Differential Mortality, Uncertain Medical Expenses, and the Saving of Elderly Singles

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

November 4, 2006

Abstract

People have heterogenous life expectancies: women live longer than men, rich people live longer than poor people, and healthy people live longer than sick people. People are also subject to heterogenous out-of-pocket medical expense risk. Using AHEAD data and the method of simulated moments, we estimate a rich structural model of saving for retired single households that accounts for this heterogeneity. We find that the risk of living long and facing high medical expenses goes a long way toward explaining the elderly's saving decisions. Specifically, medical expenses that rise quickly with both age and permanent income can explain why elderly singles, and especially the richest ones, run down their assets so slowly. We also find that social insurance has a big impact on the elderly's savings.

*We thank Marco Bassetto, Marco Cagetti, Michael Hurd, Nicola Pavoni, Luigi Pistaferri, participants at the NBER Summer Institute, the MRRC researcher workshop, the Conference on Structural Models in Labor, Aging, and Health, and seminar participants at many institutions. Olga Nartova and Annie Fang Yang provided excellent research assistance. Mariacristina De Nardi: Federal Reserve Bank of Chicago, NBER, and University of Minnesota, denardim@nber.org. Eric French: Federal Reserve Bank of Chicago, efrench@frbchi.org. John Bailey Jones: University at Albany, SUNY, jbjones@albany.edu. De Nardi gratefully acknowledges financial support from NSF grant SES-0317872. The views of this paper are those of the authors and not necessarily those of the Federal Reserve Bank of Chicago, the Federal Reserve System, or the National Science Foundation.

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 (Dynan et al. [14]). Why is this the case?

To answer this question we develop a model that is consistent with the following key facts about the U.S. data. First, women outlive men by several years. Second, there is large variation in life expectancy conditional on income and health status. Third, even in presence of health insurance, out-of-pocket medical and nursing home expenses can be large, and thus generate significant net income risk for the elderly. Fourth, these medical expenses rise with both age and income.[1]

These elements affect both individual saving behavior and the composition of the sample. The fact that income-rich people tend to live longer implies that as a cohort of people grows older it becomes increasingly composed of the rich (Shorrocks [40]). The longer lifespans of women and the income-rich imply that they need to save more in order to smooth consumption. It follows that as a cohort ages it becomes increasingly composed of frugal people. Medical expenses 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.

To study these determinants of saving, we must consider theory and econometrics jointly. We proceed in two steps. Using the Assets and Health Dynamics of the Oldest Old (AHEAD) dataset, we first estimate stochastic processes for mortality and out-of pocket medical expenditures as functions of sex, health, permanent income, and age.

Our first step estimates show that average out-of-pocket medical expenditures rise very rapidly with age. For example, average medical expenditures for a woman in bad health rise from $1,200 at age 70 to $19,000 at age 100. Also, and very importantly, medical expenditures after age 85 are very much a luxury good. While a sick 95-year-old woman at the 20th percentile of the permanent income distribution expects to spend $2,700 on out-of-pocket medical costs, an otherwise identical woman at the 80th percentile expects to spend $16,000.

[1] See Attanasio and Emmerson [3], and Deaton and Paxon [12] for evidence on permanent income and mortality. See Hurd et al. [26] for evidence on health status and mortality. For evidence on medical expenses see French and Jones [18, 19], Palumbo [34], Feenberg and Skinner [16], Cohen et al. [8], and the analysis presented below.

Our first step analysis also confirms that life expectancy can vary greatly. For example, while a sick, 70-year-old male at the 20th percentile of the permanent income distribution expects to live only 6 more years, a healthy 70-year-old woman at the 80th percentile expects to live 17 more years.[2]

In our second step, we construct a rich structural model of saving behavior for retired single households, and estimate it using the method of simulated moments. In particular, the model's preference parameters are chosen so that the permanent income-conditional age-asset profiles simulated from the model match those in the data.

Notably, while our estimated values of the coefficient of relative risk aversion and the discount factor are in line with those provided by the previous literature, the additional sources of heterogeneity that we consider allow the model to fit the data extremely well. Specifically, our estimated structural model is not rejected when we test its over-identifying restrictions, which is a feat that many structural models fail to achieve.

To gauge the importance of different saving motives, we use our estimated model to perform a number of decomposition exercises. We find that the differences in average medical expenditure by permanent income are very important in explaining heterogeneity in asset decumulation decisions, while the idiosyncratic risk associated to these expenditures, while significant, is not a key force. Our baseline model predicts that, between ages 74 and 81, median assets for those in the top permanent income quintile are approximately constant at $150,000, which is roughly consistent with the data. When we eliminate medical expense risk, but hold average medical expenses constant, we find that median assets for this group fall from $150,000 to $140,000 between ages 74 and 81. However, when we eliminate all medical expenses, median assets for this group fall from $150,000 to $90,000 between ages 74 and 81.

We find that social insurance programs such Supplemental Security Income and Medicaid have large effects on the elderly's saving behavior, including the richest ones. In the absence of social insurance median assets for those in the top permanent income quintile would rise from $150,000 to $220,000 between ages 74 and 81.

We also find that a significant portion of the higher saving of the high-permanent income elderly is due to the fact that they have a longer life-expectancy. If everyone had the survival probabilities of a healthy male

[2]These life expectancies are drawn from estimates summarized in Table 1.

at the 50th percentile of the permanent income distribution, median assets for those in the top permanent income quintile would fall from \$150,000 to \$140,000 between ages 74 and 81.

Compared to the previous literature we obtain a much better fit of the model to the data, and we find a larger effect of medical expenses and the social insurance on the elderly's saving decisions.

In an early study Kotlikoff [28] 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.

More recent studies, such as Hubbard et al. [22] and Palumbo [34], find that medical expenses have fairly small effects. In contrast, we find that our model can go a long way towards accounting for the observed lack of asset decumulation after retirement, at least for the elderly singles. The cause of these differences is that Hubbard et al. and Palumbo likely understate medical expenses, both in terms of levels and riskiness (see French and Jones [18, 19]), as well as the extent to which medical expenditures rise with age and permanent income. As an example, the average expense for a 100-year-old with some college generated by Hubbard et al.'s medical expenditure model is about 15% of the average medical expense for a 100-year generated by our model. We obtain different estimates because we use newer and better data—the AHEAD contains detailed information for a large number of very old individuals—and a more flexible specification of medical costs.

Hubbard et al. [23] argue that means-tested social insurance programs 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 a large effect on saving decisions at all levels of income. Under our model of health costs, medical expenses in old age are so large that even the saving decisions of fairly rich people are affected by insurance programs such as Medicaid.

Scholz et al. [39] find that a life cycle model, augmented with realistic income and medical expense uncertainty, can do a 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.

The rest of the paper is organized as follows. In section 2, we introduce our version of the life cycle model, and in section 3, we discuss our estimation procedure. In sections 4 and 5, we describe the data and the estimated shock processes that elderly individuals face. We also construct a very simple measure of mortality bias, and show that the bias is significant. We discuss our results in section 6, which includes some robustness checks and some decomposition exercises that gauge the forces affecting saving behavior. We conclude in section 7.

2 The model

Our analysis focuses on people that have retired already, which allows us to concentrate on saving and consumption decisions, and abstract from labor supply and retirement decisions. We restrict our analysis to elderly singles to avoid the complications of dealing with household dynamics, such as the transition from two to one family members. We also sharpen our analysis by excluding bequest motives, in order to isolate the potential effects of medical expense and mortality risk.

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 current and future consumption. Each period, the individual's utility depends on its consumption, c, and health status, h, which can be either good ($h = 1$) or bad ($h = 0$).

The within-period utility function is

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

with $\nu \geq 0$. Following Palumbo [34] the function $\delta(h)$, which determines how a person's utility from consumption depends on his or her health status, is

$$\delta(h) = 1 + \delta h, \qquad (2)$$

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

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

$$y_t = y(g, I, t). \qquad (3)$$

The individual faces several sources of risk, which we treat as exogenous. While this is of course a simplification, we believe it is a reasonable assumption, especially since we focus on older people that 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

$$\pi_{j,k,g,I,t} = \Pr(h_{t+1} = k | h_t = j, g, I, t), \quad j, k \in \{1, 0\}. \tag{4}$$

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 costs, m_t, are defined as out-of-pocket costs. Since our focus is on understanding the effects of out-of-pocket medical expenses on saving decisions, we model health costs as exogenous reductions to the household's available resources, as in Scholz et al. [39], Palumbo [34] and Hubbard et al. [22, 23]. We assume that health costs depend upon sex, health status, permanent income, age and an idiosyncratic component, ψ_t:

$$\ln m_t = m(g, h, I, t) + \sigma(g, h, I, t) \times \psi_t. \tag{5}$$

Following Feenberg and Skinner [16] and French and Jones [19], we assume that ψ_t can be decomposed as

$$\psi_t = \zeta_t + \xi_t, \quad \xi_t \sim N(0, \sigma_\xi^2), \tag{6}$$

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

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 [41].

The timing is the following: at the beginning of the period the individual's health status and medical costs are realized. Then the individual consumes and saves. Finally the survival shock hits.

Next period's assets are given by

$$a_{t+1} = a_t + y_n(ra_t + y_t, \tau) + b_t - m_t - c_t, \tag{8}$$

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.[3]

Assets have to satisfy a borrowing constraint:

$$a_t \geq 0. \tag{9}$$

Following Hubbard et al. [22, 23], we also assume that government transfers provide a consumption floor:

$$b_t = max\{0, c_{min} + m_t - [a_t + y_n(ra_t + y_t, \tau)]\}, \tag{10}$$

Equation (10) says that government transfers bridge the gap between an individual's "total resources" (the quantity in the inner parentheses) and the consumption floor. Equation (10) also implies that if transfers are positive, $c_t = c_{min}$ and $a_{t+1} = 0$.

To save on state variables we follow Deaton [11] and redefine the problem in terms of cash-on-hand:[4]

$$x_t = a_t + y_n(r\, a_t + y_t, \tau) + b_t - m_t. \tag{11}$$

Note that assets and cash-on-hand follow:

$$a_{t+1} = x_t - c_t, \tag{12}$$

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

To enforce the consumption floor, we impose

$$x_t \geq c_{min}, \quad \forall t, \tag{14}$$

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

$$c_t \leq x_t, \quad \forall t. \tag{15}$$

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

[3]We do not include received bequests as a source of income, because very few individuals aged 65 or older receive them.

[4]Using cash-on-hand allows us to combine assets and the transitory component of medical expenses into a single state variable.

definition of cash-on-hand. Letting β denote the discount factor, the value function for a single individual is given by

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

subject to equations (13) - (15).

3 Estimation procedure

3.1 The Method of Simulated Moments

To estimate the model, we adopt a two-step strategy, similar to the one used by Gourinchas and Parker [21], Cagetti [7], and French and Jones [18]. 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 $\Delta = (\delta, \nu, \beta, c_{min})$ 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 permanent income save at a higher rate, we match permanent income-conditional age-asset profiles. Our approach is similar to that of French and Jones [18].

Consider individual i of birth cohort c in calendar year t. Note that the individual's age is $t - c$. 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 c falls in the qth permanent income interval of the sample. Let $a_{cqt}(\Delta, \chi)$ be the model-predicted median asset level in calendar year t for an individual of cohort c that was in the qth permanent 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 cqt will have asset levels of

$a_{cqt}(\Delta_0, \chi_0)$ or less. This leads to a well-known moment condition:[5]

$$E\big(1\{a_{it} \leq a_{cqt}(\Delta_0, \chi_0)\} - 1/2 \,|c, q, t, \text{individual } i \text{ alive at } t\big) = 0, \qquad (16)$$

for all c, q and t. 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 cqt 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. In particular, 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 1995,[6] and each is assigned a series of health, health cost, and mortality shocks consistent with the stochastic processes described in the previous section 2.[7] Solving numerically the model described in section 2 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_{cqt}) from the artificial histories in the same way as we compute them from the real data. Finally, we adjust Δ until the difference between the data and simulated profiles—a GMM criterion function based on equation (16)—is minimized.

We discuss the asymptotic distribution of the parameter estimates, the weighting matrix and the overidentification tests in Appendix B.

---

[5]See Manski [30], Powell [36] and Buchinsky [6]. Related methodologies are applied in Cagetti [7] and Epple and Seig [15].

[6]Since we do not observe ζ_t directly, we infer it from individuals' observed medical expenditures, using the model of medical spending described below and standard projection formulae.

[7]The simulated medical expenditure shocks are monte carlo draws from a discretized version of our estimated medical expenditure process. In contrast, when simulating health and mortality shocks, we give each simulated person the entire health and mortality history realized by a person in the AHEAD data with the same initial conditions. (Although the data provide health and mortality only during interview years, we simulate it in off-years using our estimated models and Bayes' Rule.) 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.

3.2 Econometric Considerations

In estimating our model, we face two well-known econometric problems (see, for example, Shorrocks [40]). First, in a cross-section or short panel, 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 will have lower lifetime incomes. Not accounting for this effect will lead us to estimate a model that overstates the degree to which elderly people run down their assets.

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 estimate a model that understates 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.[8] If older people have lower lifetime incomes in our data, they will have lower lifetime incomes in our simulations. 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.

[8] It bears noting that we are assuming that there are no cohort effects beyond those captured in the distributions of wealth, health, gender and income by age. This simplification allows us to use the same set of decision rules for all cohorts, which significantly reduces our computational burden. Moreover, as shown below, it does not prevent the model from fitting asset profiles across a wide range of ages.

4 Data

The AHEAD is a sample of non-institutionalized individuals, aged 70 or older in 1993. A total of 8,222 individuals in 6,047 households were interviewed for the AHEAD survey in 1993 (in other words, 3,872 singles and 2,175 couples). These individuals were interviewed again in 1995, 1998, 2000, and 2002. The AHEAD data include a nationally representative core sample as well as additional samples of blacks, Hispanics, and Florida residents.

If it is discovered that a sample member dies, this is recorded and verified using the National Death Index. Fortunately, attrition for reasons other than death is relatively rare, and we can use the AHEAD data to estimate mortality rates; as we show below, the mortality rates we estimate from the AHEAD are very similar to the aggregate statistics. Because our econometric approach explicitly models exit through death, we use the full unbalanced panel, and include the life histories of people who die before our sample ends.

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 with 3,510 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 3,270 individuals. We drop 315 individuals who are missing in any period, leaving us with 2,955 individuals, of whom 561 are men and 2,394 are women. Of these 2,955 individuals, 1,430 are still alive in 2002.

We use the RAND release of the data for all variables except for medical expenses. We use our own coding of medical expenses because RAND has not coded medical expenses that people incur in their last year of life—the AHEAD data include follow-up interviews of the deceased's survivors. In addition, RAND's imputation procedure does not account for high correlation of medical expenses over time, especially in the earlier waves.

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. Our measure of total assets is the sum of these items, less mortgages and other debts. We do not include pension and Social Security wealth for four reasons. First, we wish to to maintain comparability with other studies (Hurd [24] and Attanasio and Hoynes [4] for example). Second, since it is illegal to borrow against Social security wealth and difficult to borrow against most forms of pension wealth, Social Security and pension wealth are much more illiquid than other assets. Third, their tax treatment is different from other assets. Finally, differences in Social Security and pensions are captured in our model by differences in the permanent income measure we use to predict annual income.

One problem with asset data is that the wealthy tend to underreport their wealth in all household surveys (Davies and Shorrocks [9]). This leads to understate asset levels at all ages. However, Juster et al. [27] show that the the wealth distribution of the AHEAD matches up well with aggregate values for all but the richest 1% of households. This notwithstanding, problems of wealth underreporting seem particularly severe for 1993 AHEAD wave (see Rohwedder, Haider and Hurd [37]). As a result, we do not use the 1993 wealth data in our estimation procedure. (We use other 1993 data, however, in constructing some of the profiles shown below.) 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.

In addition to constructing moment conditions, we also use the AHEAD data to construct the initial distribution of permanent income, age, sex, health, health costs, and cash-on-hand that starts off our simulations. In particular, each simulated individual is given a state vector drawn from the joint distribution of state variables observed in 1995.

5 Data profiles

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.

5.1 Asset profiles and mortality bias

We construct the permanent-income-conditional age-asset profiles as follows. We sort individuals into permanent income quintiles (see section 5.3 for details), and we track birth-year cohorts. Sample size considerations lead us to focus on 4 5-year cohorts. The first cohort consists of individuals that were ages 72-76 in 1995; the second cohort contains ages 77-81; the third ages 82-86; and the fourth cohort contains ages 87-91. We use asset data for 4 different years; 1995, 1998, 2000 and 2002. It follows that for each of the

20 cohort-permanent income cells, we observe assets 4 times over a 7-year span. To construct the profiles, we calculate cell medians each year assets are observed. Because some individuals die between 1993 and 1995, or fall outside the 4 cohorts described above, the asset profiles use a subsample of the data, with 2,482 individuals.

To fix ideas, consider Figure 1, which plots assets by age in each permanent income and cohort grouping for those that are still alive at that particular moment in time. The lines at the far left of the graph are for the youngest cohort, whose members in 1995 were aged 72-76, with an average age of 74. We observe these individuals—if still alive—again in 1998, when they were 77, and in 2000 (age 79) and 2002 (age 81). There are five lines because we have split the data into permanent income quintiles. Unsurprisingly, assets turn out to be monotonically increasing in permanent income, so that the bottom left line shows median assets for surviving individuals in the lowest permanent income quintile, while the top line shows median assets for surviving individuals in the top quintile.

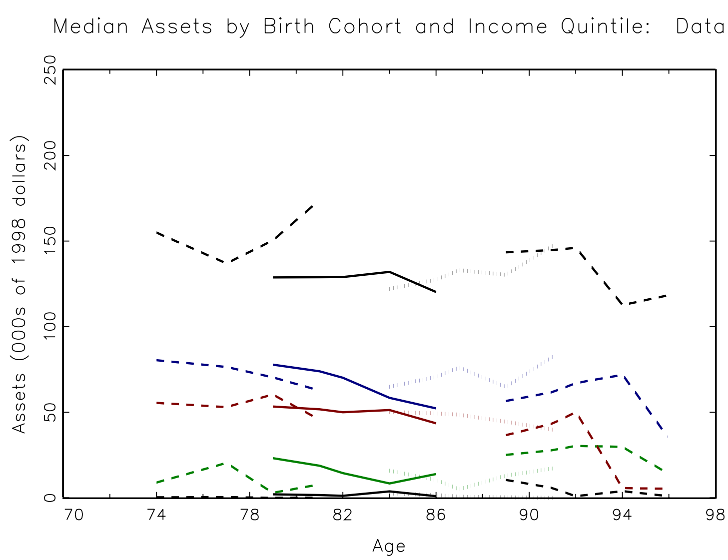

Figure 1: Median assets by cohort and PI quintile: data

For all permanent income quintiles in the youngest cohort, assets neither rise nor decline rapidly with age. If anything, those with high permanent income tend to have increases in their assets, whereas those with low permanent income tend to have declines in assets as they age.

Next, consider the lines at the far right of the graph, which are for the cohort whose members in 1995 were aged 87-91, with an average age of 89. The dynamics of assets for members of this cohort are similar to the dynamics for the youngest cohort; the only exception is that wealth in the highest permanent income quintile falls rather than rises with age.

Our finding that the income rich elderly run down their assets at a very slow rate complement and confirms those by Dynan et al. [14] who look both at younger and older households but do not have as many observations as we do on the very elderly.

It is worth stressing that the data shown in Figure 1 are drawn from an unbalanced panel: at each point in time we take the people alive at that moment to compute median assets, hence many of the individuals used to calculate the 1995 medians were deceased by 2002. Because poorer and/or less thrifty individuals have higher mortality rates, these profiles are affected by mortality bias as time goes on. To get a sense of this mortality bias, Figure 2 shows two sets of asset profiles. The first set of profiles shows median assets for every person still alive when the data are collected in a given wave; this is, what was shown in Figure 1. The second set of profiles shows median assets for the balanced panel, that is for the set of individuals that were alive in all 5 waves. The differences between the two profiles can be interpreted as mortality bias. For clarity, the picture reports profiles for two cohorts. The extent of mortality bias is similar for the other two cohorts.

Figure 2 shows that when households are sorted by permanent income, mortality bias is fairly small. This sorting, however, obscures any mortality bias caused by differential mortality across the permanent income distribution. Figure 3 compares asset profiles that are aggregated over permanent income quintiles and shows that if we do not condition on permanent income, 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 permanent income than the survivors.

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.

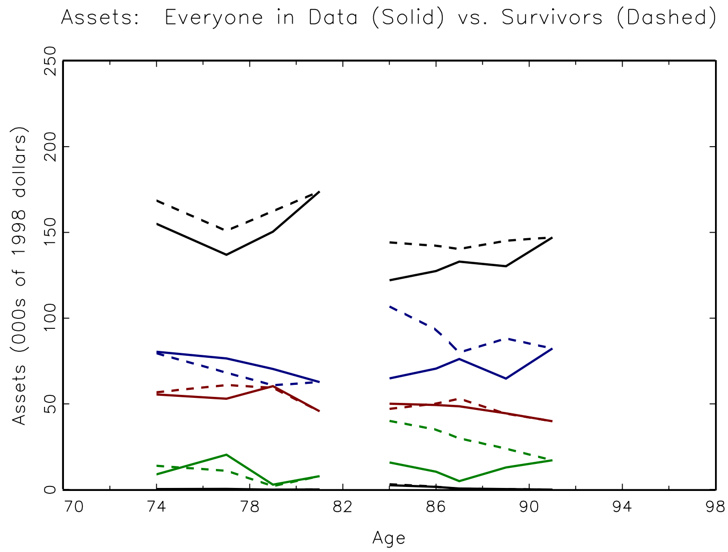

Figure 2: Median assets by birth cohort and permanent income quintile: everyone in the data (solid lines) vs. survivors (dashed lines)

5.2 Mortality and health status profiles

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; and permanent income interacted with age.

Figure 4 shows mortality rates conditional on age, sex, previous health status, and permanent income. The top panels are for women, while the bottom ones are for men. The left panels refer to those that are healthy, while the right ones refer to the unhealthy. The top left panel shows that for women in good health last year the probability of death within one year rises from 2% at age 70 to 25% at age 100.[9] The four panels together show that, conditional upon age, men, those in bad health, and those with low permanent income are more likely to die than women, those in good health, and those with high permanent income.

We find that controlling for previous health status greatly reduces the

[9] Individuals in the AHEAD dataset are surveyed every two years. Thus we estimate the two-year survival rate. We construct the one-year survival rate by taking the square root of the two-year survival rate.

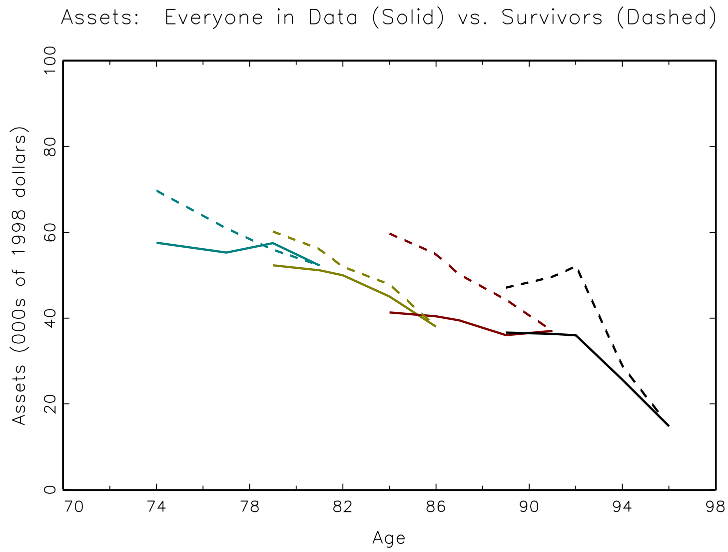

Figure 3: Median assets by birth cohort: everyone in the data (solid lines) vs. survivors (dashed lines)

estimated coefficients associated with permanent income. However, as we show below, people with high permanent income are much more likely to be in good health, even when previous health status is taken into account. Our results thus indicate that people with high permanent income have lower mortality in part because they are more likely to be healthy.[10]

Figure 5 presents health transition probabilities conditional on age, sex, previous health status, and permanent income. Consider the women first. The top left panel shows that the probability of being in bad health, conditional on being in good health one year before, is about 10% at age 70 and rises to about 25% at age 100.[11] Rich people are more likely to stay healthy: being in the 80th percentile of the permanent income distribution instead of

[10]Hurd et al. [26, 20, 25] find that other AHEAD variables, such as more sophisticated controls for health status or subjective mortality expectations, help predict mortality. Our model is a parsimonious approximation that captures much of the heterogeneity in mortality expectations.

[11]To find one-year transition rates, we first estimate the two-year Markov transition matrix, $P_{t+2|t}$. We then assume that the one-year Markov transition matrix, $P_{t+1|t}$, satisfies $P_{t+2|t} = P_{t+1|t}^2$. $P_{t+1|t}$ can then be found as the solution to a quadratic form. Details are available from the authors.

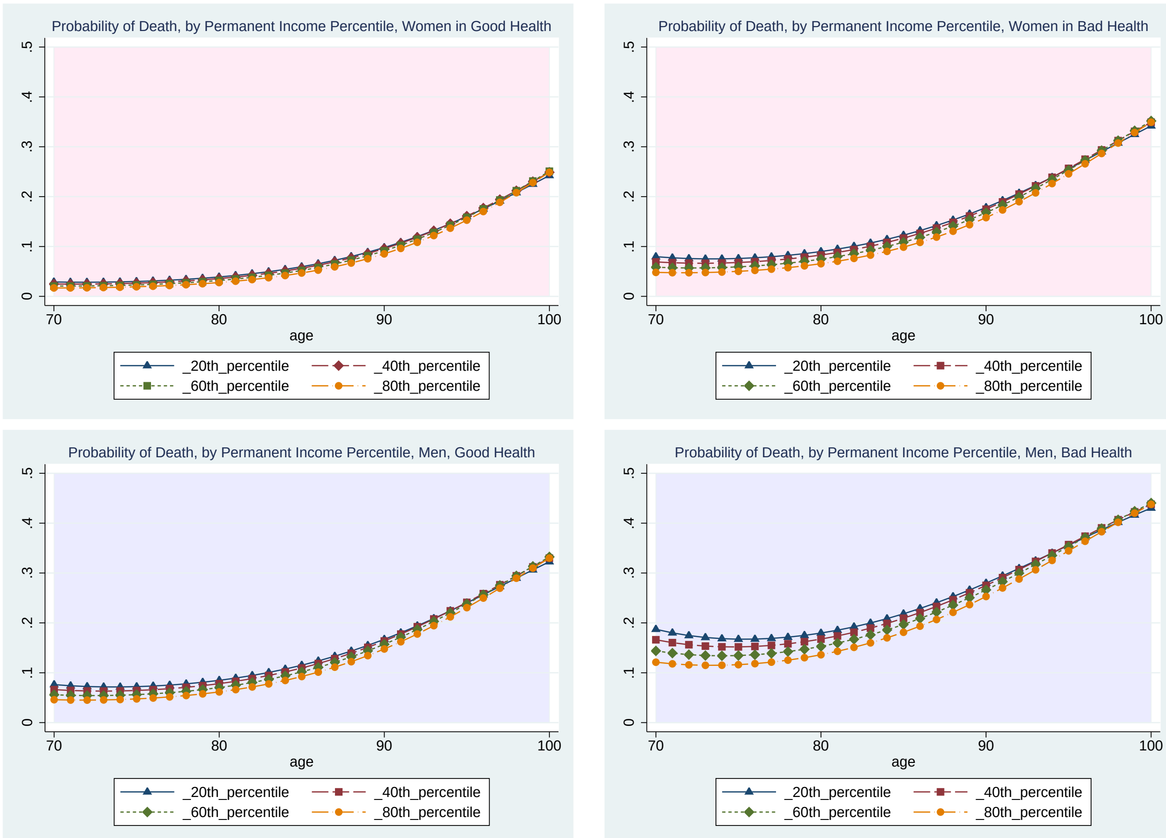

Figure 4: Mortality probabilities, by sex, permanent income percentile and health status (women on top panels, men on bottom panels, healthy on left panels, unhealthy on right panels)

the 20th percentile lowers the probability of moving into bad health by 5 to 10 percentage points. The graph on the top right shows that bad health is a very persistent state. If a 70-year-old woman was in bad health one year ago, there is almost a 90% chance that she will be in bad health this year. Surprisingly, the probability of being in bad health this year, conditional on being in bad health last year, falls with age.[12] Rich people are more likely to return to good health: having high permanent income reduces the probabil-

[12]Although this result is surprising, one should recall that we are measuring the probability of still being in bad health and surviving, conditional on being in bad health last period. The probability of either being dead or in bad health this year, conditional on being in bad health last year, remains constant at about 90% at each age.

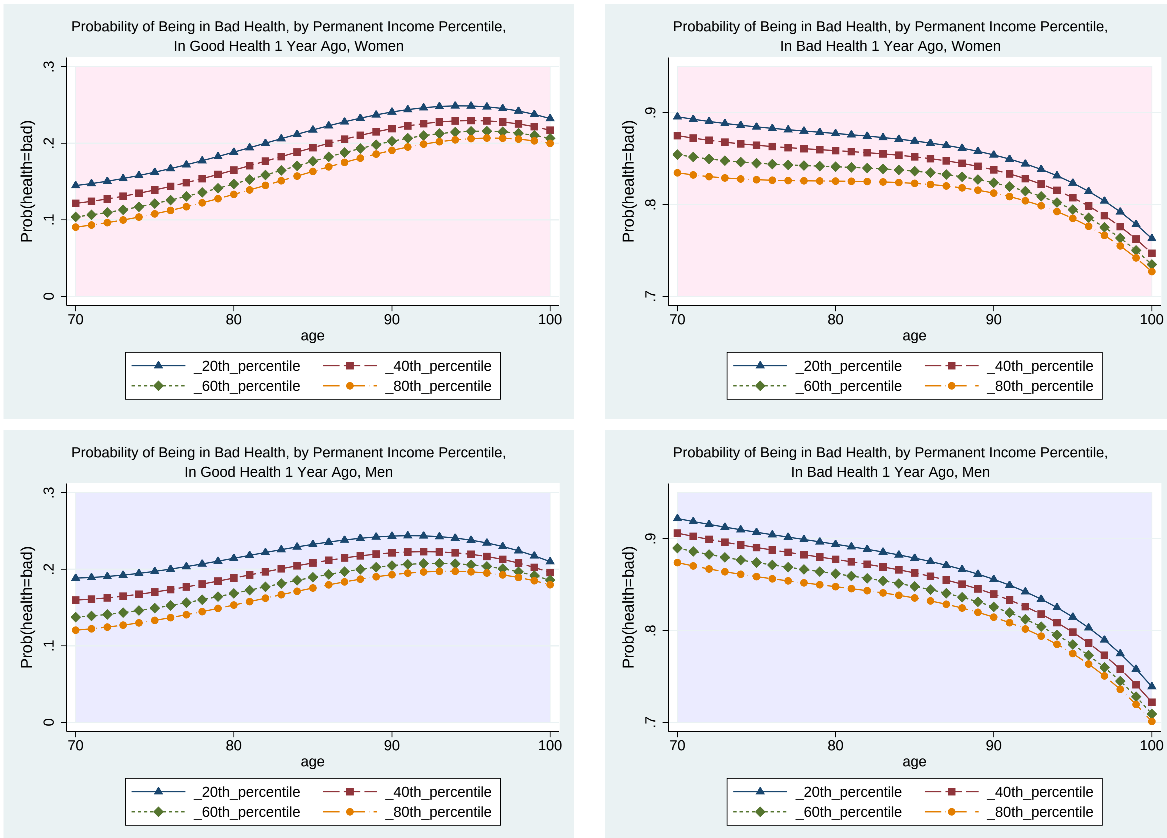

Figure 5: Health transition probabilities, by sex, permanent income percentile and health status (women on top panels, men on bottom panels, healthy on left panels, unhealthy on right panels)

ity of being in bad health in the present, conditional in being in bad health in the past.

The bottom two panels show that men are more likely to transition from good health to bad health, and to remain in bad health, than women.

Table 1 presents the life expectancies implied by our mortality and health status process. Although permanent income has only a modest effect on mortality rates, after conditioning on previous health status, it has a very strong effect on the probability of transitioning to bad health, where mortality is higher. As a result, healthy men at the 20th percentile of the permanent income distribution live 3 fewer years than healthy men at the 80th percentile, and healthy women at the 20th percentile of the permanent income distribution live 3.2 fewer years than healthy women at the 80th percentile.

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 10.2 years for men and 15.0 years for women. These differences are an artifact of 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 of a year of the aggregate statistics for both men and women.

Income Percentile	Healthy Male	Unhealthy Male	Healthy Female	Unhealthy Female	All†
20	8.2	6.2	13.8	11.9	12.0
40	9.1	7.0	14.8	12.9	13.0
60	10.1	7.9	15.9	14.1	14.1
80	11.2	9.1	17.0	15.5	15.2
By gender:‡					
Men					10.2
Women					15.0
By health status:◇					
Healthy					15.3
Unhealthy					11.9

Note: life expectancies calculated through simulations using estimated health transition and survivor functions.
† Calculations use the same (permanent-income-unconditional) gender-health distributions across all permanent income levels.
‡ 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 1: Life expectancy in years, conditional on reaching age 70

5.3 Medical expense and income profiles

Medical expenses are the sum of what the individuals spend out of pocket on insurance premia, drug costs, and costs for hospital, nursing home care, doctor visits, dental visits, and outpatient care. It does not include expenses covered by insurance, either public or private. French and Jones [19] show that the medical expense data in the AHEAD line up very well with the aggregate statistics. For our sample, mean medical expenses are $3,222 with a standard deviation of $10,339. Although this figure is large, it is not surprising, because Medicare does not cover prescription drugs, requires co-pays for services, and caps the number of nursing home and hospital nights that it pays for.

The log of medical expenses is modeled as a function of: a cubic in age; sex; sex interacted with age; current health status; health status interacted with age; a quadratic in permanent income; and permanent income interacted with age.[13]

We estimate these profiles using a fixed-effects estimator. 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.

We have also estimated specifications of equation (5) that include cohort dummy variables (i.e., we regressed the estimated fixed-effects on cohort dummies), which are statistically significant. Unfortunately, allowing for differences in medical expense and income parameters across cohorts requires that the model be solved and simulated separately for each cohort, significantly increasing the computational burden. Nevertheless, our procedure

[13] We assume that medical expenses do not affect future health and survivor probabilities. We also ignore the fact that, to some extent, the quantity of health care consumed is a choice. (See Davis [10].)

captures how medical expenses and income change with age.

Figure 6 presents average medical expenses, conditional on age, health status, and permanent income for women. Average medical expenses for men look similar to those of women, so we do not present them. We assume that medical expenses are log-normally distributed, so the predicted level of medical expenses are $\exp\left(m(g,h,I,t)+\frac{1}{2}\sigma^2(g,h,I,t)\right)$, where $\sigma^2(g,h,I,t)$ is the variance of the idiosyncratic shock ψ_t.

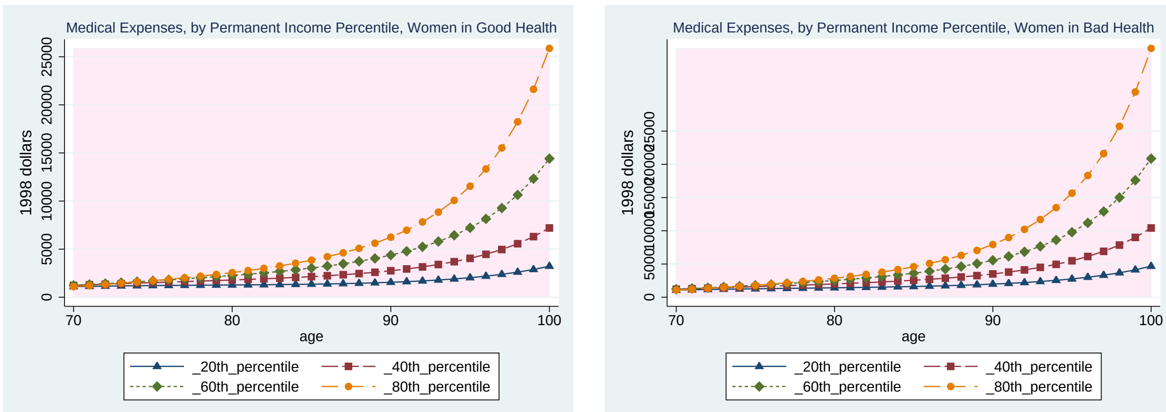

Figure 6: Average medical expenses, by permanent income percentile and health status, for women (healthy on left panel, unhealthy on right panel)

Measured health status has only a modest effect on average medical expenses, but permanent income has a large effect, especially at older ages. Average medical expenses for women in good health are \$2,000 a year at age 70, and vary little with permanent income. By age 100, they rise to \$4,000 for women at the 20th percentile of the permanent income distribution and to almost \$26,000 for women at the 80th percentile of the permanent income distribution. One might be concerned that we have few 100-year-old's in our sample, so that our predicted effects arise from using assumed functional forms to extrapolate off the support of the data. However, in our sample we have 36 observations on medical expenses for 100 year old individuals, averaging \$14,741 per year. Between ages 95 and 100, we have 483 person-year observations on medical expenses, averaging \$8,870 (with a standard deviation of \$20,783). 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 variance of log medical expenses is 2.15.[14] This implies that medical expenses for someone with a two standard deviation shock to medical expenses pays 6.41 times the average, conditional on the observables.[15]

French and Jones [19] find that a suitably-constructed lognormal distribution can match average medical expenses, as well as the far right tail of the distribution. They also find that medical expenses are highly correlated over time. Table 2 shows estimates of the persistent component ζ_t and the transitory component ξ_t found by French and Jones. 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 the persistent component of medical expenses have long-lived effects. French and Jones in fact find that most of a household's *lifetime* medical expense risk comes from the persistent component.

Parameter	Variable	Estimate
σ_ϵ^2	innovation variance of persistent component	0.0503
ρ_m	autocorrelation of persistent component	0.922
σ_ξ^2	innovation variance of transitory component	0.665

Table 2: Variance and persistence of innovations to medical expenses, as fractions of total cross-sectional variance

Our estimates of medical expense risk indicate greater risk than found in other studies (see Hubbard, Skinner, and Zeldes [22] and Palumbo [34]). However, our estimates still potentially understate the medical expense risk faced by older Americans, because our measure of medical expenditures does not include value of Medicaid contributions. Given that we explicitly model

[14] The measure of medical expenditures contained in the AHEAD is average medical expenditures over the last two years. In order to infer the standard deviation of annual medical expenditures, we multiply the two-year variance, 1.51, by 1.424. This adjustment, based on the "Standard Lognormal" Model shown in Table 7 of French and Jones [19], gives us the the variance in one-year medical expenditures that would, when averaged over two years, match the variance seen in the two-year data.

[15] Let m denote predicted log medical expenses. The ratio of the level of medical expenses two standard deviations above the mean to average medical expenses is $\frac{\exp(m+2\sigma)}{\exp(m+\sigma^2/2)} = \exp(2\sigma - \sigma^2/2) = 6.41$ if $\sigma = \sqrt{2.15}$.

a consumption floor, our conceptually preferred measure of medical expenses would includes both expenses paid by Medicaid as well as those paid out of pocket by households. Note that excluding Medicaid leads us to understate the the level of medical expenses as well.

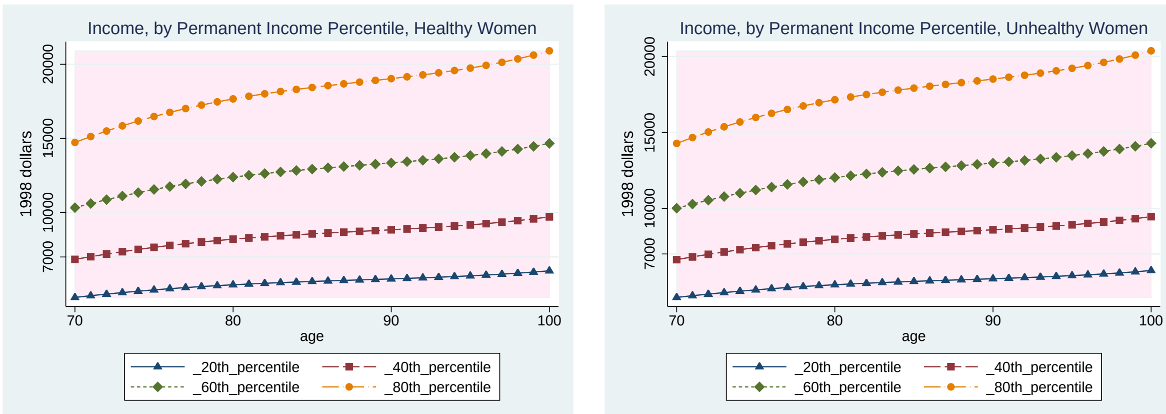

Figure 7: Average income, by permanent income percentile and health status, for women (healthy on left panel, unhealthy on right panel)

Income includes the value of Social Security benefits, defined benefit pension benefits, annuities, veterans benefits, welfare, and food stamps (the latter three are lumped together by the AHEAD). We measure permanent income as average income over all periods during which we observe the individual. 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.

We model income in the same way as medical expenses, using the same explanatory variables and the same fixed-effects estimator. Figure 7 presents average income, conditional on age, sex, health status, and permanent income for women. Given that income is largely from pensions and Social Security, which depends on previous earnings, it is unsurprising that health has a very small effect on income. Holding permanent income fixed, income for men (not shown) is only slightly higher than income for women. (Men, however, typically have more permanent income than women.) Income trends up slightly with age, which seems surprising given that most sources of income, such as Social Security benefits, should not change with age, after adjusting for inflation. However, Social Security benefits are tied to the CPI, whereas we deflate all variables by the PCE index. Between ages 70 and 100, income rises about 15%, or .5% per year. This is about the gap between the CPI and PCE.

6 Results

6.1 Preference parameter estimates and model fit

We set the interest rate to 2%. Table 3 presents preference parameter estimates under several different specifications. The first column of Table 3 refers to our "baseline specification," in which we jointly estimate all of the second stage parameters: the coefficient of relative risk aversion, the discount factor, the preference shifter due to health changes, and the consumption floor. The other columns fix one parameter at the time, that is, either the preference shifter due to health shocks, or the consumption floor.

In this section, we discuss the baseline specification. We discuss the alternative specifications in section 6.2.

Figure 8 shows how well the baseline parameterization of model fits a subset of the data profiles, using unbalanced panels. (The model fits equally well for the cells that are not shown.) The model does a very good job at matching the key features of the data that we are interested in: 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 permanent income tend to run down their wealth with age.

A more formal way to assess the goodness of fit of our model is to compute the p-value of the overidentification statistics. This value turns out to be 97.8% for our baseline specification. This is an exceptional result for a structural model, as most estimated structural models are typically rejected in overidentification tests.

Figure 9 shows how well the model fits the data when the asset profiles are aggregated over permanent income quintiles. Here too the fit is good. Among other things, the model replicates much of the large asset decumulation that occurs at very old ages. If anything, the model predicts less asset decumulation at very old ages than what is seen in the data. Previous models of consumption behavior, such as those of Hubbard [22] et al. and Palumbo [34], have predicted more asset decumulation than what is seen in the data at very old ages.

Parameter and Definition	Baseline (1)	$\delta = 0$ (2)	$c_{min} = 5,000$ (3)
ν: coeff. of relative risk aversion	4.03	4.20	7.50
	(0.97)	(0.97)	(2.34)
β: discount factor	0.965	0.966	0.923
	(0.07)	(0.07)	(0.12)
δ: preference shifter, good health	-0.197	0.0	-0.254
	(0.20)	NA	(0.39)
c_{min}: consumption floor	2791	2660	5000
	(318)	(233)	NA
Overidentification statistic	37.0	38.4	73.8
Degrees of freedom	56	57	57
P-value overidentification test	97.8%	97.2%	6.6%

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

Figure 10 shows the consumption profiles predicted by the model, namely median consumption by cohort and permanent income quintile. Figure 10 shows that the model generates flat or decreasing consumption profiles for most cohorts. This general tendency is consistent with most empirical studies of older-age consumption, which suggest that consumption falls with age (Banks, Blundell, and Tanner [5] using UK data, and Fernandez-Villaverde and Krueger [17] using US data.) For example, Fernandez-Villaverde and Krueger find that non-durable consumption declines about one percent per year between ages 70 and 90.

Figure 10, in combination with the Euler Equation, can give some intuition for the estimates presented in Table 3. Ignoring taxes, the Euler Equation is:

$$(1 + \delta h_t)c_t^{-\nu} = \beta(1 + r)s_t E_t(1 + \delta h_{t+1})c_{t+1}^{-\nu}. \tag{17}$$

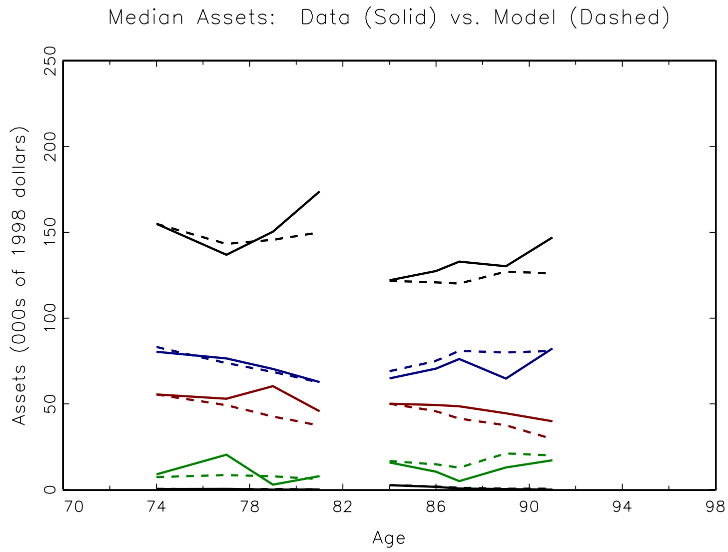

Figure 8: Median assets by cohort and PI quintile: data and model

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

$$E_t\left(\Delta \ln c_{t+1}\right) = \frac{1}{\nu}\left[\ln(\beta(1+r)s_t) + \delta E_t(h_{t+1} - h_t)\right] + \frac{\nu+1}{2} Var_t(\Delta \ln c_{t+1}). \tag{18}$$

Given that the survival rate, s_t, is often much less than 1, it follows from equation (18) that the model will generate downward-sloping, rather than flat, consumption profiles, unless the discount factor β is fairly large.

Our baseline estimated coefficient of relative risk aversion, ν, is 4.03. This parameter is identified by differences in saving rates across the permanent 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.

Our estimated coefficient of relative risk aversion falls within the range established by earlier studies. Our estimated coefficient is generally higher than the coefficients found by fitting non-retiree consumption trajectories,

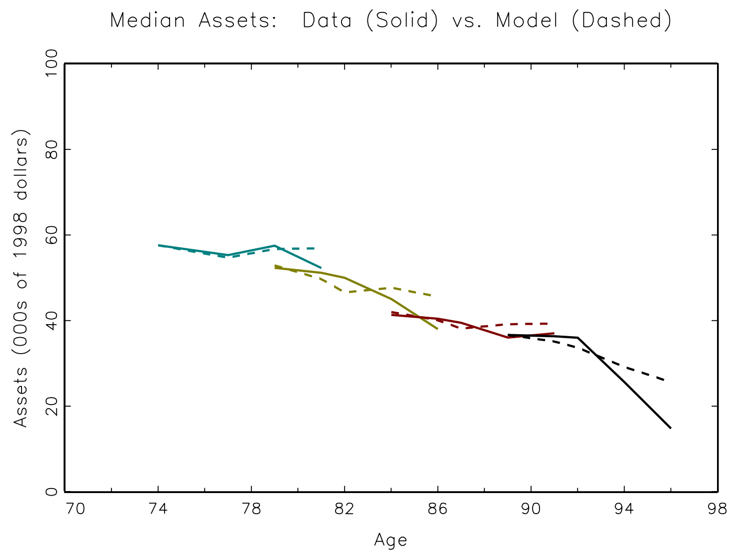

Figure 9: Median assets by birth cohort: data and model

either through Euler equation estimation (e.g., Attanasio, Banks, Meghir, Weber [2]) or through the method of simulated moments (Gourinchas and Parker [21]). Our estimated values are very much in line with those found by Cagetti [7] who matched wealth profiles with the method of simulated moments over the whole life cycle. Our estimated coefficient is lower than those produced by Palumbo [34], who matched consumption data using maximum likelihood estimation.[16] Given that our out-of-pocket medical expenditure data indicate more risk than that found by Palumbo, it is not surprising that we find less risk aversion.

We estimate that $\delta = -0.20$: holding consumption fixed, being in good health lowers the marginal utility of consumption by 20%, although we cannot reject that this parameter is significantly different to zero. Equation (18) shows that an anticipated change from good to bad health leads to consumption increasing by 5%. Note that as people age and health worsens, $E_t(h_{t+1} - h_t)$ becomes negative; multiplied by a negative delta, this implies that consumption growth increases as people age and become sicker. The

[16]It bears noting that most of these analyses do not contain a consumption floor. One notable exception is Palumbo: our estimated consumption floor of about $2,800 in 1998 dollars, is in real terms very close to Palumbo's floor of $2,000 in 1985 dollars.

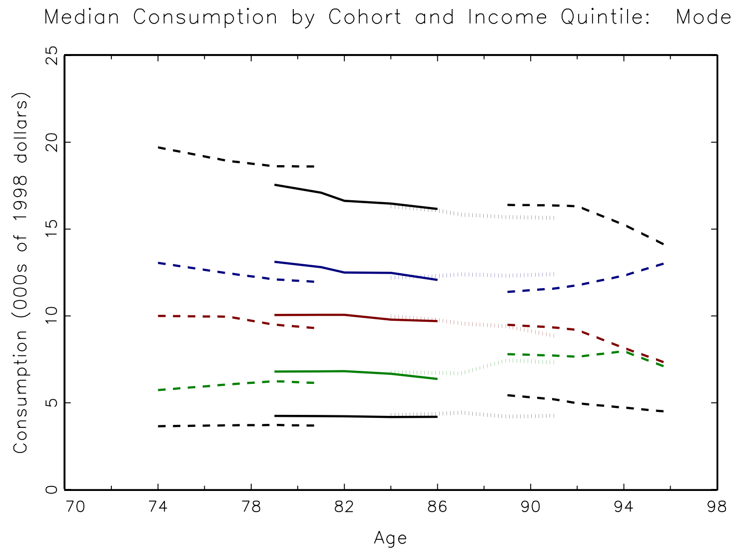

Figure 10: Consumption by cohort and PI quintile: model

data show that assets do decline more quickly at very old ages (see Figure 3), when people are most likely to be in bad health. A negative value of δ, accelerating asset decumulation at older ages, is consistent with this observation.

There is mixed evidence on whether bad health raises or lowers the marginal utility of consumption, holding consumption fixed. Lillard and Weiss [29] and Rust and Phelan [38] find that the marginal utility of consumption rises when in bad health, while Viscusi and Evans [42] find that it falls.

Given that the model uses income-, health- and sex-adjusted mortality profiles, its profiles should exhibit mortality biases similar to those found in the data. Figure 11 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 shows greater evidence of asset decumulation. A comparison of Figures 3 and 11 indicates that the size of the mortality bias generated by the model is very similar to the one in the observed data.

Figure 11: Median assets by birth cohort: everyone in the simulations (solid lines) vs. survivors (dashed lines)

6.2 Robustness checks

The remaining two columns of Table 3 present robustness checks on our benchmark estimates. Given that we do not directly measure the consumption or asset changes associated with bad health, one might question our estimate of δ. In addition, previous empirical evidence does not convincingly suggest that δ is greater than or less than 0. As a robustness check, we thus set $\delta = 0$ and re-estimate the other three parameters. These corresponding estimates are in the second column of Table 3. Setting δ to zero has very little effect on the other parameter estimates. This is consistent with our inability to reject that $\delta = 0$ in our baseline specification.

Next, we test whether our estimates are robust to our assumed consumption floor, which is meant to proxy for Medicaid health insurance (which largely eliminates medical expenses to the financially destitute) and Supplemental Security Income transfers. 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. For many individuals, however, the consumption floor is well above the SSI limits, because some individuals with income well 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, suggesting that the effective consumption floor is much lower.

In our benchmark case, we estimate our consumption floor to be about $2,800, which is similar to the value Palumbo [34] uses. However, this estimate is about half the size of the value that Hubbard et al. [22] find, and is also about half the average value of SSI benefits. Thus we may be underestimating the true consumption floor.

In the third column of Table 3, we present estimates based on a consumption floor of $5,000. Raising the consumption floor to $5,000 exposes consumers less risk: the model compensates by raising the estimated value of ν to 7.5. The corresponding estimates for the discount factor and utility shifter are basically unchanged. It bears noting that when the consumption floor is set exogenously to $5,000, the model fits the data much more poorly. The p-value for the overidentification statistic is much lower in this case, only 6.6%, compared to 97.8% for the baseline specification.

6.3 What are the important determinants of savings?

To determine the importance of the key mechanisms in our model we fix the estimated parameter values at their benchmark values and then change one feature of the model at a time. For each of these different economic environments we then 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 first shut down out-of-pocket medical expense risk, while keeping average medical expenditure (conditional on all of the relevant state variables) constant. Interestingly, and consistently with Hubbard, Skinner and Zeldes [22], we find that, conditional on constant average medical costs, the risk associated with medical expenses has only a small effect on the profiles of median wealth. Our results are also consistent with Palumbo's [34] finding that eliminating medical expense risk generates a modest increase in consumption, as a small increase in consumption translates into a small decrease in assets.

We then ask whether out-of-pocket medical expenditures of the size that we estimate from the data (and that are rising with age and permanent income) have quantitatively important effects on asset accumulation even for the elderly rich. We thus zero out medical all out-of-pocket medical expenditure for everyone and look at the corresponding profiles. This could be seen as an extreme form of insurance provided by the government.

Figure 12 shows that medical costs are a big determinant of the elderly's saving behavior, especially for those with high permanent income, for whom those costs 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 costs they expect to bear at the ends of their lives. This decomposition indicates that modeling out-of-pocket medical costs is important in evaluating policy proposals that affect the elderly, like Social Security reforms.

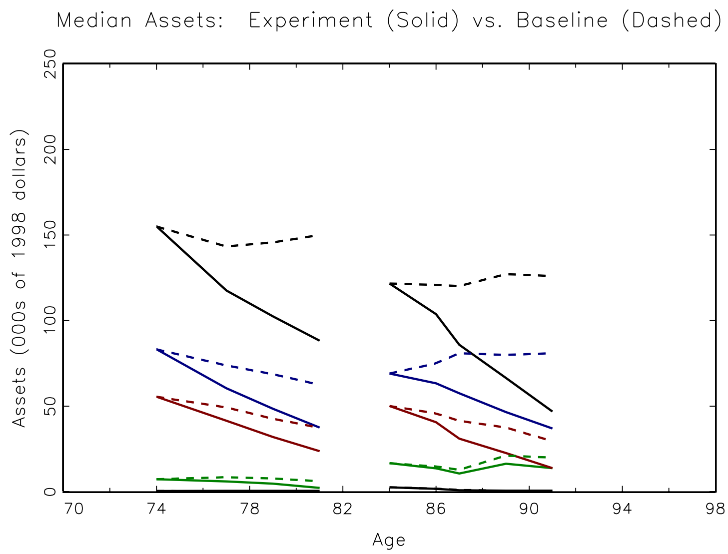

Figure 12: Median assets by cohort and PI quintile: baseline and model with no out-of-pocket medical expenditures

Next, we reduce the consumption floor to \$500. One could interpret this as a reform reducing the government-provided consumption safety net (in a partial equilibrium framework, since everything else is held constant). The effects of this change are large. Individuals respond to the increase in net income uncertainty by rapidly accumulating assets to self-insure. Figure 13 shows that this change affects the savings profiles of both low- and high-permanent-income singles. This indicates that the consumption floor matters for wealthy individuals as well as poor ones. This is perhaps unsurprising given the size of our estimated medical expenses; even wealthy households can be financially decimated by medical expenses.

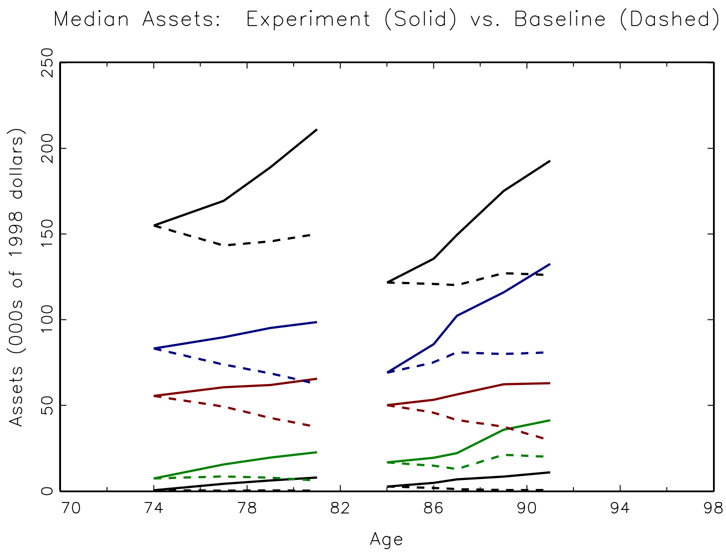

Figure 13: Median assets by cohort and PI quintile: baseline and model with a \$500 consumption floor

Finally, we turn to understanding the effect of differential life expectancy. As we have shown in Table 1, there are large differences in life expectancy by sex, permanent income, and health status. To understand the effect of this source of heterogeneity we generate asset profiles assuming that everyone faces the survival probability of a healthy male at the 50th percentile of the permanent income distribution. Figure 14 shows that, even over the short time period we are looking at, this difference in life expectancy would create a noticeable effect on asset accumulation, especially at the top end of the permanent income distribution.

What would happen if we were to assume that everyone has survival probabilities that depend only on age, but not on sex, health, or permanent income? Interestingly, we find that this would have negligible effects on the savings profiles, at least for a few years. This might indicate that there are countervailing forces that affect survival probabilities, and that these wash out for most people, even the rich. For example, males tend to be richer, so

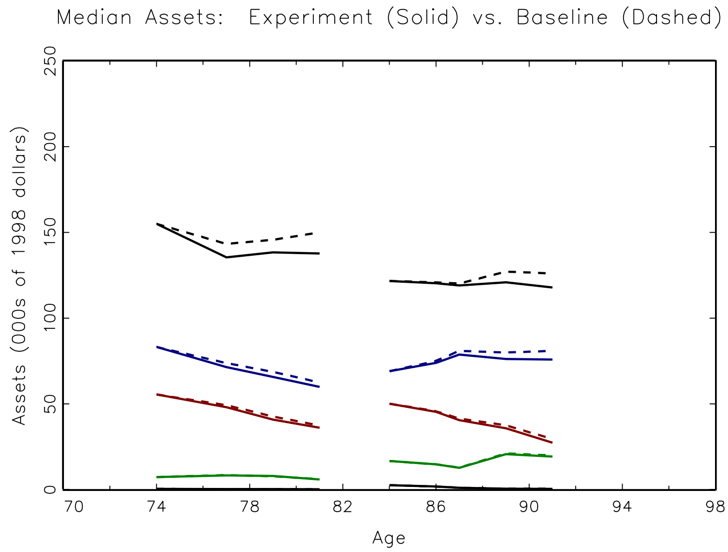

Figure 14: Median assets by cohort and PI quintile: baseline and model in which everyone faces the survival probability of a healthy male at the 50th percentile of the permanent income distribution

even if, controlling for permanent income, their expected survival is lower, the effect is counterbalanced by their higher permanent income. Figure 15 shows that the model fits the data very well even when we assume that age is the only variable affecting survival.

7 Conclusions

Our paper provides several contributions.

First, it estimates medical expenses and medical risk faced by the elderly using a better data set and a more flexible functional form. As a result, we find that medical expenses are much higher and more volatile than previously estimated, that they rise very fast with age, and that 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 elderly with higher permanent income.

Second, our paper carefully estimates mortality probabilities by age as a function of health, sex, and permanent income and finds large variations along all three dimensions.

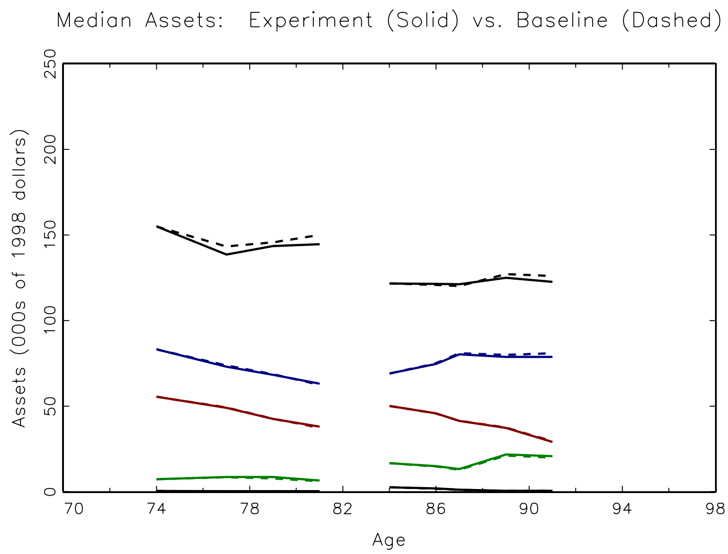

Figure 15: Median assets by cohort and PI quintile: baseline and model in which everyone faces the average survival probability

Third, our paper constructs and estimates a rich structural model of saving by using the method of simulated moments. As a result of our careful first step-estimation and of the richer sources of heterogeneity that we allow for in our model, we find that our parameter estimates are very reasonable, and, importantly, that our model provides a much better fit to the data than that previously obtained in the literature. In particular, our estimated structural model fits very well the saving profiles across the permanent income distribution, reproducing the observation that the dissaving rate of the elderly with higher permanent income is much smaller than the one of the elderly with lower permanent income.

Fourth, we find that the sources of heterogeneity that we consider have a significant role in explaining the elderly's saving behavior, with a very high level of medical expenses at very advanced ages being a key factor. Basically, if the single households live to very advanced ages, they are almost sure to face very large out-of-pocket medical costs, and they thus need to keep a large amount of assets (an amount increasing in permanent income, as medical expenses also increase) to self-insure against this risk.

Finally, we find that a publicly-provided consumption floor has a large effect on the asset profiles for all people, even those with high permanent income.

Our main conclusion is that to correctly evaluate any policy reform affecting the elderly's saving decisions, one needs to model accurately the consumption floor and, at a minimum, the average level of medical expenses by age and by permanent income.

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] 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.

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

[4] Orazio P. Attanasio and Hilary Williamson Hoynes. Differential mortality and wealth accumulation. *Journal of Human Resources*, 35(1):1–29, 2000.

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

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

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

[8] Marc A. Cohen, Eileen J. Tell, and Stanley S. Wallack. The lifetime risks and costs of nursing home use among the elderly. *Medical Care*, 24:1161–1172, 1986.

[9] 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.

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

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

[12] 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.

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

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

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

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

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

[18] Eric French and John Bailey Jones. The effects of health insurance and self-insurance on retirement behavior. Center for Retirement Research Working Paper WP 2004-12, 2004.

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

[20] Li Gan, Michael Hurd, and Daniel McFadden. Individual subjective survival curves. Working Paper 9480, National Bureau of Economic Research, 2003.

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

[22] 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.

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

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

[25] Michael D. Hurd. Subjective survival curves and life cycle behavior. In David Wise, editor, *Inquiries of Economics of Aging*. University of Chicago Press, 1998.

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

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

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

[29] 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.

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

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

[32] 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.

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

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

[35] 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.

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

[37] 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.

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

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

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

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

[42] 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.

Appendix A: Solving the model

We compute the value functions by backward induction. We discretize the persistent component and the transitory components of the health shock into Markovs Chain following Tauchen and Hussey (1991), and we assume that all other state variables lie on a finite grid. We solve the value function (and find the corresponding policy functions) at all of the points in our state space. We use linear interpolation within the grid and linear extrapolation outside of the grid to evaluate the value function at points that we do not directly compute.

The value function that we solve for can be written explicitly as

$$\begin{aligned} V_t(x_t, g, h_t, I, \zeta_t) = \max_{c_t} \Bigg\{ & u(c_t, h_t) + \beta s(g, h, I, t) \times \\ & \sum_{k=1}^{d_h} \sum_{l=1}^{d_\zeta} \sum_{n=1}^{d_\xi} \Big[\Pr(h_{t+1} = h_k | h_t, g, I, t) \Pr(\zeta_{t+1} = \zeta_l | \zeta_t) \Pr(\xi_{t+1} = \xi_n) \\ & \times V_{t+1}\big(x_{t+1}(k, l, n), g, h_{t+1}(k), I, \zeta_{t+1}(l)\big)\Big] \Bigg\}, \end{aligned}$$

subject to:

$$\begin{aligned} x_{t+1}(k, l, n) &= \max\Big\{x_t - c_t + y_n\big(r(x_t - c_t) + y_{t+1}, \tau\big) - m_{t+1}(k, l, n),\ c_{\min}\Big\}, \\ y_{t+1} &= y(g, I, t+1), \\ x_t &\geq c_{min}, \\ c_t &\leq x_t, \\ \ln\big(m_{t+1}(k, l, n)\big) &= m\big(g, h_{t+1}(k), I, t+1\big) + \sigma\big(g, h_{t+1}(k), I, t+1\big)\psi_{t+1}(l, n), \\ \psi_{t+1}(l, n) &= \zeta_{t+1}(l) + \xi_{t+1}(n), \end{aligned}$$

where $k \in \{1, ..., d_h\}$ indexes health status, $l \in \{1, ..., d_\zeta\}$ indexes persistent health cost shocks, and $n \in \{1, ..., d_\xi\}$ indexes transitory health cost shocks.

Appendix B: Moment conditions and the asymptotic distribution of parameter estimates

Our estimate, $\hat{\Delta}$, of the "true" 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 $t \in \{t_0, ..., t_T\} = \{1995, 1998, 2000, 2002\}$, we match median assets for 5 permanent income quintiles in 4 birth year cohorts. The 1995 (period-t_0) distribution of simulated assets, however, is bootstrapped from the 1995 data distribution, and is thus independent of the model's parameters. In the end we have a total of $20T = 60$ moment conditions.

The way in which we construct these moment conditions is similar to the approach described in French and Jones [18]. Let $q \in \{1, 2, ..., 5\}$ index permanent income quintiles. In this study, we convert permanent income, I, into a ordinal ranking lying in the $0 - 1$ interval. This transformation removes any sampling uncertainty over the boundaries of the permanent income quintiles, as the first quintile contains households with permanent income between 0 and 0.2, and so on. Suppose that individual i belongs to birth cohort c, and his permanent income level falls in the qth permanent income quintile. Let $a_{cqt}(\Delta, \chi)$ denote the model-predicted median asset level for individuals in individual I's group at time t. Assuming that observed assets have a continuous conditional density, a_{cqt} will satisfy

$$\Pr\big(a_{it} \leq a_{cqt}(\Delta_0, \chi_0)\,|c, q, t, \text{individual } i \text{ observed at } t\big) = 1/2.$$

As is well-known (see Manski [30], Powell [36] and Buchinsky [6]), the preceding equation can be rewritten as a moment condition. In particular, applying the indicator function produces

$$E\big(1\{a_{it} \leq a_{cqt}(\Delta_0, \chi_0)\} - 1/2\,|c, q, t, \text{individual } i \text{ observed at } t\big) = 0. \quad (19)$$

Equation (19) is merely equation (16) in the main text, adjusted to allow for "missing" as well as deceased individuals, as in French and Jones [19]. Continuing, we can convert this conditional moment equation into an unconditional one:

$$E\Big([1\{a_{it} \leq a_{cqt}(\Delta_0, \chi_0)\} - 1/2] \times 1\{c_i = c\} \\ \times 1\Big\{\frac{q-1}{Q} \leq I_i < \frac{q}{Q}\Big\} \times 1\{\text{individual } i \text{ observed at } t\} \;\Big|\; t\Big) = 0, \quad (20)$$

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

Suppose we have a data set of I independent individuals that are each observed at T separate calendar years. Let $\varphi(\Delta; \chi_0)$ denote the $20T$-element vector of moment conditions described immediately above, and let $\hat{\varphi}_I(.)$ denote its sample analog. Letting $\widehat{\mathbf{W}}_I$ denote a $20T \times 20T$ weighting matrix, the MSM estimator $\hat{\Delta}$ is given by

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

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 [33] and Duffie and Singleton [13], the MSM estimator $\hat{\theta}$ 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;

$$\mathbf{D} = \frac{\partial \varphi(\Delta; \chi_0)}{\partial \Delta'}\Big|_{\Delta = \Delta_0} \tag{21}$$

is the $20T \times 4$ gradient matrix of the population moment vector; and $\mathbf{W} = \text{plim}_{\rightarrow\infty}\{\widehat{\mathbf{W}}_I\}$. Moreover, Newey [31] shows that if the model is properly specified,

$$\frac{I}{1+\tau} \hat{\varphi}_I(\hat{\Delta}; \chi_0)' \mathbf{R}^{-1} \hat{\varphi}_I(\hat{\Delta}; \chi_0) \rightsquigarrow \chi^2_{20T-4},$$

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

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

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

$\mathbf{W} = \mathbf{S}^{-1}$, $\mathbf{V}$ simplifies to $(1+\tau)(\mathbf{D}'\mathbf{S}^{-1}\mathbf{D})^{-1}$, and $\mathbf{R}$ is replaced with $\mathbf{S}$. But even though the optimal weighting matrix is asymptotically efficient, it can be severely biased in small samples. (See, for example, Altonji and Segal [1].) We thus use a "diagonal" weighting matrix, as suggested by Pischke [35]. The diagonal weighting scheme uses the inverse of the matrix that is the same as $\mathbf{S}$ along the diagonal and has zeros off the diagonal of the matrix.

We estimate $\mathbf{D}$, $\mathbf{S}$ and $\mathbf{W}$ with their sample analogs. For example, our estimate of $\mathbf{S}$ is the $20T \times 20T$ estimated variance-covariance matrix of the sample data. When estimating preferences, we use sample statistics, so that $a_{cqt}(\Delta, \chi)$ is replaced with the sample median for group cqt. When computing the chi-square statistic and the standard errors, we use model predictions, so that the sample median for group cqt is replaced with its simulated counterpart, $a_{cqt}(\hat{\Delta}, \hat{\chi})$.

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 (20). 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 [33], Newey and McFadden [32] (section 7) and Powell [36].

To find $\mathbf{D}$, it is helpful to rewrite equation (20) as

$$\Pr\Big(c_i = c \;\&\; \frac{q-1}{Q} \leq I_i \leq \frac{q}{Q} \;\&\; \text{individual } i \text{ observed at } t\Big) \times \left[\int_{-\infty}^{a_{cqt}(\Delta_0, \chi_0)} f\Big(a_{it} \mid c, \frac{q-1}{Q} \leq I_i \leq \frac{q}{Q}, t\Big)\, da_{it} \;-\; \frac{1}{2}\right] = 0, \tag{22}$$

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

$$\Pr\Big(c_i = c \;\&\; \frac{q-1}{Q} \leq I_i \leq \frac{q}{Q} \;\&\; \text{individual } i \text{ observed at } t\Big) \times f\Big(a_{cqt} \mid c, \frac{q-1}{Q} \leq I_i \leq \frac{q}{Q}, t\Big) \times \frac{\partial a_{cqt}(\Delta_0; \chi_0)}{\partial \Delta'}. \tag{23}$$

In practice, we find $f\big(a_{cfqt} | c, q, t\big)$, the conditional p.d.f. of assets evaluated at the median a_{cqt}, with a kernel density estimator written by Ruud Koenig.