;\theta_0,\chi_0) \mid H_{it} = H, t\right) = 0, \tag{34}$$

for $t \in \{1,...,T\}$, $H \in \{0,1\}$. Equations (33) and (34), once again converted into unconditional form, yield $4T$ moment conditions, for a grand total of $21T$ moment conditions.

Combining all the moment conditions described here is straightforward: we simply stack the moment conditions and estimate jointly.

Suppose we have a data set of I independent individuals that are each observed for T periods. Let $\varphi(\theta;\chi_0)$ denote the $21T$-element vector of moment conditions that was described in the main text and immediately above, and let $\hat{\varphi}_I(.)$ denote its sample analog. Note that we can extend our results to an unbalanced panel, as we must do in the empirical work, by simply allowing some of the individual's contributions to $\varphi(.)$ to be "missing", as in French and Jones (2004a). Letting $\widehat{\mathbf{W}}_I$ denote a $21T \times 21T$ weighting matrix, the MSM estimator $\hat{\theta}$ is given by

$$\arg\min_{\theta} \frac{I}{1+\tau}\, \hat{\varphi}_I(\theta,\chi_0)' \widehat{\mathbf{W}}_I \hat{\varphi}_I(\theta,\chi_0), \tag{35}$$

where τ is the ratio of the number of observations to the number of simulated observations.

To find the solution to equation (35), we proceed as follows:

1. We aggregate the sample data into life cycle profiles for hours, participation, exit rates and assets.

2. Using the same data used to estimate the profiles, we generate an initial distribution for health, health insurance status, wages, medical expenses, AIME, and assets. See Appendix G for details. We also use these data to estimate many of the parameters contained in the belief vector χ, although we calibrate other parameters. The initial distribution also includes preference type, assigned using our type prediction equation.

3. Using χ, we generate matrices of random health, wage, mortality and medical expense shocks. The matrices hold shocks for 90,000 simulated individuals.

4. We compute the decision rules for an initial guess of the parameter vector θ, using χ and the numerical methods described in Appendix E.

5. We simulate profiles for the decision variables. Each simulated individual receives a draw of preference type, assets, health, wages and medical expenses from the initial distribution, and is assigned one of the simulated sequences of health, wage and medical expense shocks. With the initial distributions and the sequence of shocks, we then use the decision rules to generate that person's decisions over the life cycle. Each period's decisions determine the conditional distribution of the next period's states, and the simulated shocks pin the states down exactly.

6. We aggregate the simulated data into life cycle profiles.

7. We compute moment conditions, i.e., we find the distance between the simulated and true profiles, as described in equation (35).

8. We pick a new value of θ, update the simulated distribution of preference types, and repeat steps 4-7 until we find the value of θ that minimize that minimizes the distance between the true data and the simulated data. This vector of parameter values, $\hat{\theta}$, is our estimated value of θ_0.[29]

Under the regularity conditions stated in Pakes and Pollard (1989) and Duffie and Singleton (1993), the MSM estimator $\hat{\theta}$ is both consistent and asymptotically normally distributed:

$$\sqrt{I}(\hat{\theta} - \theta_0) \rightsquigarrow N(0, \mathbf{V}),$$

[29]Because the GMM criterion function is discontinuous, we search over the parameter space using a Simplex algorithm written by Honore and Kyriazidou. It usually takes 2-4 weeks to estimate the model on a 48-node cluster, with each iteration (of steps 4-7) taking around 15 minutes.

with the variance-covariance matrix $\mathbf{V}$ given by

$$\mathbf{V} = (1+\tau)(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1},$$

where: $\mathbf{S}$ is the $21T \times 21T$ variance-covariance matrix of the data;

$$\mathbf{D} = \left.\frac{\partial\varphi(\theta,\chi_0)}{\partial\theta'}\right|_{\theta=\theta_0} \tag{36}$$

is the $21T \times 39$ Jacobian matrix of the population moment vector; and $\mathbf{W} = \text{plim}_{\to\infty}\{\widehat{\mathbf{W}}_I\}$. Moreover, Newey (1985) shows that if the model is properly specified,

$$\frac{I}{1+\tau}\hat{\varphi}_I(\hat{\theta},\chi_0)'\mathbf{R}^{-1}\hat{\varphi}_I(\hat{\theta},\chi_0) \rightsquigarrow \chi^2_{21T-39},$$

where $\mathbf{R}^{-1}$ is the generalized inverse of

$$\begin{aligned}\mathbf{R} &= \mathbf{P}\mathbf{S}\mathbf{P},\\ \mathbf{P} &= \mathbf{I} - \mathbf{D}(\mathbf{D}'\mathbf{W}\mathbf{D})^{-1}\mathbf{D}'\mathbf{W}.\end{aligned}$$

The asymptotically efficient weighting matrix arises when $\widehat{\mathbf{W}}_I$ converges to $\mathbf{S}^{-1}$, the inverse of the variance-covariance matrix of the data. When $\mathbf{W} = \mathbf{S}^{-1}$, $\mathbf{V}$ simplifies to $(1+\tau)(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}$, and $\mathbf{R}$ is replaced with $\mathbf{S}$. But even though the optimal weighting matrix is asymptotically efficient, it can be severely biased in small samples. (See, for example, Altonji and Segal, 1996.) We thus use a "diagonal" weighting matrix, as suggested by Pischke (1995). The diagonal weighting scheme uses the inverse of the matrix that is the same as $\mathbf{S}$ along the diagonal and has zeros off the diagonal of the matrix.

We estimate $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{W}$ with their sample analogs. For example, our estimate of $\mathbf{S}$ is the $21T \times 21T$ estimated variance-covariance matrix of the sample data. That is, one diagonal element of $\widehat{\mathbf{S}}_I$ will be the variance estimate $\frac{1}{I}\sum_{i=1}^{I}[1\{A_{it} \le Q_{\pi_j}(A_{it},t)\} - \pi_j]^2$, while a typical off-diagonal element is a covariance. When estimating parameters, we use sample statistics,

so that $Q_{\pi_j}(A_{it}, t)$ is replaced with the sample quantile $\widehat{Q}_{\pi_j}(A_{it}, t)$. When computing the chi-square statistic and the standard errors, we use model predictions, so that Q_{π_j} is replaced with its simulated counterpart, $g_{\pi_j}(t; \hat{\theta}, \hat{\chi})$. Covariances between asset quantiles and hours and labor force participation are also simple to compute.

The gradient in equation (36) is straightforward to estimate for most moment conditions; we merely take numerical derivatives of $\hat{\varphi}_I(.)$. However, in the case of the asset quantiles and quantile-conditional labor force participation, discontinuities make the function $\hat{\varphi}_I(.)$ non-differentiable at certain data points. Therefore, our results do not follow from the standard GMM approach, but rather the approach for non-smooth functions described in Pakes and Pollard (1989), Newey and McFadden (1994, section 7) and Powell (1994). We find the asset quantile component of $\mathbf{D}$ by rewriting equation (28) as

$$F(g_{\pi_j}(t; \theta_0, \chi_0)|t) - \pi_j = 0,$$

where $F(g_{\pi_j}(t; \theta_0, \chi_0)|t)$ is the empirical c.d.f. of time-t assets evaluated at the model-predicted π_j-th quantile. Differentiating this equation yields

$$\mathbf{D}_{jt} = f(g_{\pi_j}(t; \theta_0, \chi_0)|t) \frac{\partial g_{\pi_j}(t; \theta_0, \chi_0)}{\partial \theta'}, \tag{37}$$

where $\mathbf{D}_{jt}$ is the row of $\mathbf{D}$ corresponding to the π_j-th quantile at year t. In practice we find $f(g_{\pi_j}(t; \theta_0, \chi_0)|t)$, the p.d.f. of time-$t$ assets evaluated at the π_j-th quantile, with a kernel density estimator. We use a kernel estimator for GAUSS written by Ruud Koning.

To find the component of the matrix $\mathbf{D}$ for the asset-conditional labor force participation rates, it is helpful to write equation (30) as

$$\Pr(I_{t-1} = I) \times \int_{g_{\pi_{j-1}}(t;\theta_0,\chi_0)}^{g_{\pi_j}(t;\theta_0,\chi_0)} \left[E(P_{it}|A_{it}, I, t) - \overline{P}_j(I, t; \theta_0, \chi_0) \right] f(A_{it}|I, t) dA_{it} = 0,$$

which implies that

$$\begin{aligned}\mathbf{D}_{jt} = \Big[& - \Pr(g_{\pi_{j-1}}(t;\theta_0,\chi_0) \leq A_{it} \leq g_{\pi_j}(t;\theta_0,\chi_0)|I,t)\frac{\partial \overline{P}_j(I,t;\theta_0,\chi_0)}{\partial \theta'} \\ & + [E(P_{it}|g_{\pi_j}(t;\theta_0,\chi_0),I,t) - \overline{P}_j(I,t;\theta_0,\chi_0)]f(g_{\pi_j}(t;\theta_0,\chi_0)|I,t)\frac{\partial g_{\pi_j}(t;\theta_0,\chi_0)}{\partial \theta'} \\ & - [E(P_{it}|g_{\pi_{j-1}}(t;\theta_0,\chi_0),I,t) - \overline{P}_j(I,t;\theta_0,\chi_0)]f(g_{\pi_j-1}(t;\theta_0,\chi_0)|I,t)\frac{\partial g_{\pi_{j-1}}(t;\theta_0,\chi_0)}{\partial \theta'}\Big] \\ & \times \Pr(I_{t-1} = I), \end{aligned} \tag{38}$$

with $f(g_{\pi_0}(t;\theta_0,\chi_0)|I,t)\frac{\partial g_{\pi_0}(t;\theta_0,\chi_0)}{\partial \theta'} = f(g_{\pi_{J+1}}(t;\theta_0,\chi_0)|I,t)\frac{\partial g_{\pi_{J+1}}(t;\theta_0,\chi_0)}{\partial \theta'} \equiv 0$.

Appendix G: Data and Initial Joint Distribution of the State Variables

Our data are drawn from the HRS, a sample of non-institutionalized individuals aged 51-61 in 1992. The HRS surveys individuals every two years; we have 8 waves of data covering the period 1992-2006. We use men in the analysis.

We dropped respondents for the following reasons. First, we drop all individuals who spent over 5 years working for an employer who did not contribute to Social Security. These individuals usually work for state governments. We drop these people because they often have very little in the way of Social Security wealth, but a great deal of pension wealth, a type of heterogeneity our model is not well suited to handle. Second, we drop respondents with missing information on health insurance, labor force participation, hours, and assets. When estimating labor force participation by asset quantile and health insurance for those born 1931-35 for the estimation sample [and 1936-41 for the validation sample], we begin with 21,376 [36,702] person year observations. We lose 3,872 [6,919] observations because of missing labor force participation, 2,109 [2,480] observations who worked over 5 years for firms that did not contribute to Social Security, 602 [1,074] observations due to missing wave 1 labor force participation (needed to construct the preference index), and 2,103 [3,023] observations due to missing health insurance data. In the end, from a potential sample of 21,376 [36,702] person-year observations for those between ages 51 and 69, we keep 12,870 [23,206] observations.

The labor market measures used in our analysis are constructed as follows. Hours of work are the product of usual hours per week and usual weeks per year. To compute hourly wages, we use information on how respondents are paid, how often they are paid, and how much they are paid. For salaried workers, annual earnings are the product of pay per period and the number of pay periods per year. The wage is then annual earnings divided by annual hours. If the worker is hourly, we use his reported hourly wage. We treat a worker's hours for the non-survey (e.g. 1993) years as missing.

For survey years the individual is considered in the labor force if he reports working over 300 hours per year. The HRS also asks respondents retrospective questions about their work history. Because we are particularly interested in labor force participation, we use the work history to construct a measure of whether the individual worked in non-survey years. For example, if an individual withdraws from the labor force between 1992 and 1994, we use the 1994 interview to infer whether the individual was working in 1993.

The HRS has a comprehensive asset measure. It includes the value of housing, other real estate, autos, liquid assets (which includes money market accounts, savings accounts, T-bills, etc.), IRAs, stocks, business wealth, bonds, and "other" assets, less the value of debts. For non-survey years, we assume that assets take on the value reported in the preceding year. This implies, for example, that we use the 1992 asset level as a proxy for the 1993 asset level. Given that wealth changes rather slowly over time, these imputations should not severely bias our results.

Medical expenses are the sum of insurance premia paid by households, drug costs, and out-of-pocket costs for hospital, nursing home care, doctor visits, dental visits, and outpatient care. As noted in the text, the proper measure of medical expenses for our model includes payments made by Medicaid. Although individuals in the HRS report whether they received Medicaid, they do not report the payments. The *2000 Green Book* (Committee on Ways and Means, 2000, p. 923) reports that in 1998 the average Medicaid payment was $10,242 per beneficiary aged 65 and older, and $9,097 per blind or disabled beneficiary. Starting with this average, we then assume that Medicaid payments have the same volatility as the medical

care payments made by uninsured households. This allows us to generate a distribution of Medicaid payments.

To measure health status we use responses to the question: "would you say that your health is excellent, very good, good, fair, or poor?" We consider the individual in bad health if he responds "fair" or "poor", and consider him in good health otherwise.[30] We treat the health status for non-survey years as missing. Appendix H describes how we construct the health insurance indicator.

We use Social Security Administration earnings histories to construct AIME. Approximately 74% of our sample released their Social Security Number to the HRS, which allowed them to be linked to their Social Security earnings histories. For those who did not release their histories, we use the procedure described below to impute AIME as a function of assets, health status, health insurance type, labor force participation, and pension type.

The HRS collects pension data from both workers and employers. The HRS asks individuals about their earnings, tenure, contributions to defined contribution (DC) plans, and their employers. HRS researchers then ask employers about the pension plans they offer their employees. If the employer offers different plans to different employees, the employee is matched to the plan based on other factors, such as union status. Given tenure, earnings, DC contributions, and pension plan descriptions, it is then possible to calculate pension wealth for each individual who reports the firm he works for. Following Scholz et al. (2006), we use firm reports of defined benefit (DB) pension wealth and individual reports of DC pension wealth if they exist. If not, we use firm-reported DC wealth and impute DB wealth as a function of wages, hours, tenure, health insurance type, whether the respondent also has a DC plan, health status, age, assets, industry and occupation. We discuss the imputation procedure below.

Workers are asked about two different jobs: (1) their current job if working or last job if not working; (2) the job preceding the one listed in part 1, if the individual worked at that job for over 5 years. Pension wealth from both of these jobs are included in our measure of pension

[30] Bound et al. (2003) consider a more detailed measure of health status.

wealth. Below we give descriptives for our estimation sample (born 1931-1935) and validation sample (born 1936-1941). 41% of our estimation sample [and 52% of our validation sample] are currently working and have a pension (of which 56% [57% for the validation sample] have firm-based pension details), 6% [5%] are not working, and had a pension on their last job (of which 62% [62%] have firm-based pension details), and 32% [32%] of all individuals had a pension on another job (of which 35% [29%] have firm-based pension details).

To generate the initial joint distribution of assets, wages, AIME, pensions, participation, health insurance, health status and medical expenses, we draw random vectors (i.e., random draws of individuals) from the empirical joint distribution of these variables for individuals aged 57-61 in 1992, or 1,701 observations. We drop observations with missing data on labor force participation, health status, insurance, assets, and age. We impute values for observations with missing wages, medical expenses, pension wealth, and AIME.

To impute these missing variables, we follow David et al. (1986) and Little (1988) and use the following predictive mean matching regression approach. First, we regress the variable of interest y (e.g., pension wealth) on the vector of observable variables x, yielding $y = x\beta + \epsilon$. Second, for each sample member i we calculate the predicted value $\hat{y}_i = x_i\hat{\beta}$, and for each member with an observed value of y_i we calculate the residual $\hat{\varepsilon}_i = y_i - \hat{y}_i$. Third, we sort the predicted value $\hat{y}_i$ into deciles. Fourth, for missing observations, we impute ε_i by finding a random individual j with a value of $\hat{y}_j$ in the same decile as $\hat{y}_i$, and setting $\varepsilon_i = \hat{\varepsilon}_j$. The imputed value of y_i is $\hat{y}_i + \hat{\varepsilon}_j$.

As David et al. (1986) point out, our imputation approach is equivalent to hot-decking when the "x" variables are discretized and include a full set of interactions. The advantages of our approach over hot-decking are two-fold. First, many of the "x" variables are continuous, and it seems unwise to discretize them. Second, we have very few observations for some variables (such as pension wealth on past jobs), and hot-decking is very data-intensive. A small number of "x" variables generate a large number of hot-decking cells, as hot-decking uses a full set of interactions. We found that the interaction terms are relatively unimportant, but adding extra variables were very important for improving goodness of fit when imputing

pension wealth.

If someone is not working (and thus does not report a wage), we use the wage on their last job as a proxy for their current wage if it exists, and otherwise impute the log wage as a function of assets, health, health insurance type, labor force participation, AIME, and quarters of covered work. We predict medical expenses using assets, health, health insurance type, labor force participation, AIME, and quarters of covered earnings.

Lastly, we must infer the persistent component of the medical expense residual from medical expenses. Given an initial distribution of medical expenses, we construct ζ_t, the persistent medical expense component, by first finding the normalized log deviation ψ_t, as described in equations (7) and (10), and then applying standard projection formulae to impute ζ_t from ψ_t.

Appendix H: Measurement of Health Insurance Type and Labor Force Participation

Much of the identification in this paper comes from differences in medical expenses and job exit rates between those with *tied* health insurance coverage and those with *retiree* coverage. Unfortunately, identifying these health insurance types is not straightforward. The HRS has rather detailed questions about health insurance, but the questions asked vary from wave to wave. Moreover, in no wave are the questions asked consistent with our definitions of *tied* or *retiree* coverage. Fortunately, our estimated health insurance specific job exit rates are not very sensitive to our definition of health insurance, as we show below.

In all of the HRS waves (but not AHEAD waves 1 and 2), the respondent is asked whether he has insurance provided by a current or past employer or union, or a spouse's current or past employer or union. If he responds no to this question, we code his coverage as *none*. We assume that this question is answered accurately, so that there is no measurement error when individual reports that his insurance category is *none*. All of the measurement error problems arise when we allocate individuals with employer-provided coverage between the *retiree* and *tied* categories.

If an individual has employer-provided coverage in waves 1 and 2 he is asked "Is this

health insurance available to people who retire?" In waves 3-8 the analogous question is "If you left your current employer now, could you continue this health insurance coverage up to the age of 65?". For individuals younger than 65, the question asked in waves 3-8 is a more accurate measure of whether the individual has *retiree* coverage. In particular, a "yes" response in waves 1 and 2 might mean only that the individual had *tied* coverage, but could acquire *COBRA* coverage if he left his job. Thus the fraction of individuals younger than 65 who report that they have employer-provided health insurance but who answer "no" to the follow-up question roughly doubles between waves 2 and 3. On the other hand, for those older than 65, the question used in waves 3-8 is meaningless.

Our preferred approach is to use the wave 1 response to determine who has *retiree* coverage. It is possible, however, to estimate the probability of response error to this variable. Consider first the problem of distinguishing the *retiree* and *tied* types for those younger than 65. As a matter of notation, let I denote an individual's actual health insurance coverage, and let I^* denote the measure of coverage generated by the HRS questions. To simplify the notation, assume that the individual is known to have employer-provided coverage—$I = tied$ or $I = retiree$—so that we can drop the conditioning statement in the analysis below. Recall that many individuals who report *retiree* coverage in waves 1 and 2 likely have *tied* coverage. We are therefore interested in the misreporting probability $\Pr(I = tied | I^* = retiree, wv < 3, t < 65)$, where wv denotes HRS wave and t denotes age. To find this quantity, note first that by the law of total probability:

$$\begin{aligned}\Pr(I = tied | wv < 3, t < 65) = \\ \Pr(I = tied | I^* = tied, wv < 3, t < 65) \times \Pr(I^* = tied | wv < 3, t < 65) + \\ \Pr(I = tied | I^* = retiree, wv < 3, t < 65) \times \Pr(I^* = retiree | wv < 3, t < 65). \qquad (39)\end{aligned}$$

Now assume that all reports of *tied* coverage in waves 1 and 2 are true:

$$\Pr(I = tied | I^* = tied, wv < 3, t < 65) = 1.$$

Assume further that for individuals younger than 65 there is no measurement error in waves 3-8, and that the share of younger individuals with *tied* coverage is constant across waves:

$$\begin{aligned}\Pr(I = tied|wv < 3, t < 65) &= \Pr(I = tied|wv \geq 3, t < 65) \\ &= \Pr(I^* = tied|wv \geq 3, t < 65).\end{aligned}$$

Inserting these assumptions into equation (39) and rearranging yields the mismeasurement probability:

$$\begin{aligned}&\Pr(I = tied|I^* = retiree, wv < 3, t < 65) \\ &\quad = \frac{\Pr(I^* = tied|wv \geq 3, t < 65) - \Pr(I^* = tied|wv < 3, t < 65)}{\Pr(I^* = retiree|wv < 3, t < 65)}.\end{aligned} \tag{40}$$

To account for mismeasurement in waves 1 and 2 for those 65 and older, we again assume that all reports of *tied* health insurance are true. We assume further that $\Pr(I = tied|I^* = retiree, wv < 3, t \geq 65) = \Pr(I = tied|I^* = retiree, wv < 3, t < 65)$: the fraction of *retiree* reports in waves 1 and 2 that are inaccurate is the same across all ages. We can then apply the mismeasurement probability for people younger than 65, given by equation (40), to *retiree* reports by people 65 and older.

The second misreporting problem is that the "follow-up" question in waves 3 through 8 is completely uninformative for those older than 65. Our strategy for handling this problem is to treat the first observed health insurance status for these individuals as their health insurance status throughout their lives. Since we assume that reports of *tied* coverage are accurate, older individuals reporting *tied* coverage in waves 1 and 2 are assumed to receive *tied* coverage in waves 3 through 8. (Recall, however, that if an individual with *tied* coverage drops out of the labor market, his health insurance is *none* for the rest of his life.) For older individuals reporting *retiree* coverage in waves 1 and 2, we assume that the misreporting probability—when we choose to account for it—is the same throughout all waves. (Recall that our preferred assumption is to assume that a "yes" response to the follow-up question

in waves 1 and 2 indicates *retiree* coverage.)

A related problem is that individuals' health insurance reports often change across waves, in large part because of the misreporting problems just described. Our preferred approach for handling this problem is classify individuals on the basis of their first observed health insurance report. We also consider the approach of classifying individuals on the basis of their report from the previous wave.

Figure 8 shows how our treatment of these measurement problems affects measured job exit rates. The top two graphs in Figure 8 do not adjust for measurement error. The bottom two graphs account for the measurement error problems, using the approached described by equation (40). The two graphs in the left column use the first observed health insurance report whereas the graphs in the right column use the previous period's health insurance report. Figure 8 shows that the profiles are not very sensitive to these changes. Those with *retiree* coverage tend to exit the labor market at age 62, whereas those with *tied* and *no* coverage tend to exit the labor market at age 65.

Another, more conceptual, problem is that the HRS has information on health insurance outcomes, not choices. This is an important problem for individuals out of the labor force with no health insurance; it is unclear whether these individuals could have purchased $COBRA$ coverage but elected not to do so.[31] To circumvent this problem we use health insurance in the previous wave and the transitions implied by equation (10) to predict health insurance options. For example, if in the previous wave an individual reports working and having health insurance that is tied to his job, that individual's choice set is *tied* health insurance and working or $COBRA$ insurance and not working.[32]

[31] For example, the model predicts that all HRS respondents younger than 65 who report having tied health insurance two years before the survey date, work one year before the survey date, and are not currently working should report having COBRA coverage on the survey date. However, 19% of them report having no health insurance.

[32] We are assuming that everyone eligible for COBRA takes up coverage. In practice, only about $\frac{2}{3}$ of those eligible take up coverage (Gruber and Madrian, 1996). In order to determine whether our failure to model the COBRA decision is important, we shut down the COBRA option (imposed a 0% take-up rate) and re-ran the model. Eliminating COBRA had virtually no effect on labor supply.

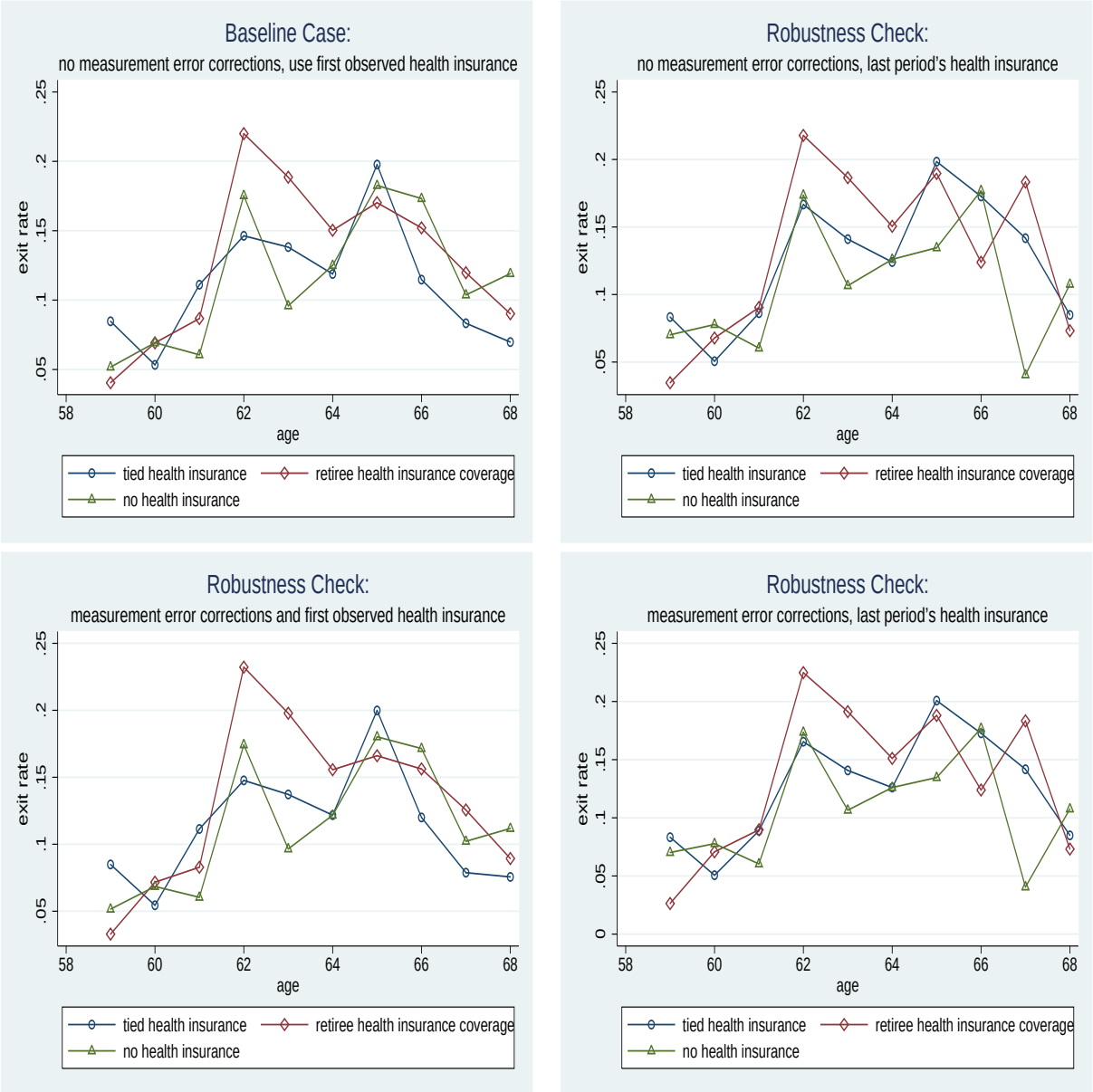

Figure 8: Job Exit Rates Using Different Measures of Health Insurance Type

Our preferred specification, which we use in the analysis, is to use the first observed health insurance report, and to not use the measurement error corrections.

Because agents in our model are forward-looking, we need to know the health-insurance-conditional process for medical expenses facing the very old. The data we use to estimate medical expenses for those over age 70 comes from the Assets and Health Dynamics of the Oldest Old survey. French and Jones (2004a) discuss some of the details of the survey, as well as some of our coding decisions. The main problem with the AHEAD is that there is no question asked of respondents about whether they would lose their health insurance if they left their job, so it is not straightforward to distinguish those who have *retiree* coverage from those with *tied* coverage. In order to distinguish these two groups, we do the following. If the individual exits the labor market during our sample, and has employer-provided health insurance at least one full year after exiting the labor market, we assume that individual has *retiree* coverage. All individuals who have employer-provided coverage when first observed, but do not meet this criterion for having *retiree* coverage, are assumed to have *tied* coverage.

Our measure of labor force participation is based on the values reported at the time of the interview. We also use the age at the time of the interview. For this reason, some of our "65-year-olds" are 65 years and 0 days old, whereas others are 65 years and 364 days old. Blau (1994) shows that most age-65 job exits occur within a few months of the 65th birthday. Thus, we may be understating the decline in labor supply at age 65, because our participation measure combines individuals who are exactly 65, who may not have yet left the labor force, with those who are almost 66, who may have left the labor force market months before.

To investigate how this timing issue affects our estimated job exit rates, we use HRS labor force histories, which provide the dates at which individuals leave the labor force, to construct three different measures of participation by age. Figure 9 presents job exit rates derived with the different measures.

The top left panel of Figure 9 shows job exit rates derived with the measure of participation that we use in the paper (participation at the time of the interview). In the top right panel, participation is measured at the time of the respondent's birthday, so that the job exit rate

at age 65 measures the probability that an individual was working on his 64th birthday but not on his 65th birthday. Relative to the baseline case, the peaks in exit rates at ages 62 and 65 are now less pronounced. The reason for this is that people who report leaving in the months after a 65th birthday are coded as having left at age 66. For example, an individual leaving the labor market at age 65 and 1 day would be classified as exiting the labor market at age 66. As a result, measuring labor force participation at birthdays leads to a higher estimated job exit rate at 66 and a lower rate at 65 than our baseline approach.

In the bottom left panel of Figure 9, participation is measured at the midpoint between the respondents' birthdays. For example, participation at age 65 is measured at age 65 $\frac{1}{2}$, so that the job exit rate at age 65 measures the probability that an individual was working at age 64 $\frac{1}{2}$ but was not at age 65 $\frac{1}{2}$. This panel looks very similar to the baseline case. In both cases job exit rates are near 20 percent at ages 62 and 65, and are lower at other ages. Furthermore, in both cases job exit rates for those with *retiree* coverage are highest at age 62, whereas job exit rates for those with *tied* coverage are highest at age 65.

Because it seems extreme to treat an individual who leaves the labor force at age 65 and 1 day as exiting at age 66, we think measuring participation 6 months after a birthday yields more plausible results. Because measuring participation on survey dates gives similar results and drops fewer observations than measuring participation 6 months after a birthday, we use participation on survey dates as our measure of participation throughout.

Another measurement issue is the treatment of the self-employed. Our preferred approach is to include the self-employed in our analysis, and treat them as working with no health insurance. The lower lower right panel of Figure 9 shows job exit rates when we drop the self employed, but measure health insurance as in the baseline case. The main difference caused by dropping the self-employed is that those with *no* health insurance have much higher job exit rates, especially at age 65. Nevertheless, those with *retiree* coverage are still most likely to exit at age 62 and those with *tied* and *no* health insurance are most likely to exit at age 65.

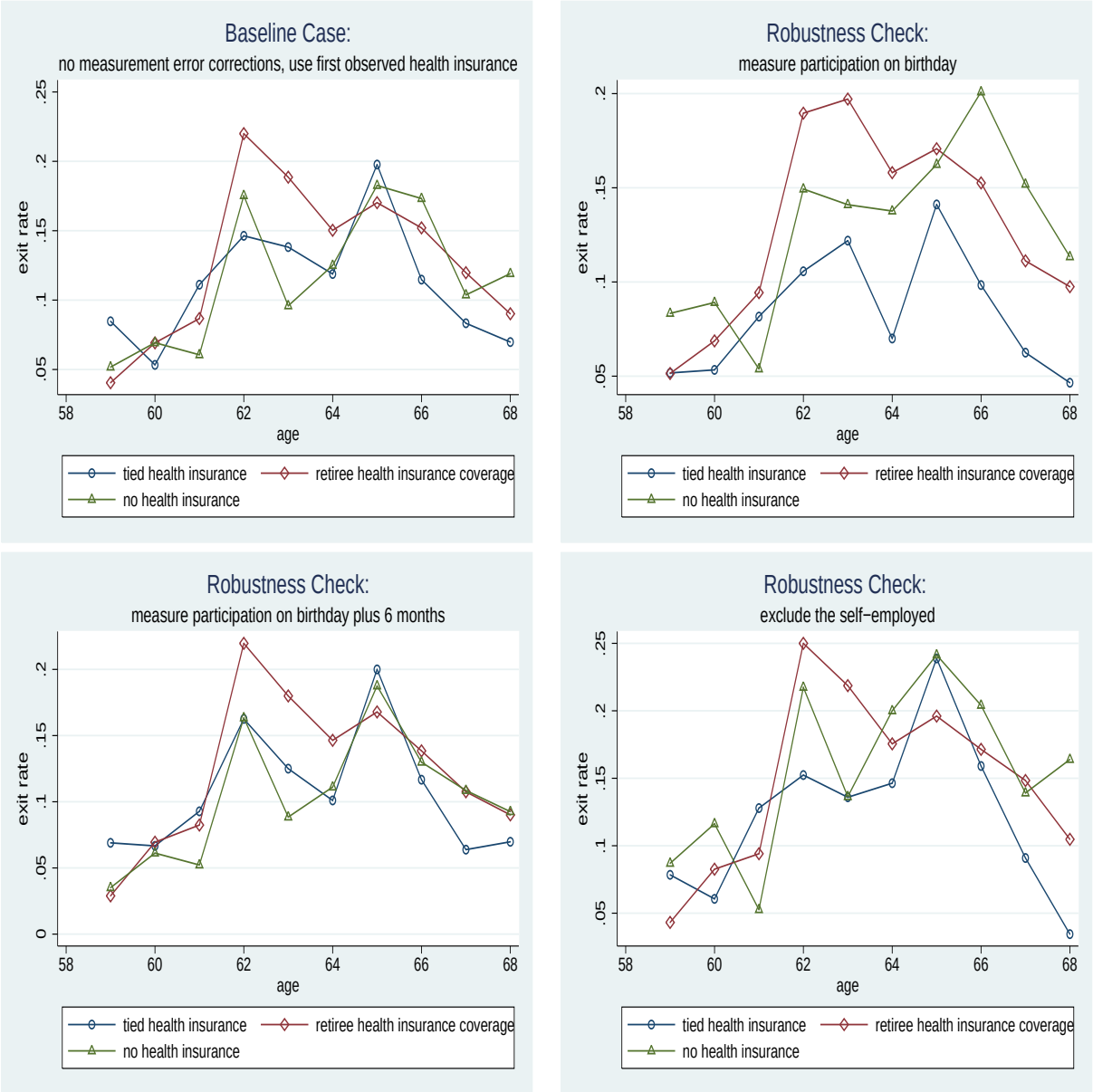

Figure 9: Job Exit Rates Using Different Measures of Labor Force Participation

Appendix I: The Medical Expense Model

Recall from equation (7) that health status, health insurance type, labor force participation and age affect medical expenses through the mean shifter $m(.)$ and the variance shifter $\sigma(.)$. Health status enters $m(.)$ and $\sigma(.)$ through 0-1 indicators for bad health, and age enters through linear trends. On the other hand, the effects of Medicare eligibility, health insurance and labor force participation are almost completely unrestricted, in that we allow for an almost complete set of interactions between these variables. This implies that mean medical expenses are given by

$$m(H_t, I_t, t, P_t) = \gamma_0 H_t + \gamma_1 t + \sum_{h \in I} \sum_{P \in \{0,1\}} \sum_{a \in \{t<65, t \geq 65\}} \gamma_{h,P,a}.$$

The one restriction we impose is that $\gamma_{none,0,a} = \gamma_{none,1,a}$ for both values of a, i.e., participation does not affect health care costs if the individual does not have insurance. This implies that there are 10 $\gamma_{h,P,a}$ parameters, for a total of 12 parameters apiece in the $m(.)$ and the $\sigma(.)$ functions.

To estimate this model, we group the data into 10-year-age (55-64, 65-74, 75-84) $\times$ health status $\times$ health insurance $\times$ participation cells. For each of these 60 cells, we calculate both the mean and the 95th percentile of medical expenses. We estimate the model by finding the parameter values that best fit this 120-moment collection. One complication is that the medical expense model we estimate is an annual model, whereas our data are for medical expenses over two-year intervals. To overcome this problem, we first simulate a panel of medical expense data at the one-year frequency, using the dynamic parameters from French and Jones (2004a) shown in Table 2 of this paper and the empirical age distribution. We then aggregate the simulated data to the two-year frequency; the means and 95th percentiles of this aggregated data are comparable to the means and 95th percentiles in the HRS. Our approach is similar to the one used by French and Jones (2004a), who provide a detailed description.

Table 1 shows some of the summary statistics implied by our medical expense model.

Relative to other research on the cross sectional distribution of medical expenses, we find higher medical expenses at the far right tail of the distribution. For example, Blau and Gilleskie (2006) use different data and methods to find average medical expenses that are comparable to our estimates. However, they find that medical expenses are less volatile than our estimates suggest. For example, they find that for households in good health and younger than 65, the maximum expense levels (which seem to be slightly less likely than 0.5% probability events) were $69,260 for those without coverage, $6,400 for those with retiree coverage, and $6,400 for those with tied coverage. Table 1 shows that our estimates of the 99.5th percentile (i.e., the top 0.5 percentile of the distribution) of the distributions for healthy workers are $86,900 for those with no coverage, $32,700 for those with *retiree* coverage, and $30,600 for those with *tied* coverage.

Berk and Monheit (2001) use data from the MEPS, which arguably has the highest quality medical expense data of all the surveys. Analyzing total billable expenses, which should be comparable to our data for the uninsured, Berk and Monheit find that those in the top 1% of the medical expense distribution have average medical expenses of $57,900 (in 1998 dollars). Again, this is below our estimate of $86,900 for the uninsured. This discrepancy is not surprising. Berk and Monheit's estimates are for all individuals in the population, whereas our estimates are for older households (many of which include two individuals). Furthermore, Berk and Monheit's estimates exclude all nursing home expenses, while the HRS, although initially consisting only of non-institutionalized households, captures the nursing home expenses these households incur in later waves.

Appendix J: The Preference Index

We construct the preference index for each member of the sample using the wave 1 variables V3319, V5009, V9063. All three variables are self-reported responses to questions about preferences for leisure and work. In V3319 respondents were asked if they agreed with the statement (if they were working): "Even if I didn't need the money, I would probably keep on working." In V5009 they were asked: "When you think about the time when you [and your (husband/wife/partner)] will (completely) retire, are you looking forward to it, are you

uneasy about it, or what?" In V9063 they were asked (if they were working): "On a scale where 0 equals dislike a great deal, 10 equals enjoy a great deal, and 5 equals neither like nor dislike, how much do you enjoy your job?"

Because it is computationally intensive to estimate the parameters of the type probability equations in our method of simulated moments approach, we combine these three variables into a single index that is simpler to use. To construct this index, we regress labor force participation on current state variables (age, wages, assets, health, etc.), squares and interactions of these terms, the wave 1 variables V3319, V5009, V9063, and indicators for whether these variables are missing. We then partition the $x\hat{\beta}$ matrix from this regression into: $x_1\hat{\beta}_1$, where the x_1 matrix consists V3319, V5009, V9063, and indicators for these variables being missing; and $x_2\hat{\beta}_2$, where the x_2 matrix contains all the other variables. Our preference index is $x_1\hat{\beta}_1$.

Individuals who were not working in 1992 were not asked any of the preference questions, and are not included in the construction of our index. Because everyone who answered the preference questions worked in 1992, we estimate the regression models with participation data from 1998-2006.

Finally, we discretize the index into three values: *out*, for those not employed in 1992; *low*, for workers with an index in the bottom half of the distribution; and *high* for the remainder.

Appendix K: Additional Parameter Estimates

We assume that the probability of belonging to a particular type follows a multinomial logit function. Table 15 shows the coefficients of the preference type prediction equation. One interesting feature of this equation is that wealthy individuals who have no health insurance coverage have a high probability of being Type-2 agents. Given that many of these individuals are entrepreneurs, it is not surprising that they are often placed in the "motivated" group.

Table 16 shows the parameter estimates for the robustness checks. In the no-saving case, shown in column (2), β and θ_B are both very weakly identified. We therefore follow Rust and Phelan (1997) and Blau and Gilleskie (2006, 2008) by fixing β, in this case to its baseline values of 0.95, 0.86, and 1.12 (for types 0, 1 and 2, respectively). Similarly, we fix θ_B to zero.

	Preference Type 1		Preference Type 2	
	Parameters	Std. Errors	Parameters	Std. Errors
	(1)	(2)	(3)	(4)
Preference Index = *out*	-5.33	1.00	-7.33	7.77
Preference Index = *low*	4.79	2.03	0.18	1.32
Preference Index = *high*	2.35	0.80	4.09	0.73
No Insurance Coverage	3.35	1.46	-2.45	1.76
Retiree Coverage	-0.98	0.87	-0.32	0.47
Initial Health†	-1.04	0.35	-0.37	0.27
Initial Wages†	2.74	0.75	-1.01	0.44
Assets/Wages†	-0.48	0.83	0.97	0.61
AIME/Wages†	-0.25	0.70	-0.21	0.44
Health Cost Shock (ψ)	-1.16	0.65	0.22	0.47
Age – 60	-0.56	0.96	1.72	1.67
Assets$^{\dagger}\times$(No Ins. Coverage)	-0.53	0.37	1.41	0.54
†Variables expressed as fraction of average				

Table 15: Preference type prediction coefficients

Since the asset distribution is degenerate in this no-saving case, we no longer match asset quantiles or quantile-conditional participation rates, matching instead participation rates for each health insurance category.

Column (4) shows the parameter estimates that result when we remove the preference index described in Appendix J from our the type prediction equations; we also remove the preference index-conditional moment conditions from the GMM criterion function. Although the coefficients of the type prediction equations change dramatically, the estimated preference parameters change very little.

Parameter and Definition	Baseline (1)	No Saving (2)	Homogeneous Preferences (3)	No Preference Index (4)
γ: consumption weight				
Type 0	0.412	0.302	NA	0.405
	(0.045)	(0.026)		(0.041)
Type 1	0.649	0.583	0.550	0.647
	(0.007)	(0.007)	(0.008)	(0.008)
Type 2	0.967	0.9999	NA	0.986
	(0.203)	NA		(0.393)
β: time discount factor				
Type 0	0.945	0.945	NA	0.962
	(0.074)	(NA)		(0.064)
Type 1	0.859	0.859	0.970	0.858
	(0.013)	(NA)	(0.009)	(0.014)
Type 2	1.124	1.124	NA	1.143
	(0.328)	(NA)		(0.580)
ν: coefficient of relative risk aversion, utility	7.49	6.35	5.78	7.61
	(0.311)	(0.174)	(0.448)	(0.289)
θ_B: bequest weight†	0.0223	0.00	0.0132	0.0221
	(0.0012)	(NA)	(0.0007)	(0.0009)
κ: bequest shifter, in thousands	444	0.00	786	443
	(31.7)	(NA)	(27.9)	(21.4)
C_{min}: consumption floor	4,380	4,440	5,000	4,430
	(167)	(154)	(169)	(214)
L: leisure endowment, in hours	4,060	4,130	4,700	4,090
	(44)	(67)	(63)	(434)
ϕ_H: hours of leisure lost, bad health	506	939	303	509
	(20.9)	(42.2)	(25.9)	(29.9)
ϕ_{P0}: fixed cost of work: intercept, in hours	826	880	1,146	827
	(20.0)	(24.5)	(36.3)	(29.2)
ϕ_{P1}: fixed cost of work: (age – 60), in hours	54.7	36.5	16.9	52.7
	(2.58)	(3.10)	(1.27)	(2.49)
ϕ_{RE}: hours of leisure lost, re-entering market	94.0	77.0	155.9	95.6
	(8.63)	(12.1)	(13.6)	(24.2)
χ^2 statistic	775	398	607	904
Degrees of freedom	171	86	169	145

Diagonal weighting matrix used in calculations. See Appendix F for details.
Standard errors in parentheses.
†Parameter expressed as marginal propensity to consume out of final-period wealth.

Table 16: Robustness Checks