Why do the Elderly Save? The Role of Medical Expenses

Mariacristina De Nardi, Eric French, and John Bailey Jones*

December 9, 2009

Abstract

This paper constructs a model of saving for retired single people that includes heterogeneity in medical expenses and life expectancies, and bequest motives. We estimate the model using AHEAD data and the method of simulated moments. Out-of-pocket medical expenses rise quickly with age and permanent income. The risk of living long and requiring expensive medical care is a key driver of saving for many higher income elderly. Social insurance programs such as Medicaid rationalize the low asset holdings of the poorest, but also benefit the rich, by insuring them against high medical expenses at the ends of their lives.

*This paper was previously circulated under the title "Differential Mortality, Uncertain Medical Expenses, and the Saving of Elderly Singles". For helpful comments and suggestions, we thank an editor and two referees, Jerome Adda, Kartik Athreya, Gadi Barlevy, Marco Bassetto, Marco Cagetti, Jeff Campbell, Chris Carroll, Michael Hurd, Helen Koshy, Nicola Pavoni, Monika Piazzesi, Luigi Pistaferri, Victor Rios-Rull, Tom Sargent, Karl Scholz, and seminar participants at many institutions. Olga Nartova, Kenley Pelzer, Phil Doctor, Charles Doss, and Annie Fang Yang provided excellent research assistance.
Mariacristina De Nardi: Federal Reserve Bank of Chicago and NBER. Eric French: Federal Reserve Bank of Chicago. John Bailey Jones: University at Albany, SUNY.
De Nardi gratefully acknowledges financial support from NSF grant SES-0317872. Jones gratefully acknowledges financial support from NIA grant 1R03AG026299. French thanks the Social Security Administration for hospitality while writing the paper. The views expressed in this paper are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago, the Federal Reserve System, the Social Security Administration, the National Science Foundation, or the National Institute on Aging.

1 Introduction

Many elderly keep large amounts of assets until very late in life. Furthermore, the income-rich run down their assets more slowly than the income-poor. Why is this the case?

To answer this question, we estimate a life cycle model of saving on a sample of single, retired elderly individuals. The key elements in our framework are risky and heterogenous medical expenses, risky and heterogenous life expectancy, a government-provided minimum consumption level (or "consumption floor", which is income and asset-tested), and bequest motives.

Our main result is that medical expenditures are important in explaining the observed savings of the elderly, especially the richer ones. For example, our baseline model predicts that, as in the data, between ages 74 and 84 median assets for those in the top permanent income quintile fall from \$170,000 to \$130,000. When we eliminate all out-of-pocket medical expenses, median assets for this group are predicted to fall much more, from \$170,000 at age 74 to \$60,000 at age 84.

This result is due to an important feature of out-of-pocket medical expenses data that we estimate: average medical expenditures rise very rapidly with age and income. For example, our model predicts that average annual out-of-pocket medical expenditures rise from \$1,100 at age 75 to \$9,200 at age 95. While a 95-year-old in the bottom quintile of the permanent income distribution expects to spend \$1,700 a year on medical expenses, a person of the same age in the top quintile expects to spend \$15,800. Medical needs that rise with age provide the elderly with a strong incentive to save, and medical expenses that rise with permanent income encourage the rich to be more frugal.

We show that the consumption floor also has an effect on saving decisions at all levels of income. When we reduce old-age consumption insurance by 20%, median assets for 90-year-olds in the highest permanent income quintile increase from \$100,000 to \$120,000, while median assets for 90-year-olds in the second highest quintile increase from \$40,000 to \$50,000. The net worth of those in the third and fourth income quintiles also increases. The consumption floor thus matters for wealthy individuals as well as poorer ones. This is perhaps unsurprising given the size of our estimated medical needs for the old and income-rich; even wealthy households can be financially decimated by medical needs in very old age.

We further find that heterogeneity in mortality is large and is important for understanding the savings patterns of the elderly. In particular, differential mortality gives rise to a bias that makes the surviving elderly seem more thrifty than they actually are. While a 70-year-old man in poor health in the bottom income quintile expects to live only six more years, a 70-year-old woman in good health and in the top income quintile expects to live 17 more years. Failure to account for the mortality bias would lead us to understate asset decumulation by over 50% for the 74 year-old people in our sample. Consistently with the data, our model allows people who are rich, healthy, and female to live longer[1] and generates asset profiles consistent with these observations.

Finally, we find that bequests are luxury goods, and that bequest motives are potentially quite important for the richest retirees. Our estimates of the bequest motive, however, are very imprecise and for most of our sample, savings barely change when the bequest motive is eliminated. One reason why the bequest motive is weakly identified is that even in the top permanent income quintile, median assets in our sample of elderly singles never exceed $200,000; hence we do not have enough "super-rich" individuals to pin down the bequest motive.

The above results are based on a life cycle model that takes out-of-pocket medical expenses as exogenous. That is, we first use the Assets and Health Dynamics of the Oldest Old (AHEAD) dataset, to estimate stochastic processes for mortality and out-of pocket medical expenditures as functions of sex, health, permanent income, and age. We then estimate our model using the method of simulated moments, where the model's preference parameters are chosen so that the permanent income-conditional median age-asset profiles simulated from the model match those in the data, cohort by cohort.

The additional sources of heterogeneity that we consider allow the model to match important aspects of the data: our estimated structural model is not rejected when we test its over-identifying restrictions. In addition, the distribution of deceased persons' estates generated by our model matches up closely with that observed in the data.

Despite the good fit of the model to the data, one might think that some of the results, such as the responses of savings to changes in government insurance, might not be robust to allowing the retirees to adjust both savings

[1]See Attanasio and Emmerson [4], and Deaton and Paxon [19] for evidence on permanent income and mortality. See Hurd et al. [41] for evidence on health status and mortality.

and medical expenses. As a robustness check, we construct a model in which retirees choose savings as well as medical expenses.

In this version of the model, retirees derive utility from consuming medical goods and services. The magnitude of this utility depends on "medical needs" shocks, which in turn depend on age and health. Retirees optimally choose medical and non-medical consumption, while taking into account the cost-sharing provided by insurance and their own resources. We estimate this version of the model by requiring it to fit observed out-of-pocket medical expenditures as well as observed asset holdings.

Importantly, although our medical needs shocks do not depend on permanent income, in our model out-of-pocket medical expenditures rise with both permanent income and age, as in the data. Medicaid pays for most of the medical expenses of the poor, who thus have little or no out-of-pocket medical expenses, but not for those of the rich, who pay out-of-pocket for their own, higher quality medical care. For this reason, Medicaid plays an important role in generating the correlation between out-of-pocket medical expenses and income. Moreover, we find that medical expenditures and the consumption floor have large effects on saving even when medical expenses are a choice variable.

The intuition for why medical needs are so important, even when people can adjust their medical expenditure, is that out-of-pocket medical expenditures found in the model have to match those in the data. This implies that high-income 70-year-olds anticipate that if they live into their nineties, they will probably choose to make large medical expenditures – like the 90-year-olds in our sample – and will probably save to pay for them. Once this feature of the data is taken into account, it is not surprising that medical expenditures have large effects on savings, whether they are exogenous or chosen.

Making medical expenses endogenous reduces the effects of social insurance on savings. The effects that remain, however, are still stronger at higher income levels than at lower ones, because the out-of-pocket medical expenses of the richest people are much higher than those of the poorest.

The rest of the paper is organized as follows. In section 2, we review the most closely related literature. In section 3, we introduce our version of the life cycle model, and in section 4, we discuss our estimation procedure. In sections 5 and 6, we describe the data and the estimated shock processes that elderly individuals face. We discuss parameter estimates and model fit in section 7. Section 8 contains some decomposition exercises that gauge the forces affecting saving behavior. In section 9 we develop a version of the model where medical expenses are a choice variable, estimate the model, and use it to perform some robustness checks. We conclude in section 10.

2 Related Literature

Our paper is related to a number of papers in the savings literature that consider either uncertain medical expenditures or bequest motives.

In an early study, Kotlikoff [48] finds that out-of-pocket medical expenditures are potentially an important driver of aggregate saving. However, Kotlikoff also stresses the need for better data on medical expenses and for more realistic modeling of this source of risk.

Subsequent works by Hubbard et al. [38] and Palumbo [56] find that medical expenses have fairly small effects compared to the ones we find. Their effects are smaller because their data understate the extent to which medical expenditures rise with age and income. As an example, the average medical expense for a 100-year-old generated by Hubbard et al.'s medical expenditure model is about 16% of the average medical expense for a 100-year-old found in our data. Our data set contains detailed information for a large number of very old individuals. This richness allows us to provide a more precise picture of how medical expenses rise at very advanced ages. Furthermore, our data have information on nursing home expenses in addition to other forms of medical expenses. Previous studies of medical expense risk had to impute nursing home expenses.

Hubbard et al. [39] and Scholz et al. [62] argue that means-tested social insurance programs (in the form of a minimum consumption floor) provide strong incentives for low income individuals not to save. Their simulations, however, indicate that reducing the consumption floor has little effect on the consumption of college graduates. In contrast, we find that the consumption floor has an effect on saving decisions at all levels of income. Because out-of-pocket medical expenditures rise rapidly with income, rich individuals value social insurance as a safeguard against catastrophic expenses, even if they often end up not using it. This finding is consistent with Brown and Finkelstein's work [10], which finds Medicaid has large effects on the decisions of fairly rich people.

Scholz et al. [62] find that a life cycle model, augmented with realistic income and medical expense uncertainty, does good job of fitting the distribution of wealth at retirement. We add to their paper by showing that a realistic life cycle model can do a good job of fitting the patterns of asset decumulation observed after retirement.

In his seminal paper Hurd [40] estimates a simple structural model of savings and bequest motives in which bequests are normal goods, and does not find support for large bequest motives. De Nardi's [17] calibration exercise shows that modeling bequests as a luxury good is important to explain the savings of the richest few. Kopczuk and Lupton [51] find that a majority of elderly singles have a bequest motive. However, whether the motive is active or not depends on the individual's financial resources because, consistently with De Nardi, they estimate bequests to be luxury goods. While none of the preceding papers accounted for medical expenses, Dynan et al. [24] argue that the same assets can be used to address both precautionary and bequest concerns. Using responses from an attitudinal survey to separate bequest and medical expense motives, Ameriks et al. [2] find that bequests are important for many people. In this paper we allow bequests to be luxury goods, and we let the AHEAD data speak to both the intensity of bequest motives and the level of wealth at which they become operative.

3 The model

Our analysis focuses on people who have retired already. This choice allows us to concentrate on saving and consumption decisions, and to abstract from labor supply and retirement decisions. We restrict our analysis to elderly singles to avoid dealing with household dynamics, such as the transition from two to one family members.

Consider a single person, either male or female, seeking to maximize his or her expected lifetime utility at age t, $t = t_{r+1}, ..., T$, where t_r is the retirement age. These individuals maximize their utility by choosing consumption c. Each period, the individual's utility depends on its consumption and health status, h, which can be either good ($h = 1$) or bad ($h = 0$).

The flow utility from consumption is

$$u(c, h) = \delta(h) \frac{c^{1-\nu}}{1-\nu}, \tag{1}$$

with $\nu \geq 0$. Following Palumbo [56] we model the dependence of utility on health status as

$$\delta(h) = 1 + \delta h, \tag{2}$$

so that when $\delta = 0$, health status does not affect utility.

When the person dies, any remaining assets are left to his or her heirs. We denote with e the estate net of taxes. The utility the household derives from leaving the estate e is

$$\phi(e) = \theta \frac{(e+k)}{1-\nu}^{(1-\nu)}, \tag{3}$$

where θ is the intensity of the bequest motive, while k determines the curvature of the bequest function and hence the extent to which bequests are luxury goods.

We assume that non-asset income y_t, is a deterministic function of sex, g, permanent income, I, and age t:

$$y_t = y(g, I, t). \tag{4}$$

The individual faces several sources of risk, which we treat as exogenous. While this is of course a simplification, we believe that it is a reasonable one, because we focus on older people who have already shaped their health and lifestyle.

1) Health status uncertainty. We allow the transition probabilities for health status to depend on previous health, sex, permanent income and age. The elements of the health status transition matrix are

$$\pi_{j,k,g,I,t} = \Pr(h_{t+1} = k | h_t = j, g, I, t), \quad j, k \in \{1, 0\}. \tag{5}$$

2) Survival uncertainty. Let $s_{g,h,I,t}$ denote the probability that an individual of sex g is alive at age $t+1$, conditional on being alive at age t, having time-t health status h, and enjoying permanent income I.

3) Medical expense uncertainty. Medical expenses, m_t, are defined as out-of-pocket expenses. Since our focus is on understanding the effects of out-of-pocket medical expenses on saving decisions, this version of the model takes medical expenses as exogenous shocks to the household's available resources for our benchmark model, as in Scholz et al. [62], Palumbo [56] and Hubbard et al. [38, 39]. In section 9 we study whether endogenizing medical expenses affects our key results.

We assume that the mean and the variance of the log of medical expenses depend upon sex, health status, permanent income, and age:

$$\ln m_t = m(g, h, I, t) + \sigma(g, h, I, t) \times \psi_t. \tag{6}$$

Following Feenberg and Skinner [28] and French and Jones [34], we assume that the idiosyncratic component ψ_t can be decomposed as

$$\psi_t = \zeta_t + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2), \tag{7}$$
$$\zeta_t = \rho_m \zeta_{t-1} + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2), \tag{8}$$

where ξ_t and ϵ_t are serially and mutually independent. In practice, we discretize ξ and ζ, using quadrature methods described in Tauchen and Hussey [64].

The timing is the following: at the beginning of the period the individual's health status and medical medical expenses are realized. Then the individual consumes and saves. Finally the survival shock hits. Households who die leave any remaining assets to their heirs.

Next period's assets are given by

$$a_{t+1} = a_t + y_n(ra_t + y_t, \tau) + b_t - m_t - c_t, \tag{9}$$

where $y_n(ra_t + y_t, \tau)$ denotes post-tax income, r denotes the risk-free, pre-tax rate of return, the vector τ describes the tax structure, and b_t denotes government transfers.

Assets have to satisfy a borrowing constraint:

$$a_t \geq 0. \tag{10}$$

Following Hubbard et al. [38, 39], we also assume that government transfers provide a consumption floor:

$$b_t = \max\{0, \underline{c} + m_t - [a_t + y_n(ra_t + y_t, \tau)]\}, \tag{11}$$

Equation (11) says that government transfers bridge the gap between an individual's "total resources" (i.e., assets plus income less medical expenses) and the consumption floor. To be consistent with logic of asset and means-tested transfers present in public insurance programs, we impose that if transfers are positive, $c_t = \underline{c}$ and $a_{t+1} = 0$.

To save on state variables we follow Deaton [18] and redefine the problem in terms of cash-on-hand, x_t. Defining cash-on-hand allows us to combine assets and the transitory component of medical expenses into a single state variable,

$$x_t = a_t + y_n(r\, a_t + y_t, \tau) + b_t - m_t. \tag{12}$$

Note that assets and cash-on-hand follow:

$$a_{t+1} = x_t - c_t, \tag{13}$$
$$x_{t+1} = x_t - c_t + y_n\big(r(x_t - c_t) + y_{t+1}, \tau\big) + b_{t+1} - m_{t+1}. \tag{14}$$

To enforce the consumption floor, we impose

$$x_t \geq \underline{c}, \quad \forall t, \tag{15}$$

and to ensure that assets are always non-negative, we require

$$c_t \leq x_t, \quad \forall t. \tag{16}$$

Note that all of the variables in x_t are given and known at the beginning of period t. We can thus write the individual's problem recursively, using the definition of cash-on-hand. Letting β denote the discount factor, the value function for a single individual is given by

$$V_t(x_t, g, h_t, I, \zeta_t) = \max_{c_t, x_{t+1}} \Big\{ u(c_t, h_t) + \beta s_{g,h,I,t} E_t V_{t+1}(x_{t+1}, g, h_{t+1}, I, \zeta_{t+1}) + \beta(1 - s_{g,h,I,t})\phi(e_t) \Big\}, \tag{17}$$

subject to

$$e_t = (x_t - c_t) - \max\{0, \tilde{\tau} \cdot (x_t - c_t - \tilde{x})\}. \tag{18}$$

and equations (14) - (16). The parameter $\tilde{\tau}$ denotes the tax rate on estates in excess of $\tilde{x}$, the estate exemption level.

4 Estimation procedure

4.1 The Method of Simulated Moments

To estimate the model, we adopt a two-step strategy, similar to the one used by Gourinchas and Parker [36], Cagetti [12], and French and Jones [35]. In the first step we estimate or calibrate those parameters that can be cleanly identified without explicitly using our model. For example, we estimate mortality rates from raw demographic data. Let χ denote the collection of these first-step parameters.

In the second step we estimate the vector of parameters $\Delta = (\delta, \nu, \beta, \underline{c}, \theta, k)$ with the method of simulated moments (MSM), taking as given the elements of χ that were estimated in the first step. In particular, we find the vector $\hat{\Delta}$ yielding the simulated life-cycle decision profiles that "best match" (as measured by a GMM criterion function) the profiles from the data. Because our underlying motivations are to explain why elderly individuals retain so many assets, and to explain why individuals with high income save at a higher rate, we match permanent income-conditional age-asset profiles.

Consider individual i of birth cohort p in calendar year t. Note that the individual's age is $t - p$. Let a_{it} denote individual i's assets. Sorting the sample by permanent income, we assign every individual to one of Q quantile-based intervals. In practice, we split the sample into 5 permanent income quintiles, so that $Q = 5$. Suppose that individual i of cohort p falls in the qth interval of the sample income distribution. Let $a_{pqt}(\Delta, \chi)$ be the model-predicted median asset level in calendar year t for an individual of cohort p that was in the qth income interval. Assuming that observed assets have a continuous density, at the "true" parameter vector (Δ_0, χ_0) exactly half of the individuals in group pqt will have asset levels of $a_{pqt}(\Delta_0, \chi_0)$ or less. This leads to a well-known set of moment conditions:[2]

$$E\big(1\{a_{it} \leq a_{pqt}(\Delta_0, \chi_0)\} - 1/2 \,|p, q, t, \text{individual } i \text{ alive at } t\big) = 0, \qquad (19)$$

for each p, q and t triple. In other words, for each permanent income-cohort grouping, the model and the data have the same median asset levels. Our decision to use conditional medians, rather than means, reflects sample size considerations; in some pqt cells, changes in one or two individuals can lead to sizeable changes in mean wealth. Sample size considerations also lead us to combine men and women in a single moment condition.

The mechanics of our MSM approach are fairly standard. We compute life-cycle histories for a large number of artificial individuals. Each of these individuals is endowed with a value of the state vector $(t, x_t, g, h_t, I, \zeta_t)$ drawn from the data distribution for 1996, and each is assigned a series of health, medical expense, and mortality shocks consistent with the stochastic processes described in section 2 above. We give each simulated person the entire health and mortality history realized by a person in the AHEAD data with

[2]See Manski [52], Powell [59] and Buchinsky [11]. Related methodologies are applied in Cagetti [12] and Epple and Sieg [27].

the same initial conditions.[3] Since the data provide health and mortality only during interview years, we simulate it in off-years using our estimated models and Bayes' Rule. The simulated medical expenditure shocks ζ and ξ are Monte Carlo draws from discretized versions of our estimated shock processes.

We discretize the asset grid and, using value function iteration, we solve the model numerically. This yields a set of decision rules, which, in combination with the simulated endowments and shocks, allows us to simulate each individual's assets, medical expenditures, health and mortality. We then compute asset profiles (values of a_{pqt}) from the artificial histories in the same way as we compute them from the real data.

We adjust Δ until the difference between the data and simulated profiles—a GMM criterion function based on equation (19)—is minimized.

We discuss the asymptotic distribution of the parameter estimates, the weighting matrix and the overidentification tests in Appendix A.[4]

The codes solving for the value functions and simulating households' histories are written in C. We use GAUSS for the econometrics. The GAUSS programs call the C programs, send them the necessary inputs - including parameter values and initial values of the state variables - and retrieve the simulated histories. The GAUSS programs then use the simulated histories and the data to compute the GMM criterion function, and/or to produce output items such as graphs and tables.

4.2 Econometric Considerations

In estimating our model, we face two well-known econometric problems. First, in a cross-section, older individuals will have earned their labor income in earlier calendar years than younger ones. Because wages have increased over time (with productivity), this means that older individuals are poorer at every age, and the measured saving profile will overstate asset decumulation over the life cycle. Put differently, even if the elderly do not run down their assets, our data will show that assets decline with age, as older individuals

[3]This approach ensures that the simulated health and mortality processes are fully consistent with the data, even if our parsimonious models of these processes are just an approximation. We are grateful to Michael Hurd for suggesting this approach.

[4]Major theoretical contributions to the method of simulated moments include Pakes and Pollard [55] and Duffie and Singleton [23]. Other useful references on asymptotic theory include Newey [53], Newey and McFadden [54] and Powell [59].

will have lower lifetime incomes and assets at each age. Not accounting for this effect will lead us to estimate a model that overstates the degree to which elderly people run down their assets (Shorrocks [63]).

Second, wealthier people tend to live longer, so that the average survivor in each cohort has higher lifetime income than the average deceased member of that cohort. This "mortality bias" tends to overstate asset growth in an unbalanced panel. In addition, as time passes and people die, the surviving people will be, relative to the deceased, healthy and female. These healthy and female people, knowing that they will live longer, will tend to be more frugal than their deceased counterparts, and hence have a flatter asset profile in retirement. Not accounting for mortality bias will lead us to understate the degree to which elderly people run down their assets.

A major advantage of using a structural approach is that we can address these biases directly, by replicating them in our simulations. We address the first problem by giving our simulated individuals age, wealth, health, gender and income endowments drawn from the distribution observed in the data. If older people have lower lifetime incomes in our data, they will have lower lifetime incomes in our simulations. Furthermore, we match assets at each age, conditional on cohort and income quintile. We address the second problem by allowing mortality to differ with sex, permanent income and health status. As a result our estimated decision rules and our simulated profiles incorporate mortality effects in the same way as the data.

5 Data

The AHEAD is part of the Health and Retirement Survey (HRS) conducted by the University of Michigan. It is a survey of individuals who were non-institutionalized and aged 70 or older in 1994. A total of 8,222 individuals in 6,047 households (in other words, 3,872 singles and 2,175 couples) were interviewed for the AHEAD survey in late 1993/early 1994, which we refer to as 1994. These individuals were interviewed again in 1996, 1998, 2000, 2002, 2004, and 2006. The AHEAD data include a nationally representative core sample as well as additional samples of blacks, Hispanics, and Florida residents.

We consider only single retired individuals in the analysis. This leaves us with 3,259 individuals, of whom 592 are men and 2,667 are women. Of these 3,259 individuals, 884 are still alive in 2006. Appendix B gives more details on the data.

Our measure of net worth (or assets) is the sum of all assets less mortgages and other debts. The AHEAD has information on the value of housing and real estate, autos, liquid assets (which include money market accounts, savings accounts, T-bills, etc.), IRAs, Keoghs, stocks, the value of a farm or business, mutual funds, bonds, and "other" assets. We do not use 1994 assets because they were underreported (Rohwedder et al. [60]).

Non-asset income includes the value of Social Security benefits, defined benefit pension benefits, annuities, veterans benefits, welfare, and food stamps. We measure permanent income (PI) as the individual's average income over all periods during which he or she is observed. Because Social Security benefits and (for the most part) pension benefits are a monotonic function of average lifetime labor income, this provides a reasonable measure of lifetime, or permanent income.[5]

Medical expenses are the sum of what the individual spends out of pocket on insurance premia, drug costs, and costs for hospital, nursing home care, doctor visits, dental visits, and outpatient care. It includes medical expenses during the last year of life. It does not include expenses covered by insurance, either public or private. French and Jones [34] show that the medical expense data in the AHEAD line up with the aggregate statistics. For our sample, mean medical expenses are \$3,712 with a standard deviation of \$13,429 in 1998 dollars. Although this figure is large, it is not surprising, because Medicare did not cover prescription drugs for most of the sample period, requires co-pays for services, and caps the number of reimbursed nursing home and hospital nights.

In addition to constructing moment conditions, we also use the AHEAD data to construct the initial distribution of permanent income, age, sex, health, medical expenses, and cash-on-hand that starts off our simulations. Each simulated individual is given a state vector drawn from the joint distribution of state variables observed in 1996.

[5]Because annuity income often reflects the earnings of a deceased spouse, our measure of permanent income is not so much a measure of the individual's own lifetime income as it is a measure of the income of his or her household.

6 Data profiles and first step estimation results

In this section we describe the life cycle profiles of the stochastic processes (e.g., medical expenditures) that are inputs to our dynamic programming model, and the asset profiles we want our model to explain.

6.1 Asset profiles

We construct the permanent-income-conditional age-asset profiles as follows. We sort individuals into permanent income quintiles, and we track birth-year cohorts. We work with 5 cohorts. The first cohort consists of individuals that were ages 72-76 in 1996; the second cohort contains ages 77-81; the third ages 82-86; the fourth ages 87-91; and the final cohort, for sample size reasons, contains ages 92-102. We use asset data for 6 different years; 1996, 1998, 2000, 2002, 2004 and 2006. To construct the profiles, we calculate cell medians for the survivors for each year assets are observed.

To fix ideas, consider Figure 1, which plots median assets by age and income quintile for the members of two birth-year cohorts that are still alive at each moment in time. The solid lines at the far left of the graph are for the youngest cohort, whose members in 1996 were aged 72-76, with an average age of 74. The dashed set of lines are for the cohort aged 82-86 in 1996.

There are five lines for each cohort because we have split the data into permanent income quintiles. However, the fifth, bottom line is hard to distinguish from the horizontal axis because households in the lowest permanent income quintile hold few assets.

The members of the first cohort appear in our sample at an average age of 74 in 1996. We then observe them in 1998, when they are on average 76 years old, and then again every two years until 2006. The other cohorts start from older initial ages, and are followed for ten years, until 2006. The graph reports median assets for each cohort and permanent-income grouping for six data points over time.

Unsurprisingly, assets turn out to be monotonically increasing in income, so that the bottom line shows median assets in the lowest income quintile, while the top line shows median assets in the top quintile. For example, the top left line shows that for the top PI quintile of the cohort age 74 in 1996, median assets started at $170,000 and then stayed rather stable over time:

$150,000 at age 76, $160,000 at age 78, $180,000 at ages 80 and 82, and $190,000 at age 84.

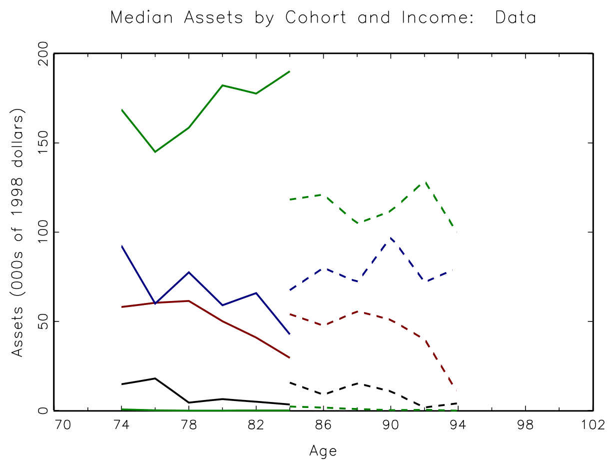

Figure 1: Median assets by cohort and PI quintile: data. Solid line: cohort aged 74 in 1996. Dashed line, cohort aged 85 in 1996.

For all permanent income quintiles in these cohorts, the assets of surviving individuals neither rise rapidly nor decline rapidly with age. If anything, those with high income tend to have increases in their assets, whereas those with low income tend to have declines in assets as they age. The profiles for other cohorts, which are shown in Appendix B, are similar.

Our finding that the income-rich elderly dissave more slowly complements and confirms those by Dynan et al. [25].

Figure 2 compares asset profiles that are aggregated over all the income quintiles. The solid line shows median assets for everyone observed at a given point in time, even if they died in a subsequent wave, i.e., the unbalanced panel. The dashed line shows median assets for the subsample of individuals who were still alive in the final wave, i.e., the balanced panel. It shows that the asset profiles for those that were alive in the final wave—the balanced panel—have much more of a downward slope. The difference between the two sets of profiles confirms that the people who died during our sample period tended to have lower assets than the survivors.

The first pair of lines in Figure 2 shows that failing to account for mortality bias would lead us to understate the asset decumulation of those who

Figure 2: Median assets by birth cohort: everyone in the data (solid lines) vs. survivors (dashed lines).

were 74 years old in 1996 by over 50%. In 1996 median assets of the 74-year-olds who survived to 2006 were \$84,000. In contrast, in 1996 median assets for all 74 year olds alive in that year were \$60,000. Median assets of those in the same cohort who survived to 2006 were \$44,000. The implied drops in median assets between 1996 and 2006 for that cohort are therefore vastly different depending on what population we look at: only \$16,000 if we look at the unbalanced panel, but \$40,000 for the balanced panel of the survivors who made it to 2006. This is consistent with the findings of Love et al. [50]. Sorting the data by permanent income reduces, but does not eliminate, mortality bias.

Since our model explicitly takes mortality bias and differences in permanent income into account, it is the unbalanced panels that we use in our MSM estimation procedure. This greatly increases the size of our estimation sample.

6.2 Medical expense profiles

The mean of logged medical expenses is modeled as a function of: a quartic in age, sex, sex interacted with age, current health status, health status interacted with age, a quadratic in the individual's permanent income ranking, and permanent income ranking interacted with age. We estimate these profiles using a fixed-effects estimator.[6]

We use fixed effects, rather than OLS, for two reasons. First, differential mortality causes the composition of our sample to vary with age. In contrast, we are interested in how medical expenses vary for the same individuals as they grow older. Although conditioning on observables such as permanent income partly overcomes this problem, it may not entirely. The fixed-effects estimator overcomes the problem completely. Second, cohort effects are likely to be important for both of these variables. Failure to account for the fact that younger cohorts have higher average medical expenditures than older cohorts will lead the econometrician to understate the extent to which medical expenses grow with age. Cohort effects are automatically captured in a fixed-effect estimator, as the cohort effect is merely the average fixed effect for all members of that cohort.

The combined variance of the medical expense shocks $(\zeta_t + \xi_t)$ is modeled with the same variables and functional form as the mean (see equation 6).

Our estimates indicate that average medical expenses for men are about 20% lower than for women, conditional on age, health and permanent income. Average medical expenses for healthy people are about 50% lower than for unhealthy people, conditional on age, sex and permanent income. These differences are large, but the differences across permanent income groups are even larger.

To better interpret our estimates, we simulate medical expense histories for the AHEAD birth-year cohort whose members were ages 72-76 (with an average age of 74) in 1996. We begin the simulations with draws from the joint distribution of age, health, permanent income and sex observed in 1996.

Figure 3 presents average medical expenses, conditional on age and permanent income quintile for a balanced sample of people. We rule out attrition in these simulations because it is easier to understand how medical expenses evolve over time when tracking the same individuals. The picture with mortality bias, however, is similar. Permanent income has a large effect on average medical expenses, especially at older ages. Average medical expenses are less than $1,000 a year at age 75, and vary little with income. By age 100, they rise to $2,900 for those in the bottom quintile of the income distribution, and to almost $38,000 for those at the top of the income

[6]Parameter estimates for the data generating process for medical expenses, income, health, and mortality, and a guide to using these data, are available at: http://www.chicagofed.org/economic_research_and_data/economists_preview.cfm?autID=29.

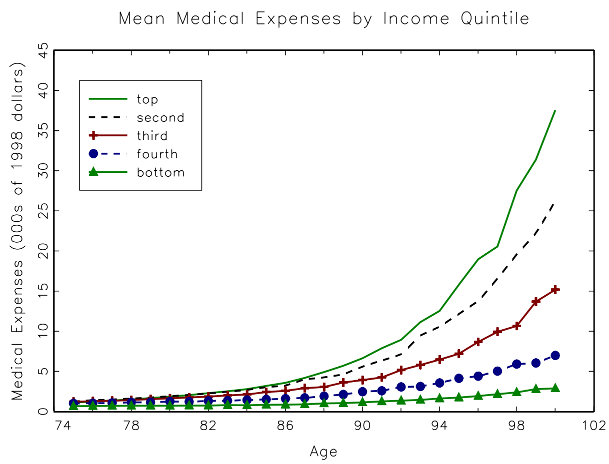

Figure 3: Average medical expenses, by permanent income quintile.

distribution. Mean medical expenses at age 100 are $17,700.

Mean medical expenses implied by our estimated processes line up with the raw data. We have 58 observations on medical expenses for 100-year-old individuals, averaging $15,603 (with a standard deviation of $33,723 and a standard error of $4,428) per year, with 72% of these expenses coming from nursing home care. Between ages 95 and 100, we have 725 person-year observations on medical expenses, averaging $9,227 (with a standard deviation of $19,988 and standard error of $737). Therefore, the data indicate that average medical expenses for the elderly are high.

Medical expenses for the elderly are volatile as well as high. We find that the average variance of log medical expenses is 2.53. This implies that medical expenses for someone with a two standard deviation shock to medical expenses pays 6.8 times the average, conditional on the observables.[7] The variance of medical expenses rises with age, bad health, and income.

French and Jones [34] find that a suitably-constructed lognormal distribution can match average medical expenses and the far right tail of the distribution. They also find that medical expenses are highly correlated over

[7] We assume that medical expenses are log-normally distributed, so the predicted level of medical expenses are $\exp\left(m + \frac{1}{2}\sigma^2\right)$, where m denotes predicted log medical expenses and σ^2 denotes the variance of the idiosyncratic shock ψ_t. The ratio of the level of medical expenses two standard deviations above the mean to average medical expenses is $\frac{\exp(m+2\sigma)}{\exp(m+\sigma^2/2)} = \exp(2\sigma - \sigma^2/2) = 6.80$ if $\sigma = \sqrt{2.53}$.

time. Table 1 shows estimates of the persistent component ζ_t and the transitory component ξ_t. The table shows that 66.5% of the cross sectional variance of medical expenses are from the transitory component, and 33.5% from the persistent component. The persistent component has an autocorrelation coefficient of 0.922, however, so that innovations to the persistent component of medical expenses have long-lived effects. Most of a household's *lifetime* medical expense risk comes from the persistent component.

Parameter	Variable	Estimate (Standard Error)
ρ_m	autocorrelation, persistent component	0.922 (0.010)
σ_ϵ^2	innovation variance, persistent component	0.050 (0.008)
σ_ξ^2	innovation variance, transitory component	0.665 (0.014)

Table 1: Persistence and variance of innovations to medical expenses (variances as fractions of total cross-sectional variance).

Our estimates of medical expense risk indicate greater risk than found in other studies (see Hubbard, et al. [38] and Palumbo [56]). However, our estimates still potentially understate both the level and risk of the medical expenses faced by older Americans, because our measure of medical expenditures does not include the value of Medicaid contributions. As equation (11) shows, some of the medical expenses (m_t) in our model may be paid for by the government through the provision of the consumption floor. Therefore, the ideal measure of m_t drawn from the data would include both the out-of-pocket expenditures actually made by the consumer and the expenditures covered by Medicaid. The AHEAD data, however, do not include Medicaid expenditures. In this respect, the medical expense process we feed into our benchmark model is a conservative one.

6.3 Income profiles

We model mean income in the same way as mean medical expenses, using the same explanatory variables and the same fixed-effects estimator.

Figure 4 presents average non-asset income profiles, conditional on permanent income, computed by simulating our model. For those in the top permanent income quintile, annual income averages $20,000 per year. Fig-

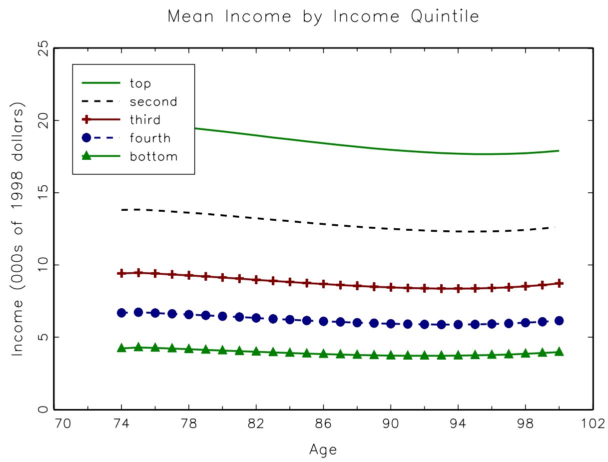

Figure 4: Average income, by permanent income quintile.

ure 1 shows that median wealth for the youngest cohort in this income group is slightly under $200,000, or about 10 years worth of income for this group.

6.4 Mortality and health status

We estimate the probability of death and bad health as logistic functions of a cubic in age; sex; sex interacted with age; previous health status; health status interacted with age; a quadratic in permanent income rank; and permanent income rank interacted with age. A detailed description of our estimates can be found in De Nardi et al. [21].

Using the estimated health transitions, survival probabilities, and the initial joint distribution of age, health, permanent income and sex found in our AHEAD data, we simulate demographic histories. Table 2 presents predicted life expectancies.[8] Rich people, women, and healthy people live much longer than their poor, male, and sick counterparts. Two extremes illustrate this point: an unhealthy male at the bottom quintile of the permanent income

[8] Our predicted life expectancy is lower than what the aggregate statistics imply. In 2002, life expectancy at age 70 was 13.2 years for men and 15.8 years for women, whereas our estimates indicate that life expectancy for men is 9.7 years for men and 14.3 years for women. These differences stem from using data on singles only: when we re-estimate the model for both couples and singles we find that predicted life expectancy is within 1/2 year of the aggregate statistics for both men and women.

distribution expects to live only 6 more years, that is, to age 76. In contrast, a healthy woman at the top quintile of the permanent income distribution expects to live 17 more years, thus making it to age 87. Such significant differences in life expectancy should, all else equal, lead to significant differences in saving behavior. In complementary work, (De Nardi et al. [22]), we show this is in fact the case.

We also find that for rich people the probability of living to very old ages, and thus facing very high medical expenses, is significant. For example, using the same simulations used to construct Table 2, we find that a healthy 70-year-old woman in the top quintile of the permanent income distribution faces a 14% chance of living 25 years, to age 95.

7 Second step estimation results

Table 3 presents preference parameter estimates for several specifications. The first column presents results for a parsimonious model with no bequest motives and no health preference shifter. The second column reports estimates for a model in which health can shift the marginal utility of consumption. In the third column, the bequest motive is activated, and in the final column, both the bequest motive and the preference shifter are active. In all cases, we set the interest rate to 2%.

Table 3 shows that the bequest parameters are never statistically significant and, as shown by the overidentification statistics, have little effect on the model's fit.

When considered in isolation, the health preference parameter is not significant either. However, the final column shows that this parameter is statistically significant when bequest motives are included. An appropriate test, however, is the joint test based on the change in the overidentification statistic (Newey and McFadden [54], section 9). Comparing the first and last columns of Table 3 shows that the test statistic decreases by 4.8, while 3 degrees of freedom are removed. With a $\chi^2(3)$ distribution, this change has a p-value of 18.7%, implying that we cannot reject the parsimonious model.

In short, the bequest and health preference parameters are (collectively) not statistically significant, do not help improve the fit of the model, and, moreover, have little effect on any of the other parameter estimates. We thus use the parsimonious model as our benchmark specification and only briefly discuss the other configurations.

7.1 The benchmark model

The first column of table 3 shows that the estimated coefficient of relative risk aversion is 3.8, the discount factor is .97, and the consumption floor is $2,663. These estimates are within the range of parameter estimates provided in the previous literature.

Our estimated coefficient of relative risk aversion, 3.8, is higher than the coefficients found by fitting non-retiree consumption trajectories, either through Euler equation estimation (e.g., Attanasio et al. [3]) or through the method of simulated moments (Gourinchas and Parker [36]). It is, however, at the lower end of the estimates found by Cagetti [12], who matched wealth profiles with the method of simulated moments over the whole life cycle. It is much lower than those produced by Palumbo [56], who matched consumption data for retirees using maximum likelihood estimation. Given that our out-of-pocket medical expenditure data indicate more risk than that found by Palumbo, it is not surprising that we can match observed precautionary and life-cycle savings with a lower level of risk aversion.

The consumption floor that we estimate ($2,700 in 1998 dollars) is similar to the value that Palumbo [56] uses ($2,000 in 1985 dollars). However, our estimate is about half the size of the value that Hubbard et al. [38] use, and is also about half the average value of Supplemental Security Income (SSI) benefits.

Our consumption floor proxies for Medicaid health insurance (which almost eliminates medical expenses to the financially destitute) and SSI. Given the complexity of these programs, and the fact that many potential recipients do not fully participate in them, it is tricky to establish a priori what the consumption floor should be. Individuals with income (net of medical expenses) below the SSI limit are usually eligible for SSI and Medicaid. However, some individuals with income above the SSI level can receive Medicaid benefits, depending on the state they live in. On the other hand, many eligible individuals do not draw SSI benefits and Medicaid, suggesting that the value of the consumption floor is much lower than the statutory benefits.[9] Our estimates likely provide an "effective" consumption floor, one which combines the complexity and variety of the statutory rules with people's perceptions and attitudes toward welfare eligibility. In appendix C, we show that fixing the consumption floor at $5,000 significantly worsens the model's fit.

[9] For example, Elder and Powers [26] (Table 2), find that less than 50 percent of those eligible for SSI receive benefits.

The Euler equation (20) and asset accumulation equation (9) show that the asset profiles generated by our model depend on expected and realized interest rates. Our parameter estimates, especially those of the discount factor β, therefore depend on our assumptions about the real interest rate r. As a robustness check, we solve a model with i.i.d. interest rate shocks, change the simulations to use the asset returns realized over the 1996-2006 period, and re-estimate the model. Appendix C shows the results. Although the realized returns are on average higher than our benchmark assumption of 2%, and our estimated discount factors accordingly lower, our main findings hold.

The Euler Equation can give some intuition for the estimates and their identification. Ignoring taxes and bequest motives, the Euler Equation is given by

$$(1 + \delta h_t)c_t^{-\nu} = \beta(1 + r)s_t E_t(1 + \delta h_{t+1})c_{t+1}^{-\nu}. \tag{20}$$

Log-linearizing this equation shows that expected consumption growth follows:

$$\begin{aligned} E_t\left(\Delta \ln c_{t+1}\right) &= \frac{1}{\nu}\left[\ln(\beta(1+r)s_t) + \delta E_t(h_{t+1} - h_t)\right] \\ &\quad + \frac{\nu+1}{2} Var_t(\Delta \ln c_{t+1}). \end{aligned} \tag{21}$$

The coefficient of relative risk aversion is identified by differences in saving rates across the income distribution, in combination with the consumption floor. Low income households are relatively more protected by the consumption floor, and will thus have lower values of $Var_t(\Delta \ln c_{t+1})$ and thus weaker precautionary motives. The parameter ν helps the model explain why individuals with high permanent income typically display less asset decumulation. Appendix C discusses the identification of the coefficient of relative risk aversion and the consumption floor in more detail.

Figure 5 shows how the baseline model fits a subset of the data profiles, using unbalanced panels. (The model fits equally well for the cells that are not shown.) Both in the model and in the observed data individuals with high permanent income tend to increase their wealth with age, whereas individuals with low income tend to run down their wealth with age.

The visual evidence shown in Figure 5 is consistent with the test statistics shown in Table 3. The p-value of the overidentification statistic for our baseline specification is 87.3%. Hence, our model is not rejected by the overidentification test at any standard level of significance. Both in the model

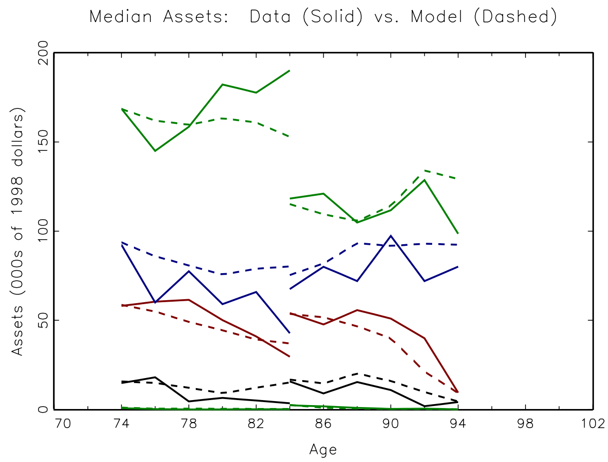

Figure 5: Median assets by cohort and PI quintile: data and model.

and the data, individuals with high permanent income do not run down their wealth with age, whereas those with low income do.

Turning to the mortality bias, Figure 6 shows simulated asset profiles, first for all simulated individuals alive at each date, and then for the individuals surviving the entire simulation period. As in the data, restricting the profiles to long-term survivors reveals much more asset decumulation. The mortality bias generated by the model is large, reflecting heterogeneity in both saving behavior and mortality patterns.

Given that the survival rate, s_t, is often much less than 1, it follows from equation (21) that the model will generate downward-sloping, rather than flat, consumption profiles, unless the discount factor β is fairly large. Figure 7 shows simulated consumption profiles for ages 74-100. Except for the last two years of life, consumption falls with age. This general tendency is consistent with many empirical studies of older-age consumption. For example, Fernandez-Villaverde and Krueger [29] find that non-durable consumption declines about one percent per year between ages 70 and 90. (Also see Banks, Blundell, and Tanner [6].)

7.2 The model with health-dependent preferences

The second and fourth columns of Table 3 show point estimates of $\delta = -.21$ or $\delta = -.36$: holding consumption fixed, being in good health lowers

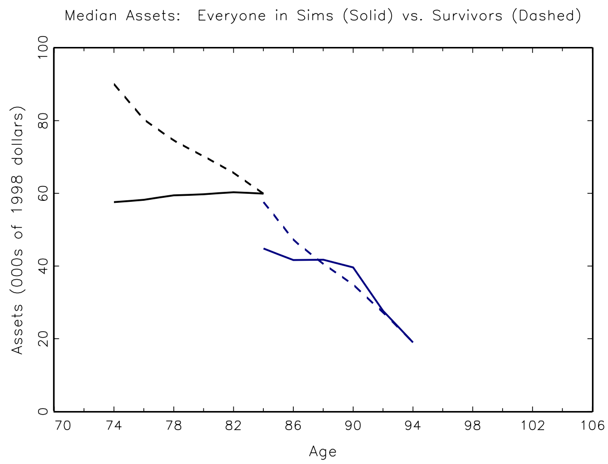

Figure 6: Median assets by birth cohort: everyone in the simulations vs. survivors.

the marginal utility of consumption by 21-36%. This implies that an anticipated change from good to bad health leads consumption to increase by 6 to 10%, depending on the specification (see equation (21)). Previous empirical studies disagree on whether δ is greater than or less than 0. (See Lillard and Weiss [49], Rust and Phelan [61], Viscusi and Evans [65] and Finkelstein et al. [30] as examples.)

This parameter, however, is not statistically significant (if considered jointly with bequests), and none of the other parameter estimates are affected by its inclusion.

7.3 The model with bequest motives

The third and fourth columns of Table 3 show parameter estimates for models that include a bequest motive. Because the two specifications deliver similar parameter estimates, we focus on the results in the third column.

The point estimates of $\theta = 2,360$ and $k = 273,000$ indicate that the consumption level above which the bequest motive becomes operative is about \$36,000 per year. (See Appendix D for a derivation.) By way of comparison, individuals in the top permanent income quintile have an average annuity income of about \$20,000 and hold less than \$200,000 of assets. This suggests that most people in our sample do not have a strong bequest motive.

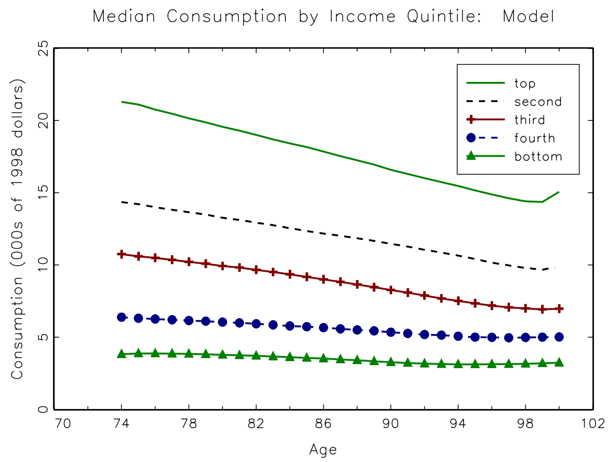

Figure 7: Median consumption by PI quintile: simulations.

Not surprisingly, we find that bequest motives are not very important for fitting our data; none of the estimated bequest parameters are statistically significant, and adding bequests does not significantly improve the model's fit.

For those sufficiently rich, however, the marginal propensity to bequeath above that consumption level is also high, with 88 cents of every extra dollar above the threshold being left as bequests. (See Appendix D.) Hence one can interpret our estimates as suggesting that the bequest motive could be present for the richest people in our sample.

To show the model's implications for bequests, Figure 8 displays the distribution of assets that individuals hold one period before their deaths.[10] Comparing the two panels of Figure 8 highlights that the models with and without the estimated bequest motive generate very similar distributions of bequests. Both versions of the model also do a good job of matching the distribution of assets before death found in the data, although both versions

[10]When AHEAD respondents die, their descendants are asked about the value of the estate. However, problems of non-response are severe – 49% of all estate values are imputed. Furthermore, it is unclear whether reported estates also include the value of the deceased individual's home. For these reasons we report the distribution of assets one year before death rather than estates. Estates are somewhat lower than assets one year before death; mean and median assets 1 year before death are $162,000 and $37,000, whereas mean and median estates are $132,000 and $20,000.

slightly under-estimate the probability of having less than \$10,000 in the year before death.

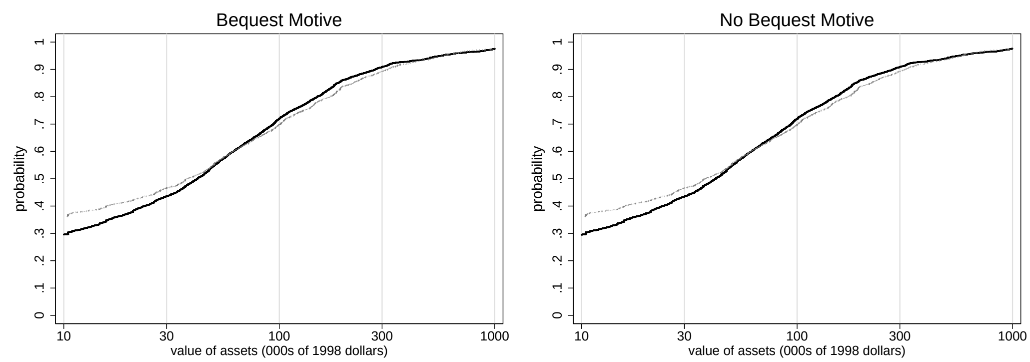

Figure 8: Cumulative distribution function of assets held 1 period before death: data and model. Model with bequest motive in the left panel, and without on the right panel. Legend: solid line is model, lighter line is data.

Our findings should not be interpreted as a rejection of bequest motives in general. Our sample is composed of elderly singles, who are poorer than couples (see for example Díaz-Giménez et al. [20]), and other evidence indicates that bequests are a luxury good (De Nardi [17], and Dynan et al. [25]). Our sample of singles may not contain enough rich households to reveal strong bequest motives. Moreover, and importantly, a significant fraction of our sample is composed of people who have already lost their partner, and it is possible that some of the estate was already split between the surviving spouse and other heirs.

7.4 Summary

Our main results from the second step estimation can be summarized as follows. First, our estimates of the coefficient of relative risk aversion, time discount factor, and consumption floor are within the range used in the previous literature. Second, we do not find that health-dependent preferences are important for understanding retirees' saving behavior. Third, we do not find that bequest motives significantly affect the savings of most households in our sample. Fourth, the model fits the data closely, both in terms of the moment conditions that we match, and in terms of bequests distribution that it generates. Finally, the model's consumption implications are consistent with previous empirical evidence. Put together, these findings give us confidence that we can use our benchmark model (without bequests and preference shocks) to study how savings depend on medical expenses and the consumption floor.

8 What are the important determinants of savings?

To determine the importance of the key mechanisms in our model we fix the estimated parameters at their benchmark values and then change one feature of the model at a time. For each of these different economic environments we compute the optimal saving decisions, simulate the model, and compare the resulting asset accumulation profiles to the asset profiles generated by the baseline model.

We display asset profiles for the AHEAD birth-year cohort whose members were ages 72-76 (with an average age of 74) in 1996. To focus on underlying changes in saving, we rule out attrition and assume every individual lives to age 100. People will face the same risks and have the same expectations as before (except for the particular aspect changed in each experiment), but in their realized lives they do not die and drop out of the simulations until age 100. None of our conclusions would change if we were to also allow for mortality bias.

First, we ask whether the out-of-pocket medical expenditures that we estimate from the data are important drivers of old age savings. To answer this question, we zero out all medical out-of-pocket medical expenditures for everyone and look at the corresponding profiles. This could be seen as an extreme form of insurance provided by the government.

Figure 9 shows that medical expenses are a big determinant of the elderly people's saving behavior, especially for those with high permanent income, for whom those expenses are especially high, and who are relatively less insured by the government-provided consumption floor. These retirees are reducing their current consumption in order to pay for the high out-of-pocket medical expenses they expect to bear later in life. For given initial wealth, if there were no out-of-pocket medical expenses, individuals in the highest

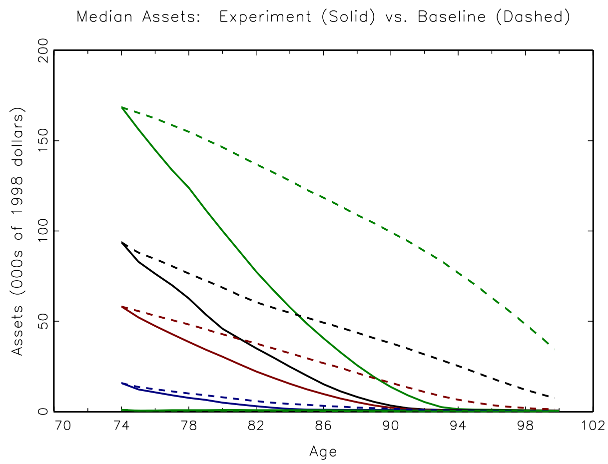

Figure 9: Median assets by cohort and PI quintile: baseline and model with no medical expenses.

permanent-income quintile would deplete their net worth by age 94. In the baseline model with medical expenses they keep almost $40,000 to pay for out-of-pocket medical expenses in the last few years of life. The risk of living long and having high medical expenses late in life increases savings. This decomposition indicates that modeling out-of-pocket medical expenses is important in evaluating policy proposals that affect the elderly.

We next shut down out-of-pocket medical expense risk (the shocks ζ and ξ), while keeping average medical expenditure constant (conditional on all of the relevant state variables). Figure 10 shows the results. Interestingly, and consistently with Hubbard, Skinner and Zeldes [38], we find that, conditional on average medical expenses, the risk associated with the volatility of medical expenses has only a small effect on the profiles of median wealth. Our results are also consistent with Palumbo's [56] finding that eliminating medical expense risk has only small effects on consumption and assets.[11]

One reason why medical expense risk might not have a large effect is that

[11]Figure 10 shows that eliminating medical expense risk sometimes causes assets to increase. One reason for the increase is that when the variance of medical expenses is reduced, the frequency of large medical expenses is also reduced, which in turn reduces the fraction of medical expenses covered by Medicaid. Eliminating volatility in m_t can thus raise the average medical expense borne by retirees. Because this cost increase will be highest at oldest ages, individuals will respond by accumulating more assets.

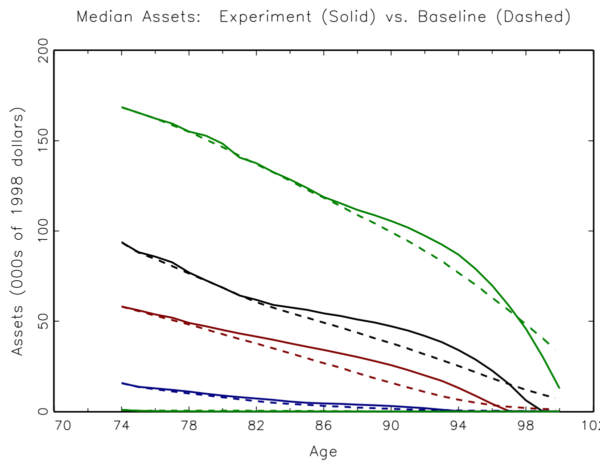

Figure 10: Median assets by cohort and PI quintile: baseline and model with no medical expense risk.

the consumption floor limits the effects of catastrophic medical expenses. To explore this effect further, we reduce the consumption floor to 80% of its value, that is from \$2,663 to \$2,100. One could interpret this as a reform reducing the government-provided consumption safety net. The effects of this change are quite evident. Individuals respond to the increase in net income uncertainty by keeping more assets to self-insure. Figure 11 shows that reducing the consumption floor affects the savings profiles of both low- and high-permanent-income singles. The assets of the 90 year old in the highest permanent income class increase from \$100,000 to \$120,000, while those of the people of the same age but in the second highest income quintile increase from \$40,000 to \$50,000. The net worth of those in the third and fourth income quintiles also displays some increases. The consumption floor thus matters for wealthy individuals as well as poor ones. This is perhaps unsurprising given the size of our estimated medical expenses for the old and income-rich; even wealthy households can be financially decimated by medical expenses in very old age.

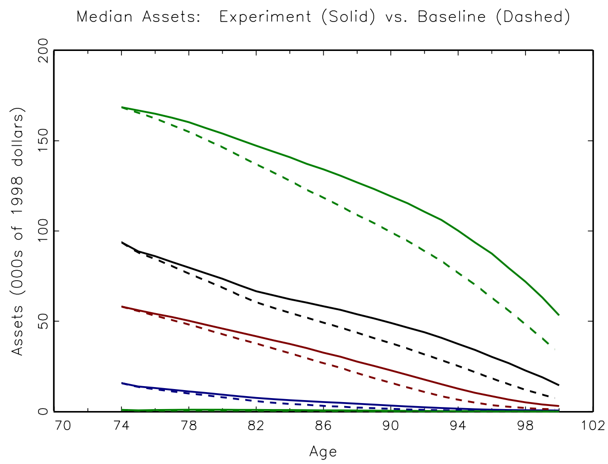

Figure 11: Median assets by cohort and PI quintile: baseline and model with 80% of the consumption floor.

9 Endogenous medical spending

To check the sensitivity of our findings to the assumption that medical expenses are exogenous, we consider a more complex model in which retirees optimally choose how much to spend on medical goods and services, as well as on non-medical consumption. Our findings are robust to this extension.

We assume that retirees derive utility from consumption of both non-medical and medical goods, with the relative weights on the two goods varying with age, health, and an idiosyncratic "medical needs" shock.

A complementary approach is that of Grossman [37], in which medical expenses represent investments in health capital, which in turn decreases mortality. While this is an appealing mechanism, the existing empirical literature suggests that these effects are particularly small for the U.S elderly for two reasons. First, the expenditures that we are considering are supplementing Medicaid, Medicare, and insurance-provided medical goods and services, which cover most life-threatening conditions. Second, the stock of health carried by an elderly person is in large part determined by the health investments that were made in the past, including those made by the person's parents in his or her childhood and even before birth. Hence, for our sample of people aged 70 and older, the effects of additional health investments are not as large as, for example, in infancy.

A key piece of empirical evidence comes from the RAND Health Insurance Experiment, where a random set of individuals were given co-payment-free health insurance over a 3- to 5-year period, while a control group faced standard co-payments. Brook et al. [9] found that even though the group with free health care utilized medical services much more intensively than the control group, the additional treatments had only a "minimal influence" on subsequent health outcomes.

Surprisingly, some empirical studies show that even programs such as Medicare, which sometimes help pay for critical treatments, do not significantly increase life expectancy. For example, Finkelstein and McKnight [31] "find no compelling evidence that, in its first 10 years, Medicare reduced overall elderly mortality." They also note that, more in general, the literature on the effects of health insurance "points strongly to no or only very modest health benefits." Fisher et al. [32] study regional variations in Medicare spending, and conclude that individuals in high-spending regions "do not have better health outcomes." More in line with what one might expect, Card et al. [13] find that Medicare caused a small reduction in mortality among 65-year-olds admitted through emergency rooms for "non-deferrable" conditions.

For these reasons some recent structural models have supplemented Grossman's mechanism with the direct utility effects that we use. For example Khwaja [45] finds that "medical utilization would only decline by less than 20% over the life cycle if medical care was purely mitigative and had no curative or preventive components." (Also see Blau and Gilleskie [8], and Davis [16].)

Given that the existing evidence indicates that the effect of additional medical spending on life expectancy is small (and especially so for the elderly) and that allowing for this additional effect would complicate the model and require the estimation of many additional parameters, we focus on the utility effects of medical expenditures.

9.1 The endogenous medical expenditure model

At the beginning of each period the individual's health status and medical needs shocks are realized and need-based transfers are given. The individual then chooses consumption, medical expenditure, and saves. Finally the survival shock hits. If the person dies, the estate is passed on to one's heir.

The flow utility from consumption is given by

$$u(c_t, m_t, h_t, \zeta_t, \xi_t, t) = \frac{1}{1-\nu} c_t^{1-\nu} + \mu(t, h_t, \zeta_t, \xi_t)\, \frac{1}{1-\omega} m_t^{1-\omega}, \tag{22}$$

where t is age, c_t is consumption of non-medical goods, m_t is total consumption of medical goods, h_t is health status, and $\mu(\cdot)$ is the medical needs shock, which affects the marginal utility of consuming medical goods and services. The consumption of both types of goods is expressed in dollar values. The substitution elasticities for the two goods, however, can differ.

As before, we allow the need for medical services to have a temporary (ξ_t) and a persistent (ζ_t) component; we recycle the variable names to save on notation. We assume that these shocks follow the same processes as in equations (7) and (8), with potentially different parameters. It is worth stressing that we not allow any of these shocks to depend on permanent income; income affects medical expenditures solely through the budget constraint.

We model two important features of the health care system:

1. Private and public insurance pay the share $1 - q(t, h_t)$ of the total medical costs incurred by the retiree. Its complement, $q(t, h_t)$, is the out-of-pocket share paid by the retiree. We estimate $q(t, h_t)$ as part of our first-stage estimation. (See Appendix B for details.)

2. Social insurance, such as Medicaid and SSI, provide monetary transfers that vary with financial resources and medical needs. We model these transfers as providing a flow *utility* floor. The transfers thus depend on the retirees' state variables, not least their medical needs shocks. For a given utility floor and state vector, we find the transfer $b^*(\cdot) = b^*(t, a_t, g, h_t, I, \zeta_t, \xi_t)$ that puts each retiree's utility at the floor. Transfers then kick in to provide the minimum utility level to retirees who lack the resources to afford it

$$b(t, a_t, g, h_t, I, \zeta_t, \xi_t) = \max\{0, b^*(t, a_t, g, h_t, I, \zeta_t, \xi_t)\}. \tag{23}$$

   As before, we impose that if transfers are positive, the individual consumes all of his resources (by splitting them optimally between the two goods), so that $a_{t+1} = 0$.

The retiree's value function is given by

$$V(t,a_t,g,h_t,I,\zeta_t,\xi_t) = \max_{c_t,m_t,a_{t+1}} \Bigg\{ \frac{1}{1-\nu} c_t^{1-\nu} + \mu(t,h_t,\zeta_t,\xi_t)\, \frac{1}{1-\omega} m_t^{1-\omega} + \beta s_{g,h,I,t} E_t \Big(V(t+1,a_{t+1},g,h_{t+1},I,\zeta_{t+1},\xi_{t+1}) \Big) \Bigg\}, \tag{24}$$

subject to equations (23), (18), (10), and

$$a_{t+1} = a_t + y_n(ra_t + y_t) + b(t,a_t,g,h_t,I,\zeta_t,\xi_t) - c_t - m_t q(t,h_t). \tag{25}$$

9.2 Estimation

The log of the medical needs shifter μ_t is modeled as a function of: a cubic in age; current health status; and health status interacted with age. The combined variance of the shocks $\zeta_t + \xi_t$ is modelled as a quadratic in age; current health status; and health status interacted with age. To identify these parameters, we expand the moment conditions described by equation (19) to include moments relating to: mean medical expenses by age and birth cohort, for each half of the permanent income distribution; the 90th percentile of medical expenses in the same cells; and the first and second autocorrelations for medical expenses in each cell.[12] Detailed moment conditions can be found in Appendix A. In all other respects our MSM procedure is the same as before.

9.3 Results

The preference parameter estimates for the endogenous medical expenditure model are shown in Appendix C. The new estimate of β, the discount factor is 0.99, is slightly higher than the benchmark estimate of 0.97. The estimate of ν, the coefficient of relative risk aversion for "regular" consumption is 2.15, while the estimate of ω, the coefficient of relative risk aversion

[12]Since the AHEAD medical expenses data are reported net of any Medicaid payment, we net out government transfers(b_t) from the medical expenses generated by the model. Hence, AHEAD medical expenses and model-generated ones are both net of Medicaid payments.

for medical goods is 3.19, so that the demand for medical goods is less elastic than the demand for consumption. Both coefficients are similar to, but somewhat smaller than, the benchmark estimate of 3.82.

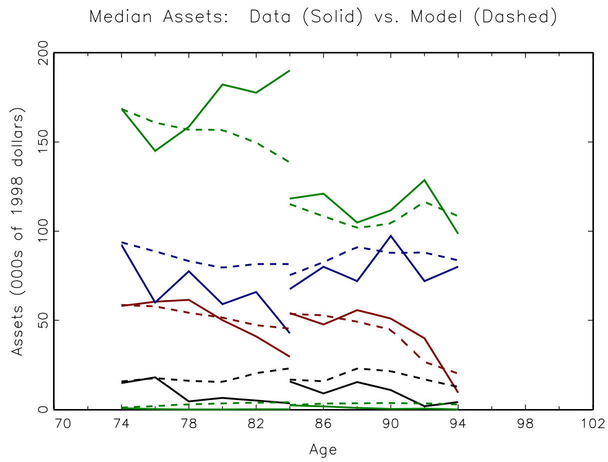

Figure 12: Median assets by cohort and PI quintile: data and endogenous medical expenditure model.

The model also requires parameter estimates for the mean of the logged medical needs shifter $\mu(t, h_t, \zeta_t, \xi_t)$ and the process for the shocks ζ_t and ξ_t. The estimates for these parameters (available from the authors on request) show that the demand for medical services rises rapidly with age.

Although the overidentification test statistic shows that the model with endogenous medical expenditure is rejected by the data, the model does match the main patterns of the asset profiles (see Figure 12). Furthermore, the model fits the medical expense distribution rather well: see Appendices B and C.

Out-of-pocket (and total) medical expenses in the endogenous medical expenditure model can be eliminated by setting the medical needs parameter μ_t to zero. Figure 13 shows that the effects of eliminating medical expenses are similar to those found in the exogenous medical expense model (Figure 9). Given that the consumer must prepare for the same pattern of medical expenditures in either model, this is not surprising. Retirees will save for high medical expenditures at old ages whether the expenditures are exogenous shocks estimated from the data or medical expenditure choices consistent with the same data.

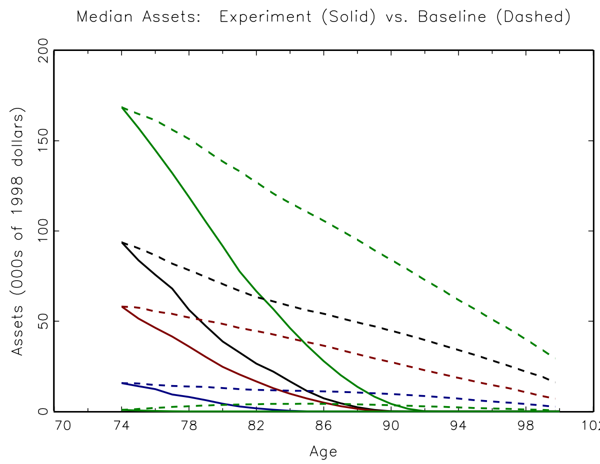

Figure 13: Median assets: baseline endogenous medical expense model, with and without medical expenses.

Figure 14 shows the effects of reducing the generosity of social insurance by 50% for both the exogenous medical expenses model and the endogenous one with a utility floor.[13] As in the exogenous medical expenditure model, a change in social insurance affects the savings of all income groups, including the richest. The effects are of course smaller, because retirees in the endogenous medical expense model can adjust medical expenditures as well as consumption.

In sum, the endogenous medical expense model confirms and reinforces our conclusion that medical expenses are a major savings motive, and that social insurance affects the saving of the income-rich as well as that of the income-poor. Our main findings appear thus robust to the way in which we model the medical expenditure decision.

10 Conclusions

In this paper, we construct, estimate, and analyze a rich model of saving for retired single people. In doing so, we provide several contributions.

[13]In the endogenous medical expense model, we index the estimated utility floor by the consumption level that provides the floor when $\mu = 1$. To run the counterfactual, we cut that consumption level in half.

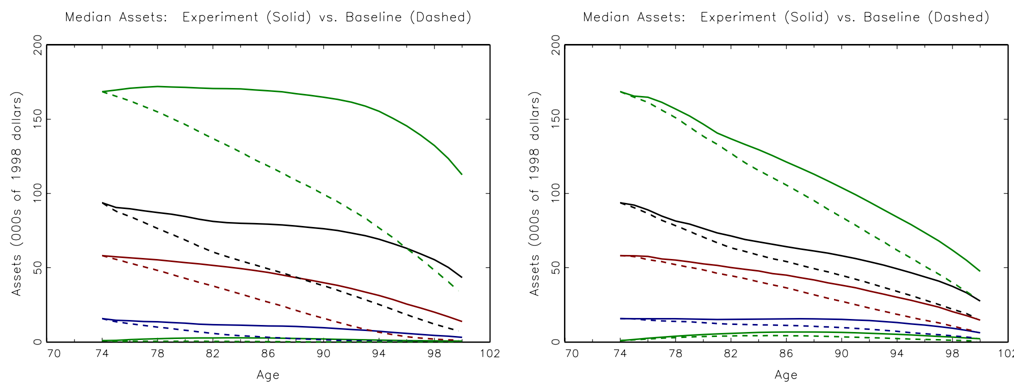

Figure 14: Median assets: baseline and model with 50% of the consumption floor for the exogenous (left panel) and endogenous (right panel) medical expense models.

First, we estimate the out-of-pocket medical expenses faced by the elderly using a larger data set (that includes nursing home expenses) and more flexible functional forms than in the previous literature. As a result, we find that medical expenses are much higher and more volatile than previously estimated, that they rise very fast with age. Also, at very advanced ages (that is starting from about age 80), medical expenses are very much a luxury good; i.e., they are much higher for individuals with higher permanent income.

Second, we estimate mortality probabilities by age as a function of health, sex, and permanent income and find large variations along all three dimensions. We find that in an unbalanced panel, mortality bias—the tendency of rich people to live longer—is significant.

Third, we construct and estimate a structural model of saving using the method of simulated moments. As a result of our improved data and richer specifications, we find that not only our parameter estimates are consistent with the existing estimates from the literature, but also that our model can explain many important features of the data, unlike previously found. In particular, our estimated structural model fits saving profiles across the entire income distribution, and reproduces the observation that elderly people with high permanent income have a smaller dissaving rate than elderly people with low permanent income.

Our model and estimates imply that the pattern of out-of-pocket medical expenses by age and permanent income is a key determinant of savings. If single people live to very advanced ages, they are almost sure to need very expensive medical care, and they thus choose to keep a large amount of assets (an amount increasing in permanent income, as medical expenses also increase) to self-insure against this risk. We also find that a publicly-provided consumption floor has a large effect on the asset profiles of all people, even those with high income. Our findings are robust to endogenizing medical expenditures in an empirically realistic way.

In short, we find that out-of-pocket medical expenditures, and the way in which they interact with the consumption floor, go a long way toward explaining the elderlys' saving decisions and are very important elements that should be included in models evaluating old-age policy reforms.

Following up on our findings, several recent papers have incorporated medical expenses and government insurance programs in general equilibrium, heterogeneous agent, life-cycle settings, and used these models to study the effects of various policy reforms. Kopecky and Koreshkova [47] find that nursing home expenses are a major savings motive for the wealthy, and estimate that Medicaid and other social insurance programs for the elderly crowd out a large fraction of savings. Imrohoroglu and Kitao [42], Jeske and Kitao [43] and Attanasio, et al. [5] use similar models to study the effects of Social Security reforms, tax policies on employer-provided health insurance and Medicare, respectively. Paschenko [57] adopts a model similar to ours to study why few retirees purchase annuities.

Applying our findings to policy evaluation also highlights the value of extending our analysis, which is currently for elderly singles, to elderly couples. One way in which couples may differ from singles is in the importance of bequests. Our estimates imply that while bequest motives are potentially quite important for the wealthy, for most of our sample they have relatively little effect: in our data, median net worth in the top permanent income quintile is less than \$200,000. The parameter values for this bequest motive are imprecisely estimated as well, probably because we do not have enough rich people in our sample. Bequest motives may well play a larger role in married households, who are generally richer than single households, and who probably wish to ensure that if one spouse dies the surviving spouse is financially secure. Because bequests and out-of-pocket medical expenditures are both luxury goods, pinning down their relative contributions will be challenging as well as important.

References

[1] Joseph G. Altonji and Lewis M. Segal. Small sample bias in gmm estimation of covariance structures. *Journal of Business and Economic Statistics*, 14, 1996.

[2] John Ameriks, Andrew Caplin, Steven Laufer, and Stijn Van Nieuwerburgh. The joy of giving or assisted living? using strategic surveys to separate bequest and precautionary motives. NBER Working Paper 13105, 2007.

[3] Orazio P. Attanasio, James Banks, Costas Meghir, and Guglielmo Weber. Humps and bumps in lifetime consumption. *Journal of Business and Economic Statistics*, 17(1):22–35, 1999.

[4] Orazio P. Attanasio and Carl Emmerson. Mortality, health status, and wealth. *Journal of the European Economic Association*, 1(4):821–850, 2003.

[5] Orazio P. Attanasio, Sagiri Kitao, and Gianluca Violante. Financing medicare: A general equilibrium analysis. In John Shoven, editor, *Demography and the Economy*. University of Chicago Press.

[6] James Banks, Richard Blundell, and Sarah Tanner. Is there a retirement-savings puzzle? *The American Economic Review*, 88:769–788, 1998.

[7] David M. Blau. Retirement and consumption in a life cycle model. *Journal of Labor Economics*, 26(1):35–71, 2008.

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

[9] Robert Brook, John E. Ware, William H. Rogers, Emmett B. Keeler, Allyson R. Davies, Cathy A. Donald, George A. Goldberg, Kathleen N. Lohr, Patricia C. Masthay, and Joseph P. Newhouse. Does free care improve adults' health? results from a randomized trial. *New England Journal of Medicine*, 309(23):1426–1434, 1983.

[10] Jeff Brown and Amy Finkelstein. The interaction of public and private insurance: Medicaid and the long term care insurance market. *American Economic Review, forthcoming*, 2009.

[11] Moshe Buchinsky. Recent advances in quantile regression models: A practical guideline for empirical research. *Journal of Human Resources*, 33:88–126, 1998.

[12] Marco Cagetti. Wealth accumulation over the life cycle and precautionary savings. *Journal of Business and Economic Statistics*, 21(3):339–353, 2003.

[13] David Card, Carlos Dobkin, and Nicole Maestas. Does medicare save lives? NBER Working Paper 13668, 2007.

[14] Gary Chamberlain. Comment: Sequential moment restrictions in panel data. *Journal of Business & Economic Statistics,*, 10(1):20–26, 1992.

[15] James B. Davies and Antony F. Shorrocks. The distribution of wealth. In Anthony B. Atkinson and François Bourguignon, editors, *Handbook of Income Distribution*, pages 605–675. Handbooks in Economics, vol. 16. Amsterdam; New York and Oxford: Elsevier Science, North-Holland, 2000.

[16] Morris Davis. The insurance, health, and savings decisions of elderly women living alone. Working Paper, 2005.

[17] Mariacristina De Nardi. Wealth inequality and intergenerational links. *Review of Economic Studies*, 71(3):743–768, 2004.

[18] Angus Deaton. Saving and liquidity constraints. *Econometrica*, 59(5):1221–1248, 1991.

[19] Angus Deaton and Christina Paxson. Mortality, education, income, and inequality among american cohorts. In David A. Wise, editor, *Themes in the Economics of Aging*, pages 129–165. NBER Conference Report Series. Chicago and London: University of Chicago Press, 2001.

[20] Javier Díaz-Giménez, Vincenzo Quadrini, and José-Víctor Ríos-Rull. Dimensions of inequality: Facts on the U.S. distributions of earnings, income and wealth. *Federal Reserve Bank of Minneapolis Quarterly Review*, 21(2):3–21, 1997.

[21] Mariacristina De Nardi, Eric French, and John B. Jones. Differential mortality, uncertain medical expenses, and the saving of elderly singles. Working Paper 12554, National Bureau of Economic Research, 2006.

[22] Mariacristina De Nardi, Eric French, and John B. Jones. Life expectancy and old age savings. *American Economic Review: Papers and Proceedings*, 99(2):110–115, 2009.

[23] Darrell Duffie and Kenneth J. Singleton. Simulated moments estimation of markov models of asset prices. *Econometrica*, 61(4):929–952, 1993.

[24] Karen E. Dynan, Jonathan Skinner, and Stephen P. Zeldes. The importance of bequests and life-cycle saving in capital accumulation: A new answer. *American Economic Review*, 92(2):274–278, 2002.

[25] Karen E. Dynan, Jonathan Skinner, and Stephen P. Zeldes. Do the rich save more? *Journal of Political Economy*, 112(2):397–444, 2004.

[26] Todd E. Elder and Elizabeth T. Powers. Public health insurance and ssi program participation among the aged. Working Paper 2006-117, Michigan Retirement Research Center, 2006.

[27] Dennis Epple and Holger Seig. Estimating equilibrium models of local jurisdictions. *Journal of Political Economy*, 107(4):645–681, 1999.

[28] Daniel Feenberg and Jonathan Skinner. The risk and duration of catastrophic health care expenditures. *Review of Economics and Statistics*, 76:633–647, 1994.

[29] Jesus Fernandez-Villaverde and Dirk Krueger. Consumption over the life cycle: Facts from consumer expenditure survey data. 2004.

[30] Amy Finkelstein, Erzo F. P. Luttmer, and Matthew J. Notowidigdo. What good is wealth without health? the effect of health on the marginal utility of consumption. NBER Working Paper 14089, 2008.

[31] Amy Finkelstein and Robin McKnight. What did medicare do (and was it worth it)? NBER Working Paper 11609, 2005.

[32] Elliott S. Fisher, David E. Wennberg, Therese A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Etoile L. Pinder. The implications of regional

variations in medicare spending. part 2: Health outcomes and satisfaction with care. *Annals of Internal Medicine*, 138(4):288–322, 2003.

[33] Eric French, Phil Doctor, and Olesya Baker. Asset rundown after retirement: The importance of rate of return shocks. *Economic Perspectives*, 31(2):48–65, 2007.

[34] Eric French and John Bailey Jones. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics*, 19(4):705–721, 2004.

[35] Eric French and John Bailey Jones. The effects of health insurance and self-insurance on retirement behavior. Michigan Retirement Research Center Working Paper 2007-170, 2007.

[36] Pierre-Olivier Gourinchas and Jonathan A. Parker. Consumption over the life cycle. *Econometrica*, 70(1):47–89, 2002.

[37] Michael Grossman. On the concept of health capital and the demand for health. *Journal of Political Economy*, 80(2):223–255, 1972.

[38] R. Glenn Hubbard, Jonathan Skinner, and Stephen P. Zeldes. The importance of precautionary motives in explaining individual and aggregate saving. *Carnegie Rochester Series on Public Policy*, pages 59–125, 1994.

[39] R. Glenn Hubbard, Jonathan Skinner, and Stephen P. Zeldes. Precautionary saving and social insurance. *Journal of Political Economy*, 103(2):360–399, 1995.

[40] Michael D. Hurd. Mortality risk and bequests. *Econometrica*, 57(4):779–813, 1989.

[41] Michael D. Hurd, Daniel McFadden, and Angela Merrill. Predictors of mortality among the elderly. Working Paper 7440, National Bureau of Economic Research, 1999.

[42] Selahattin İmrohoroğlu and Sagiri Kitao. Social security, benefit claiming and labor force participation: A quantitative general equilibrium approach. Mimeo, 2009.

[43] Karsten Jeske and Sagiri Kitao. U.s. tax policy and health insurance demand: Can a regressive policy improve welfare? *Journal of Monetary Economics*, 56(2):210–221, 2009.

[44] Thomas F. Juster, James P. Smith, and Frank Stafford. The measurement and structure of household wealth. Mimeo, 1998.

[45] Ahmed Khwaja. Estimating willingness to pay for medicare using a dynamic life-cycle model of demand for health insurance. *Journal of Econometrics, forthcoming*, 2009.

[46] Ruud H. Koning. Kernel: A gauss library for kernel estimation, 1996. http://www.xs4all.nl/

[47] Karen Kopecki and Tatyana Koreshkova. The impact of medical and nursing home expenses and social insurance policies on savings and inequality. Mimeo, 2009.

[48] Laurence J. Kotlikoff. Health expenditures and precautionary savings. In Laurence J. Kotlikoff, editor, *What Determines Saving?* Cambridge, MIT Press, 1988, 1988.

[49] Lee Lillard and Yoram Weiss. Uncertain health and survival: Effect on end-of-life consumption. *Journal of Business and Economic Statistics*, 15(2):254–268, 1996.

[50] David A. Love, Michael Palumbo, and Paul Smith. The trajectory of wealth in retirement. *Journal of Public Economics*, 93:191–208, 2009.

[51] Joseph Lupton and Wojciech Kopczuch. To leave or not to leave: The distribution of bequest motives. *Review of Economic Studies*, 74, 2007.

[52] Charles F. Manski. *Analog Estimation Methods in Econometrics*. Chapman and Hall, 1988.

[53] Whitney K. Newey. Generalized method of moments specification testing. *Journal of Econometrics*, 29(3):229–256, 1985.

[54] Whitney K. Newey and Daniel L. McFadden. Large sample estimation and hypothesis testing. In Robert Engle and Daniel L. McFadden, editors, *Handbook of Econometrics, Volume 4*. Elsevier, Amsterdam, 1994.

[55] Ariel Pakes and David Pollard. Simulation and the aysmptotics of optimization estimators. *Econometrica*, 57(5):1027–1057, 1989.

[56] Michael G. Palumbo. Uncertain medical expenses and precautionary saving near the end of the life cycle. *Review of Economic Studies*, 66:395–421, 1999.

[57] Svetlana Paschenko. Accounting for non-annuitization. Mimeo, University of Virginia, 2009.

[58] Jorn-Steffen Pischke. Measurement error and earnings dynamics: Some estimates from the psid validation study. *Journal of Business & Economics Statistics*, 13(3):305–314, 1995.

[59] James Powell. Estimation of semiparametric models. In Robert Engle and Daniel L McFadden, editors, *Handbook of Econometrics, Volume 4*. Elsevier, Amsterdam, 1994.

[60] Susann Rohwedder, Steven J. Haider, and Michael Hurd. Increases in wealth among the elderly in the early 1990s: How much is due to survey design? Working Paper 10862, National Bureau of Economic Research, 2004.

[61] John Rust and Cristopher Phelan. How social security and medicare affect retirement behavior in a world of incomplete markets. *Econometrica*, 65:781–831, 1997.

[62] John Karl Scholz, Ananth Seshadri, and Surachai Khitatrakun. Are americans saving optimally for retirement? *Journal of Political Economy*.

[63] Antony F. Shorrocks. The age-wealth relationship: a cross-section and cohort analysis. *The Review of Economics and Statistics*, 55(3):155–163, 1975.

[64] George Tauchen and Robert Hussey. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. *Econometrica*, 59(2):371–396, 1991.

[65] Kip W. Viscusi and William N. Evans. Utility functions that depend on health status: Estimates and economic implications. *American Economic Review*, 68(4):547–560, 1978.

[66] Motohiro Yogo. Portfolio choice in retirement: Health risk and the demand for annuities, housing, and risky assets. Mimeo, University of Pennsylvania, 2009.

Appendix A: Moment conditions and asymptotic distribution of parameter estimates

Benchmark model

In the model with exogenous medical expenses, our estimate, $\hat{\Delta}$, of the "true" $M \times 1$ preference vector Δ_0 is the value of Δ that minimizes the (weighted) distance between the estimated life cycle profiles for assets found in the data and the simulated profiles generated by the model. For each calendar year $t \in \{t_0, ..., t_T\} = \{1996, 1998, 2000, 2002, 2004, 2006\}$, we match median assets for $Q = 5$ permanent income quintiles in $P = 5$ birth year cohorts.[14] The 1996 (period-t_0) distribution of simulated assets, however, is bootstrapped from the 1996 data distribution, and thus we match assets to the data for 1998, ..., 2006. In addition, we require each cohort-income-age cell have at least 10 observations to be included in the GMM criterion. In the end we have a total of $J = 101$ moment conditions.

Suppose that individual i belongs to birth cohort p, and his permanent income level falls in the qth permanent income quintile. Let $a_{pqt}(\Delta, \chi)$ denote the model-predicted median asset level for individuals in individual i's group at time t, where χ includes all parameters estimated in the first stage (including the permanent income boundaries). Assuming that observed assets have a continuous conditional density, a_{pqt} will satisfy

$$\Pr\left(a_{it} \le a_{pqt}(\Delta_0, \chi_0)\,|p, q, t, \text{individual } i \text{ observed at } t\right) = 1/2.$$

The preceding equation can be rewritten as a moment condition (Manski [52], Powell [59] and Buchinsky [11]). In particular, applying the indicator function produces

$$E\big(1\{a_{it} \le a_{pqt}(\Delta_0, \chi_0)\} - 1/2\,|p, q, t, \text{individual } i \text{ observed at } t\big) = 0. \quad (26)$$

Equation (26) is merely equation (19) in the main text, adjusted to allow for "missing" as well as deceased individuals. Continuing, we can convert this conditional moment equation into an unconditional one (e.g., Chamberlain [14]):

$$E\Big([1\{a_{it} \le a_{pqt}(\Delta_0, \chi_0)\} - 1/2] \times 1\{p_i = p\} \\ \times 1\Big\{\frac{q-1}{Q} \le I_i < \frac{q}{Q}\Big\} \times 1\{\text{individual } i \text{ observed at } t\} \;\Big|\; t\Big) = 0, \quad (27)$$

[14] Because we do not allow for macro shocks, in any given cohort t is used only to identify the individual's age.

for $p \in \{1, 2, ..., P\}$, $q \in \{1, 2, ..., Q\}$, $t \in \{t_1, t_2 ..., t_T\}$.

Suppose we have a data set of I independent individuals that are each observed at up to T separate calendar years. Let $\varphi(\Delta; \chi_0)$ denote the J-element vector of moment conditions described immediately above, and let $\hat{\varphi}_I(.)$ denote its sample analog. Letting $\widehat{\mathbf{W}}_I$ denote a $J \times J$ weighting matrix, the MSM estimator $\hat{\Delta}$ is given by

$$\underset{\Delta}{\operatorname{argmin}} \frac{I}{1+\tau} \hat{\varphi}_I(\Delta; \chi_0)' \widehat{\mathbf{W}}_I \hat{\varphi}_I(\Delta; \chi_0),$$

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

In practice, we estimate χ_0 as well, using the approach described in the main text. Computational concerns, however, compel us to treat χ_0 as known in the analysis that follows. Under regularity conditions stated in Pakes and Pollard [55] and Duffie and Singleton [23], the MSM estimator $\hat{\Delta}$ is both consistent and asymptotically normally distributed:

$$\sqrt{I}\left(\hat{\Delta} - \Delta_0\right) \rightsquigarrow N(0, \mathbf{V}),$$

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

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

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

$$\mathbf{D} = \frac{\partial \varphi(\Delta; \chi_0)}{\partial \Delta'}\bigg|_{\Delta = \Delta_0} \tag{28}$$

is the $J \times M$ gradient matrix of the population moment vector; and $\mathbf{W} = \operatorname{plim}_{I \to \infty}\{\widehat{\mathbf{W}}_I\}$. Moreover, Newey [53] shows that if the model is properly specified,

$$\frac{I}{1+\tau} \hat{\varphi}_I(\hat{\Delta}; \chi_0)' \mathbf{R}^{-1} \hat{\varphi}_I(\hat{\Delta}; \chi_0) \rightsquigarrow \chi^2_{J-M},$$

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

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

The asymptotically efficient weighting matrix arises when $\widehat{\mathbf{W}}_I$ converges to $\mathbf{S}^{-1}$, the inverse of the variance-covariance matrix of the data. When

$\mathbf{W} = \mathbf{S}^{-1}$, $\mathbf{V}$ simplifies to $(1+\tau)(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}$, and $\mathbf{R}$ is replaced with $\mathbf{S}$. This is the matrix we use for all the results shown in this paper.

But even though the optimal weighting matrix is asymptotically efficient, it can be biased in small samples. (See, for example, Altonji and Segal [1].) To check for robustness, we also use a "diagonal" weighting matrix, as suggested by Pischke [58]. This diagonal weighting scheme uses the inverse of the matrix that is the same as $\mathbf{S}$ along the diagonal and has zeros off the diagonal of the matrix. This matrix delivers parameter estimates very similar to our benchmark estimates.

We estimate $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{W}$ with their sample analogs. For example, our estimate of $\mathbf{S}$ is the $J \times J$ estimated variance-covariance matrix of the sample data. When estimating this matrix, we use sample statistics, so that $a_{pqt}(\Delta, \chi)$ is replaced with the sample median for group pqt.

One complication in estimating the gradient matrix $\mathbf{D}$ is that the functions inside the moment condition $\varphi(\Delta; \chi)$ are non-differentiable at certain data points; see equation (27). This means that we cannot consistently estimate $\mathbf{D}$ as the numerical derivative of $\hat{\varphi}_I(.)$. Our asymptotic results therefore do not follow from the standard GMM approach, but rather the approach for non-smooth functions described in Pakes and Pollard [55], Newey and McFadden [54] (section 7) and Powell [59].

To find $\mathbf{D}$, it is helpful to rewrite equation (27) as

$$\Pr\Big(p_i = p \ \& \ \frac{q-1}{Q} \le I_i \le \frac{q}{Q} \ \& \text{ individual } i \text{ observed at } t\Big) \times \\ \left[\int_{-\infty}^{a_{pqt}(\Delta_0, \chi_0)} f\Big(a_{it} \mid p, \frac{q-1}{Q} \le I_i \le \frac{q}{Q}, t\Big)\, da_{it} \ - \ \frac{1}{2}\right] = 0, \tag{29}$$

It follows that the rows of $\mathbf{D}$ are given by

$$\Pr\Big(p_i = p \ \& \ \frac{q-1}{Q} \le I_i \le \frac{q}{Q} \ \& \text{ individual } i \text{ observed at } t\Big) \times \\ f\Big(a_{pqt} \mid p, \frac{q-1}{Q} \le I_i \le \frac{q}{Q}, t\Big) \times \frac{\partial a_{pqt}(\Delta_0; \chi_0)}{\partial \Delta'}. \tag{30}$$

In practice, we find $f\big(a_{pfqt} | p, q, t\big)$, the conditional p.d.f. of assets evaluated at the median a_{pqt}, with a kernel density estimator written by Koning [46].

The model with endogenous medical expenditures

To estimate the model with endogenous medical expenses, we add several moment conditions relating to medical expenses. To use these moment conditions, we first simulate medical expenses at an annual frequency, and then take two-year averages to produce a measure of medical expenses comparable to the ones contained in the AHEAD.

Because we are interested in the possibility of catastrophic medical expenses, especially at very old ages, it is important that each cohort-income-age cell have a fairly large number of observations. To ensure this we split the income distribution in half, rather than into quintiles, so that that the quintile index q lies in $\{1, 2\}$. As before, we divide individuals into 5 cohorts, and match data from 5 waves covering the period 1998-2006.

To fit the upper tail of the medical expense distribution, we require the model to match the 90th percentile of out-of-pocket medical expenditures in each cohort-income-age cell. Let $m^{90}_{pqt}(\Delta, \chi)$ denote the model-predicted 90th percentile for individuals in cohort p and permanent income half q at time (age) t. Letting i index individuals, and proceeding as before, we have the following moment condition:

$$E\Big([1\{m_{it} \leq m^{90}_{pqt}(\Delta_0, \chi_0)\} - 0.9] \times 1\{p_i = p\} \\ \times 1\Big\{\frac{q-1}{Q} \leq I_i < \frac{q}{Q}\Big\} \times 1\{\text{individual } i \text{ observed at } t\} \;\Big|\; t\Big) = 0, \quad (31)$$

for $p \in \{1, 2, ..., P\}$, $q \in \{1, 2\}$, $t \in \{t_1, t_2..., t_T\}$.

We also require the model to match mean medical expenses (in levels, not logs) in each cell. Let $\overline{m}_{pqt}(\Delta, \chi)$ denote the model-predicted mean. The associated moment condition is

$$E\Big([m_{it} - \overline{m}_{pqt}(\Delta_0, \chi_0)] \times 1\{p_i = p\} \\ \times 1\Big\{\frac{q-1}{Q} \leq I_i < \frac{q}{Q}\Big\} \times 1\{\text{individual } i \text{ observed at } t\} \;\Big|\; t\Big) = 0, \quad (32)$$

Finally, to pin down the autocorrelation coefficient for ζ (ρ_m), and its contribution to the total variance $\zeta + \xi$, we require the model to match the first and second autocorrelations of logged medical expenses. Define the residual R_{it} as

$$R_{it} = \ln(m_{it}) - \overline{\ln m}_{pqt},$$
$$\overline{\ln m}_{pqt} = E(\ln(mit)|pi = p, qi = q, t),$$

and define the standard deviation σ_{pqt} as

$$\sigma_{pqt} = \sqrt{E\big(R_{it}^2 | p_i = p, q_i = q, t\big)}.$$

Both $\overline{\ln m}_{pqt}$ and σ_{pqt} can be estimated non-parametrically as elements of χ. Using these quantities, the autocorrelation coefficient AC_{pqtj} is:

$$AC_{pqtj} = E\left(\frac{R_{it} R_{i,t-j}}{\sigma_{pqt}\ \sigma_{pq,t-j}} \;\middle|\; p_i = p, q_i = q\right).$$

Let $AC_{pqtj}(\Delta, \chi)$ be the jth autocorrelation coefficient implied by the model, calculated using model values of $\overline{\ln m}_{pqt}$ and σ_{pqt}. The resulting moment condition for the first autocorrelation is

$$\begin{aligned} E\Bigg(\bigg[\frac{R_{it} R_{i,t-1}}{\sigma_{pqt}\ \sigma_{pq,t-1}} - AC_{pqt1}(\Delta_0, \chi_0)\bigg] \times 1\{p_i = p\} \times 1\Big\{\frac{q-1}{Q} \leq I_i < \frac{q}{Q}\Big\} \\ \times 1\{\text{individual } i \text{ observed at } t \,\&\, t-1\} \,\Bigg|\, t\Bigg) = 0, \end{aligned} \tag{33}$$

The corresponding moment condition for the second autocorrelation is

$$\begin{aligned} E\Bigg(\bigg[\frac{R_{it} R_{i,t-2}}{\sigma_{pqt}\ \sigma_{pq,t-2}} - AC_{pqt2}(\Delta_0, \chi_0)\bigg] \times 1\{p_i = p\} \times 1\Big\{\frac{q-1}{Q} \leq I_i < \frac{q}{Q}\Big\} \\ \times 1\{\text{individual } i \text{ observed at } t \,\&\, t-2\} \,\Bigg|\, t\Bigg) = 0, \end{aligned} \tag{34}$$

To summarize, the moment conditions used to estimate model with endogenous medical expenses consist of: the moments for asset medians described by equation (27); the moments for the 90th percentile of medical expenses described by equation (31); the moments for mean medical expenses described by equation (32); and the moments for the autocorrelations of logged medical expenses described by equations (33) and (34). The GMM criterion function, standard errors, and overidentification test statistic are straightforward extensions of those used in the baseline model.

Appendix B: Data and first-stage results

Overview of the AHEAD data

The AHEAD data provides high quality information on, amongst other things, mortality, medical expenses, income, and wealth.

If it is discovered that a sample member dies, this is recorded and verified using the National Death Index. Attrition for reasons other than death is relatively rare. Furthermore, the mortality rates we estimate from the AHEAD are very similar to the aggregate statistics, giving us confidence in the data.

We consider only single retired individuals in the analysis. We drop all individuals who were either married or co-habiting at any point in the analysis (so we include individuals who were never married with those who were divorced or widowed by wave 1), which leaves us with 3,498 individuals. After dropping individuals with missing wave 1 labor income data and individuals with over $3,000 in labor income in any wave, we are left with 3,259 individuals, of whom 592 are men and 2,667 are women. Of these 3,259 individuals, 884 are still alive in 2006.

We use the RAND release of the data for all variables, although we augment the RAND medical expense data. Parameter estimates and documentation for out-of-pocket medical expenses, mortality rates, and health transition probabilities are available at: http://www.chicagofed.org/economic_research_and_data/economists_preview.cfm?autID=29.

Assets

One problem with asset data is that the wealthy tend to underreport their wealth in all household surveys (Davies and Shorrocks [15]). This leads to understating asset levels at all ages. However, Juster et al. [44] show that the wealth distribution of the AHEAD matches up well with aggregate values for all but the richest 1% of households. However, problems of wealth underreporting seem particularly severe for 1994 AHEAD wave (see Rohwedder et al. [60]). As a result, we do not use the 1994 wealth data in our estimation procedure. (We use other 1994 data, however, in constructing the income, health, and mortality profiles.) Given that, and the fact that we are matching median assets (conditional on permanent income), the underreporting by the very wealthy should not significantly affect our results.

Figure 15 shows the full set of asset profiles we use in the analysis.

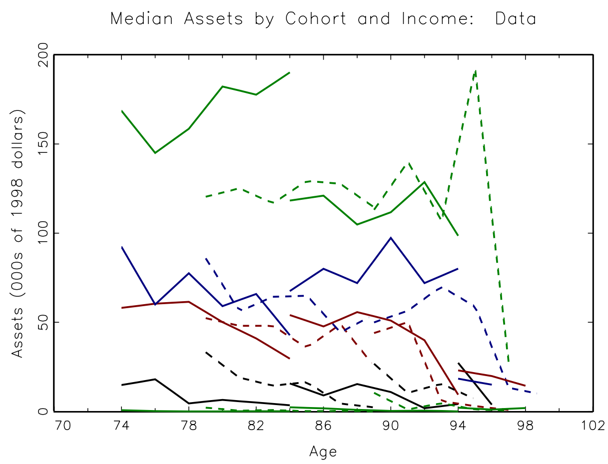

Figure 15: Median assets by cohort and PI quintile: data.

Medical expenses

The measure of medical expenditures contained in the AHEAD is average medical expenditures over the last two years. For surviving individuals, we use the RAND release of medical expense data. RAND has not coded medical expenses that people incur in their last year of life. However, the AHEAD data include follow-up interviews of the deceased's survivors. These follow-up interviews include information on medical expenses in the last year of life. Because AHEAD respondents were asked only a limited set of questions about their medical spending in 1994, we use medical expense data from 1996 onwards.

Because medical expeditures in the AHEAD are two-year averages, while our model operates at an annual frequency, we multiply the two-year variance residual variance by 1.424. This adjustment, based on the "Standard Lognormal" Model shown in Table 7 of French and Jones [34], gives us the variance in one-year medical expenditures that would, when averaged over two years, match the variance seen in the two-year data.[15]

Figure 16 compares the cumulative distribution function (CDF) of out-of-pocket medical expenditures found in the data with that produced by the

[15]To keep mean medical expenses (as opposed to their log) constant, this variance adjustment is accompanied by an adjustment to the log mean. A description of this adjustment can be found in our on-line documentation.

models with exogenous and endogenous medical expenditures. In contrast to Figure 3, which illustrates the medical expense process we feed into the model, Figure 16 shows model-generated medical expenses that are net of government transfers, as modelled by the consumption floor. Both models fit the data well.

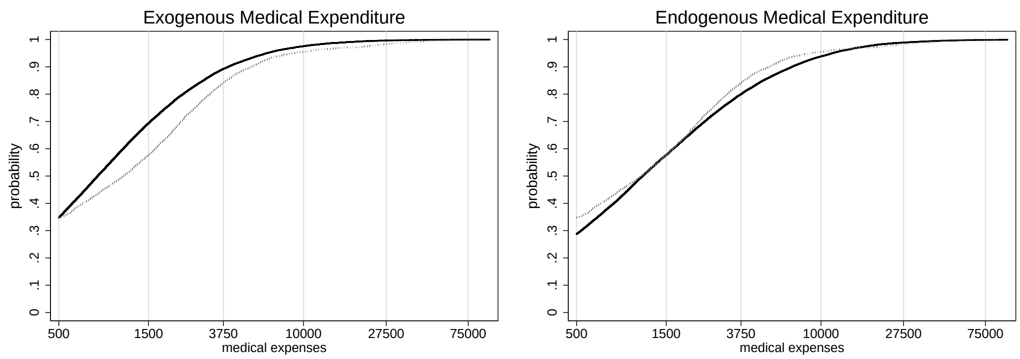

Figure 16: Cumulative distribution function of medical expenses: data and the exogenous (left panel) and endogenous (right panel) medical expenditure models. Legend: solid line is model, lighter line is data.

The CDF for the model with exogenous medical expenses lies slightly above the CDF for the data until about \$40K, and thus understates the probability of large medical expenditures. This is not surprising because, due to the nature of the AHEAD data, the exogenous medical expense process is estimated on out-of-pocket expenditures net of Medicaid payments that, ideally, should have been included.

The model with endogenous medical expenditures, in contrast, explicitly accounts for this kind of censoring. As a result, the moment conditions used to estimate the process for the "medical needs" shocks use model-generated and actual data that account for Medicaid payments in the same way. This should and in fact does increase the estimated probability of large medical expenses, relative to the exogenous medical expense model. The second panel of Figure 16 shows that the CDF for the endogenous model understates the probability of low medical expenses, but fits the upper tail of the distribution (beginning around \$20K) quite well.

Co-insurance rates

We also use the AHEAD data to estimate the co-insurance rate $q(t, h_t)$. Recall that in all waves AHEAD respondents are asked about what medical expenses they paid out of pocket. In 1998, 2000, and 2002 they were also asked about total billable medical expenses (including what is paid for by insurance). We measure the co-insurance rate as the amount spent out of pocket (less insurance premia) divided by the total billable medical expenses. Following Yogo [66] we regress the log of this on an age polynomial and health status, and health status interacted with age. We find that the co-insurance rate falls with age and for people in bad health.

Appendix C: More estimation results

The first column of results in Table 4 reports the parameter estimates from our benchmark model.

The second column of Table 4 shows estimates for a version of the benchmark model with an exogenously fixed consumption floor. This is meant to further investigate whether the risk aversion parameter and the consumption floor are separately identified. Raising the consumption floor to \$5,000 exposes consumers to less risk: the model compensates by raising the estimated value of ν to 6.0. This adjustment to ν, however, is not enough to prevent a significant worsening of the model's fit. When the consumption floor is set exogenously to \$5,000, the overidentification test statistic rises from 82.3 to 110.2. The difference of 17.9 should be compared with the critical value of a χ^2 of degree 1 (which is 3.8) because we are imposing one parameter restriction. Therefore, this restriction is rejected and shows that $\underline{c}$ is identified.

The third and fourth columns of Table 4 show the parameter estimates that arise when the model is solved for an i.i.d rate of return and the simulations use the year-by-year realized rates of returns over the sample period.

We use two measures of returns. The first one, used in column (3), is the return series calculated by Blau [7], updated to cover the period 1976-2008. Over this period, Blau's series has a mean of 0.9% and standard deviation of 2.9%. The dynamic programming problem is solved using these numbers. In the simulations, we feed in the rates of return that were realized in each given year, starting from 1997 and until 2005. (Initial cash-on-hand for 1996 is not changed.) The average rate of return over this shorter time period was 3.1%.

The second measure of returns, used in column (4), is an update of the series in French et al. [33]. Based on data from 1960 to 1995, French et al.'s series has a mean of 4.3% and standard deviation of 2.3%. The model's decision rules are found using these numbers. The average rate of return over the 1997-2005 period was 6.0%. As in the Blau case, the simulations use the rates of return realized in each of these years.

Interestingly, even using French et al.'s series, where the realized rate of return is four percentage points higher than in our benchmark model, has little effect on ν, which drops from 3.81 to 3.24, and on the consumption floor, which drops from \$2,663 to \$ 2,257. In contrast, the discount factor falls almost exactly by the increase in realized interest rates, from 0.97 to 0.93. This confirms that we cannot separately identify β and $1 + r$, but only their product. The intuition for why this is difficult comes from the Euler equation, which shows that β and $1 + r$ enter as a product.

Medical expenditures and the consumption floor continue to be important determinants of savings under the alternative rates of return. As an example, Figure 17 shows that for the models estimated with these return series, the effects of eliminating medical expenses are still large.

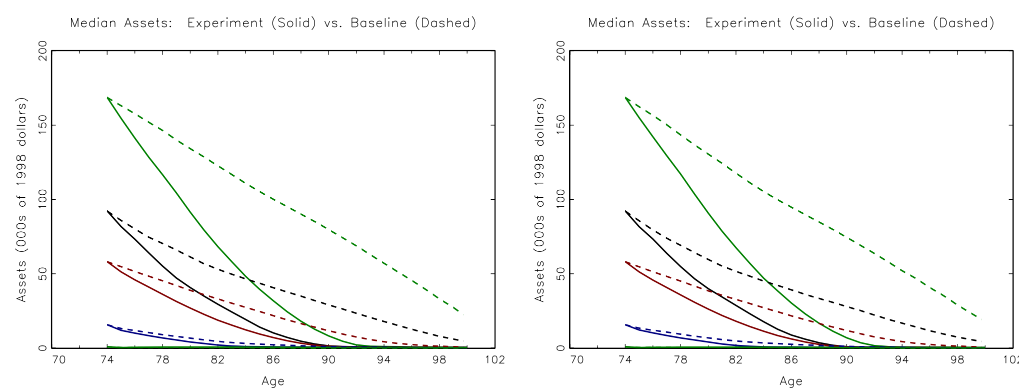

Figure 17: Median Assets, with and without medical expenses. Left panel: model using Blau's returns. Right panel: model using French et al.'s returns.

The fifth column of Table 4 shows the preference parameters for the model with endogenous medical expenses. These results have been discussed in the main text. Estimating the model with endogenous medical expenditures also requires us to find the parameters of the log of the medical needs shifter

$\mu(t, h_t, \zeta_t, \xi_t)$. Figure 18 shows that, at the estimated parameter values, the model fits the medical expenditure data well. Figure 18 compares observed and simulated profiles for the mean and the 90th percentile of the medical expenditure distribution, by cohort and permanent income. The figure shows that the model slightly under-predicts medical expenditures, but captures well the way in which expenditures rise with both age and permanent income.

Appendix D: Interpreting the size of the bequest motive

To get a sense of the size of the bequest motive, consider a person who starts the period with cash-on-hand x, dies for sure the next period without incurring any medical expenses, and is in good health. Assume further that this person's assets are well below the estate taxation exemption level. The budget constraint for such a person is given by $e = (1+r)(x-c)$, where e is the bequest (estate) left. The first order condition for an interior solution implies that the marginal utility of consumption today equals the appropriately discounted marginal utility of bequests. Using the budget constraint and our first order condition we can solve for optimal bequests:

$$e = \frac{1+r}{1+r+f}(fx-k), \tag{35}$$

where k is the shift parameter introduced in equation (3), and f is a function of several preference parameters. Since bequests cannot be negative, the expression for e reveals that x has to be large ($> k/f$) before the person will leave any bequests. If x is not sufficiently large, then $c = x$ and the solutions just derived do not apply. Assuming that x is in fact large enough, the marginal propensity to bequeath out of an extra dollar today is $\frac{\partial}{\partial x}\left(\frac{e}{1+r}\right) = \frac{f}{1+r+f}$.

Using the parameter values in the third column of Table 3, we find that marginal propensity to bequeath is 0.88, and the bequest motive becomes operative at $x = \$36,000$. In a dynamic model, where the odds of dying in any given period are low, x should be interpreted not as the total stock of wealth, but as its annuity or consumption value. The bequest motive appears to be strong, but only at very high levels of wealth.

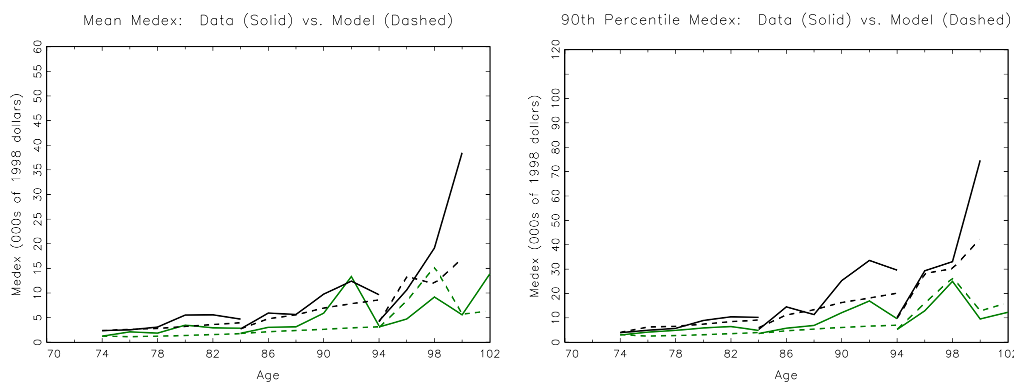

Figure 18: Medical expenditures: data and endogenous medical expenditure model. Left panel: means. Right panel: 90th percentile.

Income Quintile	Healthy Male	Unhealthy Male	Healthy Female	Unhealthy Female	All†
bottom	7.6	5.9	12.8	10.9	11.1
second	8.4	6.6	13.8	12.0	12.4
third	9.3	7.4	14.7	13.2	13.1
fourth	10.5	8.4	15.7	14.2	14.4
top	11.3	9.3	16.7	15.1	14.7
By gender:‡					
Men					9.7
Women					14.3
By health status:◇					
Healthy					14.4
Unhealthy					11.6

Notes: Life expectancies calculated through simulations using estimated health transition and survivor functions; †Calculations use the gender and health distributions observed in each permanent income quintile; ‡Calculations use the health and permanent income distributions observed for each gender; ◇Calculations use the gender and permanent income distributions observed for each health status group.

Table 2: Life expectancy in years, conditional on reaching age 70.

	Benchmark	With health-dependent preferences	With bequests	All
ν: relative risk aversion coeff.	3.81	3.75	3.84	3.66
	(0.50)	(0.47)	(0.55)	(0.55)
β: discount factor	0.97	0.97	0.97	0.97
	(0.04)	(0.05)	(0.05)	(0.04)
δ: preference shifter	0.0	-0.21	0.0	-0.36
	NA	(0.18)	NA	(0.14)
θ: bequest intensity	0.0	0.0	2,360	2,419
	NA	NA	(8,122)	(1,886)
k: bequest curvature (in 000s)			273	215
			(446)	(150)
$\underline{c}$: consumption floor	2,663	2,653	2,665	2,653
	(346)	(337)	(353)	(337)
Overidentification test	82.3	80.6	81.5	77.5
Degrees of freedom	98	97	96	95
p-value	87.4%	88.5%	85.4%	90.5%

Table 3: Estimated structural parameters. Standard errors are in parentheses below estimated parameters. NA refers to parameters fixed for a given estimation.

	Cases†				
	(1)	(2)	(3)	(4)	(5)
ν: RRA, consumption	3.81	6.04	3.72	3.24	2.15
	(0.50)	(1.01)	(0.50)	(0.37)	(0.055)
ω: RRA, medical expenses					3.19
					(0.054)
β: discount factor	0.97	0.96	0.96	0.93	0.99
	(0.04)	(0.06)	(0.05)	(0.04)	(0.01)
$\underline{c}$: consumption floor	2,663	5,000	2,658	2,257	
	(346)	NA	(367)	(397)	
$\underline{c}$: utility floor‡					202
					(23.4)
Overidentification test	82.3	110.2	81.5	80.3	667.0
Degrees of freedom	98	99	98	98	242
p-value overidentification test	87.4%	20.7%	88.5%	90.4%	0.0%

† Cases: (1) Benchmark specification; (2) $\underline{c} = 5,000$; (3) Blau return series; (4) French et al. return series; (5) Endogenous medical expenses.

‡ In the endogenous medical expense model, the estimated utility floor is indexed by the consumption level that provides the floor when $\mu = 1$. This consumption value is *not* comparable to the consumption floor.

Table 4: Estimated preference parameters. Standard errors are in parentheses below estimated parameters. NA refers to parameters fixed for a given estimation .