Are Americans Saving "Optimally" for Retirement?

John Karl Scholz
Department of Economics, the Institute for Research on Poverty, and NBER
University of Wisconsin – Madison
1180 Observatory Drive
Madison, Wisconsin 53706-1393
jkscholz@facstaff.wisc.edu

Ananth Seshadri
Department of Economics
University of Wisconsin – Madison
1180 Observatory Drive
Madison, Wisconsin 53706-1393
aseshadr@ssc.wisc.edu

Surachai Khitatrakun
ERS Group
4901 Tower Court
Tallahassee, Florida 32303
skhitatrakun@ersgroup.com

January 8, 2004

We thank colleagues at the University of Michigan for developing the Health and Retirement Survey. We also are grateful to the National Institute of Aging, the Center for the Demography of Health and Aging at UW-Madison, and the Russell Sage Foundation for financial support, Sheng-Kai Chang for sharing code that assisted in the earnings estimation, and to numerous colleagues at Wisconsin and elsewhere and to seminar participants for helpful comments.

Abstract:

This paper examines the degree to which Americans are saving optimally for retirement. Our standard for assessing optimality comes from a life-cycle model that incorporates uncertain lifetimes, uninsurable earnings and medical expenses, progressive taxation, government transfers, and pension and social security benefit functions derived from rich household data. We solve every household's decision problem from death to starting age and then use the decision rules in conjunction with earnings histories to make predictions about wealth in 1992. Ours is the first study to compare, household by household, wealth predictions that arise from a life-cycle model that incorporates earnings histories for a nationally representative sample. The results, based on data from the Health and Retirement Study, are striking – we find that the model is capable of accounting for more than 80 percent of the 1992 cross-sectional variation in wealth. Fewer than 20 percent of households have less wealth than their optimal targets, and the wealth deficit of those who are undersaving is generally small.

There is considerable skepticism in public policy discussions and in the financial press that Americans are preparing adequately for retirement. A quotation from the *Wall Street Journal* captures a popular view:

> A long time ago, New England was known for its thrifty Yankees. But that was before the baby boomers came along. These days, many New Englanders in their 30s and 40s, and indeed their counterparts all over America, have a different style: they are spending heavily and have sunk knee-deep in debt. ... A recent study sponsored by Merrill Lynch & Co. showed that the average middle-aged American had about $2,600 in net financial assets. Another survey by the financial-services giant showed that boomers earning $100,000 will need $653,000 in today's dollars by age 65 to retire in comfort – but were saving only 31 percent of the amount needed. In other words, the saving rate will have to triple. Experts say the failure to build a nest egg will come to haunt the baby boomers, forcing them to drastically lower standards of living in their later years or to work for longer, perhaps into their 70s.[1]

Assessing the adequacy or optimality of wealth accumulation is difficult, since it requires some standard against which to measure observed behavior. Several authors use augmented life-cycle models for this standard, simulating the expected distribution of wealth for representative household types (see, for example, Hubbard, Skinner, and Zeldes, 1995; and Engen, Gale, and Uccello, 1999).[2] While augmented life-cycle models provide natural benchmarks, these researchers do not fully assess the adequacy (let alone the optimality) of wealth accumulation. They derive optimal *distributions* of wealth (or wealth-income ratios). But given underlying

[1]"Binge Buyers: Many Baby Boomer Save Little, May Run Into Trouble Later On: They Don't Build Nest Eggs Nearly Rapidly Enough for an Easy Retirement," Bernard Wysocki Jr., 6/5/95, A1 Wall Street Journal. Wolff (2002) concludes "retirement wealth accumulation needs to be improved for the vast majority of households."

[2] Kotlikoff, Spivak, and Summers (1982); Moore and Mitchell (1998); and Gustman and Steinmeier (1999) examine saving adequacy by comparing data to financial planning rules of thumb. But a rule of thumb cannot describe optimal behavior for households with widely different patterns of earnings realizations, even if preferences are homogeneous. Hamermesh (1984); Banks, Blundell, and Tanner (1998); and Bernheim, Skinner, and Weinberg (2001) make inferences about adequacy from consumption changes around retirement. But, for the reasons given in Aguiar and Hurst (2003) and Hurd and Rohwedder (2003), it is difficult to make inferences about adequacy or optimality from patterns of consumption changes around retirement.

model parameters, expectations, and earnings realizations, the models have household-specific implications for optimal wealth accumulation. These household-specific implications have not been studied.

We examine the degree to which households are optimally preparing for retirement by constructing a stochastic life-cycle model that captures the key features of a household's consumption decisions. Our model incorporates many behavioral features shown by prior work to affect consumption, including precautionary savings and buffer stock behavior (Deaton, 1991; Aiyagari, 1994; Carroll, 1997). It incorporates asset-tested public transfers (Hubbard, Skinner, and Zeldes, 1995), where benefits vary over time and by household size. Our model incorporates end-of-life uncertainty and medical shocks (Palumbo, 1999). We also incorporate a stylized, time-varying progressive income tax that reflects the evolution of average effective federal income tax rates over the period spanned by our data. Households in the model form realistic expectations about social security, which depends on lifetime income; pension benefits, which depend on income in the final year of work; and earnings. We incorporate detailed data from the Health and Retirement Study (HRS) on family structure and age of retirement (treating both as exogenous and known from the beginning of working life) in calculating optimal life-cycle consumption profiles.

Our approach has other distinctive features. Most important, we calculate household-specific optimal wealth targets, using data from the HRS. A crucial input to our behavioral model is 41 years of information on earnings realizations drawn from restricted-access social security earnings records. The timing of earnings shocks can cause optimal wealth to vary substantially, even for households with identical preferences, demographic characteristics, and lifetime income. Hence, it is essential for life-cycle models of wealth accumulation to incorporate earnings realizations, at least to the extent model implications are compared to actual behavior.

We find that over 80 percent of HRS households have accumulated more wealth than their optimal targets. These targets indicate the amounts of private saving households need to solve their life-cycle planning problem, given social security and defined benefit pension expectations and realizations. For those not meeting their targets, the magnitudes of the deficits are typically small. In addition, the cross-sectional distribution of wealth in 1992 closely matches the predictions of our life-cycle model.

I. The Health and Retirement Study

The HRS is a national panel study with an initial sample (in 1992) of 12,652 persons and 7,702 households.[3] It oversamples blacks, Hispanics, and residents of Florida. The baseline 1992 study consisted of in-home, face-to-face interviews of the 1931-1941 birth cohort and their spouses, if they are married. Follow-up interviews were given by telephone in 1994, 1996, 1998, 2000, and 2002. For the analyses in this paper we exclude 379 married households where one spouse did not participate in the 1992 HRS, 93 households that failed to have at least one year of full-time work, and 908 households where the highest earner began working full time prior to 1951.[4] Our resulting sample has 6,322 households.

---

[3] An overview of the HRS is given in a Supplementary issue of the Journal of Human Resources, 1995 (volume 30). There, 22 authors discuss and assess the data quality of many dimensions of the initial wave of the HRS.

[4] We drop the first group because we do not have information on spousal, and hence household, income. We drop the second group because we do not have information on transfer payments in years prior to the HRS survey and therefore we cannot model the lifetime budget constraint. We drop households where the highest earner started working before 1951 for computational reasons. Our procedures to impute missing and top-coded data are considerably more complicated when initial values of the earnings process are missing. Details for the earnings imputations are given in Appendix I.

The survey covers a wide range of topics, including batteries of questions on health and cognitive conditions; retirement plans; subjective assessments of mortality probabilities and the quality of retirement preparation; family structure; employment status and job history; demographic characteristics; housing; income and net worth; and pension details.

Wealth Measures in the HRS

Households typically maintain living standards in retirement by drawing on their own (private) saving, employer-provided pensions, and social security wealth. To study the degree to which households optimally accumulate wealth, therefore, we need accurate measures of these wealth components.

Net worth (private saving) is a comprehensive measure that includes housing assets less liabilities, business assets less liabilities, checking and saving accounts, stocks, bonds, mutual funds, retirement accounts including defined contribution pensions, certificates of deposit, the cash value of whole life insurance, and other assets, less credit card debt and other liabilities. It excludes defined benefit pension wealth, social security wealth, and future earnings.[5] The concept of wealth is similar (and in many cases identical) to those used in other studies of wealth and saving adequacy.

We use the "Pension Present Value Database" that Bob Peticolas and Tom Steinmeier kindly made available on the HRS World Wide Web Site to calculate the value of defined benefit pensions and, as described below, estimate the household's expectations of future pension

---

[5] We account rigorously for defined benefit pensions, social security, and future earnings in the household decision problem. We do not model the purchase and service flows that result from consumer durables since we do not have data on durables. We do not include term life policies in net worth because we do not have information on premium payments and, more importantly, a comparison of data from the HRS and Assets and Health Dynamics of the Oldest Old suggests that a substantial fraction of term life policies are dropped in retirement. Those that are not dropped have a median value of less than $7,000.

benefits.[6] The program makes present value calculations of HRS pensions for wave 1 (1992) respondents for nine different scenarios, corresponding to the Social Security Administration's low, intermediate and high long-term projections for interest rates, wage growth rates, and inflation rates. We use the intermediate values when calculating DB pension wealth.[7]

HRS Earnings Data

Restricted access social security earnings data provide a direct measure of earnings realizations and lifetime income, and, as described below, they are used to estimate household's expectations of future earnings. They also allow us to simulate accurately social security benefits for the respondent and spouse or for the couple, if the benefit would be higher.[8]

Two issues arise in using earnings information. First, social security earnings records are not available for 22.8 percent of the respondents included in the analysis (we have 10,523 respondents in 6,322 households). Second, the social security earnings records are top-coded (households earn more than the social security taxable wage caps) for 16 percent of earnings observations between 1951 and 1979. Beginning in 1980, censoring is not an issue, because we have access to uncensored W-2 earnings records from 1980-1991.

We impute earnings histories for those individuals with missing or top-coded earnings records assuming the individual log-earnings process

$$y_{i,0}^{*} = x_{i,0}'\beta_0 + \varepsilon_{i,0}$$

[6] See http://www.umich.edu/~hrswww/center/rescont2.html. The programs use detailed plan descriptions along with information on employee earnings. We use self-reported defined-benefit pension information for households not included in the Peticolas and Steinmeier file. The assumptions used in the program to calculate the value of defined contribution (DC) pensions – particularly the assumption that contributions were a constant fraction of income during years worked with a given employer – are likely inappropriate. Consequently, we follow others in the literature (for example, Engen et al., 1999, p. 159) and use self-reported information to calculate DC pension wealth.

[7] The intermediate Social Security Administration assumptions are 6.3 percent for interest rates, 5 percent for wage growth, and 4 percent for inflation.

[8] Appendix II provides details of the social security calculations.

$$y^*_{i,t} = \rho y^*_{i,t-1} + x'_{i,t}\beta + \varepsilon_{i,t},\ t \in \{1,2,...,T\} \qquad (1)$$

$$\varepsilon_{i,t} = \alpha_i + u_{i,t}$$

where $y^*_{i,t}$ is the log of observed earnings of the individual *i* at time *t* in 1992 dollars, $x_{i,t}$ is the vector of *i*'s characteristics at time *t*, and the error term $\varepsilon_{i,t}$ includes an individual-specific component α_i, which is constant over time, and an unanticipated white noise component, $u_{i,t}$. We employ random-effect assumptions with homoskedastic errors to estimate equation (1).

We estimate the model separately for four groups: men without a college education, men with some college, women without a college education, and women with some college. Details of the empirical earnings model and coefficient estimates are given in Appendix I. In Appendix I we also describe how we use the model of individual earnings and a Gibbs sampling procedure to impute earnings for individuals who refuse to release their social security earnings histories to the HRS,[9] and for earnings in years when social security earnings records are top-coded. The Gibbs procedure is conceptually appealing as it allows us to use information from the entire sequence of individual earnings to impute missing and top-coded earnings.

Table 1 provides descriptive statistics for the HRS sample. Mean (median) earnings in 1991 of HRS households are \$35,263 (\$28,298), though note that 13 percent are not in the paid labor force when interviewed in 1992. The mean (median) present discounted value of lifetime household earnings is \$1,691,104 (\$1,516,931).[10] Retirement consumption will be financed out

[9] We repeated our central empirical analyses dropping individuals who refused to release their social security records and generated nearly identical results to those reported in the paper. Brief details are given in the sensitivity analysis, section 4.6.

[10] When calculating present discounted values of earnings, social security or defined benefit pension wealth, we discount the constant-dollar sum of earnings (social security, or pensions) by a real interest rate measure (prior to 1992, we use the difference between the CPI-W and the 3-month Treasury bill rate; for 1992 and after we use 4 percent).

of pension wealth (mean is $105,919, median is $17,371); social security wealth (mean is $106,714, median is $97,150); and nonpension net worth (mean is $250,513, median is $107,000).

Our empirical procedures result in social security replacement rates that are somewhat lower than those discussed in Engen *et al.* (1999). The replacement rate is defined as equaling annual social security benefits divided by the average of the final five years of earned income (prior to retirement), multiplied by 100. The median for our sample of married couples is 38.4 percent. Those with less than a high school diploma have a median of 43.7 percent. Those with a high school diploma or some college have a median rate of 39.3 percent. College graduates have a median rate of 31.6 percent, while those with more than a college degree have a median rate of 29.8 percent. Engen *et al.* cite figures from Grad (1990) who reports that an average replacement rate for couples was roughly 55 percent in 1982, but social security has become less generous since then. Because we use social security earnings records and a close approximation to the social security benefit rules, our measure compared to those in Grad (1990) shows how replacement rates have changed over time.

Figure 1, which shows the median levels of defined benefit pension wealth, social security wealth, and net worth (excluding DB pensions) in each lifetime income decile, highlights the reason we account rigorously for social security in our model. Social security exceeds the combined value of pension and nonpension net worth in the bottom three deciles of the lifetime income distribution. Private net worth significantly exceeds the value of social security only in the top two deciles of the lifetime income distribution. The metaphor of the "three-legged stool," in which retirement income security is supported by the three legs of social security, employer-provided pensions, and private wealth accumulation, appears to apply only to households in the top 70 percent of the lifetime income distribution because low-income workers lack employer-provided pension coverage.

II. A Model of Optimal Wealth Accumulation

We solve a simple life-cycle model, augmented to incorporate uncertain lifetimes, uninsurable earnings, uninsurable medical expenses, and borrowing constraints. The unit of analysis is a household, which can be married or single.[11] Individuals within a household live to a maximum age D. Between ages 0 and $S-1$ individuals are children and make no consumption decisions. Adults start working at age S, have exogenous labor supply, and give birth to as many as n children at ages B_1, B_2, ..., B_n. Earnings depend on age (which affects work experience) and a random shock that may be correlated across time. Each period, adults decide how much to consume and how much to save for the future.

Households retire exogenously at the end of age R and face a probability of death in each remaining year of life. In retirement, they start receiving health shocks that are allowed to be correlated across ages. They receive income from social security, defined benefit plans (if covered) and assets. Social security receipts depend on total earnings during the preretirement period. Defined benefit pension receipts are a function of the household's last earnings receipt before retirement.

II.1. A Household's Maximization Problem

A household derives utility $U(c)$ from period-by-period consumption in equivalent units, where $g(A_j, K_j)$ is a function that adjusts consumption for the number of adults A_j and children K_j in a household at age j. Let c_j and a_j represent consumption and assets at age j. With probability p_j the individual survives into the next period, so an individual survives until age j with probability $\prod_{k=S}^{j-1} p_k$, where $\prod_{k=S}^{j-1} p_k = 1$ if $j-1 < R$. At age D, $p_D = 0$. The discount factor on future utilities is β. Expected lifetime utility is then

$$E\left[\sum_{j=S}^{D} \beta^{j-S} U\left(c_j / g(A_j, K_j)\right)\right].$$

The expectation operator E denotes the expectation over future earnings uncertainty, uncertainty in health expenditures, and length of life.

Consumption and assets are chosen to maximize expected utility subject to the constraints,[12]

$$y_j = e_j + ra_j + T(e_j, a_j, j, n),\ \ j \in \{S, ..., R\},$$

$$y_j = SS\left(\sum_{j=S}^{R} e_j\right) + DB\left(e_R\right) + ra_j + T(e_R, \sum_{j=S}^{R} e_j, a_j, j, n),\ \ j \in \{R+1, ..., D\},$$

$$c_j + a_{j+1} = y_j + a_j - \tau\left(y_j\right), j \in \{S, ..., R\},$$

$$c_j + a_{j+1} + m_j = y_j + a_j - \tau\left(y_j\right), j \in \{R+1, ..., D\}.$$

The first two equations define taxable income for working and for retired households. The last two equations show the evolution of resources available for consumption. In these constraints e_j denotes labor earnings at age j, and $\tau(\cdot)$ is a tax function that depicts total tax payments as a

[11] We do not model marriage or divorce. Married households become single only if a spouse dies. Single households remain single forever.

[12] The model implies a borrowing constraint in the sense that asset balances are positive in every period. The intuition is the following: for the problem to be well-specified, the household should not be allowed to die with debt, regardless of the stochastic sequence of earnings (and medical) shocks. Since earnings shocks in every period can get arbitrarily close to zero, the household should be in a position to repay debt even if they get a long sequence of near-zero earnings draws – failing this, consumption goes to zero and marginal utility of consumption goes to infinity, which is clearly not optimal (since the utility function satisfies the Inada condition). Consequently, the household will maintain a non-negative asset position in every age. The same logic applies in retirement, with the exception that rather than earnings uncertainty, the individual now faces uncertainty in medical expenses and lifespan.

function of taxable income, y_j. $SS(\cdot)$ are social security receipts, which are a function of aggregate lifetime income, and $DB(\cdot)$ are defined benefit receipts, which are a function of earnings received at the last working age. $T(\cdot)$ denotes means-tested transfers, which depend on earnings, social security benefits and defined benefit pensions, assets, the year, and the number of people in the household. Medical expenditures are denoted by m_j and the interest rate is denoted by r.

II.2. Recursive Formulation of the Life-Cycle Problem

The life-cycle problem may be solved backwards from age D, given the terminal condition at that age. There are two sources of uncertainty in retirement – lifespan and medical expenses. We start by describing the problem for retired married households. The problem for single households is dealt with in a similar fashion.

II.2.1. The Retired Household's Problem[13]

A retired household between the ages R and D obtains income from social security, defined benefits, and preretirement assets.[14] The dynamic programming problem at age j for a retired, married household with both members alive at the beginning of age j is given by

$$V(e_R, E_R, a, j, m, 3) = \max_{c,a'} \left\{ \begin{array}{l} U(c/g(2,0)) + \\ \beta p_{hj} p_{wj} \int V(e_R, E_R, a', j+1, m', 3) d\Omega_{jm}(m' \mid m) + \\ \beta p_{hj}(1-p_{wj}) \int V(e_R, E_R, a', j+1, m', 1) d\Omega_{js}(m' \mid \frac{m}{2}) + \\ \beta p_{wj}(1-p_{hj}) \int V(e_R, E_R, a', j+1, m', 2) d\Omega_{js}(m' \mid \frac{m}{2}) \end{array} \right\}, \tag{2}$$

[13] To define a household's retirement date for those already retired, we use the actual retirement date for the person in the household who contributed the largest share of lifetime household earnings. For those not retired, we use the expected retirement date of the person who is expected to contribute the largest share of lifetime household earnings.
[14] To simplify notation, age j variables will be expressed without any subscripts or superscripts, and age $j-1$ variables and age $j+1$ variables will be represented with subscript "-1" and superscript "$'$", respectively.

subject to

$$y = SS(E_R) + DB(e_R) + ra + T(e_R, E_R, a, j, n),$$

$$c + a' + m = y + a - \tau\left(y\right). \tag{3}$$

In equation (2), $V(e_R, E_R, a, j, m, 3)$ denotes the present discounted value of utility from age j until the date of death, $V(e_R, E_R, a', j+1, m', 3)$ denotes the corresponding value in the following year; β is the discount factor on future utilities; and, as noted before, p_{hj} and p_{wj} denote the probability of survival between ages j and $j+1$ for the husband and the wife respectively. Medical expenses are drawn from the Markov processes $\Omega_{jm}(m' \mid m)$ and $\Omega_{js}(m' \mid m)$. These distributions are allowed to be different depending upon whether the household is married (*m*) or single (*s*). Total earnings up to the current period are denoted by E_R, while the last earnings draw at the age of retirement is e_R. Note that E_R and e_R do not change once the individual is retired. The integers in the last argument of the value function signify that only the husband survives (1), only the wife survives (2), or both the husband and wife are alive (3) at the beginning of the period.

II.2.2. The Problem at the Age of Retirement

Age R represents the last working age for the individual. At this age, the individual knows that in the next period he or she will cease working and begin receiving income from social security and defined benefit pensions. The corresponding dynamic programming problem for a married household is

$$V(e_R, E_R, a, R) = \max_{c,a'} \begin{Bmatrix} U(c/g(2,0)) + \\ \beta p_{hR} p_{wR} \int V(e_R, E_R, a', R+1, m', 3) d\Omega^*_{Rm}(m') + \\ \beta p_{hR}(1-p_{wR}) \int V(e_R, E_R, a', R+1, m', 1) d\Omega^*_{Rs}(m') + \\ \beta p_{wR}(1-p_{hR}) \int V(e_R, E_R, a', R+1, m', 2) d\Omega^*_{Rs}(m') \end{Bmatrix}, \tag{4}$$

subject to

$$y = e_R + ra + T(e_R, E_{R-1}, a, R, n),$$

$$c + a' = y + a - \tau(y),$$

and

$$E_R = E_{R-1} + e_R.$$

The last earnings draw is e_R. This draw will determine the household's defined benefit pension receipts in retirement. Further, this earnings draw together with total earnings to date, E_{R-1}, determine total lifetime earnings, E_R, which in turn determines social security benefits. In addition, the household realizes that they will start receiving medical shocks beginning in the next period from the distribution $\Omega^*_{Rm}(m')$ or $\Omega^*_{Rs}(m')$, depending on whether both spouses, or just one, survive into the next period.[15] These distributions are assumed to be the stationary distributions associated with the corresponding Markov chains.

II.2.3. The Working Household's Problem

To simplify computation, the dynamic programming problem for working households has two fewer state variables than it does for households beginning at the retirement age. Specifically, we assume households incur no out-of-pocket medical expenses prior to retirement and face no preretirement mortality risk. Between ages S and R, the individual receives an

[15] Medical expenses drawn from the distribution for single households are half those drawn from the distribution for married couples.

exogenous earnings draw e. Given earnings and savings from the previous period, the individual decides how much to consume and save. The decision problem reads

$$V(e,E_{-1},a,j)=\max_{c,a'}\left\{U\left(c/g(A_j,K_j)\right)+\beta\int_{e'}V(e',E,a',j+1)d\Phi_j\left(e'\middle|e\right)\right\}, \tag{5}$$

subject to

$$y=e+ra+T(e,a,j,n),$$

$$c+a'=y+a-\tau\left(y\right),$$

and

$$E=E_{-1}+e.$$

Note that during working years, earnings draws for the next period come from the distribution Φ conditional on the individual's age and current earnings draw. The solution to this problem yields the decision rule as a function of the individual state: denote this decision rule $a'=G(e,E_{-1},a,j)$. We assume that each household begins life with zero assets. Given the observed earnings history of a household, we compute the optimal level of assets at every age using the decision rules.

III. Model Parameterization and Estimation of Exogenous Processes

In this section we specify functional forms and parameter values that we use to solve the model.

Preferences: The utility function for consumption of final goods is assumed to be CRRA:

$$U(c)=\begin{cases}\dfrac{c^{1-\gamma}}{1-\gamma}, & \text{if } \gamma\neq 1\\ \log c, & \text{if } \gamma=1\end{cases}.$$

Following Hubbard, Skinner, and Zeldes (1995); Engen, Gale, and Uccello (1999); and Davis, Kubler, and Willen (2002), the discount factor is set as $\beta=0.97$, and the coefficient of relative risk aversion (the reciprocal of the intertemporal elasticity of substitution) is set as $\gamma = 3$. We describe sensitivity analyses on these key parameters below.

Equivalence Scale: This is obtained from Citro and Michael (1995) and takes the form $g(A_j, K_j) = (A_j + 0.7K_j)^{0.7}$, where again, A_j (K_j) indicates the number of adults (children) in the household.

Survival Probabilities: These are based on the 1992 life tables of the Centers for Disease Control and Prevention, U.S. Department of Health and Human Services (http://www.cdc.gov/nchs/data/lifetables/life92_2.pdf).

Rate of Return: We assume an annualized real rate of return of 4 percent. This assumption is consistent with McGrattan and Prescott (2003), who find that the real rate of return for both equity and debt in the United States over the last 100 years, after accounting for taxes on dividends and diversification costs, is about 4 percent.[16] We include sensitivity analysis on this parameter below.

Taxes: We model an exogenous, time-varying, progressive income tax that takes the form

$$\tau(y) = a_0\left(y - \left(y^{-a_1} + a_2\right)^{-1/a_1} \right),$$

where *y* is in thousands of dollars. Parameters are estimated by Gouveia and Strauss (1994, 1999), and characterize U.S. effective, average household income taxes between 1966 and

[16] Four percent is also the average (rounded to the nearest percentage point) of the average real stock market return between 1947 and 1996 (7.6 percent) and the average real return on 3-month Treasury bills (0.8 percent).

1989.[17] We use the 1966 parameters for years before 1966 and the 1989 parameters for 1990 and 1991.

Transfers: We model the cumulative benefits from public income transfer programs using a specification suggested by Hubbard, Skinner and Zeldes (1995). Specifically, the transfer that an individual receives while working is given by

$$T = \max\left\{0, \underline{c} - \left[e + (1+r)a\right]\right\},$$

whereas the transfer that he or she will receive upon retiring is

$$T = \max\left\{0, \underline{c} - \left[SS(E_R) + DB(e_R) + (1+r)a\right]\right\}.$$

This transfer function guarantees a pre-tax income of $\underline{c}$, which we set based on parameters drawn from Moffitt (2002).[18] Subsistence benefits ($\underline{c}$) for a one-parent family with two children increased sharply, from $5,992 in 1968 to $9,887 in 1974 (all in 1992 dollars). Benefits have trended down from their 1974 peak – in 1992 the consumption floor was $8,159 for the one-parent, two-child family. We assume through this formulation that earnings, retirement income, and assets reduce public benefits dollar for dollar.

Social Security and Defined Benefit Functions: By making use of the social security earnings records, we calculate a close approximation of each household's social security entitlement. Households in the model understand social security rules and develop expectations

[17] Estimated parameters, for example, in 1989 are $a_0 = 0.258$, $a_1 = 0.768$ and $a_2 = 0.031$. In the framework, $a_1 = -1$ corresponds to a lump sum tax with $\tau(y) = -a_0 a_2$, while when $a_1 \to 0$, the tax system converges to a proportional tax system with $\tau(y) = a_0 y$. For $a_1 > 0$ we have a progressive tax system.

[18] Moffitt (2002, http://www.econ.jhu.edu/People/Moffitt/DataSets.html) provides a consistent series for average benefits received by a family of four. We use his "modified real benefit sum" variable, which roughly accounts for the cash value of food stamp, AFDC, and Medicaid guarantees. We weight state-level benefits by population to calculate an average national income floor. We use 1960 values for years prior to 1960 and use the equivalence scale described above to adjust benefits for families with different configurations of adults and children. We confirm that the equivalence scale adjustments closely match average benefit patterns for families with different

of social security benefits that are consistent with their earnings expectations. Details concerning the social security calculations are given in the first section of Appendix II.

Defined benefit pension expectations are formed on the basis of an empirical pension function that depends in a nonlinear way on union status, years of service in the pension-covered job, and expectations about earnings in the last year of work. We estimate the function with HRS data. Details are given in the second section of Appendix II.

Earnings Process: The basic unit of analysis for our life-cycle model is the household. We aggregate individual earnings histories into household earnings histories. Earnings expectations are a central influence on life-cycle consumption decisions, both directly and through their effects on expected pension and social security receipts.

The household model of log earnings is

$$\log e_j = \alpha^i + \beta_1 AGE_j + \beta_2 AGE_j^2 + u_j,$$

$$u_j = \rho u_{j-1} + \varepsilon_j,$$

where e_j is the observed earnings of the household i at age j in 1992-dollars, α^i is the household specific constant, AGE_j is age of the head of the household, u_j is an AR(1) error term of the earnings equation, and ε_j is a zero-mean i.i.d., normally distributed error term. The estimated parameters are α^i, β_1, β_2, ρ, and σ_ε.

numbers of adults and children using data from the Green Book (1983, pp. 259-260, 301-302; 1988, pp. 410-412, 789).

We divide households into six groups according to marital status, education, and number of earners in the household, giving us six sets of household-group-specific parameters.[19] Estimates (excluding the household-specific effect) are given in Appendix Table 4.[20]

Estimates of the persistence parameters range from 0.58 for single households without college degrees to 0.76 for married households with two earners, in which the highest earner has at least a college degree. The variance of earnings shocks ranges from 0.21 for single households without college degrees to 0.08 for married households with either one or two earners and in which the highest earner has at least a college degree.

We provide sensitivity analysis, setting the persistence parameter to 0.90 for all groups, below.

Out of Pocket Medical Expenses: The specification for household medical expense profiles across retirement ages is given by

$$m_t = \beta_0 + \beta_1 AGE_t + \beta_2 AGE_t^2 + u_t,$$

$$u_t = \rho u_{t-1} + \varepsilon_t,\ \varepsilon_t \sim N(0, \sigma_\varepsilon^2),$$

where m_t is the log of the household's out-of-pocket medical expenses at time t (the medical expenses are assumed to be \$1 if the self-report is zero or if the household has not yet retired),

---

[19] The six groups are (1) single without a college degree; (2) single with a college degree or more; (3) married, head without a college degree, one earner; (4) married, head without a college degree, two earners; (5) married, head with a college degree, one earner; and (6) married, head with a college degree, two earners. A respondent is an earner if his or her lifetime earnings are positive and contribute at least 20 percent of the lifetime earnings of the household.

[20] Given the model assumption, *predicted* earnings are given by

$$\hat{e}_j = \exp(\alpha^i + \beta_1 AGE_j + \beta_2 AGE_j^2 + \rho^{(j-l)} u_l + \sigma_\varepsilon^2 / 2)$$

where the period l disturbance is calculated as

$$u_l = \log(e_l) - \alpha^i - \beta_1 AGE_l - \beta_2 AGE_l^2$$

The household *expected* earnings profile is $\{\hat{e}_j\}_{j=S}^{R}$, where S is the first year that the head of the household started working full time, R is the head's last year of work. The head of household is defined throughout the paper as the person in the household with the largest share of lifetime earnings.

AGE_t is age of the household head at time t, u_t is an AR(1) error term and ε_t is white-noise. The parameters to be estimated are β_0, β_1, β_2, ρ, and σ_ε.

We estimate the medical-expense specification for four groups of households: (1) single without a college degree, (2) single with a college degree, (3) married without a college degree, and (4) married with a college degree, using the 1998 and 2000 waves of the HRS.[21] We use the age and education of the head of household. Results are given in the third section of Appendix II. The persistence parameters for medical shocks cluster tightly at 0.87 for each group. The variance of shocks is lower for households with greater education within a given household type (married or single), presumably reflecting higher rates of insurance coverage for households with college degrees relative to others.

To facilitate exposition, we denote the distribution from which a retired individual will draw next period's (age $j+1$) medical expenses m', conditional on the current earnings draw being m at age j, by $\Omega_{js}\left(m'|m\right)$ for single households and $\Omega_{jm}\left(m'|m\right)$ for married households.

III.1. Model Solution

We solve the dynamic programming problem by linear interpolation on the value function. Recall that the state space is composed of six variables for retired households: earnings drawn at j=R, e_R; cumulative earnings at the time of retirement, E_R; assets, a; age, i; medical expenses, m; and the number of household members alive, n (as noted earlier, we assume there are no mortality risks and out-of-pocket medical expenses for working households). We begin by "discretizing" the state space. The 50-point grid for earnings is constructed using the procedure

[21] New cohorts were added to the HRS in 1998, so the later years give us a broader range of ages to estimate age–medical expense profiles. These new cohorts were not matched to their social security earnings records, so they cannot be used for our baseline analysis.

discussed in Tauchen (1986). The 100-point grid for assets is chosen to be denser at lower levels of assets and progressively coarser so as to account for nonlinearities in the decision rules for assets induced by the borrowing constraint. We start at age *D,* assumed to be 100, and compute the value function $V(e_R, E_R, a, D, m, n)$ associated with all possible states in the discretized set. (The problem at this stage is trivial, since the individual will simply consume all income.) We move backward to the previous period and solve for the value function and the decision rule for assets. If optimal assets do not lie on the grid, we linearly interpolate between the points on the grid that lie on either side. In this fashion we go all the way to the starting age *S* and consequently recover the decision rules $a' = G(e, E_{-1}, a, j)$ for all $j = S...R$ and $a' = G(e, E_{-1}, a, j, m, n)$ for all $j = R+1...D$.

To summarize, for each household in our sample we compute optimal decision rules for consumption (and hence asset accumulation) from the oldest possible age (D) to the beginning of their working life (S) for any feasible realizations of the random variables: earnings, health shocks, and mortality. These decision rules differ for each household, since each faces stochastic draws from different earnings distributions (recall that α_i is household-specific). Household-specific earnings expectations also directly influence expectations about social security and pension benefits. Other characteristics also differ across households: for example, birth years of children affect the "adult equivalents" in a household at any given age. Consequently, it is not sufficient to solve the life-cycle problem for just a few household types.

Once optimal decision rules are solved for each household, we calculate optimal consumption (and therefore wealth) each period for each household using data on the observed realizations of earnings. Specifically, we start at age *S*, the first working age, where the household is assumed to begin with zero assets. Earnings to date are also zero at *S*. Given observed earnings at age *S*, $\hat{e}_S$, wealth (saving) is given by $a_{S+1} = G(\hat{e}_S, 0, 0, S)$. In the next period, the household receives an observed earnings draw $\hat{e}_{S+1}$, so aggregate earnings are given by $\hat{E}_S = \hat{e}_S$. Consequently, wealth is given by $a_{S+1} = G(\hat{e}_{S+1}, \hat{E}_S, a_S, S+1)$. We move forward in this fashion until we reach the age at which wealth data are available for that particular household.

The model, for the sample of 6,322 households, takes roughly six days to solve on a current desktop PC. Computational constraints, therefore, limit our ability to add features to the model or estimate the fit-maximizing discount rate or coefficient of relative risk aversion.

IV. Model Predictions and Their Correspondence to HRS Data

In this section we compare the optimal wealth levels for each household to their actual wealth holdings. We start by providing information on optimal wealth for HRS households as specified by our augmented life-cycle model. We then present detailed information on the degree to which HRS households are meeting their specific targets. The third subsection examines correlations between household characteristics and the degree to which households are and are not meeting their targets. We then compare our results to wealth predictions that would arise from a naïve behavioral model, where households save an age-varying and income-varying fraction of their annual earnings, and show our model significantly outperforms several alternatives. We also show that the model matches other features of wealth and consumption data. The section concludes with sensitivity analyses on key model parameters.

IV.1 Optimal Wealth Accumulation in the HRS

Table 2 summarizes the distribution of optimal net worth holdings. These targets include resources that could be accumulated in real and financial assets, the current value of defined contribution pensions, including 401(k)s, and housing net worth (for now, we assume households are willing to reduce housing in retirement to maintain consumption standards). Recall that the mean age of households in our sample is 55.7, so the average household will work many additional years before retiring.

As in the results of Hubbard, Skinner, and Zeldes (1995), low-skilled households will optimally accumulate negligible amounts of wealth outside of social security. The optimal wealth target for the median households in the lowest decile of the lifetime income distribution is $2,941 (including housing wealth). Means-tested transfer programs, including AFDC (during the period being studied), food stamps, SSI, and other forms of assistance, have income and asset tests. Households that receive transfers or that might receive transfers in the future may optimally accumulate fewer assets than they would in the absence of a safety net, since assets above a threshold will make households ineligible for benefits.[22]

Optimal wealth targets are $69,777 for the median household and are $253,631 for the median household in the highest decile of the lifetime income distribution.[23] The targets increase monotonically with lifetime income and with educational attainment.

[22] Empirical work on the effects of asset tests and asset accumulation comes to mixed conclusions. Gruber and Yelowitz (1999) find significant negative effects of Medicaid on asset accumulation, but Hurst and Ziliak (2001) find only very small effects of AFDC and food stamp asset limits.

[23] In other studies, wealth targets are commonly given in the form of wealth-to-earnings ratios. In the second column of Table 2 we show targets as a ratio of average income earned in the last five years that the household worked in the 1992 HRS. Optimal wealth-to-earnings ratios are 2.4 at the median, and median ratios range from 1.2 to 4.0 across the educational attainment distribution and from 0.6 to 3.9 across deciles of the lifetime income distribution.

A central feature of our work that distinguishes it from earlier papers is that we can compare optimal levels of wealth with actual wealth for each household in the HRS.

IV.2. Are Households Preparing Optimally for Retirement?

Figure 2 gives a scatterplot of optimal net worth against actual net worth, for HRS households with optimal and actual wealth between \$0 and \$1,000,000. The curved line gives a cubic spline of the median values of observed and optimal net worth.[24] The figure is striking, in that households appear to cluster just below the 45° line. The scatterplot gives suggestive visual evidence that most households are saving adequately for retirement.

A second striking aspect of Figure 2 is that it illustrates how a well-specified life-cycle model can closely account for variation in cross-sectional household wealth accumulation. A linear regression of actual net worth against predicted net worth and a constant shows the model explains 84 percent of the cross-household variation in wealth (that is, the R^2 is 84 percent).[25]

Table 3 shows the fraction of HRS households with wealth deficits, broken out by educational attainment and lifetime earnings deciles. Overall, 18.6 percent of the HRS sample has deficits (their wealth is less than the optimal target). However, the median magnitude (conditional on having a deficit) is very small, averaging \$5,714. Although some households are approaching retirement with significant wealth deficits, Table 3 provides additional evidence that HRS households overwhelmingly are well prepared for retirement.

[24] The median band is smoothed by dividing households into 30 groups based on observed net worth. We use Stata's "connect(s) bands(30)" option for the figure.

[25] Hurst (2003), using an innovative test from the PSID, shows that between 10 and 20 percent of the population appears not to be following the permanent income hypothesis. In brief, he splits the sample into low-residual undersavers (the bottom 10–20 percent of wealth residuals) and other households based on a log-wealth regression estimated from the 1989 wealth supplement of the PSID. He then shows that undersavers violate Euler equation excess sensitivity tests, whereas other households do not.

The last four columns of Table 3 help put the magnitude of the wealth deficits in perspective. Column 3 repeats the optimal net worth targets shown in Table 2. Comparing columns 2 and 3, median deficits for households who have deficits are generally small relative to the optimal targets. The remaining columns show the actual nonpension, non-social-security net worth accumulated by households, their social security entitlements, and their defined benefit pension entitlements. Reinforcing the impression from Figure 2, Table 3 shows that households accumulate more than the optimal level of net worth throughout the education and lifetime income distribution.

Another striking feature of Table 3 is that the probability of failing to meet the target is 34.6 percent in the lowest lifetime income decile and falls monotonically to 6.4 percent in the highest lifetime income decile. The simple cross-tabulation suggests that to the extent undersaving is a problem, it is particularly acute for low-income households. In this case, however, the descriptive statistics are misleading.

Table 4 shows the results from a probit regression of the probability that an HRS household failed to meet its optimal wealth target. It is striking that lifetime income does not have a strong, statistically significant effect once we condition on other covariates.[26] The only factor that is strongly correlated with having a wealth deficit is being single – married households are 27.2 percentage points less likely to have a deficit than single households.

In additional regressions estimated separately for a sample of single households and for a sample of married households, no covariates are correlated with saving less than the optimal target for single households. The only covariate correlated with a wealth deficit for married couples is an indicator variable for Hispanic ethnicity. Hispanic couples are 2.8 percentage points less likely than white couples to have deficits. These results suggest undersaving is approximately randomly distributed throughout the population – it is not a phenomenon disproportionately affecting poor households or households with low levels of education. Moreover, the strong income gradient shown in Table 3 is purely a composition effect – single households are much more likely than married households not to meet their wealth targets. Since single households are more likely to have lower incomes than married households, they are disproportionately represented in the lower deciles of the lifetime earnings distribution.[27]

IV.3. Are Americans Oversaving?

To this point we have only presented figures for the median household in the population or the median household within education or lifetime income deciles. Figure 3 shows selected percentiles – 10th, 25th, 50th, 75th, and 90th – of the *distribution* of the difference between actual and optimal wealth targets by lifetime income deciles. Two things are striking from this figure. First, only very small percentages of households accumulated less than their optimal wealth target. Undersavers are concentrated in the bottom half of the lifetime income distribution. And the magnitudes of the shortfalls, conditional on having a shortfall, are small. Second, the most striking aspect of Figure 3 is the degree to which people are saving too much. We probe this result in the remaining part of this subsection.

---

[26] The income coefficients are not jointly significant at even the 15 percent level of confidence, nor are the coefficients for the highest three lifetime income deciles.

[27] We were also concerned that the results for singles could be driven by divorce. If one-earner, married households divorced prior to the HRS survey, we would likely treat single earner as undersaving (they had income that before was supporting a family, yet following the divorce they would be expected to have only half the assets). Similarly, the nonworking partners would appear to have oversaved – they earned no income but are observed to have half the previously accumulated assets. However, this concern appears to be misplaced: the fraction of singles failing to meet their wealth targets is stable as we drop recently divorced individuals from the sample, or when we drop ever-divorced individuals.

There is some question about the degree to which the elderly are willing to reduce housing equity to sustain consumption in retirement. Venti and Wise (2001), for example, write "these results suggest that in considering whether families have saved enough to maintain their preretirement standard of living after retirement, housing equity should not be counted on to support general non-housing consumption." This conclusion is controversial. Hurd (2003) shows that the elderly households in the AHEAD data (the Study of Assets and Health Dynamics Among the Oldest Old) decumulate housing wealth as they age. Engen, Gale, and Uccello (1999) make a forceful case for including at least a significant portion of housing wealth when measuring resources households can draw on to maintain living standards in retirement.[28] Nevertheless, we do not want the first conclusion from Figure 3 – that a substantial majority of Americans are preparing well for retirement – to be driven by our treatment of home equity.

To explore the consequences of altering the treatment of housing in our calculations, we also examine the distribution of wealth deficits excluding half of housing from the resources available to meet the wealth target. Excluding half of housing equity, 57.9 percent of all households meet or exceed their wealth targets. The 25th percentile of the saving surplus distribution (net worth minus optimal targets) has a deficit of \$10,296, implying that 75 percent of households are exceeding or within \$10,300 of their optimal (nonpension, non-social-security) wealth target, even excluding half their net home equity. The lowest decile of the distribution has a deficit of \$40,371. The full distribution is given in Figure 4. Whereas results in the paper

[28] Engen, Gale, and Uccello (1999) make four points. First, existing work suggesting the elderly do not decumulate is flawed; housing should be the last asset to tap since it is illiquid and tax-preferred, and because some evidence is based on cohorts that were considerably less mobile than the HRS cohort. Second, households have vigorously extracted equity from houses in the 1980s and 1990s. Third, tax consequences of selling housing have fallen in recent years making it difficult to make inferences about people's willingness to downsize from earlier data. Fourth, housing provides consumption services and thus represents wealth. Conceptually and from a policy perspective, it seems odd to ignore one important source of wealth when considering economic well-being among households in retirement.

are qualitatively similar if we exclude half of housing equity, we report results when using all net worth for the remainder of the paper.

There are at least three features of the model that could account for the fact that many households appear to be accumulating significantly more than their optimal life-cycle targets.[29] First, we assume households expected and received a real rate of return of 4 percent. To the extent perceptions or realizations of real rates of return differ from our assumption, households will accumulate less or more than the target. One way to probe the importance of this factor is to make use of the limited geographic information available in the HRS.[30] House prices on the East and West Coasts increased substantially more over this period than house prices elsewhere.[31] If rate of return assumptions are systematically accounting for oversaving, we expect that households living in the New England Division (MA, NH, VT, MA, RI, and CN) or the Pacific Division (WA, OR, CA, AK, and HI) will be more likely than other households to significantly exceed their targets.

Second, households may intend to leave bequests. Dynan, Skinner, and Zeldes (2003), for example, argue that bequest intentions are the best way to reconcile wealth-income patterns in several nationally representative datasets. We use HRS questions probing the subjective likelihood of households leaving bequests of $10,000 and $100,000 to explore whether

[29] A fourth would be that our chosen preference parameters are incorrect. We explore the model sensitivity to preference parameters in a later subsection.
[30] HRS rules prohibit using restricted-access geocoded data with the restricted-access earnings data. Given the centrality of the earnings data to this project, we can only use the publicly available geographic information, which breaks the United States into nine regions.
[31] The repeat-sales house price index increased 381 percent between 1975 and 1992 in the Pacific, 278 percent in New England, and less than 120 percent in the East South Central, West South Central, and West North Central Divisions. See http://www.huduser.org/periodicals/ushmc/fall02/histdat10.htm.

households with a high (or certain) likelihood of leaving a bequest of these magnitudes are more likely than other households to exceed their wealth targets.[32]

Third, households might expect to live longer than suggested by the life tables published by the Centers for Disease Control and Prevention. Households with expectations of greater longevity will optimally accumulate more resources than predicted by our model. We use HRS questions probing subjective expectations of life expectancy to explore the importance of this factor in explaining oversaving.[33]

We examine the importance of rate of return differences, bequest intentions, and longevity risk along with factors in an empirical median regression model of "saving adequacy," defined as the difference between actual net worth (excluding DB pensions and social security wealth) and optimal net worth. The results are shown in Table 5. There is a sharply increasing, positive relationship between wealth accumulation and lifetime income starting in the sixth lifetime income decile. There is also a strong positive relationship between net worth and social security wealth and being self-employed and married. Households headed by men and black households have lower wealth than white, female-headed households. Saving adequacy also declines with the number of children in the family. There is no evidence that region of the country or subjective life expectancies have any relationship with saving adequacy (or oversaving). Bequest intentions, however, are positively, significantly related to acquiring more wealth than the optimal target. This result is consistent with purposeful bequest intentions affecting life-cycle wealth accumulation.

---

[32] The specific questions come from the 1994 wave of the HRS and read, "What are the chances that you [or your husband/wife/partner] will leave an inheritance totaling $10,000 (or $100,000) or more?

[33] The specific questions come from the 1992 wave of the HRS and read, "What do you think are the chances that you will live to be 75 (or 85) or more?"

IV.4. Alternative, "Naïve" Models

In the conclusion of their paper on variation in retirement wealth, Bernheim, Skinner, and Weinberg (2001) write, "the empirical patterns in this paper are more easily explained if one steps outside the framework of rational, far-sighted optimization. If, for example, households follow heuristic rules of thumb to determine saving prior to retirement…." Indeed, naïve or rule-of-thumb models of consumption have had an important place in the consumption literature at least since the Keynesian consumption function.

The exceptionally rich data we have on household earnings contain a great deal of information. Health shocks prior to retirement, unemployment, changes in labor demand and supply, among other things, will be reflected in the 41-year series of earnings we have for most households. Given the rich earnings data, it is natural to ask how much of the variation in HRS wealth can be explained by applying simple, rule-of-thumb saving behavior to the household-specific earnings trajectories. In doing this analysis, we continue to assume a 4 percent real interest rate, as we have done in the baseline life-cycle simulations. Our results are summarized in Table 6.

The simplest model we examined assumes that households save a constant fraction of their income, independent of their income or age. We iteratively sought the saving rate that maximized the goodness of fit measure, R^2. The fit-maximizing saving rate is 6.9 percent and the model explains 7.1 percent of the 1992 cross-sectional distribution of wealth in the HRS. A naïve model with age-varying and income-varying saving rates, in this case drawn from the parameters estimated in Dynan, Skinner, and Zeldes (2003, Table 3), explains 11.4 percent of the variation in retirement wealth. The original formulation of the life-cycle model (Modigliani and

Brumberg, 1954) where households save a constant fraction of permanent income (in this case, 5.8 percent) explains 16.1 percent of the variation in retirement wealth. It is clear that the augmented life-cycle model presented in this paper, which explains 83.7 percent of the cross-sectional variation in wealth, does a vastly better job matching the cross-sectional distribution of wealth in the 1992 HRS than the rule-of-thumb models we examine.

Our augmented life-cycle model includes many more parameters than the rule-of-thumb models. With more parameters, model fit should improve. Our augmented life-cycle model has a household-specific intercept, α_i, of the household age-earnings profiles, in effect adding 6,322 parameters to the model. We think this is a sensible way to model earning expectations – households presumably have a reasonable understanding of their place in the ability distribution, given observable characteristics such as educational attainment and age. Nevertheless, we also consider an alternative, more parsimonious version of the baseline model in which we assume that all households possess identical $\alpha's$. To do this we re-estimate the AR(1) process (assuming identical $\alpha's$ within household types) and simulate the optimal decision rules. We find that the model can account for 43.6 percent of the observed variation in 1992 wealth. Thus, our model, even with relatively few parameters, does a fairly good job matching the 1992 cross-sectional distribution of wealth in the HRS.

Anther useful benchmark for our augmented life-cycle model is to compare its model fit to reduced-form regression models. We regress wealth in 1992 against earnings and a host of other household characteristics, in which each of the 41 years of earnings observations appear separately.[34] This regression accounts for 25.3 percent of the variation in wealth. If we add

[34] The model also includes a quadratic in age, indicator variables for race and ethnicity, marital status, educational attainment, region, 17 occupations, 13 industries, two-earner household, retired, self-employment, unemployment,

quadratic terms in each year of earnings the resulting empirical model accounts for 29.7 percent of the variation in wealth. Even a parsimonious parameterization of the augmented life-cycle model that includes fewer parameters does a better job of explaining the observed variation in wealth than a regression that incorporates separately the linear and quadratic terms of annual earnings.

Our last experiment with alternative models attempts to clarify the importance of the augmented life-cycle model's decision rules in explaining the model fit, relative to the unusually rich earnings histories we use. To examine this, consider the following thought experiment. Once we solve the model for each individual, we have the optimal decision rules. Now, rather than using the actual earnings draws to obtain wealth predictions for 1992, imagine that we were to obtain, for each individual, 10,000 (sequences of) draws from the empirical earnings distribution and then use these draws to obtain predictions for wealth in 1992. The predicted wealth level for each individual is the average value implied by the potential realization of all such sequences. One can do the same for all the individuals in the sample and obtain the goodness of fit between the model and the data. The resulting R^2 is 41.1 percent. This suggests that although it is important to have the earnings realizations, the decision rules arising from the augmented life-cycle model are equally critical in arriving at such a close correspondence between model and data.

IV.5. Other Model Features

In this section we briefly examine three other features of the baseline model and show that they are consistent with several well-established facts about consumption. First, a well-known

---

health status, pension coverage, past marital history, and counts of the number of children, number of children between 12 and 17, and the number of children between 18 and 21.

fact from the consumption literature is that consumption is hump-shaped over the lifecycle (see, for example, Carroll, 1997). Figure 5 reports mean optimal consumption and income by age for our sample after we have netted out children's consumption so as to keep family size constant. Consumption is hump-shaped and peaks around age 46, whereas the peak in income occurs around age 51. The patterns in Figure 5 mirror those reported in Gourinchas and Parker (2002).[35]

A second issue has to do with how our augmented life-cycle model can match the well-known skewness of the wealth distribution. Our model predicts that the top 1 percent of our sample holds about 17 percent of the wealth. The corresponding figure in the data is 24 percent. So, just as is often claimed, the lack of bequest motives undermines the ability of the model to match the wealth holdings of the very rich. But the model does capture a large part of the skewness.

Third, a more stringent test of our model is how well it can match the change in wealth between 1992 and 2000. To put this in context, it is instructive to compare our model fit with a regression of the difference in wealth levels between 1992 and 2000 against household characteristics and earnings at every age.[36] The resulting R^2 is 3.6 percent. Introducing a quadratic term for earnings (at each age) increases the R^2 to 6.6 percent. In contrast, the baseline model (with a 4 percent real interest rate between 1992 and 2000) generates an R^2 of 28 percent.

[35] Prior to age 45 households in the model engage in "buffer stock" saving (Carroll, 2001). Carroll argues that a version of the life-cycle model where $\left((1+r)\beta\right)^{\frac{1}{\gamma}} = 1.0029 < g$, where g represents that growth rate of income does a very good job of capturing consumption-saving behavior during working ages. Our baseline parameters satisfy this condition. Setting $\beta = 1$, for example, leads to households beginning retirement wealth accumulation earlier in life and a monotonically increasing age-consumption profile during working years.

[36] The regression model also includes the covariates listed in footnote 34.

We conclude that the model presented does a good job of capturing observed behavior, not just of a cross-section at a point in time, but also changes over time across households.

IV.6. Sensitivity Analysis

There are three model parameters that we specify exogenously before solving the model – the discount factor (β), the coefficient of relative risk aversion (γ), and the real interest rate (r). In this subsection we analyze the sensitivity of the results to our choices of β, γ, and r. Table 7 shows the results. As expected, increases in β and r increase the incentive to save more in the future. In the life-cycle model this raises the optimal (or "target") level of wealth. When these targets are matched to the observed HRS data, more households fail to save adequately for retirement. For example, raising the real interest rate from 4 percent to 7 percent increases the fraction of households with wealth less than the optimal from 18.6 percent to 38.9 percent. An increase in γ has a similar effect because, as households become more risk averse, precautionary saving increases, increasing the optimal (or target) level of wealth accumulation and consequently undersaving. Nevertheless, the degree of undersaving is not particularly high – assuming that $\gamma = 5$, for example, increases the fraction of households with wealth less than the optimal from 18.6 percent to 35.7 percent in the baseline results.

Another parameter that plays an important role is the persistence in earnings across ages. Recall that these persistence parameters vary by type; they range from 0.58 for the single household without a college degree to 0.76 for the married, two-earner household, in which the head has a college degree. These parameters were estimated directly from the 41 years of Social Security earnings data. But many life-cycle models assume more persistence in the earnings process. The last line of Table 7 arbitrarily sets all persistence parameters in earnings expectations to 0.9. Households understand that this dramatically increases the odds of retaining a bad or good draw if one is received. The resulting R^2 is 68.9 percent, and 27.6 percent of households fail to meet their optimal targets. As expected, increasing the persistence in earnings increases optimal wealth accumulation at retirement.

Our sensitivity analysis leads us to conclude that within the range of values considered, most households in the HRS appear to have saved adequately for retirement.[37] Moreover, within a reasonably broad range of parameter values, the model can explain at least 68.9 percent of the cross-sectional variation in wealth in the 1992 HRS. These results do not depend at all on the inclusion of households in the sample with fully imputed earnings histories. When we drop households that did not allow the HRS to have access to their social security earnings records, the results are nearly identical using our baseline parameters: 18.4 percent of households accumulate less wealth than their optimal targets. Conditional on having a deficit, the median shortfall is $5,028. And the model accounts for 84 percent of the cross-sectional distribution of wealth in this subsample.

V. Conclusions

In this paper we develop a rigorous approach for assessing the degree to which a representative sample of households nearing retirement have prepared financially for that event. We find strikingly little evidence that HRS households have undersaved. And because consumption requirements likely fall when households reach retirement (if for no other reason than work expenses fall), our standard may overstate required wealth. We also note that our primary data come from 1992, well before the exceptionally strong stock market performance of

[37] If social security benefits are cut by 25 percent for all households, we find that 36.1 percent of households under-save, almost twice as many as in the baseline model. But we think such benefit cuts are unlikely for the HRS cohort.

the 1990s. Because 81 percent of households meet or exceed their wealth targets (and most of those who are below miss by a relatively small amount), we are skeptical that the consumption changes around retirement documented by Bernheim, Skinner, and Weinberg (2001) are due to inadequate retirement wealth accumulation.

We also find it striking how much of the variation in observed wealth accumulation can be explained by our life-cycle model. We explain over 85 percent of the variation in wealth for married households, and over 70 percent for single, never-married households. And the results presented reflect no tweaking or prior fitting of the model. If we had found major deviations from the model and behavior, it would be difficult to determine whether Americans were preparing poorly for retirement, or we had constructed a poor behavioral benchmark. The fact that our predictions and data closely align suggests two things. First, as mentioned above, Americans are saving enough to maintain living standards in retirement. And second, the life-cycle model provides a very good representation of behavior related to the accumulation of retirement wealth. Of course, we still admit the possibility that Americans are preparing poorly for retirement, our underlying behavioral model is poor, and the errors, coincidentally, offset.

Although the specific measures of undersaving and model fit clearly depend on parameter values, our two main results – that the life-cycle model is capable of closely matching the cross-sectional distribution of wealth in the HRS and that most HRS households are saving more than their optimal targets – are not affected significantly by parameter choices that are within the range commonly found in the related literature. We also find the life-cycle model does a much better job of matching the cross-sectional distribution of wealth in 1992 than a naïve model in which households save an income- and age-varying fraction of income.

Turning to the question posed in the title of the paper: are Americans saving optimally for retirement? The HRS covers a specific cohort of Americans – households age 51 to 61 in 1992. Consequently, we need to be careful in generalizing our results for the HRS cohort to younger households. This is particularly true if the generosity of social security is reduced in the future. Moreover, saving too much has efficiency costs in the sense that, absent preferences about intergenerational transfers or charitable contributions, reallocating consumption across time could increase lifetime utility. Because we cannot determine whether the systematic oversaving of HRS households reflects bequest motives, the expectation that social security will be reduced in the future, other failures in our characterization of the economic environment, or reflects nonoptimal behavior on the part of HRS households, we cannot definitively answer the question posed in the paper title. But the paper provides new, strong support for the life-cycle model as a good characterization of the process governing retirement wealth accumulation. And more important, it adds considerably to our confidence that Americans are preparing well for retirement.

References

Aguiar, Mark and Erik Hurst, 2003, "Consumption vs. Expenditure," working paper, University of Chicago, http://gsbwww.uchicago.edu/fac/erik.hurst/research/draft_7_july_2003.pdf.

Aiyagari, S. Rao, 1994, "Uninsured Idiosyncratic Risk and Aggregate Saving," Quarterly Journal of Economics, 109 (3), 659-84.

Banks, James, Richard Blundell, and Sarah Tanner, 1998, "Is There a Retirement-Savings Puzzle?" American Economic Review, September 88(4), 769-88.

Bernheim, B. Douglas, Jonathan Skinner and Steven Weinberg, 2001, "What Accounts for the Variation in Retirement Wealth Among U.S. Households?" American Economic Review, September 91(4), 832-57.

Carroll, Chris, 1997, "Buffer Stock Saving and the Life-Cycle/Permanent Income Hypothesis," Quarterly Journal of Economics, February 112(1), 1-55.

Carroll, Chris, 2001, "A Theory of the Consumption Function, With and Without Liquidity Constraints," Journal of Economic Perspectives, 15(3), 23-45.

Chang, Sheng-Kai, 2002, "Simulation Estimation of Dynamic Panel Tobit Models with an Application to the Convergence of the Black-White Earnings Gap and Income Dynamics," mimeo, Wayne State University.

Citro, Constance F. and Robert T. Michael, 1995, *Measuring Poverty: A New Approach*, Washington, D.C.: National Academy Press.

Davis, Steven J., Felix Kubler, and Paul Willen, 2002, "Borrowing Costs and the Demand for Equity Over the Life Cycle," NBER Working Paper #9331.

Deaton, Angus, 1991, "Saving and Liquidity Constraints," Econometrica, 59 (5), 1221-48.

Dynan, Karen E., Jonathan Skinner, and Stephen P. Zeldes, 2003, "Do the Rich Save More?" working paper, The Federal Reserve Board, Dartmouth College, Columbia, http://www.dartmouth.edu/~jskinner/rich_08_11_03.pdf.

Engen, Eric M., William G. Gale and Cori R. Uccello, 1999, "The Adequacy of Retirement Saving," Brookings Papers on Economic Activity, 2, 65-165.

Geweke, John, Michael P. Keane, and David Runkle, 1994, "Alternative Computational Approaches to Inference in the Multinomial Probit Model," Review of Economics and Statistics, 76 (4), 609-32.

Goldberger, Arthur S., 1991, *A Course in Econometrics*, Cambridge: Harvard University Press.

Gourinchas, Pierre-Olivier and Jonathan A. Parker, 2002. "Consumption Over the Life Cycle," Econometrica, 70 (1), 47-89.

Gouveia, Miguel and Robert R. Strauss, 1994, "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis," National Tax Journal, June 47(2), 317-39.

Gouveia, Miguel and Robert R. Strauss, 1999, "Effective Tax Functions for the U.S. Individual Income Tax: 1966-89," Proceedings of the 92nd Annual Conference on Taxation, National Tax Association, 155-65.

Grad, Susan, 1990, "Earnings Replacement Rates of New Retired Workers," Social Security Bulletin, October 53(10), 2–19.

Green Book, 1983, Background Material and Data on Programs Within the Jurisdiction of the Committee on Ways and Means, Committee on Ways and Means, U.S. House of Representatives, Washington D.C.

Green Book, 1988, Background Material and Data on Programs Within the Jurisdiction of the Committee on Ways and Means, Committee on Ways and Means, U.S. House of Representatives, Washington D.C.

Gruber, Jonathan and Aaron Yelowitz, 1999, "Public Health Insurance and Private Saving," Journal of Political Economy, December 107(6), 1249-74.

Gustman, Alan L. and Thomas L. Steinmeier, 1999, "Effects of Pensions on Savings: Analysis with Data from the Health and Retirement Survey," Carnegie-Rochester Conference Series on Public Policy, 50, 271-324.

Hamermesh, Daniel S., 1984, "Consumption During Retirement: The Missing Link in the Life Cycle," Review of Economics and Statistics, February 66(1), 1-7.

Hubbard, Glen R., Jonathan Skinner, and Stephen P. Zeldes, 1995, "Precautionary Saving and Social Insurance," Journal of Political Economy, April 103(2), 360–99.

Hurd, Michael D., 2003, "Bequests: By Accident or by Design?" in *Death and Dollars: the Role of Gifts and Bequests in America*, Alicia H. Munnell and Annika Sunden (eds.), pp. 93-118, Washington D.C.: Brookings Institution Press.

Hurd, Michael and Susann Rohwedder, 2003, "The Retirement-Consumption Puzzle: Anticipated and Actual Declines in Spending at Retirement," NBER Working Paper #9586.

Hurst, Erik, 2003, "Grasshoppers, Ants and Pre-Retirement Wealth: A Test of Permanent Income Consumers," mimeo, University of Chicago Graduate School of Business.

Hurst, Erik and James P. Ziliak, 2001, "Welfare Reform and Household Saving," Institute for Research on Poverty Discussion Paper no. 1234-01, University of Wisconsin-Madison.

Kotlikoff, Laurence J., Avia Spivak, and Lawrence H. Summers, 1982, "The Adequacy of Savings," American Economic Review, December 72(5), 1056-1069.

McGrattan, Ellen R. and Edward C. Prescott, 2003, "Taxes, Regulations, and the Value of U.S. and U.K. Corporations" Federal Reserve Bank of Minneapolis, Research Department, Staff Report Number 309.

Modigliani, Franco and R. Brumberg, 1954, "Utility Analysis and the Consumption Function: An Interpretation of Cross Section Data," In *Post-Keynesian Economics*, K.K. Kurihara (ed.), pp. 388-436. New Brunswick, NJ: Rutgers University Press.

Moffitt, Robert, 2002, "Documentation for Moffitt Welfare Benefits File," mimeo, February 22, http://www.econ.jhu.edu/People/Moffitt/DataSets.html.

Moore, James F. and Olivia S. Mitchell, 1998, "Projected Retirement Wealth and Savings Adequacy in the Health and Retirement Survey," Pension Research Council Working Paper 98-1, Wharton School, University of Pennsylvania, http://prc.wharton.upenn.edu/prc/prc.html

Palumbo, Michael G., 1999, "Uncertain Medical Expenses and Precautionary Saving Near the End of the Life Cycle," Review of Economic Studies, 66(2), 395-421.

Tauchen, George, 1986, "Finite State Markov-Chain Approximations to Univariate and Vector Autoregressions," Economic Letters, 20: 177-81.

Venti, Steven F. and David A. Wise, 2001, "Aging and Housing Equity: Another Look," manuscript prepared for the NBER conference on the Economics of Aging, July.

Wolff, Edward N., 2002, "Retirement Insecurity: The Income Shortfalls Awaiting the Soon-to-Retire," Washington, D.C.: Economic Policy Institute.

Appendix I: Imputing Earnings in the HRS

We have two problems that we address with the earnings data. For 77 percent of the 1992 HRS sample, we have access to each individual's social security earnings records from 1951 to 1991. The social security earning records report wage, salary, and self-employment income up to the earnings maximum (the earnings thresholds at which social security taxes are no longer taken from income). For 92 percent of the respondents with Social Security earnings records, we also have W-2 earnings records from 1980 to 1991. These W-2 records provide complete earnings information for wage and salary earners and the self-employed. The first difficulty is that 16 percent of positive social security earnings records are top-coded, and 41 percent of respondents with social security earnings records have at least one top-coded observation.

Our second problem is that 23 percent of observations refused to grant access to social security earnings records. For these households we have self-reported earnings information for their current job (or the most recent job if not employed) and as many as three previous jobs. We need to estimate their earning profiles based on their self-reported earnings information.

The goal is to use all available information to impute top-coded and missing earning observations, and as a result obtain complete individual earnings histories. For the imputation, we proceed in two steps. First, based on the social security and W-2 records, we estimate a dynamic-panel Tobit model to obtain individual earning processes. Then, conditional on all available earnings information, we use the estimates to impute the top-coded and missing observations.

Estimation

We start by describing our approach to estimating earnings for individuals with top-coded earnings.

For simplicity, suppose that we have earnings records of N individuals from time $t = 0$ to T, where 0 is the first period that these individuals started working full time. Assume for the moment that earnings are positive in each time period.[38] Denote the logarithmic value of individual i's actual (latent) and observed (by the researcher) earnings as $y^*_{i,t}$ and $y_{i,t}$, respectively. The relationship between the observed and actual (latent) earnings can be described as

$$y_{i,t} = \begin{cases} y^*_{i,t} & \text{if } y^*_{i,t} < y^{tc}_t \\ y^{tc}_t & \text{if } y^*_{i,t} \geq y^{tc}_t \end{cases}$$

where y^{tc}_t is the logarithmic value of the social security maximum taxable earnings at time t.

The individual log-earnings process is specified as

$$\begin{aligned} y^*_{i,0} &= x'_{i,0}\beta_0 + \varepsilon_{i,0} \\ y^*_{i,t} &= \rho y^*_{i,t-1} + x'_{i,t}\beta + \varepsilon_{i,t}, \ t \in \{1,2,...,T\} \\ \varepsilon_{i,t} &= \alpha_i + u_{i,t} \end{aligned} \tag{6}$$

[38] Generalizing this to the case in which the earnings series begins after time 0 and the case in which some earnings observations are zero is straightforward but detail-oriented, so we omit the discussion. We did treat these cases in practice, however.

where $x_{i,t}$ is the vector of i's characteristics at time t, and the error term $\varepsilon_{i,t}$ includes an individual-specific component α_i, which is constant over time and known to the individual before time 0, and the unanticipated white noise component, $u_{i,t}$. Notice that parameters β_0 and β are allowed to be different. In the following analysis, we employ random-effect assumptions with homoskedastic errors

(A1) $\alpha_i \mid \underline{x}_i \sim iid\, N(0,\sigma_\alpha^2)$

(A2) $u_{i,t} \mid \underline{x}_i \sim iid\, N(0,\sigma_u^2)\ \ \forall t$

(A3) $E[u_{i,t} \mid \underline{x}_i,\alpha_i]=0, E[u_{i,t}^2 \mid \underline{x}_i,\alpha_i]=\sigma_u^2, E[u_{i,t}u_{i,k} \mid \underline{x}_i,\alpha_i]=0,\ \ \forall t \in \{0,1,..,T\}, t \neq k$

where $\underline{x}_i \equiv (x_{i,0},x_{i,1},...,x_{i,T})$.

These three assumptions imply that

$$\varepsilon_i = (\varepsilon_{i,0},\varepsilon_{i,1},...,\varepsilon_{i,T})' \sim N(0,\Sigma) \tag{7}$$

where

$$\Sigma = \begin{bmatrix} \sigma_{0,0}^2 & \sigma_{0,1} & \cdots & \sigma_{0,T} \\ \sigma_{1,0} & \sigma_{1,1}^2 & \cdots & \sigma_{1,T} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{T,0} & \sigma_{T,1} & \cdots & \sigma_{T,T}^2 \end{bmatrix}$$

with $\sigma_{j,k}^2 = \sigma_\alpha^2 + \sigma_u^2$ for $j = k$, and $\sigma_{j,k}^2 = \sigma_\alpha^2$ otherwise. Our goal here is to obtain consistent estimates of the true parameters $\theta^* = (\beta,\beta_0,\rho,\sigma_\alpha^2,\sigma_u^2)$. We do this by maximum likelihood.

To construct the likelihood function for each individual's earnings series, notice that we can write the joint distribution function of each pair of random variables $(y_{i,t}, y_{i,t}^*)$ as $g(y_{i,t},y_{i,t}^* \mid y_{i,t-1},y_{i,t-1}^*,y_{i,t-2},y_{i,t-2}^*,.....,y_{i,0},y_{i,0}^*;\underline{x}_i,\theta)$. From the AR(1) assumption on earnings made in (6), it follows that

(A4) $g(y_{i,t},y_{i,t}^* \mid y_{i,t-1},y_{i,t-1}^*,y_{i,t-2},y_{i,t-2}^*,.....,y_{i,0},y_{i,0}^*;\underline{x}_i,\theta) = g(y_{i,t},y_{i,t}^* \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta)$

In other words, of all the information about past realized and observed earnings, only information from the previous period matters. As a special case, the conditional likelihood of the pair $(y_{i,0}, y_{i,0}^*)$ is $g_0(y_{i,0},y_{i,0}^* \mid \underline{x}_i,\theta)$ because there is no information about earnings before period 0.

Applying Bayes' rules to the density $g(\,)$, we have

$$g(y_{i,t},y_{i,t}^* \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta) = h(y_{i,t}^* \mid y_{i,t},y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta)q(y_{i,t} \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta) \tag{8}$$

where the density for the observed log-earnings conditional on the past information is a conventional Tobit likelihood function

$$q(y_{i,t} \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta) = [f(y_{i,t} \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta)]^{1\{y_{i,t}^*<y_t^{tc}\}}[\Pr(y_{i,t}^* \geq y_t^{tc} \mid y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta)]^{1\{y_{i,t}^* \geq y_t^{tc}\}}$$

where $f(\,)$ and $\Pr(\,)$ are a probability density and a cumulative density function respectively, and the conditional density $h(\,)$ for noncensored observations is the probability *mass* function

$$h(y_{i,t}^* \mid y_{i,t} < y_t^{tc},y_{i,t-1},y_{i,t-1}^*;\underline{x}_i,\theta) = \begin{cases} 1 & if\ y_{i,t}^* = y_{i,t} \\ 0 & if\ y_{i,t}^* \neq y_{i,t} \end{cases}$$

and the conditional density is simply $h(y^*_{i,t} \mid y_{i,t} = y^{tc}_t, y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta)$ for censored observations.

Similarly, we can write $g_0(y_{i,0}, y^*_{i,0} \mid \underline{x}_i, \theta) = h_0(y^*_{i,0} \mid y_{i,0}; \underline{x}_i, \theta) q_0(y_{i,0} \mid \underline{x}_i, \theta)$ where the conditional density $h_0(\,)$ for noncensored observations is the probability *mass* function $h_0(y^*_{i,0} \mid \underline{x}_i, \theta) = \begin{cases} 1 & \text{if } y^*_{i,0} = y_{i,0} \\ 0 & \text{if } y^*_{i,0} \neq y_{i,0} \end{cases}$ and the conditional density is $h_0(y^*_{i,0} \mid y_{i,0} = y^{tc}_0; \underline{x}_i, \theta)$ for censored observations. In addition, the density for the time-0 observed log-earnings conditional on past information is $q_0(y_{i,0} \mid \underline{x}_i, \theta) = [f(y_{i,0} \mid \underline{x}_i, \theta)]^{1\{y^*_{i,0} < y^{tc}_0\}} [\Pr(y^*_{i,0} \geq y^{tc}_0 \mid \underline{x}_i, \theta)]^{1\{y^*_{i,0} \geq y^{tc}_0\}}$.

From (6) and (7), it is apparent that the functions $f(y_{i,0} \mid \underline{x}_i, \theta)$ and $f(y_{i,t} \mid y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta)$, $t > 0$, are normal probability distribution functions and $\Pr(y^*_{i,0} \geq y^{tc}_0 \mid \underline{x}_i, \theta)$ and $\Pr(y^*_{i,t} \geq y^{tc}_t \mid y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta)$, $t > 0$, are normal cumulative distribution functions. For expositional convenience, define

$$\begin{aligned}
h^{tc}(0; \underline{x}_i, \theta) &\equiv h_0(y^*_{i,0} \mid y_{i,0} = y^{tc}_0; \underline{x}_i, \theta) \\
h^{tc}(t; \underline{x}_i, \theta) &\equiv h(y^*_{i,t} \mid y_{i,t} = y^{tc}_t, y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta), \quad t > 0 \\
q(0; \underline{x}_i, \theta) &\equiv q_0(y_{i,0}; \underline{x}_i, \theta) \\
q(t; \underline{x}_i, \theta) &\equiv q(y_{i,t} \mid y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta), \quad t > 0
\end{aligned}$$

The likelihood function for *i*'s series of observed log-earnings is

$$\begin{aligned}
&L_i(y_{i,T}, y_{i,T-1}, \dots, y_{i,1}, y_{i,0}; \underline{x}_i, \theta) \\
&= \int_{y^*_{i,T}}^{\infty} \cdots \int_{y^*_{i,0}}^{\infty} \left\{ \prod_{t=0}^{T} q(t; \underline{x}_i, \theta) \right\} \cdot \left\{ h(y^*_{i,0}; \underline{x}_i, \theta) \prod_{t=1}^{T} h(y^*_{i,t} \mid y_{i,t}, y_{i,t-1}, y^*_{i,t-1}; \underline{x}_i, \theta) \right\} dy^*_{i,0} \cdots dy^*_{i,T} \qquad (9) \\
&= \int_{y^*_{i,c_n}}^{\infty} \cdots \int_{y^*_{i,c_1}}^{\infty} \left\{ \prod_{t=0}^{T} q(t; \underline{x}_i, \theta) \right\} \cdot \left\{ \prod_{k=c_1}^{c_n} h^{tc}(k; \underline{x}_i, \theta) \right\} dy^*_{i,c_1} \cdots dy^*_{i,c_n} \quad = E_{y^*_{i,c_1}, \dots, y^*_{i,c_n}} \left[\prod_{t=0}^{T} q(t; \underline{x}_i, \theta) \right]
\end{aligned}$$

where c_1, c_2, ..., c_n are the periods where the observed log-earnings are censored, i.e., equal to their corresponding top-coded limits. Notice that, since we do not observe $y^*_{i,t}$ when it is censored, we integrate out $y^*_{i,t}$ for censored observations. Unfortunately, the integration does not yield any analytical solution, nor is direct numerical evaluation of the integral computationally feasible in this case. As an alternative, Chang (2002) proposes using a GHK (probit) simulator to deal with the computational burden of the integration.[39] The estimation results are given in Table A1.

[39] The GHK simulator gives a numerical approximation of a probit probability of interest. The GHK simulator is a popular choice of probit simulators due to its relative accuracy; see Geweke, Keane, and Runkle (1994) for details.

Appendix Table A1: Estimation Results of Individual Earning Processes §

Variables	Sample Male without College	Male with College	Female without College	Female with College
Initial Observation [t = 0]				
Constant	5.585*** (0.051)	3.858*** (0.326)	7.270*** (0.150)	4.954*** (0.427)
Race (white = 1, 0 otherwise)	0.304*** (0.025)	0.323*** (0.086)	0.132*** (0.037)	-0.013 (0.035)
Years of Schooling / Professional Post-Graduate Degree Dummy §	0.042*** (0.001)	-0.116** (0.059)	0.023*** (0.008)	0.041 (0.273)
No High School Dummy / Post-Graduate Degree Dummy §	-0.007*** (0.002)	-0.120*** (0.034)	-0.174*** (0.045)	0.122 (0.088)
Marital Status in 1992 HRS	0.113*** (0.018)	0.030 (0.019)	-0.034 (0.025)	0.108 (0.128)
Two-Earner Household Dummy	-0.071 (0.048)	-0.087 (0.067)	0.175*** (0.032)	0.143 (0.093)
Age	0.165*** (0.003)	0.294*** (0.021)	0.062*** (0.009)	0.231*** (0.028)
0.01 x Age^2	-0.189*** (0.004)	-0.336*** (0.033)	-0.066*** (0.013)	-0.284*** (0.042)
Subsequent Observations [t > 0]				
Constant	2.618*** (0.053)	2.120*** (0.069)	3.252*** (0.075)	2.932*** (0.169)
Race (white = 1, 0 otherwise)	0.114*** (0.006)	0.048** (0.019)	0.038** (0.019)	-0.025 (0.024)
Years of Schooling / Professional Postgraduate Degree Dummy §	0.019*** (0.002)	0.043*** (0.009)	0.029*** (0.004)	0.066* (0.034)
No High School Dummy / Post-Graduate Degree Dummy §	-0.028*** (0.008)	-0.009 (0.007)	-0.070*** (0.020)	0.097*** (0.021)
Marital Status in 1992 HRS	0.140*** (0.014)	0.123*** (0.026)	-0.128*** (0.017)	-0.123*** (0.031)
Two-Earner Household Dummy	-0.082*** (0.009)	-0.106*** (0.014)	0.210*** (0.014)	0.171*** (0.027)
Age	0.037*** (0.001)	0.034*** (0.005)	0.019*** (0.002)	0.014** (0.006)
0.01 x Age^2	-0.045*** (0.001)	-0.040*** (0.005)	-0.013*** (0.003)	-0.010 (0.006)
Earnings at the Previous Period	0.633*** (0.005)	0.730*** (0.014)	0.565*** (0.008)	0.667*** (0.027)
Variance of the Individual-Specific Effect (σ_α^2)	0.027*** (0.001)	0.013*** (0.002)	0.047*** (0.003)	0.028*** (0.007)
Variance of the Gross Error Term (σ_ε^2)	0.214*** (0.004)	0.227*** (0.010)	0.258*** (0.006)	0.239*** (0.014)
Number of Individual-Year Observations	92,889	21,479	48,625	8,747
Number of Respondents	3,095	724	2,670	452

§ The dependent variable is the respondents' natural-log-earnings. For samples with at most high school, the education variables are (i) Years of Schooling, and (ii) No High School Dummy. For samples with at least a bachelor's degree, the education variables are (i) Professional Postgraduate Degree (MBA, J.D., M.D., or Ph.D.) Dummy, and (ii) Postgraduate Degree Dummy. Standard errors are in parentheses. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Imputation

The idea is to impute top-coded and missing earnings observations with their conditional expectations where the conditioning variables include both the individual's characteristics and observed earnings. The conditional expectations are calculated numerically on the basis of the dynamic earnings model (6) and the distributional assumption (7). The imputation scheme is similar for top-coded and missing observations; therefore, we only discuss the scheme for top-coded observations here.[40]

To be concrete, notice that (6) implies that

$$E[y_0^* \mid \underline{y},\underline{x},\theta] = x_0'\beta_0 + E[\varepsilon_0 \mid \underline{y},\underline{x},\theta]$$
$$E[y_t^* \mid \underline{y},\underline{x},\theta] = \rho E[y_{t-1}^* \mid \underline{y},\underline{x},\theta] + x_t'\beta + E[\varepsilon_t \mid \underline{y},\underline{x},\theta],\ t \in \{1,2,...,T\} \qquad (10)$$

where $\underline{y} = (y_0, y_1,..., y_T)$ is the series of individual *i*'s observed earnings (the individual subscript *i* is omitted throughout this subsection). By construction, given information about individual's characteristics and observed earnings, $E[y_t^* \mid \underline{y},\underline{x},\theta]$ is on average the best guess for y_t^*. In other words, for top-coded observations, equation (6) suggests the imputed values

$$y_t^{imp} = E[y_t^* \mid y_t = y_t^{tc}; \underline{y},\underline{x},\theta]$$

which requires knowledge of $E[\varepsilon_t \mid y_t = y_t^{tc}; \underline{y},\underline{x},\theta]$ for every period *t* in which $y_t = y_t^{tc}$. The analytical form of $E[\varepsilon_t \mid y_t = y_t^{tc}; \underline{y},\underline{x},\theta]$ is not available in our case; therefore, we calculate this object numerically using the Gibbs sampling procedure.[41]

To facilitate the discussion about details of the procedure, denote $\underline{\varepsilon}_{<t} = (\varepsilon_0, \varepsilon_1,..., \varepsilon_{t-1})'$, $\underline{\varepsilon}_{>t} = (\varepsilon_{t+1}, \varepsilon_{t+2},..., \varepsilon_T)'$ and $\underline{\varepsilon}_{-t} = (\varepsilon_0, \varepsilon_1,..., \varepsilon_{t-1}, \varepsilon_{t+1},..., \varepsilon_T)'$ for any vector $\underline{\varepsilon} = (\varepsilon_0, \varepsilon_1,..., \varepsilon_T)'$. Here, we want to simulate *R* sets of $\underline{\varepsilon}$ that are consistent with the observed $\underline{y}$ and $\underline{x}$ given θ. The Gibbs sampling procedure does this in 2 steps for each round of simulation.

1. In the r^{th} round of simulation, *r* = 1,2, ... ,*R*, generate a "random" initial value $\underline{\varepsilon}_0^{(r)} = (\varepsilon_{0,0}^{(r)}, \varepsilon_{1,0}^{(r)},..., \varepsilon_{t,0}^{(r)},..., \varepsilon_{T,0}^{(r)})$ which satisfies (6) given $\underline{y}$, $\underline{x}$ and θ.
   Notice that $\varepsilon_{t,0}^{(r)}$ is not identified when $y_t = y_t^{tc}$. In this case, $\varepsilon_{t,0}^{(r)}$ is chosen randomly under the restriction that $y_{t,0}^{*(r)} \equiv (y_t^* \mid \underline{\varepsilon}_{<t} = \underline{\varepsilon}_{<t,0}^{(r)}, \varepsilon_t = \varepsilon_{t,0}^{(r)}; \underline{x},\theta) \geq y_t^{tc}$. If $y_{t-1} < y_{t-1}^{tc}$ and $y_t < y_t^{tc}$, $\varepsilon_{t,0}^{(r)}$ is defined by (6), i.e. it is actually not random.
2. Start with *m* = 1, draw a random number $\varepsilon_{t,m}^{(r)}$ from *t* = 0, ..., *T* from the distribution of $\varepsilon_t \mid \underline{\varepsilon}_{-t}; \underline{x},\theta$ such that $y_{t,m}^{*(r)} = (y_t^* \mid \underline{\varepsilon}_{>t} = \underline{\varepsilon}_{>t,m-1}^{(r)}, \underline{\varepsilon}_{<t} = \underline{\varepsilon}_{<t,m}^{(r)}, \varepsilon_t = \varepsilon_{t,m}^{(r)}; \underline{x},\theta) = y_t$ if $y_t < y_t^{tc}$,

[40] We can think of a missing earnings observation as an observation with top-coded value of 0, which is equivalent to saying that we know nothing about earnings in that period (as opposed to the case in which we observe top-coded earnings and know that the actual earnings is at least as large as the top-coded earnings).
[41] Briefly, Gibbs sampling is a procedure to draw a set of numbers randomly from a (valid) joint distribution. Then, the random draws are used to estimate properties of any marginal distribution of interest, which is difficult to derive analytically from the joint distribution. The procedure relies upon the law of large numbers, i.e., that moments of a distribution can be estimated consistently from a set of random draws from that distribution.

and $y^{*(r)}_{t,m} = (y^*_t \mid \underline{\varepsilon}_{>t} = \underline{\varepsilon}^{(r)}_{>t,m-1}, \varepsilon_{<t} = \varepsilon^{(r)}_{<t,m}, \varepsilon_t = \varepsilon^{(r)}_{t,m}; \underline{x}, \theta) \geq y^{tc}_t$ if $y_t = y^{tc}_t$. (This is equivalent to drawing $\varepsilon^{(r)}_{t,m}$ from $\varepsilon_t \mid \underline{\varepsilon}_{-t}; \underline{y}, \underline{x}, \theta$.) Then, continue from $m = 2$ to $m = M$.

With $\underline{\varepsilon}^{(r)}_M$, $r = 1,2, \ldots, R$, an estimate of $E[\varepsilon_t \mid \underline{y}, \underline{x}, \theta]$ is

$$\hat{E}[\varepsilon_t \mid \underline{y}, \underline{x}, \theta] = \frac{1}{R}\sum_{r=1}^{R} \varepsilon^{(r)}_{t,M} \tag{11}$$

Given the estimate $\hat{E}[\varepsilon_t \mid \underline{y}, \underline{x}, \theta]$, we calculate the imputed value of earnings as

$$\begin{aligned} y^{imp}_0 &= x'_0\beta_0 + \hat{E}[\varepsilon_0 \mid \underline{y}, \underline{x}, \theta] \\ y^{imp}_t &= \rho y^{imp}_{t-1} + x'_t\beta + \hat{E}[\varepsilon_t \mid \underline{y}, \underline{x}, \theta],\ t \in \{1,2,\ldots,T\} \end{aligned} \tag{12}$$

Notice that, by construction, $y^{imp}_0 = \frac{1}{R}\sum_{r=1}^{R} y^{*(r)}_{0,M}$ and $y^{imp}_t = \frac{1}{R}\sum_{r=1}^{R} y^{*(r)}_{t,M}$, $t \in \{1,2,\ldots,T\}$, and that $y^{imp}_t = y_t$ if $y_t < y^{tc}_t$ and $y^{imp}_t \geq y_t$ if $y_t = y^{tc}_t$.

The remaining parts of this subsection (i) construct the functional form for the conditional distribution of $\varepsilon_t \mid \underline{\varepsilon}_{-t}; \underline{x}, \theta$, and (ii) show how to draw a random number $\varepsilon^{(r)}_{t,m}$ from this conditional distribution to satisfy (6) given $\underline{y}$, $\underline{x}$ and θ. More notation is required. For any matrix Σ, denote $\Sigma_{t,t}$ as the element of Σ on the t^{th} row and t^{th} column, $\Sigma_{t,-t}$ as the t^{th} row of Σ with the element $\Sigma_{t,t}$ removed, $\Sigma_{-t,t}$ as the t^{th} column of Σ with the element $\Sigma_{t,t}$ removed, and $\Sigma_{-t,-t}$ as the matrix Σ with the t^{th} row and t^{th} column removed.

Recall the property of a joint-normal vector that

$$\begin{aligned} &\underline{\varepsilon} \sim N(E[\underline{\varepsilon}], \Sigma) \quad \Rightarrow \quad \varepsilon_t \mid \underline{\varepsilon}_{-t} \sim N(\mu_{t|-t}, \Sigma_{t|-t}) \\ &\mu_{t|-t} = E[\varepsilon_t] + \Sigma_{t,-t}\Sigma^{-1}_{-t,-t}\{\underline{\varepsilon}_{-t} - E[\underline{\varepsilon}_{-t}]\},\ \Sigma_{t|-t} = \Sigma_{t,t} - \Sigma_{t,-t}\Sigma^{-1}_{-t,-t}\Sigma_{-t,t} \end{aligned} \tag{13}$$

[42]

Recall from (7) that $E[\underline{\varepsilon}_i] = 0$ and $\Sigma = (1-\rho)\sigma^2_\varepsilon I_{T+1} + \rho\sigma^2_\varepsilon 1_{T+1}$, where 1_{T+1} is a $(T+1)$ x $(T+1)$ matrix whose elements are all 1. Thus, $\mu_{t|-t} = \Sigma_{t,-t}\Sigma^{-1}_{-t,-t}\underline{\varepsilon}_{-t}$, and $\Sigma_{t,-t}\Sigma^{-1}_{-t,-t} = \Sigma_{s,-s}\Sigma^{-1}_{-s,-s}$ and $\Sigma_{t|-t} = \Sigma_{s|-s}$ for any $t = 0, \ldots, T$ and $s = 0, \ldots, T$.

Recall that we draw a value for $\varepsilon^{(r)}_{t,m}$ randomly from the conditional distribution $\varepsilon_t \mid \underline{\varepsilon}_{-t}$ such that, given $\underline{\varepsilon}^{(r)}_{>t,m-1}, \varepsilon^{(r)}_{<t,m}, \varepsilon^{(r)}_{t,m}, \underline{x}, \theta$,

$$\begin{aligned} y^{*(r)}_{t,m} &= y_t & \text{if } y_t < y^{tc}_t \\ y^{*(r)}_{t,m} &\in [y^{tc}_t, \infty) & \text{if } y_t = y^{tc}_t \end{aligned} \tag{14}$$

In practice, it is more convenient to work with the standard-normal transformation of $\varepsilon_t \mid \underline{\varepsilon}_{-t}$

$$z_{t|-t} \equiv \frac{(\varepsilon_t \mid \underline{\varepsilon}_{-t}) - \mu_{t|-t}}{\sigma_{t|-t}} \sim N(0,1), \quad \sigma_{t|-t} = \sqrt{\Sigma_{t|-t}} \tag{15}$$

[42] See, for example, Goldberger, 1991, pp. 196-97.

From (6), $\varepsilon_0 \mid \underline{\varepsilon}_{-0} = (y_0^* \mid \underline{\varepsilon}_{-0}) - x_0'\beta_0$ and $\varepsilon_t \mid \underline{\varepsilon}_{-t} = (y_t^* - \rho y_{t-1}^* \mid \underline{\varepsilon}_{-t}) - x_t'\beta,\ t \in \{1,2,...,T\}$.
Also, since $(y_t^* \mid \underline{\varepsilon}_{-t}; \underline{x}, \theta) = (y_t^* \mid \underline{\varepsilon}_{<t}; \underline{x}, \theta)$, $(y_{t-1}^* \mid \underline{\varepsilon}_{>t} = \underline{\varepsilon}_{>t,m-1}^{(r)}, \underline{\varepsilon}_{<t} = \underline{\varepsilon}_{<t,m}^{(r)}; \underline{x}, \theta) = y_{t-1,m}^{*(r)}$.

Thus, with the transformation (15), drawing $\varepsilon_{t,m}^{(r)}$ from (13) to satisfy (14) is equivalent to drawing $z_{t|-t,m}^{(r)}$ such that

$$z_{0|-0,m}^{(r)} = \begin{cases} \{y_0 - x_0'\beta_0 - \mu_{0|-0,m}^{(r)}\} / \sigma_{0|-0} & \text{if } y_0 < y_0^{tc} \\ \Phi^{-1}\left(\xi_{0,m}^{(r)} + (1-\xi_{0,m}^{(r)})\Phi(\{y_0^{tc} - x_0'\beta_0 - \mu_{0|-0,m}^{(r)}\} / \sigma_{0|-0})\right) & \text{if } y_0 = y_0^{tc} \end{cases}$$

for $t = 0$, and

$$z_{t|-t,m}^{(r)} = \begin{cases} \{y_t - \rho y_{t-1,m}^{*(r)} - x_t'\beta - \mu_{t|-t,m}^{(r)}\} / \sigma_{t|-t} & \text{if } y_t < y_t^{tc} \\ \Phi^{-1}\left(\xi_{t,m}^{(r)} + (1-\xi_{t,m}^{(r)})\Phi(\{y_t^{tc} - \rho y_{t-1,m}^{*(r)} - x_t'\beta - \mu_{t|-t,m}^{(r)}\} / \sigma_{t|-t})\right) & \text{if } y_t = y_t^{tc} \end{cases}$$

for $t > 0$,[43] with

$$y_{0,m}^{*(r)} = x_0'\beta_0 + (\sigma_{0|-0} z_{0|-0,m}^{(r)} + \mu_{0|-0,m}^{(r)})$$
$$y_{t,m}^{*(r)} = \rho y_{t-1,m}^{*(r)} + x_t'\beta + (\sigma_{t|-t} z_{t|-t,m}^{(r)} + \mu_{t|-t,m}^{(r)}),\ t \in \{1,2,...,T\}$$

where $\mu_{t|-t,m}^{(r)} = \Sigma_{t,-t}\Sigma_{-t,-t}^{-1}\begin{pmatrix} \underline{\varepsilon}_{<t,m}^{(r)} \\ \underline{\varepsilon}_{>t,m-1}^{(r)} \end{pmatrix}$, $\varepsilon_{t,m}^{(r)} = \sigma_{t|-t} z_{t|-t,m}^{(r)} + \mu_{t|-t,m}^{(r)}$, $\sigma_{t|-t} = \sqrt{\Sigma_{t|-t}}$ and $\Sigma_{t|-t} = \Sigma_{t,t} - \Sigma_{t,-t}\Sigma_{-t,-t}^{-1}\Sigma_{-t,t}$, and $\xi_{t,m}^{(r)}$ is a random draw from a [0,1] uniform distribution. Notice that $y_{t,m}^{*(r)} = y_t$ if $y_t < y_t^{tc}$ and $y_{t,m}^{*(r)} \geq y_t$ if $y_t = y_t^{tc}$ by construction.

---

[43] To see how this works, note first that for $\varepsilon \sim N(\mu, \sigma^2)$, $f(\varepsilon) = (2\pi\sigma^2)^{-1/2}\exp(-0.5\{\varepsilon\text{-}\mu\}/\sigma)$. Define $z \equiv \{\varepsilon\text{-}\mu\}/\sigma$. It follows that $F(\varepsilon) = \Phi(z(\varepsilon))$, where Φ is the standard normal cumulative density function. Thus,

$$\Phi(z(\varepsilon^{(r)})) = F(\varepsilon^{(r)}) = \xi_t^{(r)} + (1-\xi_t^{(r)})F(\varepsilon^{tc}) = \xi_t^{(r)} + (1-\xi_t^{(r)})\Phi(z(\varepsilon^{tc}))$$

In other words, drawing $\varepsilon^{(r)}$ from a truncated distribution of $(\varepsilon|\varepsilon \geq \varepsilon^{tc})$ is equivalent to drawing $z^{(r)} = z(\varepsilon^{(r)})$ from a truncated distribution of $(z|z \geq z^{tc})$ and then transforming $z^{(r)}$ back to $\varepsilon^{(r)}$.

Appendix II: Underlying Model Processes

A1. Social Security Function

From the expected earnings profiles, we can calculate the *lifetime* summation of household earnings up to the year of retirement as $E_R \equiv \sum_{j=S}^{R} e_j$, where e_j denotes the household earnings at age *j* in a common base-year unit, and *S* and *R* denote the first and the last working ages, respectively.[44] Denote $\overline{\phi}^h$ and $\overline{\phi}^w$ as the fractions of E_R that are contributed by the husband and wife of the household, respectively.[45] Based on E_R, $\overline{\phi}^h$ and $\overline{\phi}^w$, we can approximate the household annual social security benefits as follows.

(a) Calculate *Individual* Primary Insurance Amount (PIA)[46]

Individual *i*'s annual indexed monthly earnings (AIME) can be approximated as

$$AIME^i \approx \overline{\phi}^i \, E_R / L^i \tag{16}$$

with

$$L^i = 12 \times \max\{R^i - 22, 40\}$$

where $i = h$ (husband) or w (wife), and L^i is the number of months of *i*'s covered period.[47] Without loss of generality, we set $L^w = 40$ for single-male households and $L^h = 40$ for single-female households.

Individual PIA can be calculated as

$$\begin{aligned} PIA^i = {} & 0.90 \times \min\{AIME^i, b_0\} + 0.32 \times \min\{\max\{AIME^i - b_0, 0\}, b_1 - b_0\} \\ & + 0.15 \times \max\{AIME^i - b_1, 0\} \end{aligned} \tag{17}$$

where b_0 and b_1 are the bend points. For the 1992 formula, b_0 = \$387 and b_1 = \$2,333.

---

[44] As opposed to a *discounted* present value of earnings, the summation is a straightforward summation of earnings in a common base-year currency unit which is the concept employed by the Social Security Administration.

[45] The terminologies "husband" and "wife" are not literal. In particular, we call a single male respondent "husband" and a single female respondent "wife." Without this simplification, we need separate treatments for married and single households. Under this generalization, $\overline{\phi}_i^h$ = 1 and $\overline{\phi}_i^w$ = 0 for single-male households, and $\overline{\phi}_i^h$ = 0 and $\overline{\phi}_i^w$ = 1 for single-female households.

[46] Social security benefits derived from the calculations in this section are not precise because the calculated AIME may be smaller than the actual AIME and, conditional on AIME being correctly calculated, the calculated household benefits may be larger than the actual ones. For the former, the reasons are (i) we do not exclude 5 years of lowest earnings from calculation, (ii) we use base-year (i.e. real) values of earnings after age 60 instead of nominal values, (iii) we do not take into account earnings in retirement if respondents work beyond their household retirement dates. For the latter, the reason is that we assume both husband and wife of a married household are eligible for collecting benefits at the household retirement date. If one of them is not eligible at the retirement date, the approximation will overstate the benefits. Nevertheless, by virtue of having complete earnings histories for most individuals, our calculations are considerably more accurate than those in other life-cycle simulation models of wealth accumulation.

[47] Without the lower bound of 40 years in the max operator ($\max\{R^i - 22, 40\}$), AIME would be too high for households whose members retire before age 62. In addition, notice that we use the *household* retirement date (R^i) rather than the *individual* retirement date.

(b) Calculate *Household* Annual Social Security Benefits

First, the *individual* monthly social security benefits are calculated as

$$ssb^i = \max\{d^i_{own} PIA^i, d^i_{spouse} PIA^{i\text{'s spouse}}, ssx^i\} \tag{18}$$

where *i's spouse* = *h* (*w*) if *i* = *w* (*h*), d^i_{own} is the fraction of *i*'s PIA that *i* would get if *i* collected benefits based on *i*'s PIA, d^i_{spouse} is the fraction of PIA of *i*'s spouse that *i* would get if *i* collected benefits based on PIA of *i*'s spouse, and ssx^i is the monthly benefit that *i* would get if *i* collected benefits based on PIA of *i*'s ex-spouse.[48] Without loss of generality, for single-male households, $d^h_{spouse} = d^w_{own} = d^w_{spouse} = ssx^w = 0$, and $d^w_{spouse} = d^h_{own} = d^h_{spouse} = ssx^h = 0$ for single-female households. In addition, we set $ssx^h = ssx^w = 0$ for married households because we do not have any information to determine ssx^i. Similarly, $ssx^i = 0$ for any single households without information to determine their ex-spouses' PIA.

Finally, household *i*'s *annual* social security benefits can be approximated as

$$ss_i = 12 \times (ssb^h_i + ssb^w_i) \tag{19}$$

which, for a married household, is the benefits the household would get when both the husband and wife survive. When one of the spouses in a married household dies, the *annual* social security benefits of the surviving spouse is

$$ss^{survive}_i = 12 \times \max\{d^h_{own} PIA^h, d^w_{own} PIA^w\} \tag{20}$$

In other words, we approximate the surviving spouse's benefits to be the higher of the husband's and wife's benefits that they would be able to collect on the basis of their own earning histories (which determine their PIAs) and the household retirement date (which determines the factors *d*). This approximates the actual guideline of the Social Security Administration.

AII.2. Defined Benefit pension:

The annual defined benefit (DB) pension benefit is estimated as

$$\begin{aligned} db = {} & DB^h\{\beta^h_0 + \beta^h_1 UNION^h + \beta^h_2 YRSV^h + (\gamma^h_0 + \gamma^h_1 UNION^h + \gamma^h_2 YRSV^h\}\phi^h_R e_R\} + \\ & DB^w\{\beta^w_0 + \beta^w_1 UNION^w + \beta^w_2 YRSV^w + (\gamma^w_0 + \gamma^w_1 UNION^w + \gamma^w_2 YRSV^w\}\phi^w_R e_R\} \\ & + \beta^b_0 DB^h DB^w + \xi \end{aligned} \tag{21}$$

where the superscripts *h* and *w* indicate "husband" and "wife," respectively. DB^i is a binary variable equal to 1 if *i* has a DB pension. $UNION^i$ is a binary variable equal to 1 if *i* belongs to a union at the DB job. $YRSSV^i$ is the number of years that *i* stays in the DB job up to *i*'s retirement date. e_R is the household earnings in the last period of work, and ϕ^h_R and ϕ^w_R indicate the fractions of e_R that belong to the husband and wife, respectively, with $\phi^h_R + \phi^w_R = 1$ by construction. ξ is an error term that is assumed to be distributed as $N(0, \sigma^2_\xi)$.[49] Finally, the parameters to be estimated are $\beta^b_0, \beta^h_0, \beta^h_1, \beta^h_2, \beta^w_0, \beta^w_1, \beta^w_2, \gamma^h_0, \gamma^h_1, \gamma^h_2, \gamma^w_0, \gamma^w_1, \gamma^w_2$ and σ^2_ξ.

---

[48] To recover the ex-spouse's PIA, we first compute the benefit amount that a single respondent would get based on her own earning history. Then, we compare the amount to the reported amount of social security benefits in the first wave that the respondent reported collecting the benefits. If the reported benefit amount is higher, we assume that the single respondent collected benefits based on her ex-spouse's records and the reported amount is used to recover her ex-spouse's PIA.

[49] The specification is estimated with ordinary least squares using the White formula for the standard error.

db is calculated by assuming that the household receives annual DB pension benefits that are constant in real terms from the first period of retirement until none of the recipients survive. In particular, let *dbwealth* be the observed present discounted value of *db*.

$$dbwealth = \sum_{j=R+1}^{D} \pi_j \frac{db}{\delta_j} \quad \Rightarrow \quad db = dbwealth \Bigg/ \sum_{j=R+1}^{D} \frac{\pi_j}{\delta_j}$$

where δ_j is the discount rate that converts pension benefits at age *j* into an equivalent value of 1992 dollars (i.e. having δ_j 1992-dollars at age *j* is as good as having one 1992 dollar in 1992), and π_j is the probability that the household will survive at age *j* conditional on surviving in the year that *dbwealth* was reported, *R* is the last period of work, and *D* is a terminal age (the household will not live beyond this age). The estimation results are given in Table A2.

Appendix Table A2: Coefficient Estimates for Annual DB Pension Benefits

Variable	Coefficient Estimates	Standard Errors
Husband's Estimate of Constant	1,914.9***	(701.7)
Husband's Estimate of Union Status	-483.943	(613.2)
Husband's Estimate of Years in Service	47.9	(30.0)
Husband's Estimate of His Last-Period Earnings	-0.028	(0.024)
Husband's Estimate of His Last-Period Earnings Interacting with Union Status	0.008	(0.022)
Husband's Estimate of His Last-Period Earnings Interacting with Years In Service	0.004***	(0.001)
Wife's Estimate of Constant	-245.366	(540.1)
Wife's Estimate of Union Status	1,108.1***	(334.0)
Wife's Estimate of Years in Service	66.6***	(22.6)
Wife's Estimate of Her Last-Period Earnings	0.013	(0.033)
Wife's Estimate of Her Last-Period Earnings Interacting with Union Status	0.003	(0.022)
Wife's Estimate of Her Last-Period Earnings Interacting with Years in Service	0.004***	(0.001)
Estimate of Constant if Both Husband and Wife Have a Pension	-168.864	(420.7)
R^2	0.572	
N	2,203	

AII.3. Out-of-Pocket Medical Expenses

We construct household annual medical expenses based on the (HRS-imputed) answers to the four medical expense questions asked in the 1998 and 2000 HRS. The four questions are:

> E10. About how much did you pay out-of-pocket for [nursing home/hospital/ nursing home and hospital] bills [since R's LAST IW MONTH, YEAR/in the last two years]?
>
> E18a. About how much did you pay out-of-pocket for [doctor/outpatient surgery/dental/doctor and outpatient surgery/doctor and dental/outpatient surgery and dental/doctor, outpatient surgery, and dental] bills [since R's LAST IW MONTH, YEAR/in the last two years]?
>
> E21a. On the average, about how much have you paid out-of-pocket per month for these prescriptions [since R's LAST IW MONTH, YEAR/in the last two years]?
>
> E24a. About how much did you pay out-of-pocket for [in-home medical care/special facilities or services/in-home medical care, special facilities or services] [since R's LAST IW MONTH, YEAR/in the last two years]?

We construct household annual medical expenses as one-half of E10 + E18a + 24*E21a + E24a. The 1996 and 1997 household annual medical expenses are calculated from the information from the 1998 HRS and, similarly, the 1998 and 1999 household annual medical expenses are calculated from the information from the 2000 HRS. The sample included is all households (HRS, AHEAD, CODA, and WB) that participated in and retained marital statuses between the 1998 and 2000 HRS. The estimation results are given in Table A3.

Appendix Table A3: Coefficient Estimates for the AR(1) Annual Medical Expenses

Group	Coefficient Estimates					R^2	*N*
	Group Constant	*Age*	0.01**Age*^{2}	$\hat{\rho}$	$\hat{\sigma}$		
Single, No College	0.020 (0.018)	0.123*** (0.004)	-0.075*** (0.005)	0.869	1.436	0.214	20,096
Single, College	0.111*** (0.038)	0.139*** (0.007)	-0.084*** (0.009)	0.869	1.183	0.335	3,200
Married, No College	0.051*** (0.013)	0.174*** (0.003)	-0.115*** (0.004)	0.869	0.948	0.519	18,228
Married, College	0.026 (0.017)	0.185*** (0.004)	-0.121*** (0.005)	0.871	0.696	0.693	5,824

Note: The numbers of households for these groups are 10048, 1600, 9114, and 2912 respectively. *, **, and *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

AII.4. Estimates of Household Earnings Expectations.

We construct household earnings as the summation of individual earnings for all adults in the household. The estimates for the model of household earnings described in the text are given in Table A4.

Appendix Table A4: Coefficient Estimates for the Household AR(1) Earnings Profiles

Group	Coefficient Estimates					R^2	N
	Group Constant	*Age*	0.01**Age*^{2}	$\hat{\rho}$	$\hat{\sigma}$		
Single, No College	4.758 (0.022)	0.231 (0.003)	-0.259 (0.004)	0.58	0.46	0.065	43,339
Single, College	3.787 (0.042)	0.293 (0.007)	-0.316 (0.009)	0.68	0.38	0.175	8,677
Married, No College, One-Earner	6.753 (0.018)	0.173 (0.002)	-0.195 (0.003)	0.62	0.32	0.138	65,472
Married, No College, Two-Earner	5.157 (0.038)	0.264 (0.006)	-0.282 (0.007)	0.70	0.31	0.283	15,779
Married, College, One-Earner	6.741 (0.019)	0.173 (0.003)	-0.187 (0.004)	0.67	0.28	0.183	56,482
Married, College, Two-Earner	5.003 (0.038)	0.259 (0.007)	-0.269 (0.009)	0.76	0.29	0.254	14,626

Note: The numbers of households for these groups are 1873, 351, 2076, 512, 1821 and 519 respectively.

Table 1: Descriptive Statistics for the Health and Retirement Study (dollar amounts in 1992 dollars)

Variable	Mean	Median	Standard Deviation
1991 Earnings	$35,263	$28,298	$35,264
Present Discounted Value of Lifetime Earnings	$1,691,104	$1,516,931	$1,193,969
Defined Benefit Pension Wealth	$105,919	$17,371	$191,328
Social Security Wealth	$106,714	$97,150	$65,140
Nonpension Net Worth	$250,513	$107,000	$541,164
Mean Age (years)	55.7		4.7
Mean Education (years)	12.7		3.4
Fraction Male	0.70		0.46
Fraction Black	0.11		0.31
Fraction Hispanic	0.06		0.25
Fraction Couple	0.66		0.48
No High School Diploma	0.22		0.41
High School Diploma	0.55		0.50
College Graduate	0.12		0.33
Post-College Education	0.10		0.30
Fraction Self-Employed	0.15		0.35
Fraction Retired	0.29		0.45

Source: Authors' calculations from the 1992 HRS. The table is weighted by the 1992 HRS household analysis weights.

Table 2: Optimal Net Worth (excluding DB Pensions) and Optimal Wealth-to-Earnings Ratios for HRS Households (dollar amounts in 1992 dollars)

Group	Median Optimal Wealth Target	Median Optimal Wealth-to-Income Ratios[a]
All Households	$69,777	2.4
No High School Diploma	$22,524	1.2
High School Diploma	$70,383	2.5
College Degree	$137,528	3.1
Post-College Education	$178,924	4.0
Lowest Lifetime Income Decile	$2,941	0.6
2nd Income Decile	$15,368	1.4
3rd Income Decile	$30,059	1.9
4th Income Decile	$48,200	2.1
Middle Income Decile	$60,513	2.2
6th Income Decile	$83,399	2.6
7th Income Decile	$89,488	2.4
8th Income Decile	$106,724	2.5
9th Income Decile	$140,853	2.7
Highest Lifetime Income Decile	$253,631	3.9

Notes: Authors' calculations from the life-cycle model described in the text. Calculations use the 1992 HRS household analysis weights.
[a] Income is the average of the last 5 working years.

Table 3: Percentage of Population Failing to Meet Optimal Wealth Targets and Magnitude of Wealth Deficit (dollar amounts in 1992 dollars)

Group	Percentage Failing to Meet Optimal Target	Median Deficit (conditional on deficit)	Optimal Net Worth Target	Median Net Worth	Median Social Security Wealth	Median DB Pension Wealth
All Households	18.6%	$5,714	$69,777	$107,000	$97,150	$17,371
No High School Diploma	20.9%	$2,982	$22,524	$40,000	$71,774	$0
High School Diploma	19.1%	$5,315	$70,383	$106,000	$97,086	$21,290
College Degree	14.7%	$13,696	$137,528	$217,314	$127,167	$60,752
Post-College Education	16.1%	$21,579	$178,924	$263,500	$126,691	$152,639
Lowest Lifetime Income Decile	34.6%	$2,885	$2,941	$5,288	$25,667	$0
2nd Income Decile	33.6%	$3,904	$15,368	$26,050	$41,346	$0
3rd Income Decile	26.0%	$6,599	$30,059	$48,000	$56,951	$0
4th Income Decile	24.5%	$5,500	$48,200	$80,938	$76,426	$18,428
Middle Income Decile	18.1%	$9,477	$60,513	$92,828	$95,527	$27,994
6th Income Decile	13.0%	$4,997	$83,399	$123,000	$116,054	$44,418
7th Income Decile	12.2%	$13,415	$89,488	$138,000	$133,596	$55,100
8th Income Decile	6.7%	$7,688	$106,724	$170,000	$150,893	$76,165
9th Income Decile	7.7%	$4,312	$140,853	$229,000	$163,372	$107,655
Highest Lifetime Income Decile	6.4%	$29,062	$253,631	$395,889	$200,747	$123,192

Notes: Authors' calculations as described in the text.

Table 4: Correlates of the Probability that Households Accumulate Too Little Wealth

	Probit Regression on Having a Wealth Deficit	
	dF/dx [§]	Standard Error
2nd Lifetime Income Decile	.016	.018
3rd Lifetime Income Decile	-.005	.019
4th Lifetime Income Decile	.015	.023
5th Lifetime Income Decile	-.006	.024
6th Lifetime Income Decile	-.021	.025
7th Lifetime Income Decile	-.017	.028
8th Lifetime Income Decile	-.061**	.025
9th Lifetime Income Decile	-.046	.029
10th Lifetime Income Decile	-.043	.034
Retired	.001	.011
Has Pension	-.003	.011
Social Security Wealth	-9.41e-08	1.88e-07
Age	-.002	.001
Male	-.007	.012
Black	-.006	.012
Hispanic	-.028	.015
Married	-.272***	.017
High School Diploma	.004	.012
College Degree	-.009	.018
Graduate Degree	-.000	.020
Self-Employed	-.012	.014

[§] For dummy variables, dF/dx is a discrete change. The mean probability of a deficit in the sample is .185. The pseudo R2 for the probit regression is .1562 and sample size is 6,271.
*, **, *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table 5: Correlates of the Median Wealth Surplus

	Median Regression of "Saving Adequacy" (Actual-Optimal Net Worth)	
	Coefficient Estimates	Standard Error
Constant	-16,301.4**	6,946.8
2nd Lifetime Income Decile	-1,253.3	888.9
3rd Lifetime Income Decile	-915.6	1,293.4
4th Lifetime Income Decile	-1,396.0	1,356.2
5th Lifetime Income Decile	71.6	1,779.1
6th Lifetime Income Decile	5,856.9***	2,848.7
7th Lifetime Income Decile	8,424.3*	2,912.1
8th Lifetime Income Decile	13,482.3***	5,237.6
9th Lifetime Income Decile	17,853.1***	6,860.9
10th Lifetime Income Decile	56,459.3***	5,420.9
Retired	1,584.8*	1,055.6
Has Pension	-1,849.6**	1,231.1
Social Security Wealth	0.066***	0.0
Age	249.3*	124.2
Male	-2,313.7*	1,316.1
Black	-2,017.9*	1,019.2
Hispanic	-769.6	1,668.0
Married	8,957.8***	1,620.3
High School Diploma	667.4	1,041.8
College Degree	3,849.5	2,223.6
Graduate Degree	5,223.6	3,999.4
Self-Employed	13,665.8***	2,993.9
Number of Children	-462.9**	212.4
Number of Grandchildren	162.1**	72.4
Subjective Probability of Living > 75	6.5	19.0
Subjective Probability of Living > 85	-8.4	18.0
Subjective Probability of Bequest > $10k	23.4***	9.7
Subjective Probability of Bequest > $100k	282.7***	30.6
Mid-Atlantic Division	471.8	4,344.8
East North Central Division	-206.1	4,632.2
West North Central Division	2,198.0	4,901.5
South Atlantic Division	-45.1	4,989.6
East South Central Division	121.2	4,776.4
West South Central Division	-1,805.9	4,889.2
Mountain Division	731.0	5,146.5
Pacific Division	1,336.5	5,123.1

§ For dummy variables, dF/dx is a discrete change. Standard errors are bootstrapped in the median regression. The pseudo R2 for the median regression is .0918 and the sample size is 6,271.
*, **, *** denote statistical significance at the 10 percent, 5 percent, and 1 percent levels, respectively.

Table 6: Alternative Models

Model	**R^2 (in percent)**
Naïve (save a constant fraction of Y_t)	7.1
Naïve (save an income-varying fraction of Y_t)	11.4
Modigliani (S_t=a+bY_p)	16.1
Constant Alpha	43.6
Reduced-Form Regression Including 41 Years of Earnings	25.3
Reduced-Form Regression Including Quadratic Terms for 41 Years of Earnings	29.7
Monte Carlo Draws on Earning Sequence	41.1
Base Case in Paper	83.7

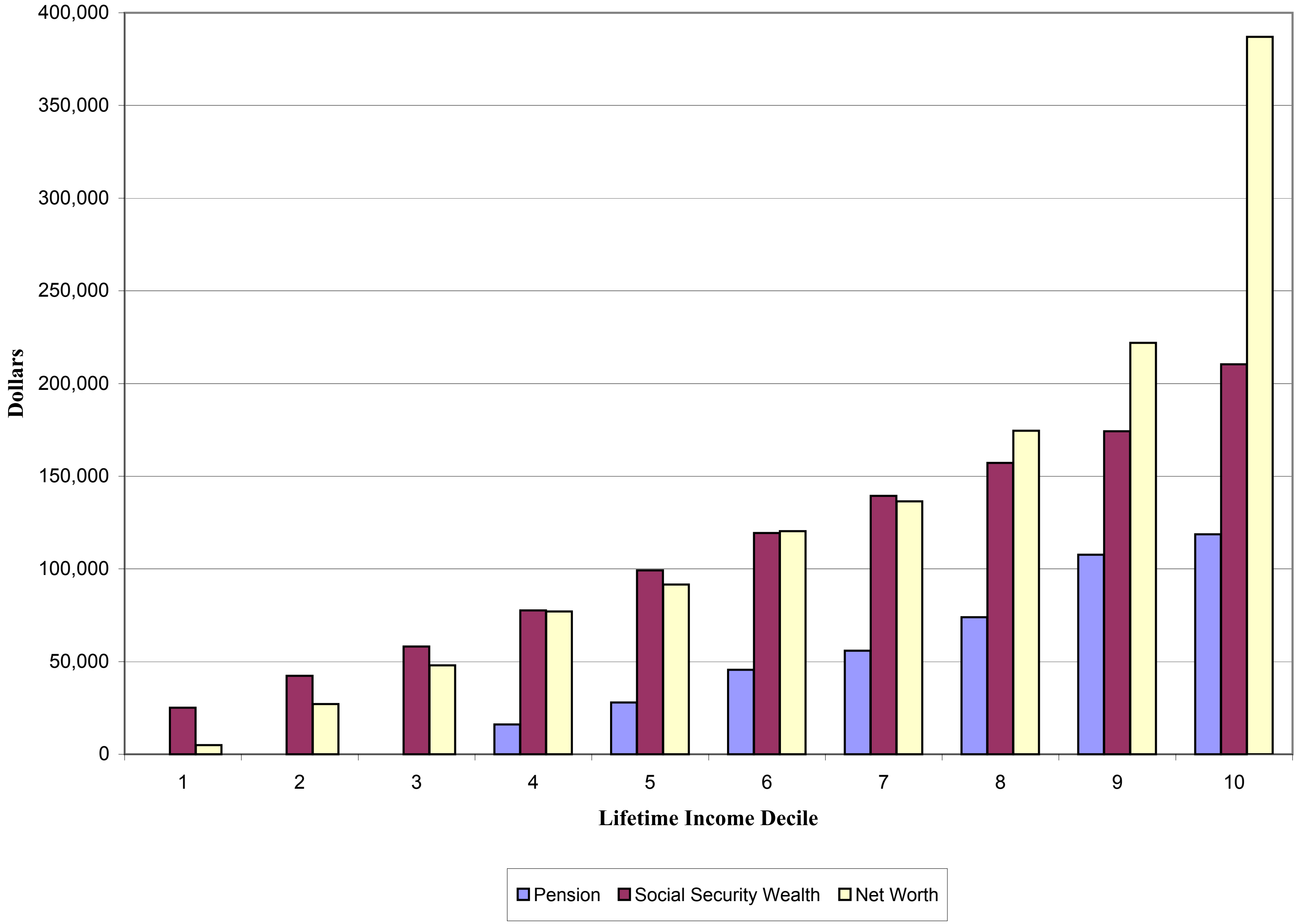
Source: Authors' calculations as described in the text.

Table 7: Sensitivity Analysis

Parameter Value	Percentage Failing to Meet Optimal Target	Measure of fit: R^2 (in %)	Deficit Conditional on Failing to Meet Optimal Target (1992$)
Baseline: $\beta = 0.97$, $\gamma = 3$, $r = 4\%$	18.6	83.7	5,714
$\beta = 1.0$	25.6	85.5	6,242
$\beta = 0.93$	14.1	81.3	6,567
$r = 5\%$	24.7	85.1	6,000
$r = 7\%$	38.9	77.3	18,752
$\gamma = 1.5$	14.5	92.3	4,656
$\gamma = 5$	35.7	84.8	11,131
$\rho = 0.9$	27.6	68.9	18,634

Source: Authors' calculations as described in the text.

Figure 1: Median DB Pension Wealth, Social Security Wealth, and Net Worth (excluding DB Pensions) by Lifetime Income Decile, (1992 dollars)



Source: Authors' calculations from the 1992 HRS

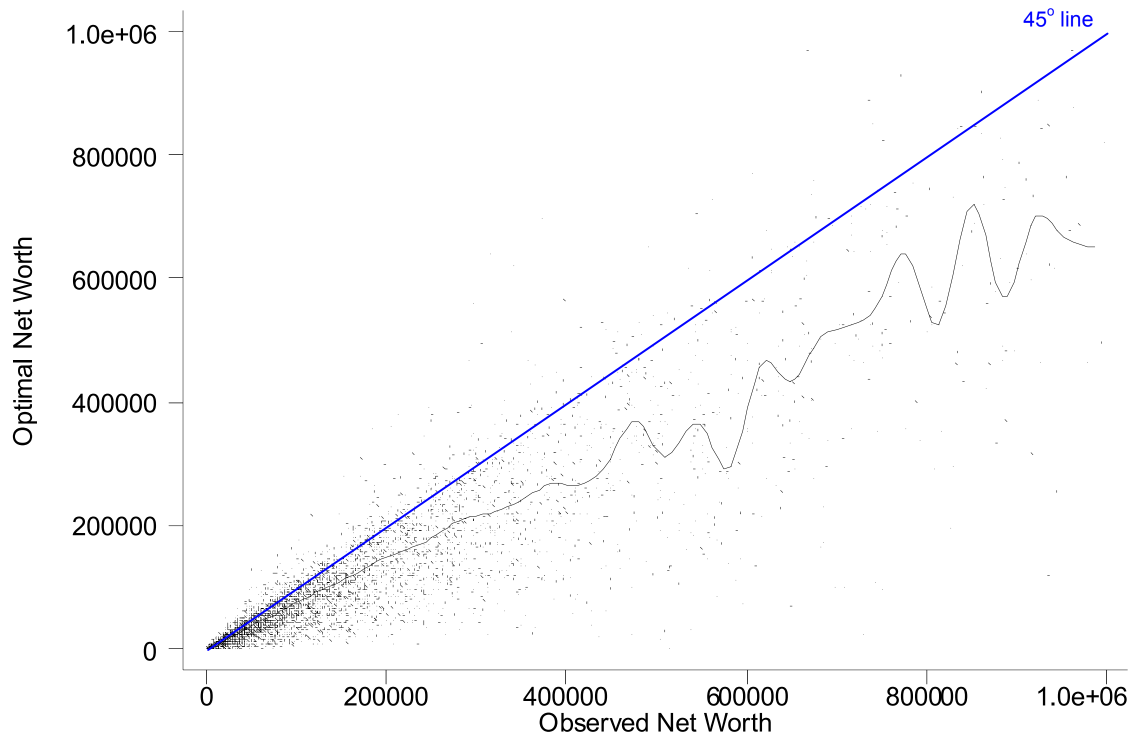
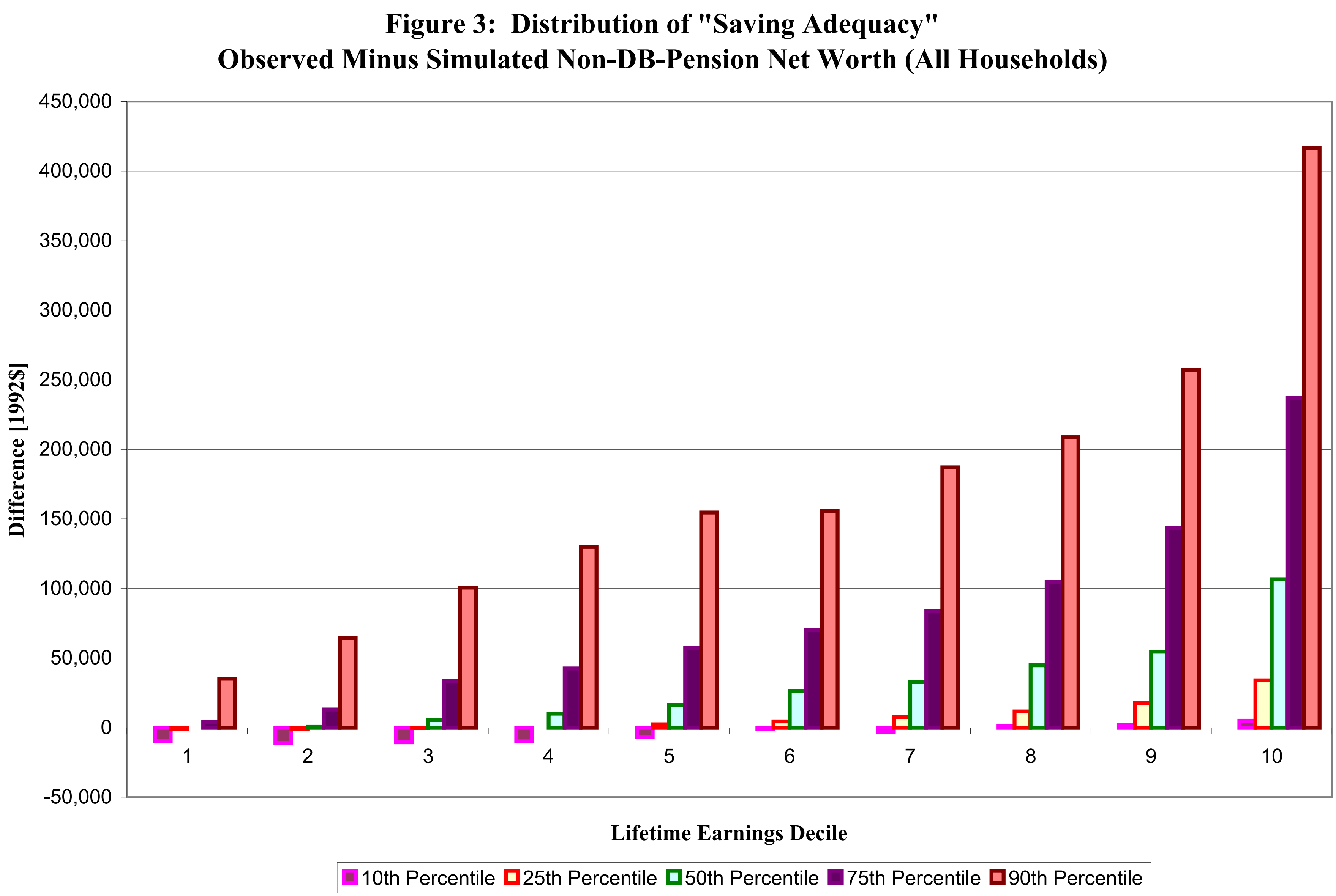


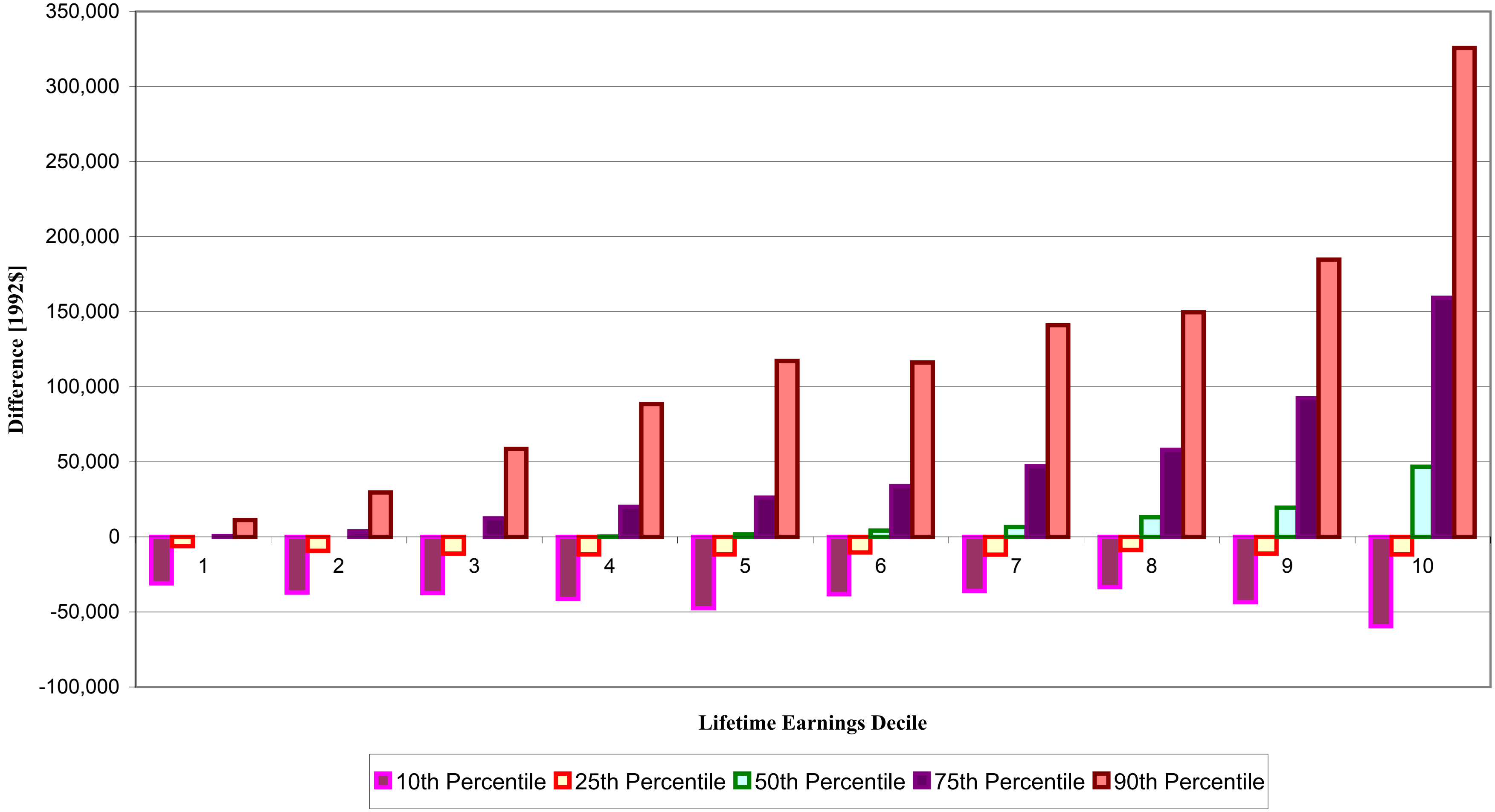
Figure 2: Scatterplot of Optimal and Actual Wealth

Source: Observed net worth is constructed from the 1992 HRS. Optimal net worth comes from solving the baseline model described in the text.



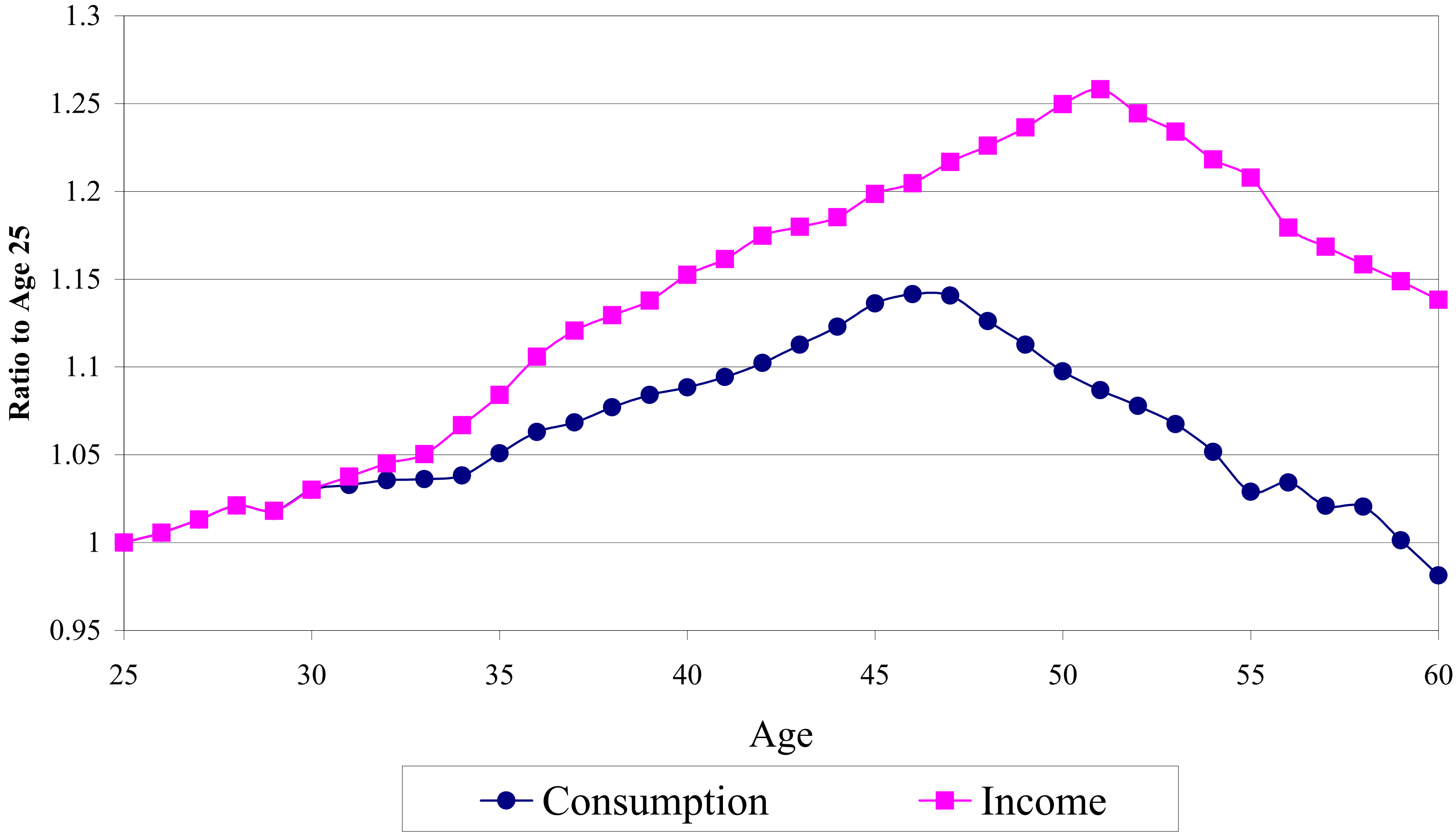
Source: Authors' calculations from the baseline model and 1992 HRS

Figure 4: Distribution of "Saving Adequacy" (Observed Minus Simulated Non-DB-Pension Net Worth), Excluding Half of First-Home Housing Wealth



Source: Authors' calculations from the baseline model and 1992 HRS

Figure 5: Consumption and Income by Age



Source: Authors' calculations from the baseline model and the 1992 HRS