Subjective Life Expectancies, Time Preference Heterogeneity, and Wealth Inequality*

Richard Foltyn† — *Norwegian School of Economics*

Jonna Olsson‡ — *Norwegian School of Economics*

April 2024

This paper examines how objective and subjective heterogeneity in life expectancy affects savings behavior of healthy and unhealthy people. Using data from the Health and Retirement Study, we first document systematic biases in survival beliefs across self-reported health: those in poor health not only have a shorter actual lifespan but also underestimate their remaining life time. To gauge the effect on savings behavior and wealth accumulation, we use an overlapping-generations model where survival probabilities and beliefs evolve according to a health and survival process estimated from data. We conclude that differences in life expectancy are important to understand savings behavior, and that the belief biases, especially among the unhealthy, can explain up to a fifth of the observed health-wealth gap.

JEL Classification: D15, E21, G41, I14
Keywords: Life expectancy, preference heterogeneity, subjective beliefs, life cycle

1 Introduction

The determinants of the wealth distribution are of fundamental interest to economists. Standard consumption/savings theory predicts that people who place a larger weight on

*We are grateful to James Banks, Timo Boppart, Per Krusell, Alexander Ludwig, Hannes Malmberg, Ayşegül Şahin, Paolo Sodini, Christian Stoltenberg, Magnus Åhl, Erik Öberg and participants in numerous conferences and seminars for helpful discussions and comments. The feedback and suggestions of the editor and four anonymous referees improved the paper considerably. The collection of data used in this study was partly supported by the National Institutes of Health under grant number R01 HD069609 and R01 AG040213, and the National Science Foundation under award numbers SES 1157698 and 1623684.

†richard.foltyn@nhh.no

‡jonna.olsson@nhh.no, Norwegian School of Economics, Helleveien 30, 5045 Bergen, Norway.

future states will be wealthier than people who are more impatient, all else equal. This paper explores one reason to put a higher weight on the future: the higher probability to survive to old age.

However, an individual's consumption/savings decision is not necessarily guided by the objective (statistical) survival probability but rather the individual's beliefs about survival. The first contribution of this paper is to document new facts about a *within-cohort steepness bias* in survival beliefs: people overestimate the health gradient of survival.

It has previously been shown (e.g., Hamermesh 1985, Elder 2013, Ludwig and Zimper 2013, Heimer, Myrseth, and Schoenle 2019) that there is a systematic flatness bias over age: younger people tend to underestimate their survival probabilities, while older people overestimate their chances of a long life. We show that within a cohort, individuals in bad health not only have a shorter expected life span but are also relatively more downward biased about their survival chances, while individuals in good health and thus with higher survival probability display an upward bias. These systematic biases exacerbate the life expectancy heterogeneity in the population.

The differences in beliefs about survival translate into time preference heterogeneity in the population. Our second contribution is to quantify this heterogeneity and its implications for savings and wealth accumulation in an overlapping-generations model. With a stochastic health and survival process, the effective discount rate varies depending on age, health and the forecast horizon. Over a one-year horizon, the effective discount rate for 50-year-olds ranges from 2% for an individual in best health to around 20% for an individual in worst health. At a 10-year horizon, this gap shrinks somewhat but still amounts to eight percentage points between best and worst health. For 70-year-olds in worst vs. best health state, the difference at the 10-year horizon is close to ten percentage points. This resulting time preference heterogeneity is in line with the dispersion (Calvet et al. 2021) and the age gradient (Kureishi et al. 2021) of the time preference distribution found in other empirical studies.

To gauge the quantitative effect of survival heterogeneity on savings behavior and wealth accumulation, we use an overlapping-generations general-equilibrium model with uninsurable idiosyncratic shocks. Agents face heterogeneous survival risk that depends on their age and current health state, and are subject to health shocks that follow a process estimated from data. The current health state also affects labor earnings and medical expenditure risk. Besides this uncertainty, we additionally include standard persistent and transitory shocks to labor productivity during working age. After agents reach a fixed retirement age, they are entitled to retirement benefits mimicking the US social security system. Finally, our model includes probabilistic bequests that feature

intergenerational persistence of income and wealth. We purposely use an otherwise standard model of consumption/savings to establish a benchmark and focus on the survival heterogeneity savings channel.

We compare three scenarios. The first scenario is a standard model in which there is no health risk: all agents face the same labor earnings risk, the same medical expenditure risk, and the same survival risk and thus have the same effective discount factor, conditional on age. In the second scenario, we introduce health risk that affects labor earnings and medical expenditures. In terms of survival risk, individuals are perfectly informed about their true survival probability conditional on health and age. In the third scenario, agents believe and act according to their subjective survival beliefs. Thus, our analysis is designed to answer the question: what if we turned off the discount factor heterogeneity implied by the biases in survival beliefs we uncovered in the empirical part? Would savings patterns look quantitatively different?

The simulations show that the survival expectation channel is important for understanding wealth accumulation. Not surprisingly, agents in bad health and thus with a shorter expected life span save less than their healthy counterparts, and the differences in savings rates are large. For example, for 60-year-olds in the middle of the wealth distribution, the total savings rates of an agent in the best and an agent in the worst health state differ by 5 percentage points when they are endowed with correct objective beliefs about survival. When we let them act according to the estimated subjective beliefs instead, the difference doubles to 10 percentage points.

These differences in savings behavior translate into large differences in accumulated wealth: in the model with subjective survival beliefs, median wealth differs by 193% between those in the worst and best health states at ages 55–59. This health-wealth gradient is very close to the magnitude we observe in the data. A fifth of this difference is driven by the erroneous survival beliefs, especially by individuals in poor health underestimating their remaining life span. Thus, the biases in survival beliefs are important to understand the health-wealth gradient in older ages.

This paper speaks to three broad strands of literature. The first is concerned with subjective survival expectations (Hamermesh 1985; Smith, Taylor, and Sloan 2001; Hurd and McGarry 2002; Ludwig and Zimper 2013; Elder 2013; Gan et al. 2015; Groneck, Ludwig, and Zimper 2016; Heimer, Myrseth, and Schoenle 2019; Bresser 2023). Many studies have documented the existence of an age bias in subjective life expectancies, and a few of the papers within this group are concerned with the implications for the consumption/savings behavior. Some predict individual survival probabilities and contrast them with elicited beliefs (Gan, Hurd, and McFadden 2005; Bissonnette, Hurd,

and Michaud 2017; Grevenbrock et al. 2021), but none of these look at the implications for within-cohort savings behavior in a structural model where beliefs change in the event of health shocks, or analyze the implications for wealth inequality.

The second strand are macroeconomic studies pointing out the importance of heterogeneity in time preferences to explain wealth inequality (e.g., Krusell and Smith Jr. 1998; Hendricks 2007; Quadrini and Ríos-Rull 2015; Krueger, Mitman, and Perri 2016) and studies documenting time preference heterogeneity in the population (Epper et al. 2020; Calvet et al. 2021). Compared to these papers, we provide a micro-foundation for one source of time preference heterogeneity — differences in life expectancy — and evaluate its importance.

The third is the literature about the general impact of health (including life expectancy) on wealth (Smith 1999; Lee and Kim 2008; Coile and Milligan 2009; De Nardi, French, and Jones 2009; Kopecky and Koreshkova 2014; Capatina 2015; De Nardi, Pashchenko, and Porapakkarm 2017; Poterba, Venti, and Wise 2017; Margaris and Wallenius 2023, to name a few). In contrast to these papers, we include heterogeneity in *subjective* life expectancy and examine its impact on savings and consumption behavior.

In the next section, we describe how we estimate the health and survival process and give details about the systematic bias in survival expectations. Section three describes the model we use to quantify the importance of the heterogeneity in survival expectations. After that, we discuss the parametrization and then we present our results. The last section concludes.

2 Empirical evidence

2.1 Data

We use the Health and Retirement Study (HRS), a representative panel of elderly US households, to investigate the evolution of health and longevity in the later stages of life. The survey includes questions about self-reported health and expectations about survival, and records the date of death, if applicable.

Our analysis is based on the survey years 1992–2014 taken from the HRS data compiled by RAND, version 2018 (V2) (Health and Retirement Study (2023)).[1] The first cohort included in the survey was between 51 and 61 years old in 1992, and thereafter new (older and younger) cohorts were added. Many of the respondents died over the sample period, making it an ideal data set for studying survival.

[1]The HRS (Health and Retirement Study) is sponsored by the National Institute on Aging (grant number NIA U01AG009740) and is conducted by the University of Michigan.

In this section, we first document the relationship between health and wealth, and between beliefs about survival and wealth. We then briefly describe how we estimate the objective survival probabilities. In section 2.4 we show how average elicited beliefs about survival are biased, and in section 2.5 we estimate a subjective life expectancy process that replicates this bias.

2.2 The health-wealth gradient and the life expectancy/savings channel

The HRS asks participants to assess their health using one of the five categories *excellent*, *very good*, *good*, *fair*, or *poor*. Figure 1 shows net total wealth over the life cycle by self-reported health state computed for the pooled sample of all respondents.[2,3] The health-wealth gradient is well documented, but the underlying causal relationship is debated (Attanasio and Hoynes 2000; Deaton 2002; Duncan et al. 2002; Attanasio and Emmerson 2003; Hajat et al. 2010). One line of argument is that low economic status leads to poor health. There could be many reasons: poor people have access to less or lower-quality medical care, do not invest enough in preventive health measures, and/or have more health-deteriorating habits. However, there are also many arguments for the reversed causality: poor health has economic consequences in itself. First, poor health may restrict the individual's earnings potential by making it more costly to work and/or by lowering the wage. Second, poor health may lead to large medical expenditures. Third, poor health may lower the savings incentives due to a lower survival expectancy. This last channel is the focus of this paper.

If individuals adjust their savings behavior based on their survival prospects, this could be either on the basis of objective (statistical) survival probabilities or subjective survival beliefs, which are also surveyed by the HRS. To assess how wealth correlates with survival beliefs, we regress net total wealth on an indicator of whether an individual believes to have above-median survival chances compared to other respondents of the same age, race and sex. Table 1 shows a positive correlation between having above-median beliefs and being wealthier.[4] The positive relationship also holds when additionally controlling for education and couple status.[5] Other empirical studies cor-

[2] Net total wealth is defined as sum of housing, other real estate, vehicles, businesses, IRA and Keogh accounts, stocks, checkings, and all other savings, net of mortgages and other debts.

[3] In online appendix section A.4, we disaggregate these wealth profiles by race, sex, household size and education. The overall picture remains unchanged.

[4] All empirical results in this paper are reported with standard errors and confidence intervals that take into account the stratification and clustering of the HRS, see online appendix section A.3.

[5] We apply an inverse hyperbolic sine transformation since assets are heavily skewed and contain zeros and negative values. All coefficients of interest are positive and significant at the 1% level when alternatively using assets in levels. Online appendix section A.5.1 contains further information and robustness checks.

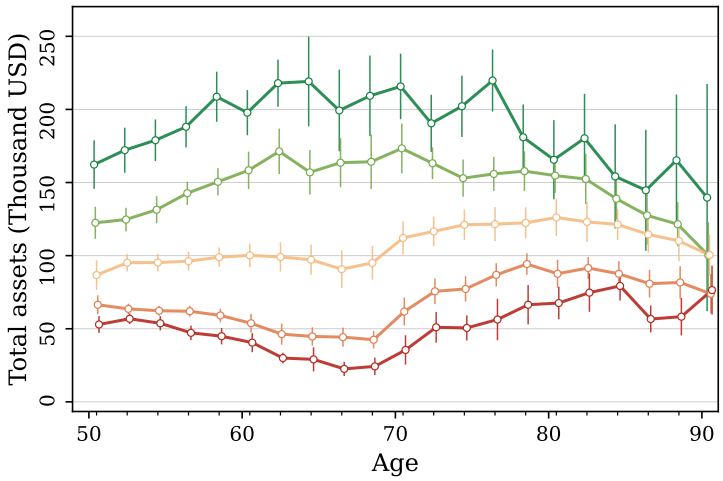


Figure 1: Median net total household wealth by self-reported health state. Pooled sample from HRS 1992–2014. Assets are adjusted for outliers, time and cohort fixed effects. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

roborate the existence of the life expectancy/savings channel and suggest a causal link. For instance, Heimer, Myrseth, and Schoenle (2019) administer a novel survey and estimate that greater survival optimism correlates with higher savings rates, not only after controlling for standard demographic characteristics such as education, marital status, and income, but also financial literacy and risk tolerance.

Another prediction of the life expectancy/savings channel is that individuals who receive a bad health shock, i.e., a plausible decrease in life expectancy, should exhibit lower asset growth. Table 2 reports the results from regressing the two-year change in net total wealth (again using an inverse hyperbolic sine transformation) on a negative health shock defined as an indicator for a deterioration in self-reported health between survey waves. As column 1 shows, men who experience a negative health shock decumulate their assets more compared to men of the same race, age and initial health who don't. Column 2 additionally controls for education, while columns 3–4 show the corresponding results for women. A negative health shock is associated with a faster decumulation (or slower accumulation) of assets in all specifications. More details are given in online appendix section A.5.2.[6]

While these results are indicative, a recent study by Kvaerner (2022) using the plausi-

[6]It is possible that the decumulation of assets associated with a health deterioration is driven by lower labor income or large medical expenditures. In online appendix section A.5.2, we show that the results also hold in the subsample aged 65 and older (who are likely to be retired and on Medicare) even after accounting for out-of-pocket medical expenditures.

	Dep. variable: net total wealth (inverse hyperbolic sine transformation)					
	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
Above median SSB	0.418***	0.230***	0.221***	0.575***	0.374***	0.353***
	(0.045)	(0.040)	(0.039)	(0.032)	(0.029)	(0.029)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Education FE		Yes	Yes		Yes	Yes
Couple FE			Yes			Yes
Observations	54,212	54,212	54,212	70,537	70,537	70,537

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1: Net total wealth and above-median subjective survival beliefs. The table shows the results of regressing net total wealth (after an inverse hyperbolic sign transformation) on an indicator of above-median survival beliefs compared to individuals of the same age, race and sex. The regression includes fully interacted fixed effects as indicated. Nonblack population.

bly exogenous timing of cancer diagnoses shows that news about a bad health shock increases the probability of an immediate inter vivos transfer, suggesting a causal link between survival prospects and wealth decumulation. Thus, if a decrease in life expectancy increases own consumption and/or increases the probability of inter vivos transfers is an open question, and one interpretation of inter vivos transfers from the giver's perspective is to view them as "consumption of gift-giving." While both inter vivos transfers and own consumption show up as a decumulation of assets and thereby affect the health-wealth gradient in older ages in the same way, they have different implications for wealth among the younger receiving generation and thus for the wealth distribution. Our structural model does not allow for inter-vivo transfers and consequently leaves this question open for future research.

2.3 Objective health and survival probabilities

In this paper, we examine the effect of heterogeneity in survival expectancy on savings behavior and its implications for wealth inequality through the lens of a structural model. Therefore, we need to formulate heterogeneity in survival expectations, both objective and subjective, in a way that can be used in such a model.

For our quantitative model, we use a Markov process for health transitions and survival at an annual frequency. We estimate this Markov process as described in Foltyn and Olsson (2021). Conceptually, the method is a straightforward maximum likelihood estimator where the probability of observing the transitions in the data is maximized. We briefly summarize the method and estimation sample in the next few paragraphs.

	Dep. variable: relative change in net total wealth			
	Men		Women	
	(1)	(2)	(3)	(4)
Negative health shock	-0.149***	-0.130***	-0.134***	-0.115***
	(0.023)	(0.023)	(0.018)	(0.018)
Age FE	Yes	Yes	Yes	Yes
Health FE	Yes	Yes	Yes	Yes
Education FE		Yes		Yes
Observations	61,528	61,528	80,615	80,615

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 2: Health shocks and changes in wealth. The table shows the results of regressing changes in net total wealth (after an inverse hyperbolic sign transformation) on an indicator for a deterioration in self-reported health between two consecutive survey waves. The regression includes fully interacted fixed effects as indicated. Nonblack population.

To put structure on the Markov process, we follow Pijoan-Mas and Ríos-Rull (2014) and use a logit model, where survival and health transitions conditional on survival are modeled as functions of the current health state and age. The probability of survival follows the usual binary logit model while, conditional on survival, health transitions are modeled using multinomial logit. For example, the one-period-ahead survival probability is given by

$$p^s_{t+1} = \frac{1}{1 + e^{-g(\boldsymbol{x}_t|\gamma)}} \tag{1}$$

where $g(\bullet)$ is a function of the covariate vector $\boldsymbol{x}_t$ which contains race, sex, age, health and and potentially other observables such as education. Survival probabilities are governed by the parameter vector γ to be estimated. Transition probabilities for health conditional on survival are defined in an analogous manner.[7]

Estimation sample. We exclude all observations with missing age, race, sex or self-reported health, as well as individuals with only a single observation (since then we do not have any transition probability to estimate). We only consider individuals aged 50 or older.[8] Furthermore, we restrict the sample to maximum age of 99 years at transition start (even though individuals can be older when we observe them in the end of a transition). This leaves us with 34,196 individuals and 219,539 observations in total.

[7] In Foltyn and Olsson (2021), we provide details about the estimation and also perform an extensive evaluation of the results. The estimated Markov process is shown to predict actual mortality very well, both short- and long-term. See online appendix section B for a brief overview.

[8] Each incoming HRS cohort is aged 51 or above, but the survey contains younger individuals who are spouses of age-eligible respondents.

We estimate the health and objective (statistical) survival process separately for the subsamples of men/women and the black/nonblack population, since it is well known that the life expectancies for these groups follow very different trajectories.[9] Table B.1 in the online appendix shows descriptive statistics and the number of individuals and observations by subgroup.

Results. From these estimates, we construct a first-order Markov process defined on five health states and the absorbing state of death which governs the objective health and survival probabilities. This process can be used to calculate objective life expectancies conditional on age, health, race, and sex. Not surprisingly, there is a substantial health gradient in life expectancy. For example, for a 70-year old nonblack men in excellent health, the predicted probability of surviving an additional 10 years is approximately 75%, while the probability is just around 35% if instead starting out in poor health. For a brief overview of the results, see online appendix section B.2.

It is important to note that even though the health and survival process is based on self-reported health — a subjective measure of how respondents perceive their health state — the result from the estimation is an objective statistical life expectancy for each combination of race, sex, age, and health. Self-reported health can be thought of as letting the respondents themselves aggregate the multidimensional information about their health (that is potentially unobservable to the econometrician) into a single categorical variable, and the variable can also capture subjective perceptions of the respondent. The estimated Markov process maximizes the probability of observing the health transitions and survival in the data conditional on self-reported health, irrespective of why a particular health state was reported.

2.4 Expectation errors in survival probabilities

In the expectations survey module of the HRS, respondents are asked about the probability they assign to certain events. One of these questions is about the probability of surviving to a certain age, for example: "Using a number from 0 to 100, what do you think are the chances that you will live to be at least 100 years?"[10]

[9] For the remainder of the paper, the "black" sample consists of respondents who identify as black or African-American, while "nonblack" is the complementary group which also includes Hispanics. The HRS is not large enough to disaggregate the nonblack group further, since the (unweighted) sample of person-year observations is approximately 72.7% white, 15.7% black/African-American, 9.4% Hispanic, with other ethnicities together contributing the remaining 2.3%.

[10] Before the respondent answers the questions about expectations, the interviewer discusses probabilities and verifies that the respondent understands the concept.

	All	Nonblack		Black	
		Male	Female	Male	Female
Sample size					
N. of individuals	32,590	12,073	14,884	2,274	3,359
N. of observations	252,299	92,013	120,481	14,914	24,891
Avg. observations/indiv.	7.7	7.6	8.1	6.6	7.4
Forecast horizon (years)					
Min.	10.0	10.0	10.0	10.0	10.0
Mean	19.4	19.5	19.2	20.2	19.8
Max.	39.0	35.0	39.0	35.0	35.0
Target age distribution					
Age 75	37.2%	37.7%	36.3%	41.0%	39.5%
Age 80	16.3%	16.3%	16.4%	15.3%	16.1%
Age 85	35.1%	35.6%	34.4%	37.1%	36.2%
Age 90	5.6%	5.3%	6.1%	3.7%	4.3%
Age 95	3.9%	3.5%	4.5%	2.0%	2.6%
Age 100	2.0%	1.7%	2.4%	0.9%	1.3%

Table 3: Descriptive statistics for elicited subjective survival beliefs. Mean forecast horizon and distribution over target ages are weighted using HRS sample weights.

The exact target age depends on the respondent's age and survey wave. For instance, in 1995, respondents below the age of 70 were asked about the probability of living until the age of 80, while respondents above the age of 85 were asked about the target age of 100. In later surveys, individuals were asked about survival beliefs for up to two target ages.

Table 3 shows the number of individuals and observations (an observation being one elicited survival belief) by subgroup. The tabulated distribution of target ages shows that the questions about survival to age 75 or 85 are by far the most common, hence we focus on these two in the main text.

Using these elicited beliefs, we compare the average probability that individuals of a certain age assign to survival until a given target age to the probability according to official life tables. As can be seen in Figure 2, there is a systematic error along the age gradient: younger individuals on average tend to underestimate, while older individuals tend to overestimate their survival probability as compared to objective life table estimates.[11]

This age-dependent error is a stylized fact in the literature on survival expectations

[11] A potential concern is that the sample of individuals answering the subjective beliefs question is not representative, and thus any differences between subjective beliefs and life tables arise due to selection. In online appendix section C.1, we show that the same pattern emerges if we instead compare subjective beliefs to objective survival probabilities estimated on the same sample.

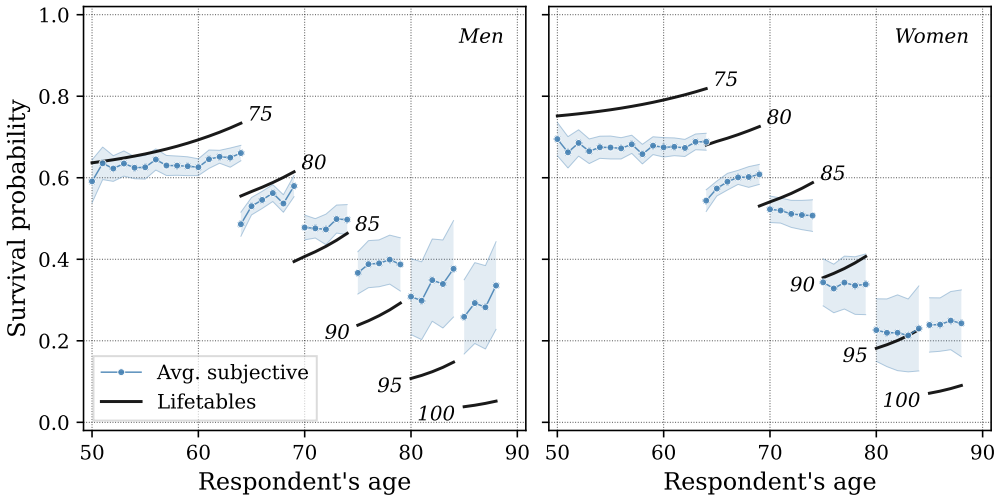
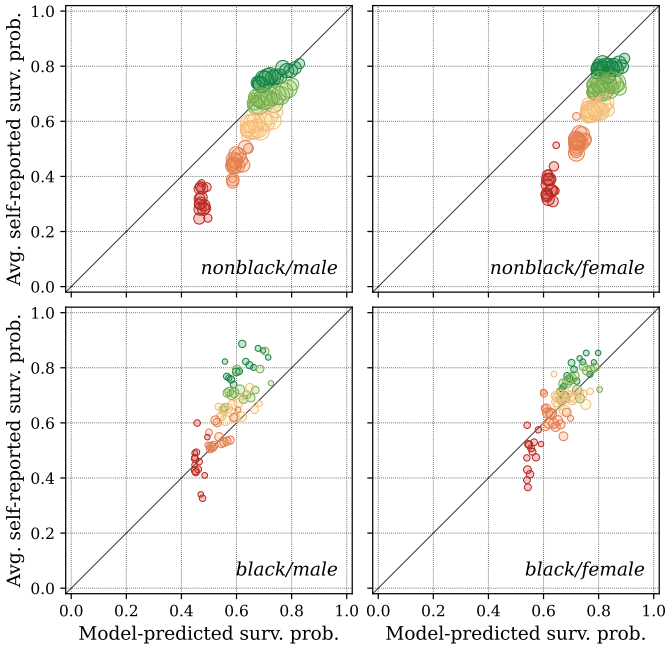


Figure 2: Objective vs. subjective survival probabilities by age and sex (controlling for year and cohort fixed effects). The number next to the black line indicates target age. The blue line shows (weighted) average expectation in the nonblack population. Shaded areas indicate 95% confidence intervals.. Only age bins with at least 50 observations are shown.

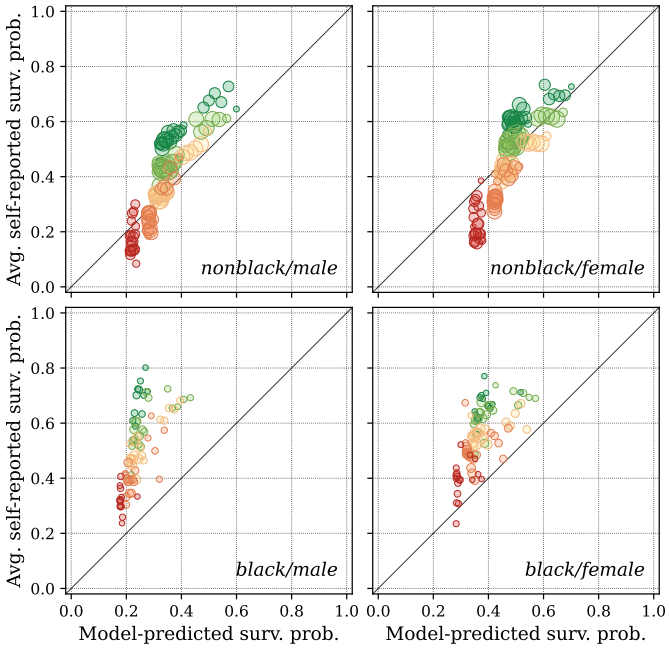
(see, e.g., Ludwig and Zimper (2013), Groneck, Ludwig, and Zimper (2016), and Heimer, Myrseth, and Schoenle (2019) and the references therein). The pattern has been used to, e.g., improve the fit of the asset profile of the canonical life cycle model with the data: due to underestimation early in life, young agents do not accumulate as much wealth, while overestimation in later years dampens the rate at which assets are decumulated.

Since our estimator conditions on health, race and sex, we can go one step further and document survival belief differences along these dimensions. Figure 3 shows these expectation errors among respondents who have answered the question about their perceived probability of survival until target ages 75 and 85. The first observation is the striking positive correlation between subjective self-reported survival probability and the predicted (objective) survival probability, which means that subjective beliefs are informative. This is in line with the consensus in the literature, which finds that subjective beliefs are highly correlated with objective survival probabilities and serve as predictors of mortality, and that expectations are updated in the event of health shocks (Smith, Taylor, and Sloan 2001; Hurd and McGarry 2002; Gan, Hurd, and McFadden 2005).

The second observation is the systematic steepness bias in beliefs over the health gradient. As was shown in Figure 2, on average individuals underestimate their probability of survival until the age of 75. Looking at Figure 3(a), and focusing on nonblack males, it



(a) Target age 75



(b) Target age 85

Figure 3: Elicited beliefs about survival vs. estimated objective (statistical) survival probabilities. Each bubble represents the average for a race/sex/age/health group. The x-axis shows the model-predicted (objective) survival probability to the indicated target age. The y-axis shows the average self-reported survival probability for that group. Colors indicate the health state: dark green is excellent while red is poor health. The size indicates the number of observations in each cell. We exclude cells with less than 20 observations.

is clear that it is mainly individuals in bad health who are underestimating their survival probability, while individuals in excellent health are on average reasonably close to their objective survival probability.

Figure 3(b) shows expectation errors for target age 85. Again, individuals in bad health underestimate their survival probability more than those in good health, even though on average the expectations are more upward biased for this target age. Figure C.2 in the appendix shows the corresponding graphs for target age 95. The figure illustrates that for those higher ages individuals in bad health have beliefs closer to their objective probability, while individuals in good or excellent health are severely overestimating their survival chances. In online appendix section C.3, we additionally disaggregate the expectation errors by education. The overall picture remains unchanged.

Next, we impose some structure on the biased survival beliefs documented in these scatter plots. To this end, we define survival bias as the elicited subjective survival belief minus the objective survival probability (thus an underestimation of the survival probability is negative bias). For example, for some target age j we define the bias as

$$\Delta_{ijgt} \equiv p_{ijgt}^{subj} - \Pr\left(\text{ alive at } j \mid g, h_{it}, age_{it}\right)$$

for individual i in demographic group g (nonblack/male, nonblack/female, black/male, or black/female) when the individual is in health state h_{it}. We first quantify the bias observed over the life cycle by estimating the regression

$$\Delta_{ijgt} = \beta_{0jg} + \beta_{1jg} \cdot \text{age}_{it} + \beta_{2jg} \cdot \text{age}_{it}^2 + \varepsilon_{ijgt} \tag{2}$$

for each target age separately, where we allow the coefficients of the age polynomial to vary by race and sex. In Figure 4, we plot the predicted survival bias for each age and race/sex group for the target ages 75 and 85. As the figure shows, men and black individuals are on average more upward biased than women and nonblack individuals. This confirms the findings by Bissonnette, Hurd, and Michaud (2017).

In a second step, we disaggregate the bias by initial health. We estimate the following regression separately for each demographic group,

$$\Delta_{ijght} = \beta_{0jgh} + \beta_{1jgh} \cdot \text{age}_{it} + \beta_{2jgh} \cdot \text{age}_{it}^2 + \varepsilon_{ijht} \tag{3}$$

where we interact the initial health state h with a quadratic polynomial in age. The predicted values for this exercise for target ages $j \in \{75, 85\}$ are shown in Figure 5 for the nonblack population: individuals in bad health generally display a more negative

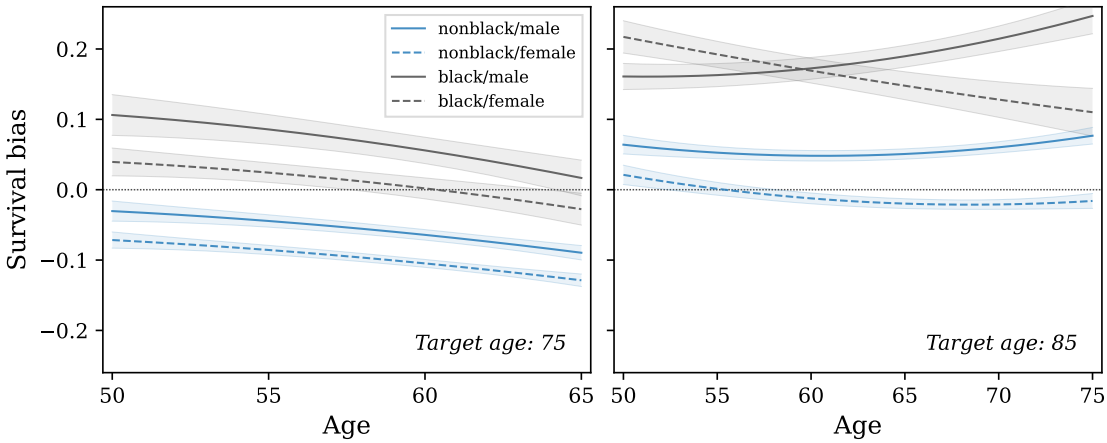


Figure 4: Survival bias by race, sex and age. The graph shows the predicted $\widehat{\Delta}_{ijgt}$ from equation (2). A positive bias of 0.1 indicates that the group is overestimating their survival probability by 10 percentage points. Error bars indicate 95% confidence intervals.

bias than individuals in better health.[12]

In a related paper, Grevenbrock et al. (2021) estimate survival based on several additional characteristics besides self-reported health, age, and sex, such as smoking and drinking behavior, and chronic diseases. Grouping individuals based on their estimated objective survival probability, they find that individuals with low objective survival probability in general overestimate, while individuals with high objective probability underestimate their survival probabilities, in other words, at a first glance the reverse pattern compared to what we find. However, there are two reasons for why our results are not directly comparable. First, they use a different estimation strategy for objective survival probabilities. Second, and more important, the grouping of individuals is different. We group based on age, sex, race and self-reported health, which is the level of heterogeneity in our economic model (and thus, for our purposes, the most appropriate one) and show that, conditional on age, sex and race, individuals with lower survival probabilities (poorer health) tend to under-predict their survival. Conversely, Grevenbrock et al. (2021) show that, when considering differences in survival due to age or permanent characteristics such as sex or race, groups with lower survival rates tend to over-predict their survival. For a more comprehensive discussion, see online appendix section C.5.

More generally, the finding that individuals with high life expectancy display an upward bias in their survival beliefs is in line with evidence about forecast errors in

[12] Figure C.6 in the online appendix shows the corresponding results for the black population. The pattern is the same, but standard errors are larger due to smaller sample size.

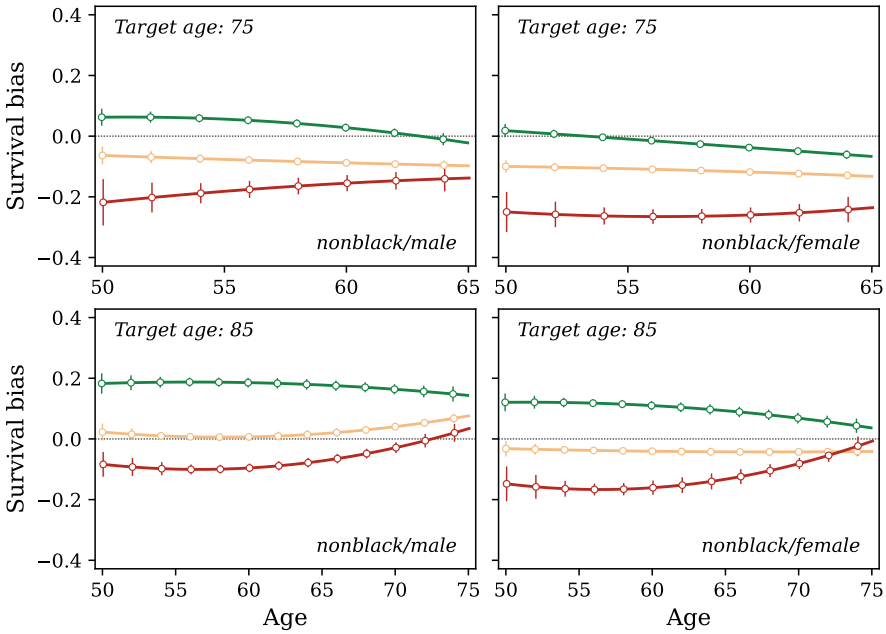


Figure 5: Survival bias by sex, age and self-reported health for the nonblack population. The graph shows the predicted $\widehat{\Delta}_{ijght}$ from equation (3). A positive bias of 0.1 indicates that the group is overestimating their survival probability by 10 percentage points. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

other domains. For example, Rozsypal and Schlafmann (2023) document that people in the upper part of the income distribution overestimate their future income growth while the opposite is true for lower-income households.

To summarize, we stress two observations: first, subjective beliefs are informative and correlated with objective probabilities. Second, subjective beliefs are biased. Subjective probabilities overestimate the health/survival gradient, with individuals in bad health underestimating their survival probability relative to individuals in good health. Hence, there is a systematic bias both along the age and health dimensions.

2.5 Estimation of the subjective life expectancy process

In this section, we use the health transitions estimated in Foltyn and Olsson (2021) as a basis for estimating a different set of parameters that govern *subjective* survival beliefs. The difference between objective probabilities and subjective beliefs could arise due to erroneous beliefs about survival conditional on health, erroneous beliefs about the health process, or a combination of the two. Without further information we cannot distinguish between those alternatives, and the HRS does not elicit any beliefs about

future health states. To keep the model reasonably parsimonious, we take as given the parameters controlling health-to-health transitions conditional on survival and estimate a different set of survival probabilities to capture the elicited beliefs.

We take an agnostic approach as to *why* the erroneous beliefs arise. In the literature, various mechanisms have been proposed, such as Choquet Bayesian learning models of survival beliefs which allow for likelihood insensitivity (Ludwig and Zimper 2013; Groneck, Ludwig, and Zimper 2016) and, closely related, age dependent cognitive weakness and relative optimism (Grevenbrock et al. 2021) as well as overweighting the likelihood of rare events (Heimer, Myrseth, and Schoenle 2019).

As explained above, the underlying data for this exercise takes the following form: HRS respondents are asked at date t to state their beliefs about surviving to a certain target age j (for example 75 or 85), which we reinterpret as the probability of being alive in T_{ijt} periods, with $T_{ijt} = j - age_{it}$. We treat multiple observations from one individual independently: say a respondent i is surveyed on survival beliefs to target ages j_1 and j_2 in calendar years t_1 and t_2. This gives rise to the data

$$\begin{aligned}&\left(\boldsymbol{x}_{it_1}, p_{ij_1t_1}, T_{ij_1t_1}\right), \left(\boldsymbol{x}_{it_1}, p_{ij_2t_1}, T_{ij_2t_1}\right)\\&\left(\boldsymbol{x}_{it_2}, p_{ij_1t_2}, T_{ij_1t_2}\right), \left(\boldsymbol{x}_{it_2}, p_{ij_2t_2}, T_{ij_2t_2}\right)\end{aligned}$$

where $\boldsymbol{x}_{it}$ are vectors of covariates including race, sex, age, health and possibly other observables, and p_{ijt} are elicited survival beliefs. We treat these as four independent observations, except when bootstrapping confidence intervals which we cluster as described in online appendix section A.3.

Assume that the i-th individual forms T-year-ahead survival beliefs based on the model

$$p_{it}^s = \phi_T\left(\boldsymbol{x}_{it}\right) \tag{4}$$

where ϕ_T is an unknown nonlinear function that maps $\boldsymbol{x}$ into $[0, 1]$.

In what follows, we partition the sample into groups indexed by k, such that each unique combination of $(\boldsymbol{x}, T)$ forms a separate group. Denote by Γ_k all individual/target age/year observations that satisfy

$$\Gamma_k = \left\{(i, j, t) \,\middle|\, \boldsymbol{x}_{it} = \boldsymbol{x}_k, T_{ijt} = T_k\right\}$$

i.e., all observations where the individuals are asked about their survival beliefs over the same horizon, are of the same age, report the same health state and share any other covariates in $\boldsymbol{x}_{it}$. Denote by $\overline{p}_k^s$ the (weighted) sample average of reported survival beliefs

conditional on $(\boldsymbol{x}_k, T_k)$, i.e.,

$$\overline{p}_k^s = \frac{\sum_{(i,j,t)\in\Gamma_k} w_{it} \cdot \phi_{T_k}(\boldsymbol{x}_k, \boldsymbol{z}_{it})}{\sum_{(i,k,t)\in\Gamma_k} w_{it}} \tag{5}$$

where w_{it} are the respondent-level sampling weights.

Now consider the logit counterpart of (5), which we denote by

$$\widehat{p}_k^s = \Pr\left(\text{ alive at } age_k + T_k \,\middle|\, \boldsymbol{x}_k, \boldsymbol{\nu}\right)$$

i.e., the predicted probability of being alive for group k, taking into account all possible health transitions to T_k. We assume exactly the same functional form as for the objective health and survival process, but allow the parameter vector $\boldsymbol{\nu}$ governing survival to differ.

The observed sample moment for each group can then be written as

$$\overline{p}_k^s = \widehat{p}_k^s + u_k$$

where u_k is the deviation from the group mean not explained by our model. Our aim is to minimize these group-specific residuals using the least-squares objective function

$$J(\boldsymbol{\nu}) = \frac{1}{W}\sum_k W_k \left[\overline{p}_k^s - \widehat{p}^s\left(\boldsymbol{x}_k, T_k \mid \boldsymbol{\nu}\right)\right]^2 \tag{6}$$

where $W_k = \sum_{(i,j,t)\in\Gamma_k} w_{it}$ is the sum of weights in group k. The estimated vector $\widehat{\boldsymbol{\nu}}$ is hence the arg min of $J(\boldsymbol{\nu})$.

Estimation sample. We use all target ages from Table 3 for the estimation of the subjective life expectancy process. In the main paper, we present the results for nonblack men, as these are later incorporated into our quantitative model.

Results. The estimated subjective survival beliefs for nonblack men are shown in pink in Figure 6, juxtaposing the objective survival probabilities estimated in Foltyn and Olsson (2021) in blue. As can be seen, the subjective belief about survival in health state *excellent* or *very good* is almost 100% for all ages. This does not mean that individuals in those health states believe that they will live forever, rather that they believe that death is necessarily preceded by a deterioration in health.

Figure 7 summarizes the results, showing the life expectancy by age and health state

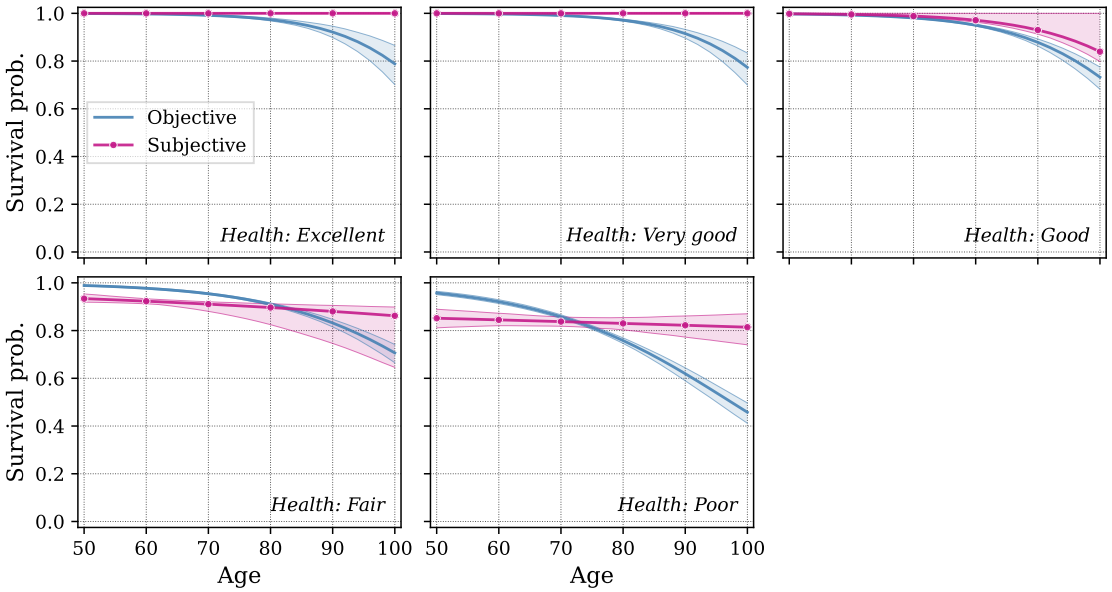


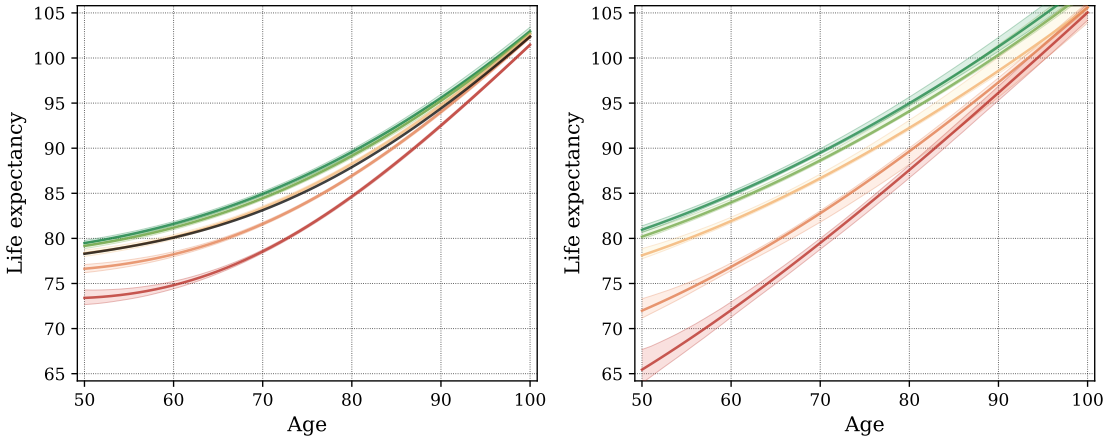
Figure 6: One-year objective and subjective survival probabilities by health state for nonblack men (model estimates). Shaded areas indicate 95% confidence intervals. For each bootstrapped sample we re-estimate the objective health process.

using the objective and the subjective survival process. At all ages, the difference in life expectancy between the best and the worst health state is larger when using subjective life expectancies. The divergence between objective and subjective life expectancies is particularly large for men in bad health, who substantially underestimate survival at younger ages. Conversely, individuals in all health states overestimate their chances of survival late in life. Figure C.8 in the appendix plots the objective and subjective life expectancy for the remaining demographic groups, which exhibit very similar patterns.

Figure C.9 in the online appendix plots the model-predicted subjective survival probabilities against their data counterparts, showing that both sets of moments are well aligned.

3 Economic model

In this section, we describe the overlapping-generations model used to quantify the implications of survival heterogeneity. Time is discrete and each period corresponds to one year. Agents derive utility from consumption and face idiosyncratic risk in the form of shocks to their labor income, medical expenditures, health and survival, as well as stochastic bequests from their parents. Markets are incomplete as agents can only save in a riskless asset while borrowing is not permitted.



(a) Objective probabilities **(b)** Subjective probabilities

Figure 7: Life expectancy by age and health for nonblack men. Colors indicate the health state: dark green is excellent while red is poor health. In panel (a), the black line depicts the weighted population average. Shaded areas indicate 95% confidence intervals.

3.1 The agent's problem

There is a unit mass of individuals distributed across N_t cohorts according to the ergodic distribution implied by the transition matrix of survival probabilities. An individual of age $t \in \{1, \ldots, N_t\}$ and health $h \in \{1, \ldots, 5\}$ has a one-period survival probability to age $t+1$ given by π^s_{ht}, with $\pi^s_{ht} = 0$ in the terminal period.[13]

Individuals are assumed to be working for the first $T_r - 1$ years of their life and exogenously retire in the period in which they attain age T_r. While working, they are hit by persistent and transitory labor productivity shocks. During retirement, individuals receive social security retirement benefits which depend on their last persistent labor productivity in working age.

Bequests are modelled via probabilistic intergenerational links along the lines of Straub (2019) so that children with higher lifetime income are more likely to have income-rich parents and thus expect to receive higher bequests.

Retirement. Let $x = (a, p, h, \eta, \mathbb{1}_b, t)$ be a retired individual's state vector, where a is cash-at-hand, p represents pre-retirement labor productivity, h is the current health state, η is the persistent component of medical expenditures, $\mathbb{1}_b$ is a bequest indicator, and t is age. In each period, an individual chooses consumption c and savings k to be invested in risk-free productive capital. Individuals earn a gross return R on their savings and receive gross retirement income $w \cdot y_r$ which is proportional to the economy-wide wage

[13] Since the model does not use calendar time, we from now on use t to denote age.

rate w, and where

$$y_r = \overline{\omega}_r \cdot \mathcal{R}_{ss}(p)$$

is a function of the average earnings profile at the time of retirement, $\overline{\omega}_r$, and $\mathcal{R}_{ss}(\bullet)$ which mimics the regressive replacement rate of the US social security system applied to the last pre-retirement labor productivity p. Retirement income is taxed using the nonlinear tax schedule $T_y(\bullet)$ so that after-tax retirement income amounts to

$$\iota = w\, y_r - T_y(w\, y_r)\ . \tag{7}$$

Agents receive an inheritance $b_* \geq 0$ at most once in their life, which we track using the indicator $\mathbb{1}_b$. They are born in state $\mathbb{1}_b = 1$ and transition to $\mathbb{1}_b = 0$ when their parents die. As long as $\mathbb{1}_b = 1$, the tuple $(b_*, \mathbb{1}_b)$ evolves according to

$$\left(b'_*,\ \mathbb{1}'_b\right) \sim \begin{cases} \left(b_*(p,h,t), 0\right) & \text{with prob. } \left(1-\overline{\pi}^s_{t_*}\right)\pi^b_{pht} \\ (0,0) & \text{with prob. } \left(1-\overline{\pi}^s_{t_*}\right)\left(1-\pi^b_{pht}\right) \\ (0,1) & \text{with prob. } \overline{\pi}^s_{t_*} \end{cases} \tag{8}$$

Once $\mathbb{1}_b = 0$, no additional bequests are expected and $(b'_*, \mathbb{1}'_b) = (0,0)$ obtains with certainty. It is not possible to directly model intergenerational links between a parent and a child as this would double the number of state variables. We therefore assume that parents are exactly 30 years older (with age $t_* = t + 30$) and survive with the age-specific population-average probability $\overline{\pi}^s_{t_*}$ to the next period. Conditional on parental death, children receive bequests with probability $\pi^b_{pht} \in (0,1)$ to reflect that many parents do not leave sizeable estates. To capture the intergenerational persistence of income and wealth, we map agents into income quintiles and use the intergenerational income quintile transition matrix from Chetty et al. (2014) to stochastically connect children to potential parents. This creates a positive sorting between children's and parents' income and wealth so that richer children are more likely to receive higher bequests. Since our mapping to income quintiles relies on the states (p,h,t), both the probability to receive a bequest π^b_{pht} and the amount received b_* are functions of (p,h,t). We describe the technical details of these linkages in section D.1 in the online appendix.

The next-period cash-at-hand is given by

$$a' = Rk + b'_* + \iota' - m\left(h', \eta', \nu', t+1\right) + \xi' \tag{9}$$

where m are out-of-pocket medical expenditures that accrue between ages t and $t+1$

which are allowed to depend on health h, a persistent component η and a transitory shock ν, similar to the approach taken in De Nardi, French, and Jones (2010).

Because these expenditure shocks can be quite large, we assume that the government guarantees a minimum consumption level $\underline{c}$ by making a transfer ξ whenever agents do not have the resources to cover the medical expenditures themselves. The required transfer is therefore defined as

$$\xi' = \max\left\{0,\ \underline{c} + m\left(h', \eta', \nu', t+1\right) - Rk - b'_* - \iota'\right\}. \tag{10}$$

In the process, whenever agents receive a positive transfer $\xi > 0$, they are not permitted to save for the next period and therefore choose $k = 0$ and $c = \underline{c}$.

Non-survivors leave their asset holdings as bequests to their offspring. Any out-of-pocket medical bills $m\left(\eta', \nu', t+1\right)$ incurred in the last period of life are deducted, and bequests are additionally subject to the estate tax $T_b\left(\bullet\right)$. Thus after-tax bequests are given by

$$b' = \max\left\{0, Rk + b'_* - m\left(\eta', \nu', t+1\right)\right\} - T_b\left(\max\left\{0, Rk + b'_* - m\left(\eta', \nu', t+1\right)\right\}\right) \tag{11}$$

where we assume that descendants are not liable for any medical bills exceeding a deceased individual's assets.[14]

Finally, we impose a warm-glow bequest motive as in De Nardi (2004),

$$V_b\left(b\right) = \phi_1 \frac{\left(b + \phi_2\right)^{1-\sigma} - 1}{1-\sigma}$$

where ϕ_1 governs the weight individuals assign to leaving bequest and ϕ_2 is a parameter controlling to what extent bequests are a luxury good.

To summarize, a retired individual's maximization problem is defined by the value function

$$\begin{aligned} V_r\left(a, p, h, \eta, \mathbb{1}_b, t\right) = \max_{c \geq 0,\ k \geq 0} \Bigg\{ & \frac{c^{1-\sigma} - 1}{1-\sigma} + \beta \pi^s_{ht} \mathbb{E}\left[V_r\left(\boldsymbol{x}'\right) \,\middle|\, p, h, \eta, \mathbb{1}_b, t \right] \\ & + \beta\left(1 - \pi^s_{ht}\right) \mathbb{E}\left[V_b\left(b'\right) \,\middle|\, h, \eta, \mathbb{1}_b, t \right] \Bigg\} \end{aligned}$$

subject to $c + k \leq a$ and the laws of motion (9) and (11), where $\boldsymbol{x}' = \left(a', p, h', \eta', \mathbb{1}'_b, t+1\right)$ is the continuation state conditional on survival. Health h evolves according to the

[14] As (11) suggests, it is possible that children die in the same period as their parents so that an inheritance immediately becomes part of a child's estate.

transitions estimated from the HRS, while the survival probabilities π^s_{ht} follow either the objective or subjective survival beliefs discussed in the previous section. Lastly, the persistent component η of medical expenditures follows an AR(1) process.

Working age. Individuals of working age solve a problem almost identical to that of retirees, except that they additionally face both persistent and transitory labor earnings risk. The persistent risk component is captured by the state variable p and is assumed to follow a first-order Markov process, while the transitory shock ϵ is i.i.d. over time. Together with the age-health earnings profile ω_{ht} they pin down an individual's labor productivity y which is allowed to depend on health h,

$$y = \omega_{ht} p \, \epsilon \, . \tag{12}$$

Moreover, workers are subject to payroll taxes $T_{ss}(\bullet)$ which we model as a function of productivity, so that their after-tax labor income is

$$\iota = \Big[y - T_{ss}(y) \Big] w - T_y \left(\Big[y - T_{ss}(y) \Big] w \right) . \tag{13}$$

The remaining problem is the same as for retirees, including the medical expenditure shocks, the consumption floor, government transfers, and the intergenerational linkages and bequests.

3.2 Technology

The production side of the model is standard. Competitive firms employ labor and capital hired from households to produce a homogeneous final good which is used for both consumption and investment. The aggregate production function is assumed to be Cobb-Douglas, $F(A, K, L) = AK^{\alpha_k} L^{1-\alpha_k}$. Capital depreciates at the rate δ_k.

3.3 Government

We assume that the government runs a PAYGO social security system that has to balance in each period, and that transfers as well as any remaining (wasteful) government expenditures have to be fully financed by income and inheritance taxes. We first describe the social security system and thereafter the general government budget.

3.3.1 Social security system

We use a stylized version of the actual retirement income formula used in the US Social Security system. It captures the main features, such as a regressive replacement rate based on pre-retirement income and a cap for maximum benefits. In the model, we define retirement benefits to be a product of the economy-wide wage w, the average life cycle profile component from the last year before retiring $\overline{\omega}_r$, and a function that mimics the regressive replacement rate,

$$\iota(p) = w \cdot y_r(p) = w \cdot \overline{\omega}_r \cdot \mathcal{R}_{ss}(p) \ .$$

The replacement function $\mathcal{R}_{ss}(\bullet)$ is given by

$$\mathcal{R}_{ss}(p) = \begin{cases} \rho_1 p & \text{if } p \leq p_1^* \\ \rho_1 p_1^* + \rho_2 \left(p - p_1^*\right) & \text{if } p_1^* < p \leq p_2^* \\ \rho_1 p_1^* + \rho_2 \left(p_2^* - p_1^*\right) + \rho_3 \Big(\min\Big\{ p_{max}^*, p \Big\} - p_2^* \Big) & \text{else} \end{cases}$$

where p_1^* and p_2^* are bend points and p_{max}^* is the contribution and benefit base (CBB) in the Social Security income formula, expressed in terms of the individual's last pre-retirement persistent labor state p which becomes permanent once retired. Online appendix section D.2.1 describes in detail how we map the dollar quantities taken from Social Security regulations to their model counterparts.

The government expenditures on retirement are financed by a payroll tax. The payroll tax function is defined as

$$T_{ss}\left(y\right) = \tau_{ss} \cdot \min\{y_{max}, y\}$$

where y_{max} represents maximum taxable earnings. The derivation of total payroll taxes raised in each period can be found in online appendix section D.2.2. To balance the social security system, we find τ_{ss} such that total expenditures on Social Security benefits equal total payroll taxes.

3.3.2 Government budget

The government needs to finance lump-sum transfers ξ to households defined in (10). We denote the aggregate transfers by Ξ and provide details on how these are computed in online appendix section D.3.1. Additionally, the government finances non-discretionary expenditures that amount to a constant fraction g of output, $G = gY$.

We adopt the same income tax function as in Heathcote, Storesletten, and Violante

(2017), which is defined as

$$T_y(\iota) = \iota - \lambda \iota^{1-\tau} \tag{14}$$

where ι is either earnings (net of payroll taxes) or retirement income, and we denote total income tax revenue by T_{inc}. Additionally, the government collects T_{estate} in estate taxes. We provide details on how to compute T_{estate} and T_{inc} in online appendix section D.3.2 and D.3.3.

We assume that the progressivity parameter τ in (14) is fixed, and we pin down λ such that the government budget is balanced in each period, i.e., $\Xi + G = T_{estate} + T_{inc}(\lambda)$.

3.4 Equilibrium

The equilibrium definition is mostly standard and can be found in online appendix section D.4. The one noteworthy addition is that we require that for each cohort, after-tax estates left by parents are consistent with the bequests expected and received by children given the stochastic intergenerational links. This introduces computational complications which we discuss in online appendix section H.

4 Calibration

4.1 Preferences

We assume log preferences, i.e., $\sigma = 1$ and thus $u(c) = \log c$. The common discount factor $\beta = 0.979$ is set to obtain a capital-to-output ratio of 3.0 in the scenario with subjective survival beliefs. In our benchmark calibration we shut down the warm-glow bequest motive by setting $\phi_1 = 0$, and hence all bequests are accidental. We discuss alternative scenarios in section 5.4.

4.2 Externally calibrated parameters

Demographics. Agents are assumed to enter the economy at age 20, which corresponds to model age 1, and retire at the age of 65, implying that $T_r = 46$. The maximum attainable age is 109, and hence we let $N_t = 90$.[15]

[15] The reason for imposing such a high maximum age is that otherwise the scope for upward bias in beliefs about survival in old age is limited: if agents know for sure that they are going to die at the age of 100, say, any gap between objective and subjective beliefs shrinks by construction, even at younger ages. However, setting a high maximum age has no effect on the age distribution: as shown in Figure E.2, the mass of individuals aged 100+ in the economy is only 0.06%.

Earnings. We assume that the logarithm of labor earnings follows a process with transitory and persistent shocks,

$$\log y_{ht} = \log \omega_{ht} + \log p_t + \log \epsilon_t \qquad\qquad t \in \{1, \ldots, T_r - 1\}$$

where ω_{ht} is the age-health profile, p_t is the persistent component and ϵ_t is the transitory component of earnings. The persistent component is assumed to follow an AR(1) process,

$$\log p_t = \rho_p \log p_{t-1} + v_t$$

with autocorrelation ρ_p and innovation $v_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_v^2)$. The transitory shock is log-normally distributed with $\log \epsilon_t \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\epsilon^2)$. The stochastic part of the wage process is therefore characterized by the parameters $(\rho_p, \sigma_v^2, \sigma_\epsilon^2)$ which we set to $(0.9695, 0.0384, 0.0522)$, following Krueger, Mitman, and Perri (2016). We use the Rouwenhorst procedure to discretize the persistent part of the process into an five-state Markov chain, and we discretize the transitory shock into three states.

The age-health profile for labor earnings is estimated for nonblack men aged 20 to 65 using PSID data (PSID (2023)). Not surprisingly, there is a strong health gradient. This is partly driven by lower wages conditional on working, and partly by a larger fraction of individuals in bad health not working at all. Since we abstract from the labor supply decision, our estimates of the age-health earnings profile captures both margins. More details can be found in online appendix section E.3.

Medical expenditures. Following French and Jones (2004) and De Nardi, French, and Jones (2010), we estimate the medical expenditure shocks from the out-of-pocket medical expenses reported at biennial frequency in the HRS for the sample of nonblack men aged 50 and above. Since the HRS includes hardly any individuals below the age of 50, we assume that agents do not face any out-of-pocket medical costs at these ages.

We impose that both the mean and variance of log medical expenditures are state-dependent and given by the following process:

$$\begin{aligned}
\log m_{it} &= \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_{it}'\boldsymbol{\gamma} + \sigma(\boldsymbol{x}_{it})\,(\eta_{it} + \nu_{it}) \\
\eta_{it} &= \rho_m \eta_{it-1} + \zeta_{it} \qquad (15) \\
\zeta_{it} &\overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\zeta^2) \\
\nu_{it} &\overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\nu^2) \qquad (16)
\end{aligned}$$

The vector $\boldsymbol{x}_{it}$ contains a third-order polynomial in age, health, as well as health interacted with age. Additionally, $\mathbf{z}_{it}$ includes controls not present in the economic model such as marital status, education level, 5-year cohort dummies, and time fixed effects, as well as interactions of these terms. We run a fixed effects estimator on the level of log medical expenditures and recover the parameters governing the variances and covariances from the residuals using GMM. Once we have identified the parameters for medical expenditures over a two-year period, we use a simulated method of moments procedure to recover the implied parameters at annual frequency which yields $\rho_m = 0.920$, $\sigma^2_\zeta = 0.084$ and $\sigma^2_\nu = 0.457$. More details can be found in the online appendix section E.4.

For the purpose of including medical expenditure shocks in the OLG model, we discretize the persistent component (15) using the Rouwenhorst procedure with seven states, and we discretize the transitory component (16) into five possible realizations.

Bequests. We assume that estates are tax exempt up to the amount χ_b and subject to a proportional tax τ_b thereafter. We set $\chi_b = 19.75$ so that in equilibrium 2% of estates are subject to estate taxes, while τ_b is set to 30%.[16] The intergenerational income quintile transition matrix used to link parents to children is taken from Chetty et al. (2014, Table II) and reproduced in Table D.1 in the online appendix. Lastly, we allow the probability to receive a bequest conditional on parental death to differ by income quintile. To this end, we use the Survey of Consumer Finances (SCF) waves 1998–2007 and compute the fraction of respondents aged 60–70 who report having ever received an inheritance by each income quintile, which gives the probabilities 20.5%, 25.2%, 27.4%, 33.3% and 40.4% for the lowest to highest quintile.

Remaining externally calibrated parameters. The remaining parameters are listed in Table 4. The bend points and the contribution and benefit base are reported in US dollars to facilitate the interpretation. The value for the consumption floor is similar to the levels used by De Nardi, French, and Jones (2010) or Palumbo (1999).

4.3 Health and survival process

We use the processes for health transitions and survival probabilities described in section 2.3 (health transitions and objective survival probabilities) and section 2.5 (subjective survival probabilities) for nonblack men. Agents enter the model at the age of 20, but

[16] The top marginal tax rate in 2023 was 40% (see https://www.irs.gov/pub/irs-pdf/i706.pdf); however, not all taxable estates fall into the top category. We choose 30% as an approximation.

Parameter	Description	Value	Source
Production technology parameters			
α_k	Capital share	36%	Krueger, Mitman, and Perri (2016)
δ_k	Depreciation rate	9.6%	Krueger, Mitman, and Perri (2016)
A	Total factor productivity	0.896	Fixes equilibrium wages at unity
Social security			
ρ_1	Replacement rate bracket 1	90%	2000 SS rules
ρ_2	Replacement rate bracket 2	32%	2000 SS rules
ρ_3	Replacement rate bracket 3	15%	2000 SS rules
$b_1^{\$}$	Bendpoint 1	\$6,384	2000 SS rules
$b_2^{\$}$	Bendpoint 2	\$38,424	2000 SS rules
$e_{max}^{\$}$	Contribution and benefit base (CBB)	\$76,200	2000 SS rules
$\underline{c}$	Consumption floor	\$2,325	5% of average annual earnings
Government budget			
g	Gov. spending (share of GDP)	6%	Brinca et al. 2016
τ	Tax progressivity	0.137	Brinca et al. 2016
τ_b	Marginal tax on estates	30%	Authors' approximation

Table 4: Calibrated parameters

the health and survival processes we estimate based on the HRS data starts at the age of 50. We therefore estimate a health process for the ages 20 to 50 using PSID data. We use data from the years 1984 to 2019 and the subsample of nonblack male household heads, and assume that survival is certain during this age span (further details can be found in the online appendix section E.1). From the age of 50, we use our estimated process based on the HRS data, and agents start facing a positive probability of death. The resulting cohort sizes and distribution of health states are shown in Figure E.2 in the online appendix.

While the model is solved with five health states, in what follows we report results only for the best, middle and worst health states to reduce visual clutter.

5 Results

We solve the model under three distinct assumptions about health heterogeneity and survival expectations:

1. *No health heterogeneity (NHH):* In this scenario, all agents of the same age face the same earnings profile, the same medical expenditure risk and the same survival risk. We eliminate health heterogeneity and use the average survival rates (depicted by the black line in Figure 7(a)), an average earnings profile, and the average medical expenditure process. The probability of receiving an inheritance and the

amount received are also averaged across health.[17]

2. *Objective survival heterogeneity (OSH):* In the second scenario, we use the objective process for health transitions and survival probabilities described in section 2.3. In this case, individuals are perfectly informed about their true survival probability conditional on health and age. Medical expenditures and labor earnings are allowed to differ by age and health.
3. *Subjective survival heterogeneity (SSH):* In the third scenario, agents conversely form beliefs and act according to the subjective survival process estimated in section 2.5. However, this subjective process does not correspond to the true survival process, which we use when simulating the model.

In the remainder of this section, we first contrast the effective discount rates that arise in the objective vs. subjective survival belief scenarios. These are important drivers of savings behavior, which we discuss next. We then turn to the implications for wealth accumulation across health. Lastly, we briefly discuss general equilibrium effects, which are mostly unchanged across all three scenarios.

5.1 Effective discount rates

An individual's effective discount rate (determined by the common discount factor β and the survival probabilities) in the model is time-varying and depends on the horizon. Following a bad health shock, the discount rate immediately rises since it implies a shorter expected life span, while the opposite happens in the event of a good health shock.

The fact that the effective discount rate depends on the whole sequence of future age- and health-dependent transition and survival probabilities makes it difficult to compare across individuals. We therefore use the following measure of effective average discounting: for an individual of age t in health h, we implicitly define the effective average discount rate ϱ at a horizon of T years as

$$\beta^T \cdot \Pr\left(\text{ alive at } T+t \mid t, h\right) = \left(\frac{1}{1+\varrho}\right)^T$$

where the probability to be alive at time $T+t$ is evaluated using either the objective or the subjective survival probabilities. This measure also connects our framework to the discount rate in a standard infinite-horizon model without survival risk where the

[17] These averages are computed using the age-specific health distribution implied by our estimated health transition probabilities and the initial distribution over health at age 20 observed in the PSID.

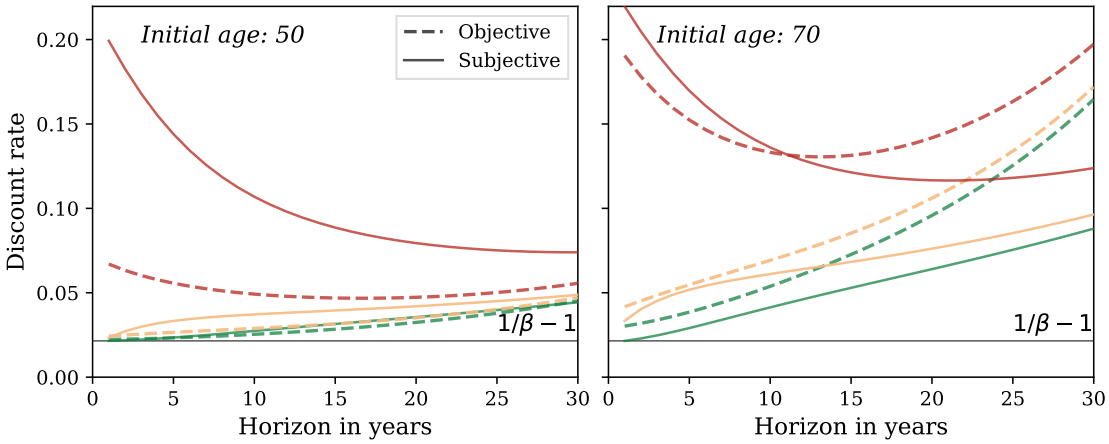


Figure 8: Effective discount rate (geometric average) as a function of initial (current) age (50 or 70), forecast horizon (on the x-axis) and objective or subjective survival probabilities (dashed vs. solid lines). Colors indicate the health state: dark green is excellent while red is poor health. The black line represents the discount rate net of any survival effect.

geometric mean discount rate is simply given by $\varrho = \beta^{-1} - 1$ irrespective of the forecast horizon.

Figure 8 plots the effective average discount rates for different time horizons. As can be seen, the effective discount rate varies substantially in the population. Using the subjective probabilities, the one-year horizon discount rate for a 50-year-old agent in good health is observationally equivalent to $1/\beta - 1 = 2.18\%$ (since the one-year ahead survival for this agent is perceived to be almost certain), while the one-year horizon discount rate for an equally old agent in bad health is almost 20%. This gap shrinks at longer horizons but is still 7.9 percentage points at a 10-year horizon. For 70-year-old agents in worst vs. best health state, the difference at the 10-year horizon is 9.5 percentage points.

The magnitude of the dispersion of subjective discount rates is consistent with the findings by Calvet et al. (2021) who estimate the cross-sectional distribution of time preference rates based on Swedish micro data, assuming a common survival probability conditional on age for all agents. They estimate the standard deviation of the time preference rate to be 7.0 percentage points around a mean of 5.2 percent.

The magnitude of the *differences* between discount rates based on subjective beliefs and objective survival probabilities across health states are well in line with the discount factor heterogeneity estimated by De Nardi, Pashchenko, and Porapakkarm (2017). In a model with rational survival expectations they find that individuals in bad health on average have a higher discount rate (i.e., are more impatient). Our results show that the

difference between the subjective and objective discount rate is larger for individuals in bad health, and the magnitude of these differences is consistent with their estimates (see online appendix section F.1 for details).

The pattern of subjective and objective discount rates are also consistent with an average downward-sloping discount rate along the age gradient, as found by Kureishi et al. (2021). Netting out the average objective (statistical) survival probability at all ages, the belief bias in survival probabilities gives rise to a downward sloping residual discount rate, since older people on average display an upward bias in their beliefs about survival compared to the young (see Figure 2).

5.2 Effect on savings behavior

The effect of subjective survival beliefs on savings varies depending on age and health. Working-age individuals in good health have on average higher earnings (see the earnings profiles in Figure E.3 on the online appendix) and slightly higher consumption. The savings rate in this group is thus primarily driven by the desire to smooth consumption over the life cycle (which includes more years in retirement due to higher life expectancy) and a precautionary savings motive (to insure against adverse health shocks). Whether they act according to subjective or objective survival probabilities has little impact since agents in good health just prior to retirement do not have a very pronounced survival bias (see Figure 5). Working-age individuals in poor health, on the other hand, save substantially less when endowed with subjective survival beliefs since these agents underestimate their survival prospects.

These differences in savings behavior are illustrated in Figure 9. For selected health states, the hatched bars show the differences in total savings rates for the model with objective survival heterogeneity compared to the model with no health heterogeneity (NHH) disaggregated by deciles of the cash-at-hand distribution. The bars in lighter color show the additional effect of imposing subjective beliefs. In these graphs, we define the *total savings rate* as the fraction of cash-at-hand, i.e., beginning-of-period assets plus current income, which the individual saves for the next period.[18]

[18] These plots are generated by computing weighted averages of differences in policy functions. We plot underlying policy functions for savings in the appendix section F.2. The differences in the figures show *partial equilibrium* effects of introducing health-related risk or objective/subjective beliefs at the individual level. We impose the equilibrium aggregates (interest rates, tax rates and transfers, bequests, and the cash-at-hand distribution) computed for the economy with subjective beliefs throughout this analysis. Note that the difference between the OSH and NHH models is the result of not only differences in survival, but also labor earnings profiles, differences in medical expenditure risk, and the resulting endogenous distribution of households.

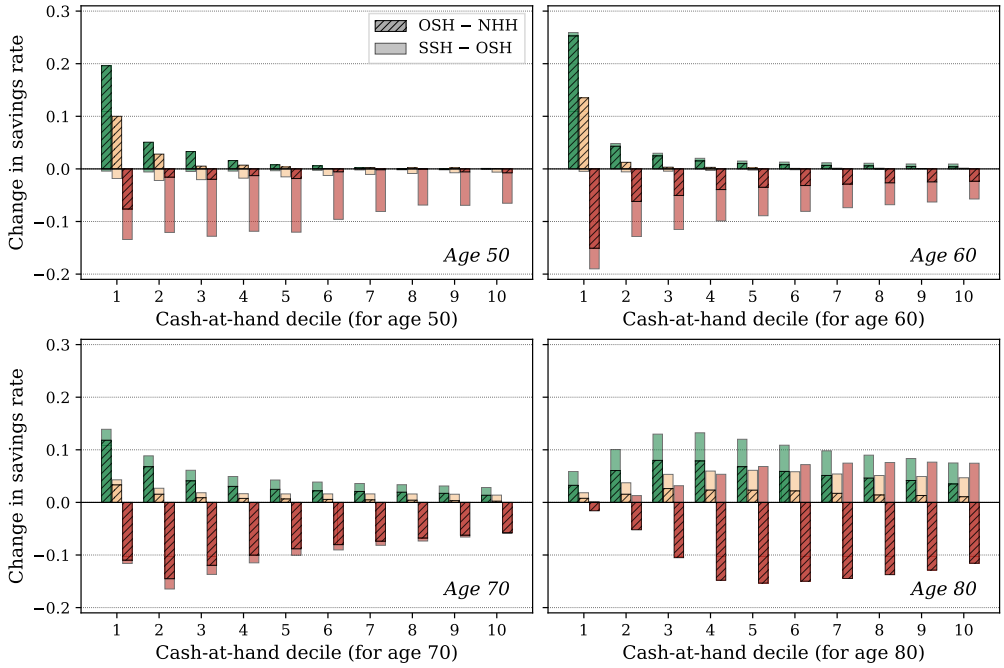
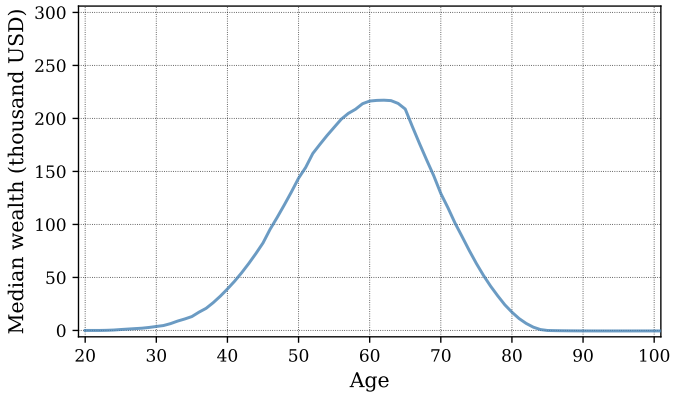


Figure 9: Differences in savings rates (defined as the fraction of current cash-at-hand the agent saves for the next period). The hatched bars show the difference between the objective survival heterogeneity and no health heterogeneity models conditional on age, health and cash-at-hand decile. The lighter bars show the additional effect of imposing subjective survival beliefs. Colors indicate health state: green is excellent while red is poor health.

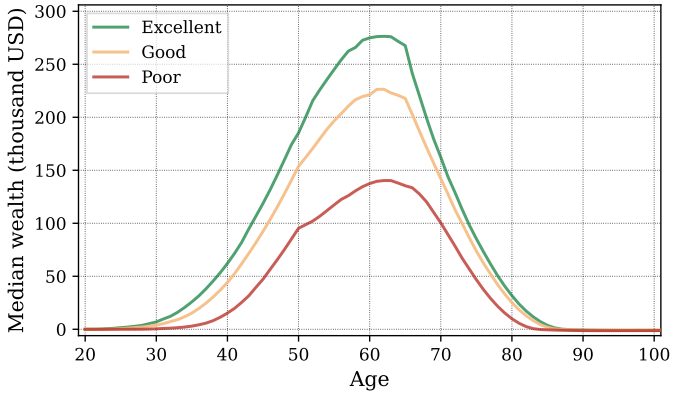
As can be seen, the 50-year-olds in poor health save substantially less across the cash-at-hand distribution, and the difference is mainly driven by the bias in survival beliefs. When acting according to subjective beliefs, 50-year old agents in poor health severely underestimate their remaining life span and therefore save approximately 10 percentage points less than in the no-health-heterogeneity model.

The same pattern can be seen in the graph for 70-year-olds, even though the additional effect from adding the subjective beliefs is slightly weaker (still focusing on the agents in poor health). The reason for this can be found in Figure 5: at the age of 70, the subjective survival beliefs of the agents in bad health are close to the objective probabilities. Moreover, at the age of 70, the effect of the subjective beliefs starts showing up more substantially for agents in good health as indicated by the lighter green parts of the green bars.

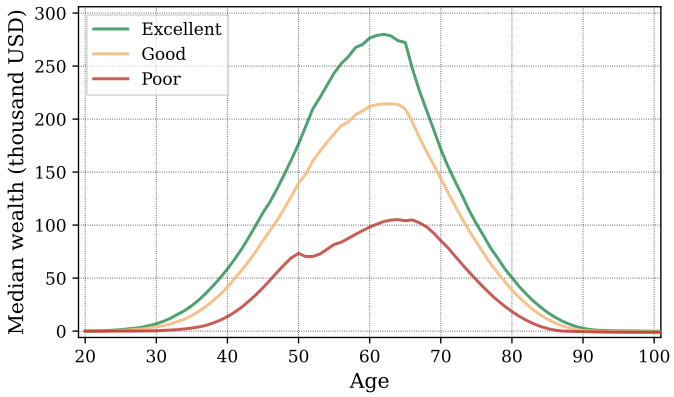
As can be seen from the graph of 80-year-olds, at higher ages the general optimism dominates. The addition of subjective beliefs at this age increases the savings rates for agents both in bad and good health.



(a) No health heterogeneity (NHH)



(b) Objective survival heterogeneity (OSH)



(c) Subjective survival heterogeneity (SSH)

Figure 10: Life cycle profiles for median wealth for different scenarios. Colors indicate the health state: dark green is excellent while red is poor health. In panel (a) there is no health heterogeneity.

5.3 Effect on wealth accumulation

Figure 10 shows the resulting life cycle profiles for wealth for the three different scenarios: no health heterogeneity (NHH), objective survival heterogeneity (OSH), and subjective survival heterogeneity (SSH). We average out all states other than assets, health and age.

For all three models, the life cycle profile for wealth peaks at the age of 64 which is the last year before retirement. When agents enter retirement, they start drawing down their wealth and the median individual who survives until the age of 90 has drawn down all of his savings (remember that all individuals receive retirement benefits and, if needed, government transfers, so they are not risking zero consumption).

The profile for the OSH model illustrates that agents in bad health accumulate less wealth than agents in good or excellent health when endowed with objective beliefs about survival. It should be noted that the mass of agents in bad health is not static but consists of individuals who have been in bad health for many periods, as well as individuals who recently drew a bad health shock.

Next, the profile for the SSH model shows that the difference in wealth between agents in best and worst health increases when the agents act according to their subjective beliefs. Note also that compared to the NHH and OSH models, the asset holdings at very high ages increase slightly, since old agents, on average, over-estimate their remaining lifespan.

In the SSH model, the median wealth among 50–54-year-olds in poor health is 36% of the median wealth held by those in excellent health of the same age. This health-wealth gap is very close to what we observe in the HRS (where the corresponding figure for nonblack males in the age group 50-54 is 34%).[19] A substantial part of this gap is due to biased survival beliefs: in the OSH model, where agents have rational expectations about their survival, the corresponding figure is 49%. Thus, the biases in survival beliefs explain approximately a fifth of the health-wealth gap in this age group.[20]

5.3.1 General equilibrium effects and economy-wide inequality

We now turn to economy-wide wealth inequality and general equilibrium effects. The first model in Table 5 (SSH-GE) shows the Gini coefficient for wealth (0.70) from the benchmark model with subjective beliefs in general equilibrium. The second model (OSH-PE-1) is a partial equilibrium exercise in which we keep aggregate variables

[19] The empirical health-wealth gap for nonblack males is shown in Figure A.5 in the online appendix.

[20] Based on the calculation $(49.2 - 35.8)/(100 - 34.1) \approx 20.3\%$. The corresponding figures for the next four 5-year age bins are 21%, 17%, 15% and 14%.

	Beliefs	GE	r	w	λ	Bequests	Gini
Subjective survival probabilities							
SSH-GE	subjective	yes	2.37%	1.0000	0.9222	SSH	0.701
Objective survival probabilities							
OSH-PE-1	objective	no	2.37%	1.0000	0.9222	SSH	0.694
OSH-PE-2	objective	no	2.19%	1.0000	0.9222	SSH	0.701
OSH-PE-3	objective	no	2.19%	1.0086	0.9222	SSH	0.700
OSH-PE-4	objective	no	2.19%	1.0086	0.9233	SSH	0.700
OSH-GE	objective	yes	2.19%	1.0086	0.9233	OSH	0.702
No health heterogeneity							
NHH-GE	average	yes	2.30%	1.0032	0.9189	NHH	0.715

Table 5: Gini coefficient for different model scenarios in general and partial equilibrium (as indicated by column "GE"). Row OSH-PE-1 imposes the same aggregates as SSH-GE but introduces objective survival beliefs. Rows OSH-PE-2 to OSH-PE-4 additionally incrementally impose the equilibrium interest rate, wages and tax rate from the OSH model. The "Bequests" columns indicates which equilibrium bequests were used.

from the SSH-GE scenario (interest rate, wage rate, taxes, and distribution of bequests received) but let agents act according to objective survival probabilities. In the OSH-PE-1 model, the Gini coefficient for wealth decreases slightly to 0.69. With objective survival beliefs, agents in the age group around 55–65 no longer have a downward bias in expected longevity and therefore higher incentives to save. This effect is strongest among the agents in poor health with low asset holdings, and thus the result is fewer poor individuals close to retirement.

The models labeled OSH-PE-2 to 4 incrementally replace the aggregate variables with the ones from the general equilibrium model with objective beliefs. The equilibrium interest rate with objective survival beliefs is lower than with subjective beliefs. Removing the average downward survival bias among the agents close to retirement (which is when life cycle assets peak) results in higher demand for assets in the model with objective beliefs. For the capital market to clear, the interest rate needs to fall. The OSH-PE-2 model uses the lower interest rate, but otherwise the same setting as the OSH-PE-1. The Gini coefficient increases slightly, from 0.69 to 0.70. There are two reasons. First, with a lower interest rate agents save less in general. Since agents with low productivity and bad health receive lower bequests, the decrease in savings have a relatively stronger effect on their total wealth.[21] Second, a lower interest rate increases frontloading of

[21] By the same logic, the health-wealth gap is slightly less pronounced in the OSH-PE-1 model than in the equilibrium version with objective beliefs (OSH-GE), even though the magnitude of the differences is small (median wealth held by those in poor health as a fraction of median wealth held by those in excellent health differs by around 2 percentage points depending on age group). Thus, comparing the impact of objective vs. subjective survival beliefs based on the general equilibrium results is a

consumption over the life cycle, and therefore results in slightly more poor elderly in the economy.

As the results from the OSH-PE-3, OSH-PE-4 and OSH-GE models show, the shifts of wage rate, taxes and distribution of bequests have virtually no effect on the economy-wide Gini coefficient for wealth.

The last row in Table 5 shows the general equilibrium results from the model with no health heterogeneity. The Gini coefficient for wealth is slightly higher than in any of the models with health heterogeneity. The driver of this is an increased number of large accidental bequests. The largest amounts bequethed stem from deaths in relatively young ages (between ages 50 and 65). In a model with no health heterogeneity, these deaths are equally likely to happen to agents in the top of the asset distribution as to agents in the bottom. With health-dependent survival on the other hand, it is more likely that deaths in relatively young ages happen to agents in poor health, who are on average poorer. Thus, the lack of health heterogeneity gives rise to larger bequests which in turn gives rise to slightly larger wealth inequality.

5.4 A model with an active bequest motive

There are many drivers of savings that could vary across health but are not included in our model: the existence of (employer-tied) health or life insurance, human capital investment, endogenous retirement decisions, portfolio composition, private pensions, and permanent characteristics such as patience, to name a few (see for instance De Nardi, French, and Jones (2010), Capatina (2015) or De Nardi, Pashchenko, and Porapakkarm (2017) for studies taking a broader perspective including several channels). These mechanisms could all add realism to the model and make the life cycle profiles more in line with the data. One mechanism often introduced to capture the slow decumulation of asset in older ages is a warm-glow bequest motive. In this section we therefore use a model calibration with an active bequest motive ($\phi_1 > 0$) to show how it interacts with (subjective) survival heterogeneity.

5.4.1 Calibration

Where applicable, we use the same calibration as in the main text. We determine the discount factor β, the preference parameters governing the bequest motive (ϕ_1 and ϕ_2) and the estate tax exemption threshold χ_b using the method of simulated moments, i.e.,

conservative choice.

Parameter	Description	Value
Model with bequest		
β	Discount factor	0.942
ϕ_1	Bequest weight	11.157
ϕ_2	Bequest shifter	0.001
χ_b	Estate tax exemption	18.333

Table 6: Parameters for the model with a bequest motive.

we minimize the weighted sum of squared distances between targeted and simulated moments from the model with subjective beliefs about survival.

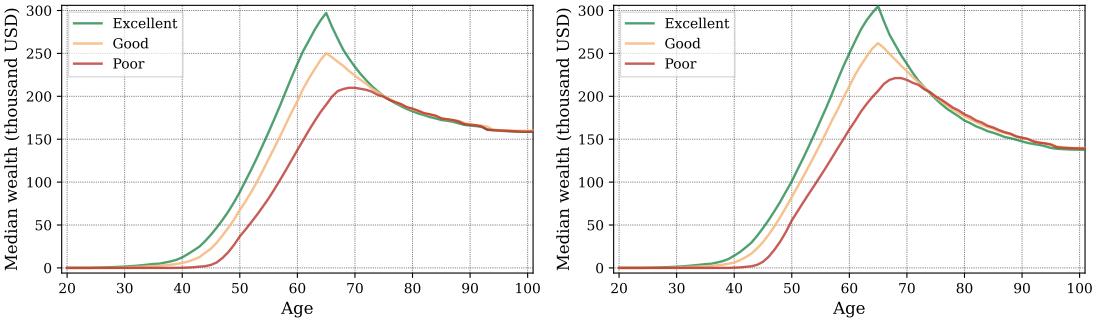
We again target a capital-to-output ratio of 3 and that 2% of estates should be subject to estate taxes. Additionally, we try to match the old-age life cycle profile of assets. To this end, we use the median wealth levels at ages 60, 65, 70, 75, 80 and 85 observed in the HRS, relative to median wealth at age 55. We choose this approach as it is quantitatively not possible to match wealth in levels and at the same time impose a capital-output-ratio of 3 in a model with productive capital as the only asset (after all, most of the wealth in the data is held in residential real estate). The capital-to-output ratio and the fraction of estates subject to estate tax are perfectly matched, while the asset holdings by age and their data counterparts are shown in Table G.1 in the online appendix. Some aspects of our model are too simplistic to match the data moments exactly. For example, because we impose an exogenous retirement age of 65, the life cycle profile of assets peaks exactly at this age, whereas this is not the case in the data.

The estimated parameters are listed in Table 6. As the table shows, the bequest luxury shifter is small. The reason is that we try to match the *median* asset holdings late in life. If bequests were a luxury good, the median asset level would fall quickly towards zero.

5.4.2 Results

Cross-section and life cycle. Figure 11 shows the life cycle profiles for the scenarios with objective survival heterogeneity (OSH) and subjective survival heterogeneity (SSH). Overall, due to the bequest motive, older agents do not decumulate their wealth, and the resulting median asset profile is more in line with the data in both scenarios. Before retirement, agents in excellent health have more wealth than agents in worse health. The main reason for this is the higher labor income of the former group.

However, the health-wealth gradient is substantially smaller prior to retirement than in the baseline calibration without a bequest motive (compare to Figure 10), and is even reversed after the age of 75, with agents in poor health being richer. This shows that the effect of combining survival heterogeneity with a bequest motive of this type is



(a) Objective survival heterogeneity (OSH) **(b)** Subjective survival heterogeneity (SSH)

Figure 11: Median life cycle profiles for wealth, model with an active bequest motive.

not entirely straightforward. The expected utility from leaving a bequest is not only a function of the amount expected to be handed over to the descendants, but also of the survival probability: agents with low (objective or subjective) survival prospects put more weight on the bequest motive. Hence, there are two effects from lower life expectancy that work in opposite directions: a shorter expected life span makes agents want to save less for their own consumption in old age, but a stronger bequest motive induces them to save more. The net effect varies depending on the calibration of bequest parameters, but the second mechanism is always present with a warm-glow bequest motive of this type: a shorter life span makes agents want to save more to leave bequests.[22]

Thus, in both the OSH and the SSH scenario, the model misses the cross-sectional correlation between wealth and health in older ages that is present in the data. The slightly stronger reversal of the health-wealth gradient in the SSH model follows directly from the biases that amplify the health-survival belief gradient.

Note also that in very high ages, agents keep less assets in the SSH model than in the OHS model. Agents in the subjective belief model overestimate the probability of a long life, and therefore put a lower weight on warm-glow bequests, resulting in lower savings.

Dynamic responses to health shocks. Next, we compare changes in wealth following negative health shocks in the data to their model counterparts. To this end, we regress the change in net total wealth (using an inverse hyperbolic sine transformation) on a negative health shock defined as an indicator for a deterioration in self-reported health

[22]To aid intuition, in section G.1 in the online appendix we show the mechanisms at play in a simple two-period model example.

	Dep. variable: relative change in net total wealth			
	Data		Model	
	(1)	(2)	No bequest	Bequest
Negative health shock	-0.116***	-0.105***	-0.1081***	0.0001***
	(0.026)	(0.026)	(0.001)	(0.000)
Age FE	Yes	Yes	Yes	Yes
Health FE	Yes	Yes	Yes	Yes
Education FE		Yes		
Observations	35,821	35,821	1,297,804	1,297,804

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Health shocks and changes in wealth. The table shows the results of regressing changes in net total wealth (after an inverse hyperbolic sign transformation) on an indicator for a deterioration in self-reported health between two consecutive survey waves. The regression includes fully interacted fixed effects as indicated. Columns 1 and 2 are the same as the first two columns in Table A.5 in the online appendix. For the model columns, we use the SSH scenario where agents act according to their subjective beliefs. Sample restricted to nonblack males age 65 and above.

between survey waves. We restrict the sample to ages 65 and above in order to focus on the part of the life cycle where life expectancy and bequest considerations are important drivers of savings. We simulate a panel of 100,000 agents and collapse the data to two-year frequency to replicate the biennial HRS.[23] The changes in assets and health are defined analogously to the data, taking differences over two-year periods.

The results in Table 7 show that the model with a bequest motive produces dynamic savings responses that are difficult to square with the data. The estimated savings response to a negative health shock in the model with an active bequest motive is effectively zero, whereas the model without an active bequest motive produces savings responses that are well in line with the data.

In sum, in a model with health heterogeneity, both the cross-sectional implications and the dynamic responses to health shocks make the model with a bequest motive calibrated to match median asset holdings in old age difficult to align with data. The relative strength of the different effects we point to in this section of course varies depending on calibration, but should be kept in mind when combining a warm-glow bequest motive with dynamically evolving expected longevity.

[23] We start the simulation with 100,000 agents at the age of 20, of which 86,809 are still alive at the age of 65.

6 Conclusions

This paper explores how variation in objective and subjectively perceived life expectancy affects savings behavior of healthy and unhealthy people. Using HRS data, we show that there exists a *within-cohort steepness bias* in survival beliefs: individuals in bad health not only have a shorter expected life span, but are also relatively more downward biased about their survival chances, while individuals in good health and thus with higher survival probability display a more upward bias. These systematic biases exacerbate the survival expectancy heterogeneity in the population.

The differences in beliefs about survival translate into time preference heterogeneity and consequently savings behavior. We show that biases in beliefs about survival can explain approximately one fifth of the differences in accumulated wealth between those in excellent vs. poor health, mostly because the latter group underestimates their life expectancy.

This paper ties into a strand of current research investigating preference heterogeneity and its importance for individual choices and aggregate outcomes. We provide an intuitively plausible and micro-founded source of heterogeneity: the perceived probability of surviving to future states of the world. Our quantification of this channel shows that life expectancy heterogeneity is important and should be included in the list of potential sources of heterogeneity that we need to consider in our analyses. Investigating the importance of the steepness bias for within-cohort differences in terms of portfolio allocations, demand for financial products, or retirement behavior is left for future research.

References

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

Attanasio, Orazio P, and Hilary Williamson Hoynes. 2000. Differential mortality and wealth accumulation. *Journal of Human Resources*, 1–29.

Bissonnette, Luc, Michael D. Hurd, and Pierre-Carl Michaud. 2017. Individual survival curves comparing subjective and observed mortality risks. *Health Economics* 26 (12): e285–e303.

Bresser, Jochem de. 2023. Evaluating the accuracy of counterfactuals: Heterogeneous survival expectations in a life cycle model. *Review of Economic Studies*, rdad088.

Brinca, Pedro, Hans A. Holter, Per Krusell, and Laurence Malafry. 2016. Fiscal multipliers in the 21st century. *Journal of Monetary Economics* 77:53–69.

Calvet, Laurent E, John Y Campbell, Francisco Gomes, and Paolo Sodini. 2021. *The cross-section of household preferences.* Technical report. National Bureau of Economic Research.

Capatina, Elena. 2015. Life-cycle effects of health risk. *Journal of Monetary Economics* 74:67–88.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics* 129 (4): 1553–1623.

Coile, Courtney, and Kevin Milligan. 2009. How household portfolios evolve after retirement: the effect of aging and health shocks. *Review of Income and Wealth* 55 (2): 226–248.

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

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

———. 2010. Why do the elderly save? The role of medical expenses. *Journal of Political Economy* 118 (1): 39–75.

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. *The lifetime costs of bad health.* Technical report. National Bureau of Economic Research.

Deaton, Angus. 2002. Policy implications of the gradient of health and wealth. *Health affairs* 21 (2): 13–30.

Duncan, Greg J, Mary C Daly, Peggy McDonough, and David R Williams. 2002. Optimal indicators of socioeconomic status for health research. *American Journal of Public Health* 92 (7): 1151–1157.

Elder, Todd E. 2013. The predictive validity of subjective mortality expectations: evidence from the Health and Retirement Study. *Demography* 50 (2): 569–589.

Epper, Thomas, Ernst Fehr, Helga Fehr-Duda, Claus Thustrup Kreiner, David Dreyer Lassen, Søren Leth-Petersen, and Gregers Nytoft Rasmussen. 2020. Time discounting and wealth inequality. *American Economic Review* 110 (4): 1177–1205.

Foltyn, Richard, and Jonna Olsson. 2021. *Health dynamics and heterogeneous life expectancies.* Working Papers 2021-17. Business School - Economics, University of Glasgow.

French, Eric, and John Bailey Jones. 2004. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics* 19 (6): 705–721.

Gan, Li, Guan Gong, Michael D. Hurd, and Daniel McFadden. 2015. Subjective mortality risk and bequests. *Journal of Econometrics* 188 (2): 514–525.

Gan, Li, Michael D. Hurd, and Daniel McFadden. 2005. Individual subjective survival curves. Chap. 12 in *Analyses in the economics of aging,* edited by David A. Wise, 377–412. University of Chicago Press.

Grevenbrock, Nils, Max Groneck, Alexander Ludwig, and Alexander Zimper. 2021. Cognition, optimism, and the formation of age-dependent survival beliefs. *International Economic Review* 62 (2): 887–918.

Groneck, Max, Alexander Ludwig, and Alexander Zimper. 2016. A life-cycle model with ambiguous survival beliefs. *Journal of Economic Theory* 162:137–180.

Hajat, Anjum, Jay S. Kaufman, Kathryn M. Rose, Arjumand Siddiqi, and James C. Thomas. 2010. Long-term effects of wealth on mortality and self-rated health status. *American Journal of Epidemiology* 173 (2): 192–200.

Hamermesh, Daniel S. 1985. Expectations, life expectancy, and economic behavior. *The Quarterly Journal of Economics* 100 (2): 389–408.

Health and Retirement Study. 2023. *RAND HRS longitudinal file 2018 (V2) public use dataset.* Produced and distributed by the University of Michigan with funding from the National Institute on Aging (grant number NIA U01AG009740). Ann Arbor, MI.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2017. Optimal tax progressivity: an analytical framework. *The Quarterly Journal of Economics* 132 (4): 1693–1754.

Heimer, Rawley Z., Kristian Ove R. Myrseth, and Raphael S. Schoenle. 2019. Yolo: mortality beliefs and household finance puzzles. *Journal of Finance* 74 (6): 2957–2996.

Hendricks, Lutz. 2007. How important is discount rate heterogeneity for wealth inequality? *Journal of Economic Dynamics and Control* 31 (9): 3042–3068.

Hurd, Michael D., and Kathleen McGarry. 2002. The predictive validity of subjective probabilities of survival. *Economic Journal* 112 (482): 966–985.

Kopecky, Karen A, and Tatyana Koreshkova. 2014. The impact of medical and nursing home expenses on savings. *American Economic Journal: Macroeconomics* 6 (3): 29–72.

Krueger, D., K. Mitman, and F. Perri. 2016. Chapter 11 - Macroeconomics and household heterogeneity, edited by John B. Taylor and Harald Uhlig, 2:843–921. Handbook of Macroeconomics. Elsevier.

Krusell, Per, and Anthony A. Smith Jr. 1998. Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy* 106 (5): 867–896.

Kureishi, Wataru, Hannah Paule-Paludkiewicz, Hitoshi Tsujiyama, and Midori Wakabayashi. 2021. Time preferences over the life cycle and household saving puzzles. *Journal of Monetary Economics* 124:123–139.

Kvaerner, Jens Soerlie. 2022. How large are bequest motives? Estimates based on health shocks. *The Review of Financial Studies* 36 (8): 3382–3422.

Lee, Jinkook, and Hyungsoo Kim. 2008. A longitudinal analysis of the impact of health shocks on the wealth of elders. *Journal of Population Economics* 21 (1): 217–230.

Ludwig, Alexander, and Alexander Zimper. 2013. A parsimonious model of subjective life expectancy. *Theory and Decision* 75 (4): 519–541.

Margaris, Panos, and Johanna Wallenius. 2023. Can wealth buy health? a model of pecuniary and non-pecuniary investments in health. *Journal of the European Economic Association*, jvad044.

Palumbo, Michael G. 1999. Uncertain medical expenses and precautionary saving near the end of the life cycle. *The Review of Economic Studies* 66 (2): 395–421.

Pijoan-Mas, Josep, and José-Víctor Ríos-Rull. 2014. Heterogeneity in expected longevities. *Demography* 51 (6): 2075–2102.

Poterba, James M., Steven F. Venti, and David A. Wise. 2017. The asset cost of poor health. *The Journal of the Economics of Ageing* 9:172–184.

PSID. 2023. *Panel Study of Income Dynamics, public use dataset.* Produced and distributed by the Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

Quadrini, Vincenzo, and José-Víctor Ríos-Rull. 2015. Inequality in macroeconomics. In *Handbook of income distribution*, 2:1229–1302. Elsevier.

Rozsypal, Filip, and Kathrin Schlafmann. 2023. Overpersistence bias in individual income expectations and its aggregate implications. *American Economic Journal: Macroeconomics* 15 (4): 331–71.

Smith, James P. 1999. Healthy bodies and thick wallets: the dual relation between health and economic status. *Journal of Economic Perspectives* 13 (2): 145–166.

Smith, V. Kerry, Donald H. Taylor, and Frank A. Sloan. 2001. Longevity expectations and death: Can people predict their own demise? *The American Economic Review* 91 (4): 1126–1134.

Straub, Ludwig. 2019. Consumption, savings, and the distribution of permanent income. *Unpublished manuscript, Harvard University.*

Subjective Life Expectancies, Time Preference Heterogeneity, and Wealth Inequality

Online appendix

Richard Foltyn
Norwegian School of Economics

Jonna Olsson
Norwegian School of Economics and CEPR

April 2024

Table of contents

A Detailed data description

A.1 Data used for the analysis

The RAND HRS Longitudinal File 2018 (V2) includes 14 waves administered over the years 1992–2018. The first cohort included in the survey was between 51 and 61 years old in 1992, and thereafter new (older and younger) cohorts were added, as illustrated in Figure A.1. Figure A.2 shows the number of respondents with positive sampling weights by wave and cohort.

Figure A.3 shows the fraction of respondents in each wave who are marked as non-respondents in all subsequent waves but do not have a death date on record. For example, in wave 11 (administered in 2012), approximately four percent of participants did not respond to any of the later waves 12–14. Since no death date is recorded for these individuals, we cannot use these observations to estimate survival probabilities. Based on these non-response patterns, we only include waves up to and including the year 2014, as the last two waves exhibit unusually high non-response rates compared to the historical averages.

Up until RAND version O (covering waves until 2012), the survey was complemented with death dates taken directly from the National Death Index (NDI), but this data was later removed from the public files. Our analysis of death dates in the releases following version O shows that without the NDI data, death dates are sometimes recorded with considerable lag. Using the RAND 2018 (V2) files, but restricting the sample only up to the year 2012 produces almost identical results to the ones obtained with the original version O data that included the NDI death dates.

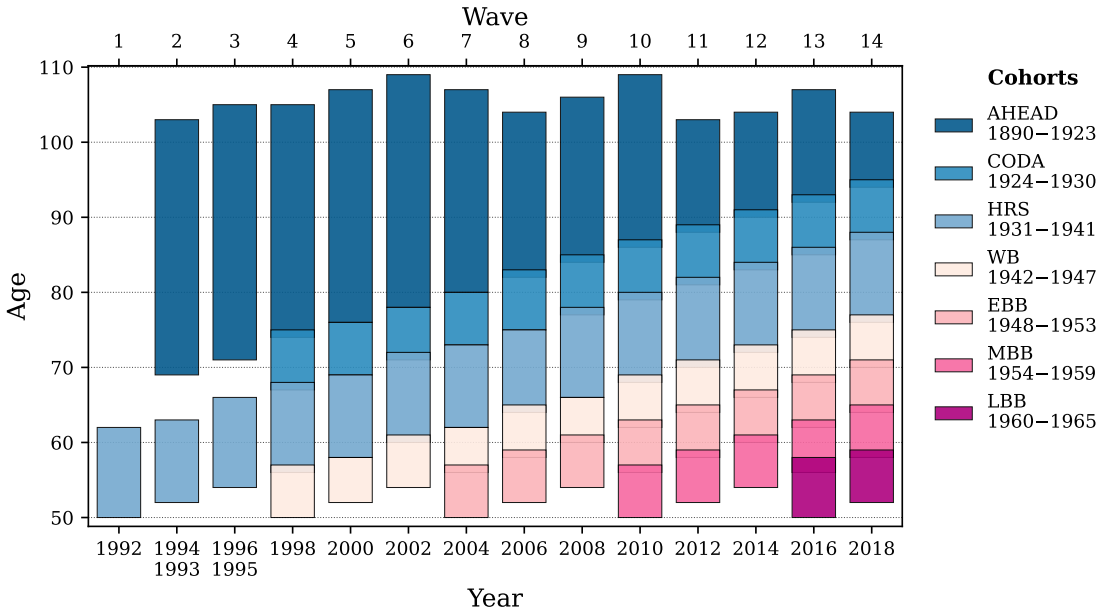


Figure A.1: Longitudinal survey design of the HRS. The y-axis shows respondents' age by cohort and wave, ignoring spouses who are not age eligible. The legend lists all birth cohorts included in the HRS (using their official acronyms) as well as their birth years. AHEAD was initially a separate survey conducted in 1993 and 1995.

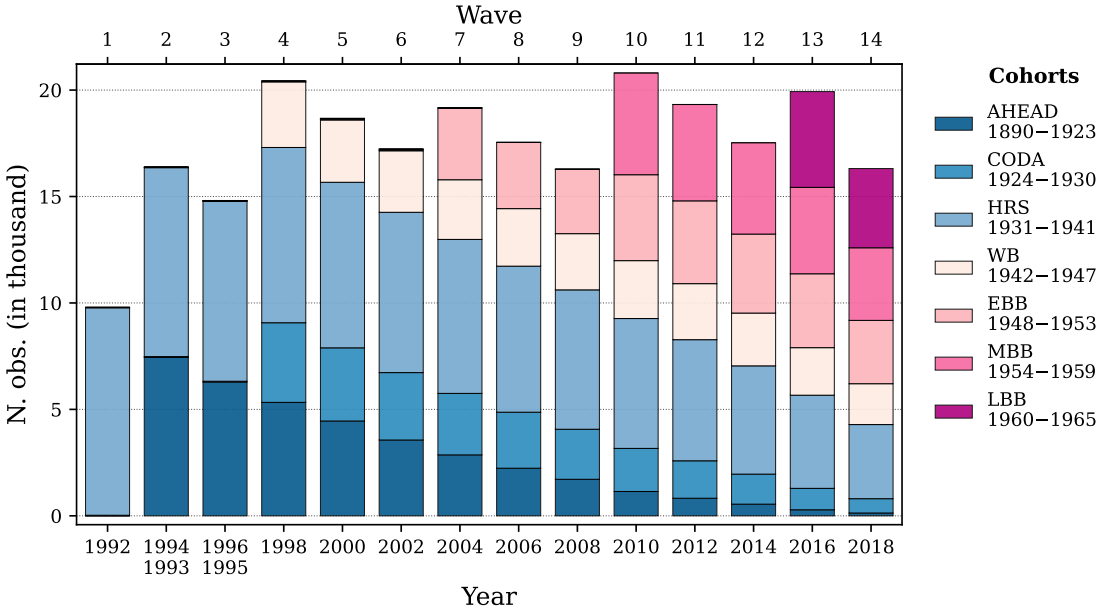


Figure A.2: Number of observations by wave and cohort. Only observations with positive weight are included. AHEAD was initially a separate survey conducted in 1993 and 1995.

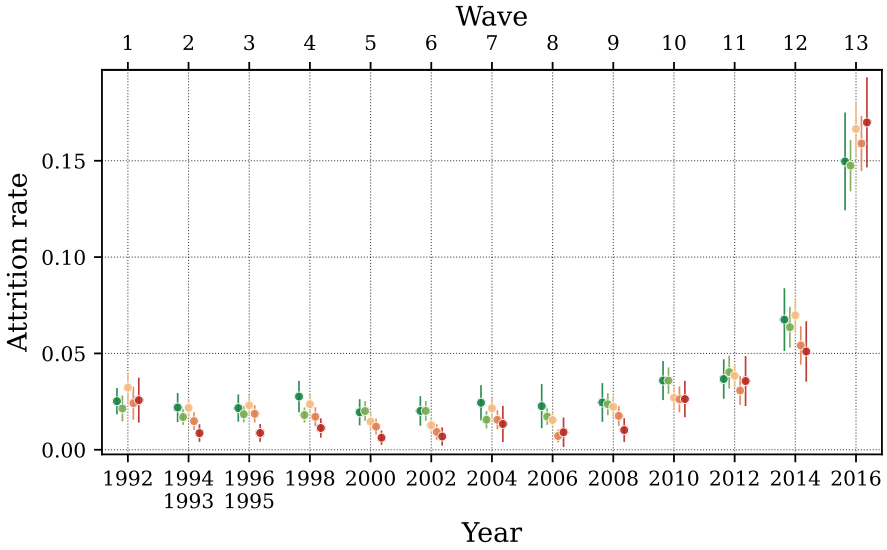


Figure A.3: HRS attrition by wave: fraction of participants who do not respond to any of the subsequent survey waves. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

A.2 Description of main variables

We list the central variables used in our analysis in Table A.1. Some of them are computed from several variables in the RAND HRS version, a process we detail below.

Variable	Description	RAND HRS variable(s)
Health	Self-reported health with possible answers (1) excellent, (2) very good, (3) good, (4) fair, or (5) poor	RwSHLT
Age	Respondent's age at time of interview end date	RwAGEY_E
Birth date	Respondent's date of birth (year, month)	RABYEAR, RABMONTH
Race	Respondent's race, recoded to *black* indicator variable	RARACEM
Sex	Respondent's sex, recoded to *female* indicator variable	RAGENDER
Education	Educational attainment, recoded to three groups	RAEDUC
Survival belief	Self-reported probability to live until a certain age (75, 85 or some other age)	RwLIV75, RwLIV85, RwLIV10, RwLIV10A
Wealth	Net total wealth, adjusted for household size, outliers, time and cohort fixed effects	HwATOTB, HwCPL
Weight	Respondent-level sampling weights	RwWTREP, RwWTR_NH, RwWTCRNH
Sampling error stratum ID	Stratum identifier used for stratified cluster bootstrapping of confidence intervals	RAESTRAT
Samping error cluster ID	Primary stage unit (cluster) identifier used for stratified cluster bootstrapping of confidence intervals	RAEHSAMP

Table A.1: Variables from RAND HRS version 2018 (V2) used in the analysis

Race. The variable RARACEM takes on three values: *White/Caucasian*, *Black/African-American* or *Other*. For our purposes, we combine the first and third groups to obtain the indicator variable $black \in \{0, 1\}$.

Wealth. We take net total wealth from the variable HwTOTB and convert all values to 2000 USD (using U.S. Bureau of Labor Statistics (2024a)). We adjust this wealth variable for cohort (using the HRS birth cohort of the oldest individual in a household) and year fixed effects, using the cohort born between 1931–1941 and the year 2000 as base categories. Furthermore, for respondents living in couple households (HwCPL = 1), we divide the wealth measure by two to get a respondent-level variable that is comparable across household sizes.

Self-reported survival probability. Depending on the survey wave, respondents are asked about their belief to live to one or two target ages. For sufficiently young respondents, this target age is 75 and the survival probability is stored in RLIV75. For

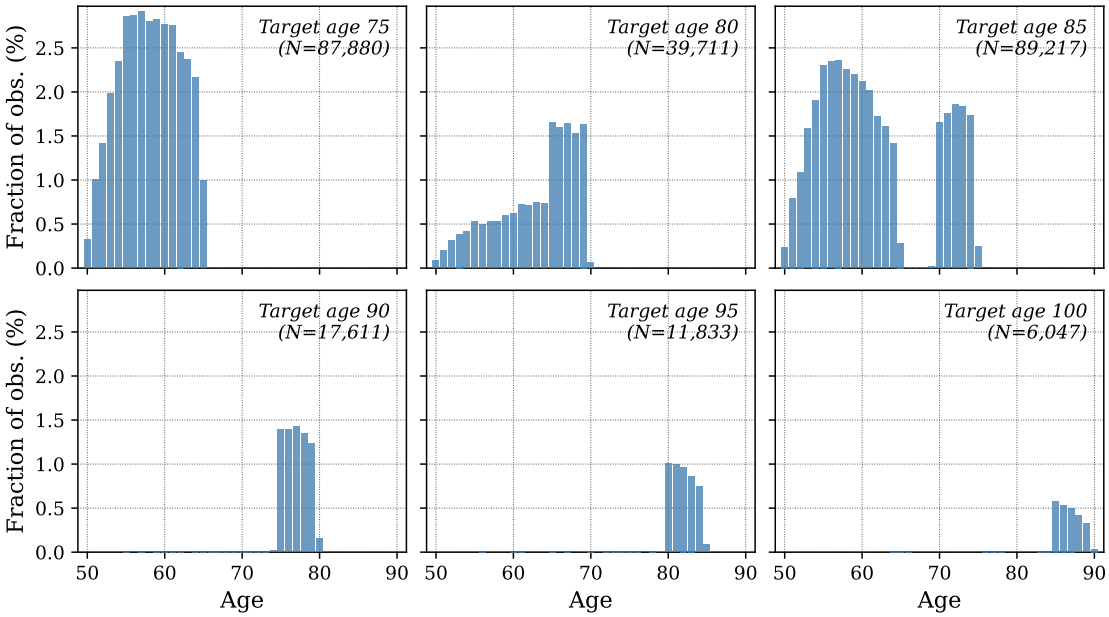


Figure A.4: Joint distribution of age and target ages used in survival probability questions. Each panel reports the total number of observations for each target age.

older respondents, the target age is 85 and the answer is stored in `RLIV85`. In waves 2 and 3, and from wave 5 onward, respondents who are too old to be asked about their survival beliefs to age 75 are instead asked to state the probability to live at least 10 more years (with possible target ages 80, 85, 90, 95 and 100), and the target age and survival probability are reported in `RLIV10` and `RLIV10A`, respectively.

Figure A.4 plots the distribution of elicited survival beliefs over the target age and the respondent's age at which the question as asked. As the figure shows, respondents were generally not asked about a target age if it was less than 10 years away. The gap at ages 66–69 for target age 85 is due to changes in questionnaires across waves.

Education. The HRS variable `RAEDUC` records five categories: *less than high school*, *GED*, *high school graduate*, *some college*, and *college and above*. We combine the three middle groups into one and use the classification *no high school*, *high school*, and *college*.

Table A.2 shows the distribution of individuals and person-year observations by education. As can be seen, only 10.8% of the black males and 12.6% of the black females have a college degree in our sample, while the corresponding figures for the nonblack population is 23.5% (males) and 16.3% (females).

	All	Nonblack		Black	
		Male	Female	Male	Female
Distribution of individuals					
By education					
No high school	27.2%	24.6%	25.5%	39.0%	36.1%
High school	54.6%	51.9%	58.3%	50.1%	51.3%
College	18.2%	23.5%	16.2%	10.8%	12.6%
Distribution of observations					
By education					
No high school	25.0%	21.8%	23.2%	40.6%	36.5%
High school	56.0%	52.7%	60.4%	48.9%	51.2%
College	19.0%	25.4%	16.4%	10.5%	12.3%

Table A.2: Distribution of individuals and person-year observations by education and permanent income tercile in the estimation sample (unweighted).

Sampling weights. We use the HRS respondent-level sampling weights. These weights are time varying, so for each transition we use a respondent's weight at transition start. Prior to wave 5, the HRS did not provide weights for individuals who moved to nursing homes, in which case the variable `RwWTRESP` was zero. From wave 5 onward, nursing home weights are provided in `RwWTR_NR`, and the combined weight is stored in `RwWTCRNH` (only one of `RwWTRESP` or `RwWTR_NR` is non-zero, depending on the nursing home status). For individuals who moved to nursing homes prior to wave 5, we back-fill their nursing-home weight from wave 5 to earlier waves if they were alive in wave 5. For respondents who moved to nursing homes prior to wave 5 but died before wave 5 and thus were never assigned a nursing-home weight, we forward-fill any missing weights using the last non-zero value of `RwWTRESP` as long as the respondent is alive.

A.3 Confidence intervals

We compute confidence intervals or standard errors in all graphs and tables presented in the paper in line with recommendations that take into account the stratified and clustered survey design of the HRS (see HRS Staff and Co-Investigators (2006) and Heeringa, West, and Berglund (2017, Chapter 4)). This involves specifying *sampling error strata* and the *sampling error cluster* (also called *sampling error computation unit* or SECU) in addition to the sampling weights. For this purpose, the RAND HRS data contains the variable `RAESTRAT` which designates sampling error strata and the variable `RAEHSAMP` which splits each stratum into exactly two sampling error clusters. Note that these clusters need not correspond one-to-one to the primary stage sampling units (PSUs) that

are part of the HRS survey design, even though this is the case for nonself-representing PSUs. Self-representing PSUs are split into two for the purpose of variance estimation (HRS Staff 2008).

To compute confidence intervals for the Maximum Likelihood and Non-linear Least Squares estimates for health transitions and objective as well as subjective survival beliefs, we draw 1000 replication samples by randomly selecting one of the two sampling error clusters from each of the 61 sampling error strata. This is a special case of the bootstrap of Rust and Rao (1996) and Rao and Wu (1988) to a setting of only two sampling error cluster per sampling error stratum.

A.4 Disaggregated wealth profiles by health

This section contains additional plots of net total wealth by health for various subsamples. Wealth is computed using the variable `HwATOTB`, which is the sum of housing, other real estate, vehicles, businesses, IRA and Keogh accounts, stocks, checkings, and all other savings, net of mortgages and other debts.

In Figure A.5, we plot the wealth profiles by self-reported health for the black/nonblack and male/female subsamples. In Figure A.6, we instead disaggregate them by household size (single vs. couple) and sex. The graphs illustrate that the health gradient of wealth is present in all subsamples, even though the confidence intervals for the single and black sub-groups are much larger due to the smaller sample size.

In Figure A.7, we disaggregate the nonblack sample by sex and education level. Again, the health gradient of wealth is present in all subsamples, even though the confidence intervals are are larger especially for the older population, despite grouping the observation into 5-year bins.

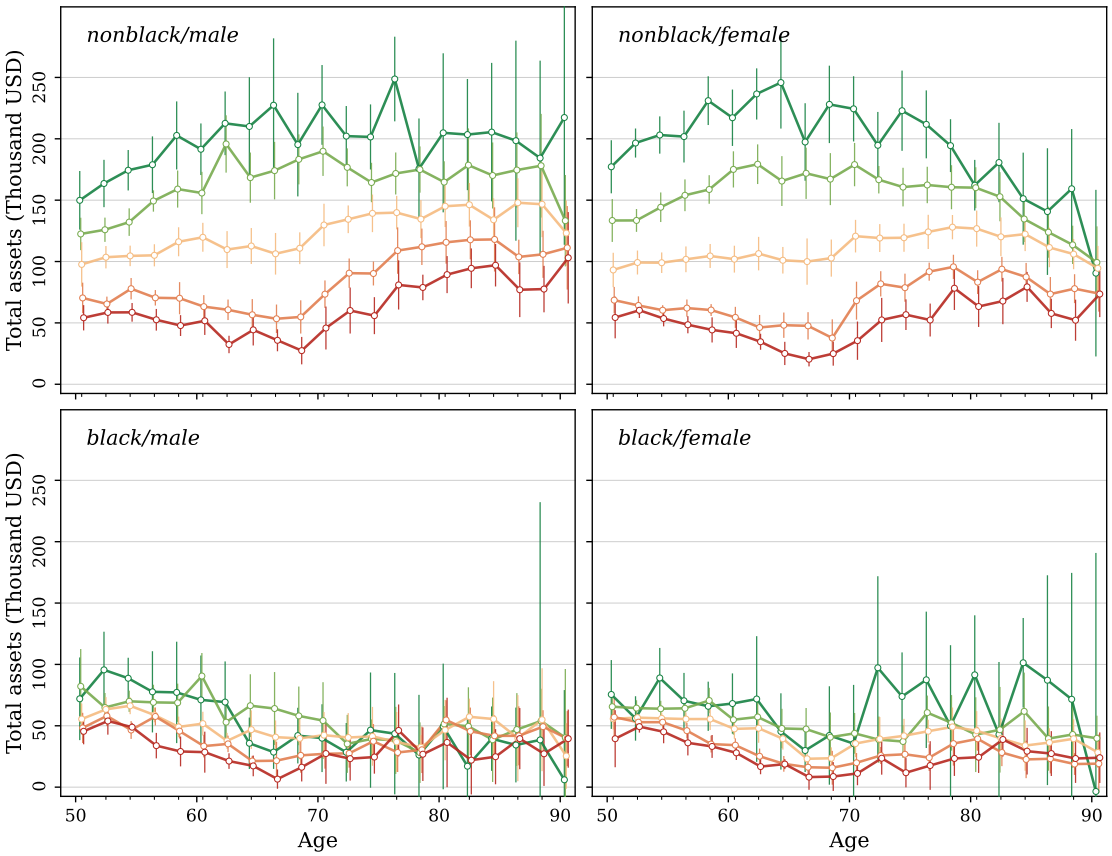


Figure A.5: Median net total household wealth by self-reported health state, race and sex. Pooled sample from HRS 1992–2014. Assets are adjusted for outliers, time and cohort fixed effects, and household size. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

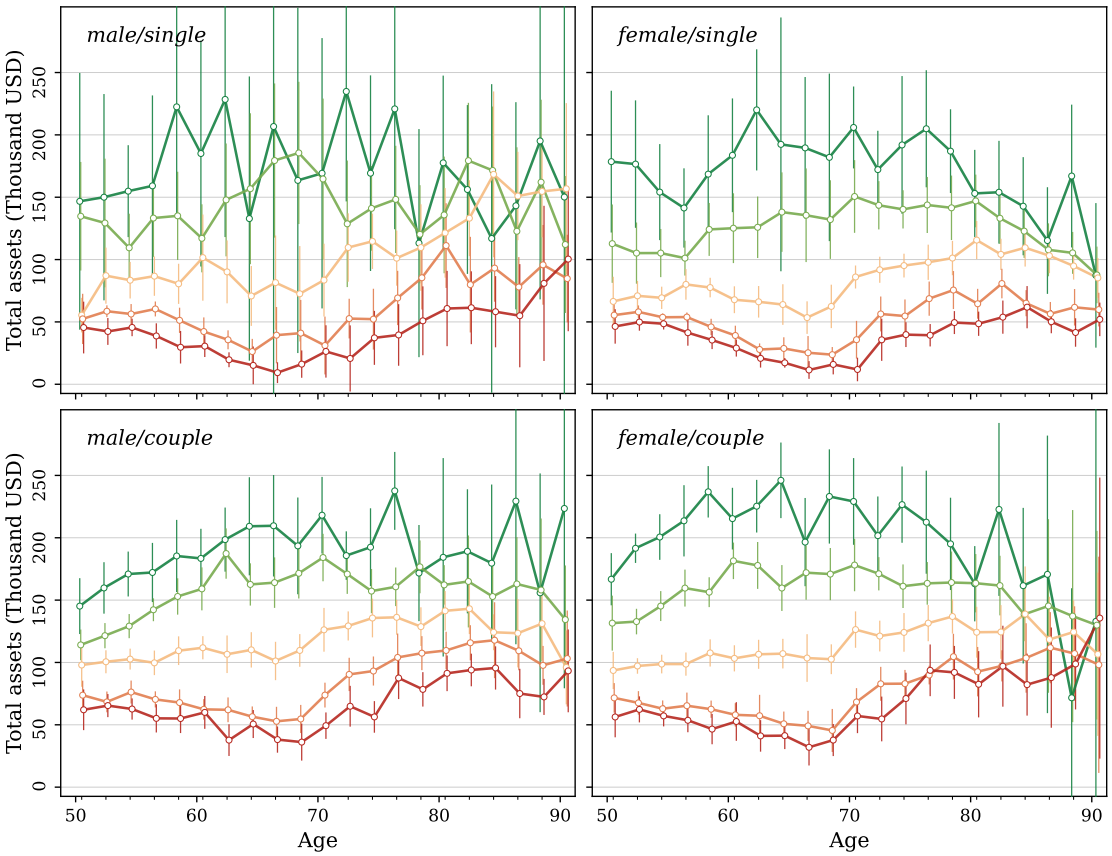


Figure A.6: Median net total household wealth by self-reported health state, household size and sex. Pooled sample from HRS 1992–2014. Assets are adjusted for outliers, time and cohort fixed effects, and household size. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

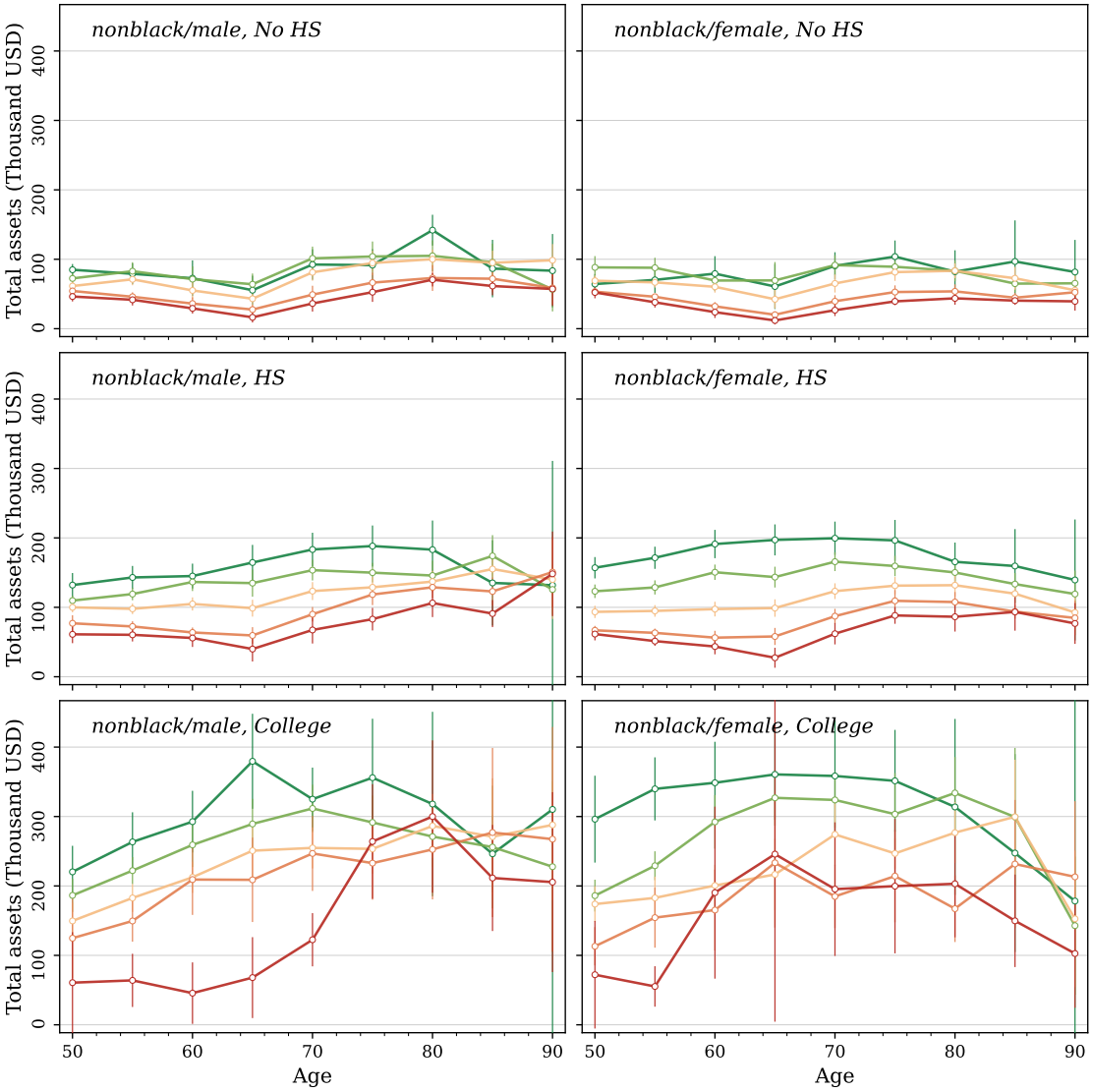


Figure A.7: Median net total household wealth by self-reported health state, sex and education level. Nonblack population. Pooled sample from HRS 1992–2014. Assets are adjusted for outliers, time and cohort fixed effects, and household size. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

A.5 Additional empirical results

A.5.1 Correlation between subjective survival beliefs and wealth

We split the sample into sex/age/race cells, and within each cell we use the elicited subjective belief about surviving until the target age for which we have the most responses within a cell (individuals are often asked about two target ages). We define an indicator variable for "above median SSB" and set it to one whenever individuals report a subjective survival belief above the median (within age/race/sex cell).

We regress assets on the "above median SSB" indicator separately for all demographic groups, controlling for age fixed effects, education and couple status. We take assets measured as adjusted net total wealth in thousands of year 2000 US dollars and then apply the inverse hyperbolic sine transformation since assets are heavily skewed and contains zeros and negative values. All coefficients of interest are positive and significant on the 1% level also using a non-transformed asset measure.

The results for the transformed asset variable are shown in Table 1 in the main text. Being above median in elicited SSB is correlated with more wealth, even after controlling for education and couple status. As a robustness check, we also use survival beliefs as a continuous variable and report the results in Table A.3.

A.5.2 Negative health shocks and asset decumulation

We identify individuals who experience a negative health shock as those who report a deterioration in their self-reported health state by at least one step (e.g., from "excellent"

	Dep. variable: net total wealth (inverse hyperbolic sine transformation)					
	Men			Women		
	(1)	(2)	(3)	(4)	(5)	(6)
Survival belief	0.978***	0.652***	0.625***	1.423***	0.971***	0.919***
	(0.092)	(0.082)	(0.080)	(0.067)	(0.061)	(0.062)
Age FE	Yes	Yes	Yes	Yes	Yes	Yes
Education FE		Yes	Yes		Yes	Yes
Couple FE			Yes			Yes
Observations	54,212	54,212	54,212	70,537	70,537	70,537

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.3: Net total wealth and subjective survival beliefs. The table shows the results of regressing net total wealth (after an inverse hyperbolic sign transformation) on subjective survival beliefs. The regression includes fully interacted fixed effects as indicated. Nonblack population.

	Dep. variable: change in net total wealth (in thousand USD)			
	Men		Women	
	(1)	(2)	(3)	(4)
Negative health shock	-12.720***	-10.786***	-5.705***	-4.609***
	(1.778)	(1.831)	(1.582)	(1.549)
Age FE	Yes	Yes	Yes	Yes
Health FE	Yes	Yes	Yes	Yes
Education FE		Yes		Yes
Observations	61,528	61,528	80,615	80,615

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4: Health shocks and absolute changes in wealth. The table shows the results of regressing changes in net total wealth (in thousand USD) on an indicator for a deterioration in self-reported health between two consecutive survey waves. The regression includes fully interacted fixed effects as indicated. Nonblack population.

to "very good" or worse) between two consecutive survey waves. Then we look at how such a negative health shock is correlated with the individuals' asset accumulation or decumulation.

We split the sample by demographic subgroup (black/nonblack and male/female) and run the following regression:

$$\Delta a_{it+1} = \alpha + \beta \mathbb{1}\left\{health\ shock\right\}_{it+1} + \boldsymbol{x}_{it}'\boldsymbol{\gamma} + \varepsilon_{it} \tag{A.1}$$

with Δa_{it+1} denoting the change in (an inverse hyperbolic sine transformation of) net total wealth between time t and $t+1$, and $\boldsymbol{x}_{it}$ being controls (age fixed effects, education fixed effects, health state in previous wave and interaction terms).

Table 2 in the main text shows the estimates of the coefficient of interest, β, for the nonblack population. Individuals who experience a negative health shock decumulate their assets more compared to individuals (of the same race, age, education level, and initial health) who do not experience a negative health shock.

Table A.4 shows the results for the corresponding regressions using a non-transformed asset variable. The patterns are the same.

One concern is that the decumulation (or slower accumulation) of assets associated with a health deterioration is driven by lower labor income or medical expenditures. To address this issue, we first restrict the sample to nonblack men age 65 or older so that the importance of labor income is small. As a second step, we consider the change in assets removing any changes due to medical expenditures, that is, we define the change

	Dep. variable: change in net total wealth			
	Men age 65+		Men age 65+, net of medical exp.	
	(1)	(2)	(3)	(4)
Negative health shock	-0.116***	-0.105***	-0.114***	-0.101***
	(0.026)	(0.026)	(0.029)	(0.029)
Age FE	Yes	Yes	Yes	Yes
Health FE	Yes	Yes	Yes	Yes
Education FE		Yes		Yes
Observations	35,821	35,821	28,107	28,107

Clustered standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.5: Health shocks and changes in wealth. The table shows the results of regressing changes in net total wealth (after an inverse hyperbolic sign transformation) on an indicator for a deterioration in self-reported health between two consecutive survey waves. The regression includes fully interacted fixed effects as indicated. Nonblack men.

in assets from one wave to the next as:

$$\Delta a_{t+1} = \text{arcsinh}(a_{t+1} + m_{t+1}) - \text{arcsinh}(a_t)$$

with m_{t+1} denoting the out-of-pocket medical expenditures accrued between survey wave t and $t+1$. Table A.5 reports the results from running regression (A.1) on these alternative samples. Columns 1 and 2 show the results for the standard asset definition, while columns 3 and 4 show the results when we remove the effect of medical expenditures. As the table shows, the main result — that a negative health shock is associated with a faster decumulation (or slower accumulation) of assets — remains, even though the magnitudes become smaller.

B Objective life expectancies

B.1 Descriptive statistics

Table B.1 shows descriptive statistics for the sample used to estimate the objective health and survival transitions.

	All	Nonblack		Black	
		Male	Female	Male	Female
Sample size					
N. of individuals	34,196	12,744	15,461	2,424	3,567
N. of observations	219,539	81,278	103,880	13,245	21,136
Avg. observations/indiv.	6.4	6.4	6.7	5.5	5.9
Age distribution					
[50, 60)	34.4%	36.0%	32.3%	40.7%	36.8%
[60, 70)	30.6%	31.8%	29.3%	32.9%	31.2%
[70, 80)	21.7%	21.4%	22.4%	18.3%	19.9%
[80, 90)	11.3%	9.5%	13.3%	7.0%	10.0%
90+	2.1%	1.3%	2.8%	1.0%	2.1%
Mean	66.1	65.3	67.0	64.0	65.4
Self-reported health distribution (all ages)					
(1) Excellent	12.5%	13.6%	12.7%	8.8%	5.1%
(2) Very good	30.0%	30.9%	31.0%	22.1%	20.3%
(3) Good	30.6%	30.8%	29.9%	31.6%	34.0%
(4) Fair	18.6%	17.1%	18.1%	25.9%	28.1%
(5) Poor	8.4%	7.6%	8.3%	11.7%	12.5%

Table B.1: Descriptive statistics for sample used to estimate objective health and survival transitions. Mean age and distributions are weighted using HRS sample weights.

B.2 Estimation results

This section is a brief summary of some of the key results from the estimation of the objective health and survival process. A full description can be found in Foltyn and Olsson (2021).

Figure B.1 plots the life expectancy conditional on the initial health state at the age of 50 and 70 for the different demographic subgroups. As can be seen, the health gradient is substantial: the difference in expected life length between a 50-year-old nonblack man in the best and in the worst health state is 6.1 years. The figure also shows that the life expectancy is lower for men compared to women, and for the black compared to the nonblack subpopulation, as is well known.

Even conditional on education, there is a strong health gradient in life expectancy.

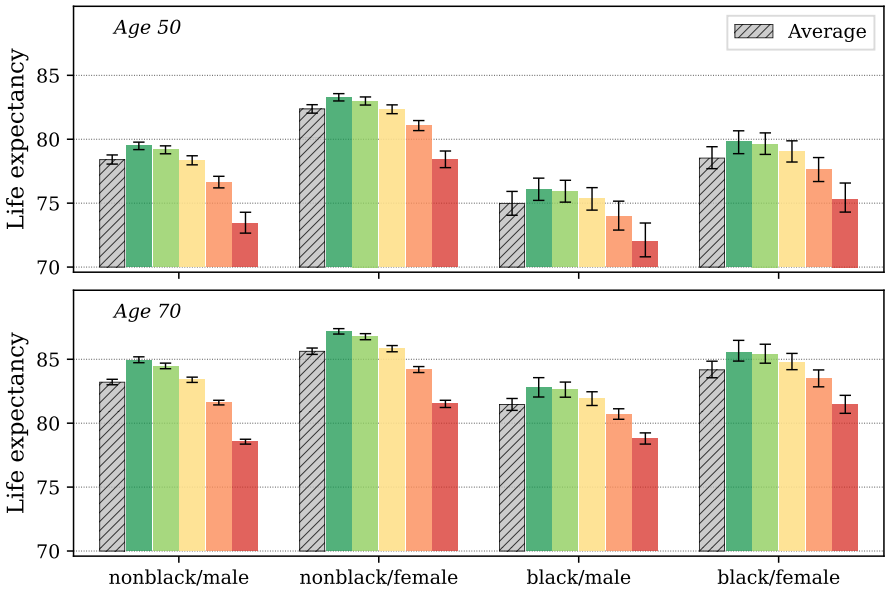


Figure B.1: Life expectancy by race, sex and health state for at age 50 and 70. Average life expectancy is computed as the weighted mean over health states at ages 50–51 (top) or 70–71 (bottom). Error bars indicate 95% confidence intervals.

Figure B.2 shows the life expectancy at age 50 by race, sex, education and health state. For instance, the difference between a college-educated nonblack male in the best health state and an equally educated nonblack male in the worst health state is 4.5 years, while the corresponding difference for college-educated nonblack women is smaller, 2.6 years (but still statistically significant). The corresponding figures for individuals with high-school is 2.8 years (men) and 2.2 years (women). For this analysis, we collapse the data into three health groups, merging health states 1 (excellent) and 2 (very good) into "best" health, and health states 4 (fair) and 5 (poor) into "worst" health, as otherwise the number of observations for some education-health groups becomes too small.

B.3 Comparing model estimates to data moments

We validate our estimated health and survival process and the resulting objective life expectancy calculations in various ways to ensure the estimated process correctly captures the observed data.

First, to compare model predictions to raw data moments, we compute the two-year transition probabilities implied by our annual model. We then plot these together with the fraction of individuals with a particular outcome in the subsample restricted to two-year transitions, which is the large majority of observations (84% of the full sample).

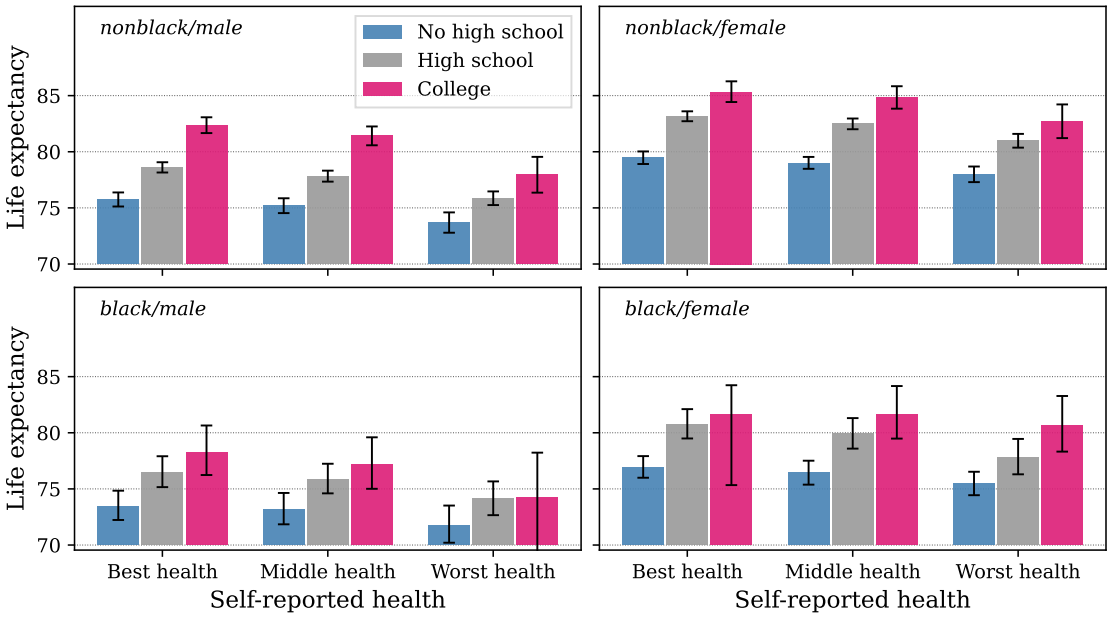


Figure B.2: Life expectancy at age 50 by race, sex, education and health state. Error bars indicate bootstrapped 95% confidence intervals.

Figure B.3 shows the results for the nonblack men. As the figure shows, the estimated probabilities and the data are remarkably close.

Second, to assess how well our model predicts long-run outcomes, we compare actual survival rates as observed in the HRS with model predictions over a time horizon of up to 22 years. Figure B.4 plots the model-predicted survival probabilities for all individuals observed in the survey in 1994 against the fraction actually surviving until 2014. Each dot represents a two-year age bin, and we discard age bins with less than 20 observations. As can be seen, the estimated model captures the long-term survival probabilities well.

Third, we compare the resulting life expectancies from our estimated model to what is reported in the period life tables by the National Vital Statistics System (NVSS) during the years 1992–2014.

Our model gives an average life expectancy of 78.4 years for 50-year-old nonblack men and 82.4 for nonblack women. This is well in line with what is reported by the NVSS during this period. For white men, the NVSS life expectancy at the age of 50 is between 77.0 and 79.9 during the sample period, while for white women it ranges from 81.7 to 83.4. For black men, our model predicts 74.9 years, while NVSS reports between 72.8 and 77.2 for the period. For black women, our model predicts 78.5, while the NVSS reports between 78.3 and 81.5. Thus, in general the model predictions are well within what is reported by NVSS, even though the prediction for life expectancy for

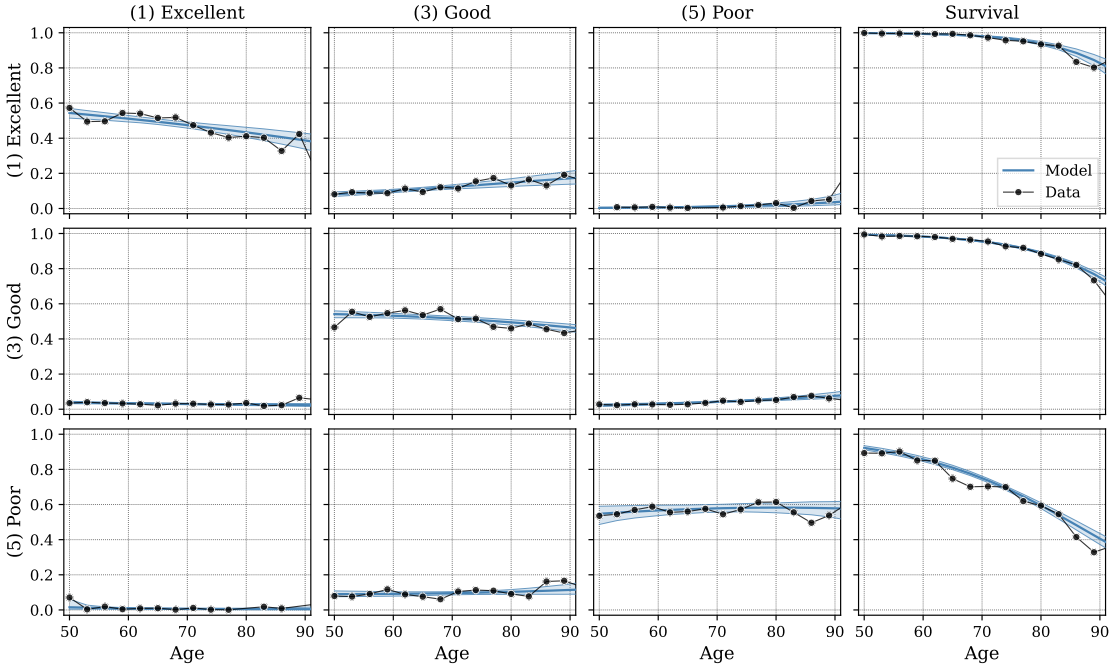


Figure B.3: Two-year transition probabilities for nonblack men. Graph show the best ("excellent"), middle ("good") and worst ("poor") health states. Health transition probabilities are conditional on survival. Right-most column shows survival probabilities. Shaded areas indicate 95% confidence intervals.

black women is on the lower end.

The conclusions are similar for life expectancy at 70. Our model predicts a life expectancy of 83.2 years for nonblack men and 85.6 for nonblack women. The corresponding life expectancies reported by NVSS during the period 1992 to 2014 range between 82.3 and 84.5 for white men, and between 85.3 and 86.6 for white women. For black men the model predicts 81.5, while the NVSS estimates range from 80.8 and 83.3, and for the black women the model prediction is 84.2, while the NVSS estimates range from 83.9 to 86.1.

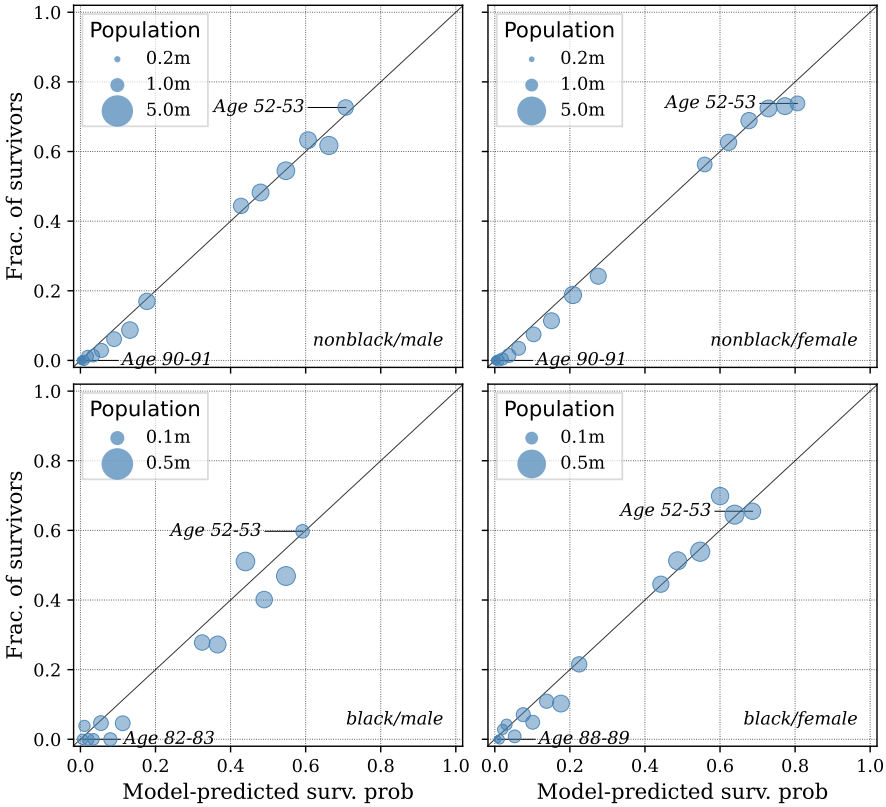


Figure B.4: Model-predicted survival probabilities (on the x-axis) against the fraction of survivors (on the y-axis) for individuals observed in 1994. Each dot represents the fraction of survivors in 2014. Dots are grouped into two-year age bins based on age in 1994.

C Survival beliefs: additional results

C.1 Objective vs. subjective survival probabilities by age and sex

A potential concern when comparing the subjective beliefs about survival to the life-table survival probabilities is that the sample of individuals answering the subjective probability questions is non-representative. Then differences between subjective survival and life tables may arise not only from biases but also from non-representativeness.

In Figure C.1, the solid black line labeled "objective survival probability" shows the weighted average of the objective survival probabilities for the *same sample* as the subjective beliefs. We use the health and survival dynamics model to estimate the objective survival for each individual, conditional on their age, race, sex and health state (using a specification with a linear time trend). We control for cohort and wave fixed effects, using the HRS cohort (birth cohort 1931–1941) and wave 6 (year 2002, in the middle of our sample) as base categories. As can be seen, the graph is very similar to Figure 3 in the main paper.

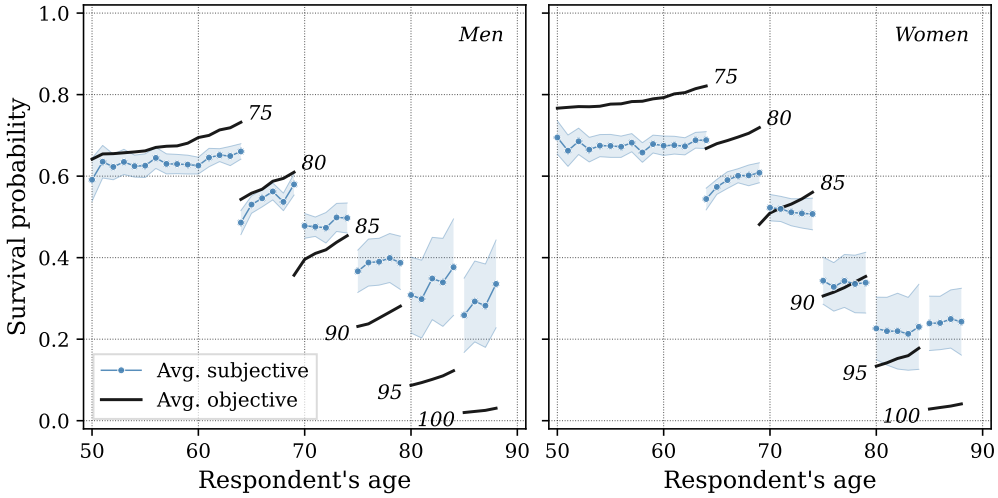


Figure C.1: Objective vs. subjective survival probabilities by age and sex (controlling for year and cohort fixed effects). **The objective probabilities are the weighted average for the individuals reporting a subjective belief.** The number next to the black line indicates target age. The blue line shows (weighted) average expectation in the nonblack population. Shaded areas indicate 95% confidence intervals..

C.2 Expectation errors in survival probabilities to age 95

Figure C.2 shows the elicited beliefs about survival vs. estimated objective (statistical) survival probabilities for target age 95. Target ages 75 and 85 are shown in Figure 3 in the main text.

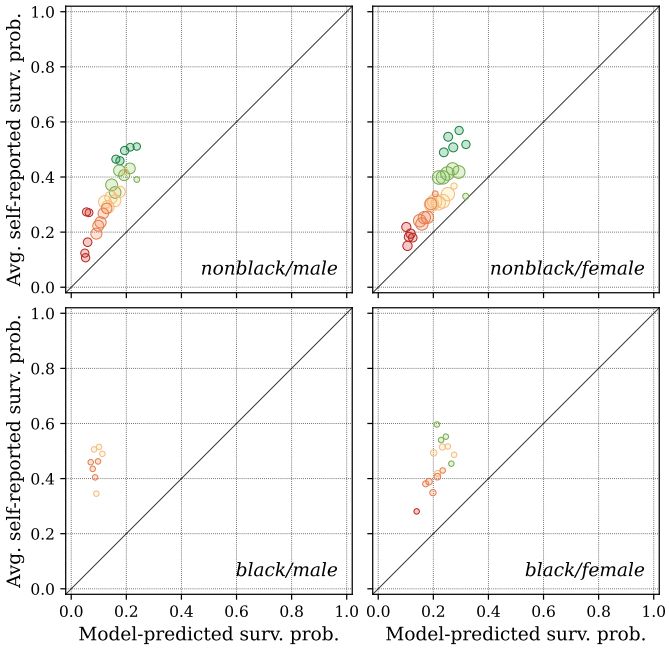
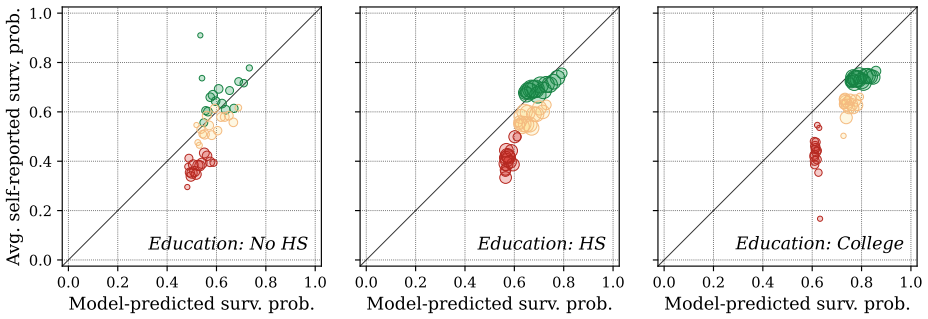


Figure C.2: Elicited beliefs about survival vs. estimated objective (statistical) survival probabilities for *target age 95*. Each bubble represents the average for a race/sex/age/health group. The x-axis shows the model-predicted (objective) survival probability to the indicated target age. The y-axis shows the average self-reported survival probability for that group. Colors indicate the health state: dark green is excellent while red is poor health. The size indicates the number of observations in each cell. We exclude cells with less than 20 observations.

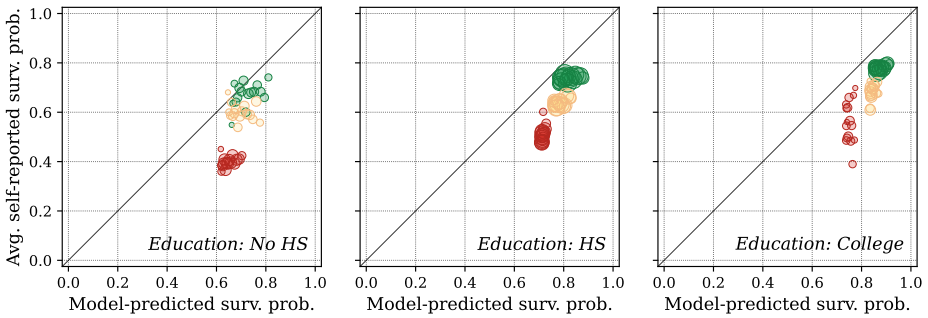
C.3 Expectation errors by education

In order to explore the robustness of the survival bias patterns documented in the main text, we use the estimates from Foltyn and Olsson (2021) for the objective health and survival process for three levels of educational attainment: no high school, high school, and college (see section A.2 for a definition of these groups). We then plot these objective survival probabilities against the elicited survival beliefs for each education level. The results are displayed in Figure C.3 for target age 75, Figure C.4 for target age 85 and in Figure C.5 for target age 95.

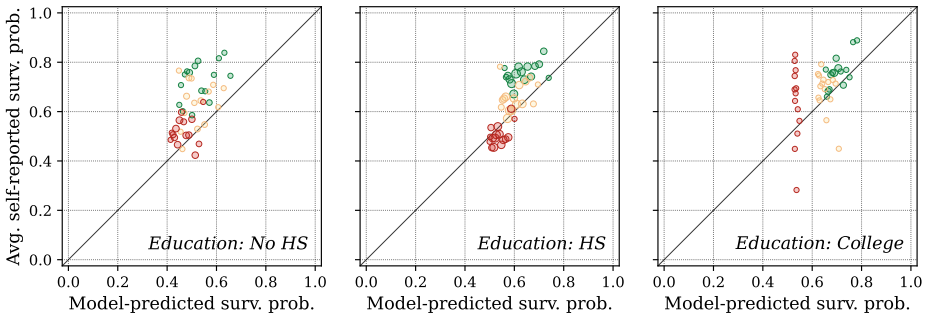
For this exercise, we collapse the original five self-reported health states to three groups, combining the two best and two worst states. This is necessary as the finer partition by education leaves very few observations for some subsamples, especially for the highly-educated black groups (see Table A.2 for the distribution over education levels). This lack of observations is particularly pronounced for college-educated blacks at high ages, as there are hardly any such respondents in the HRS who could be asked about their survival chances to age 95 (see Figure C.5, panel (c)). Note that for this reason, we include all groups with at least five observations in the scatter plots, whereas in the main text we require at least 20 observations per dot to reduce visual clutter.



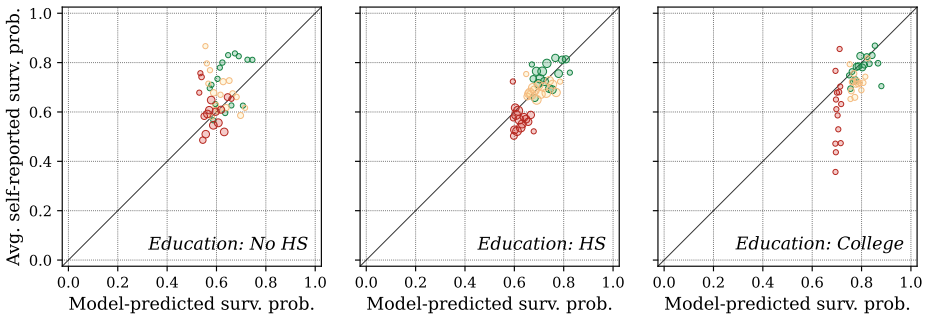
(a) Nonblack men



(b) Nonblack women

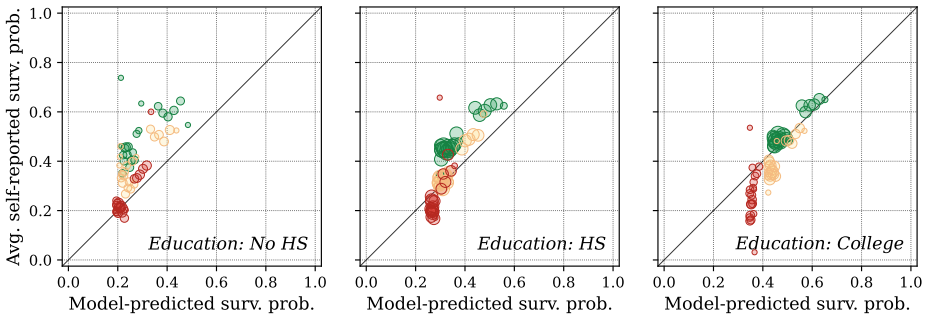


(c) Black men

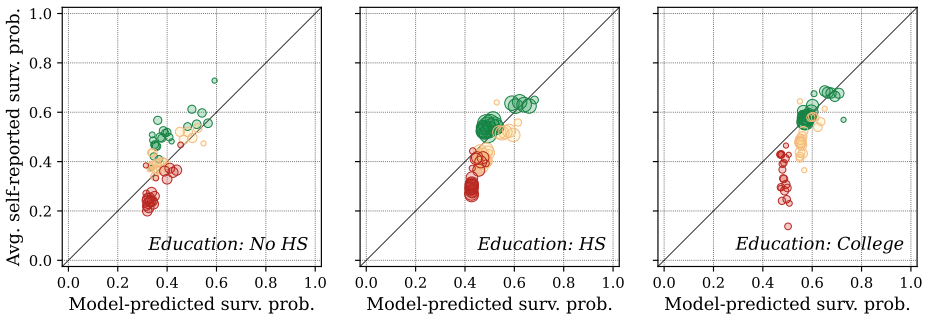


(d) Black women

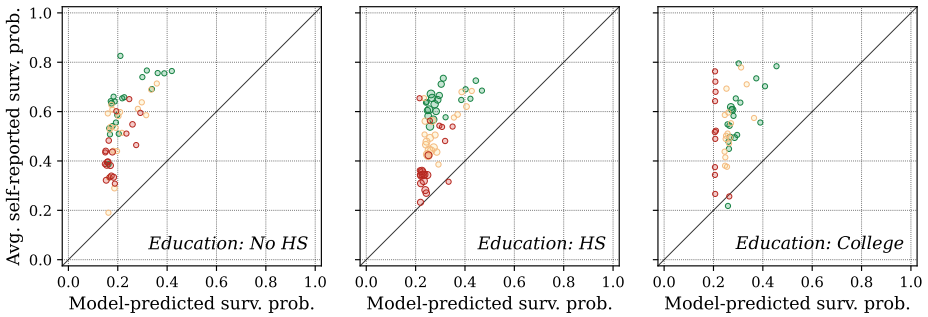
Figure C.3: Elicited beliefs about survival vs. estimated objective (statistical) survival probabilities for *target age 75*. Each bubble represents the average for a race/sex/age/-education/health group. The size indicates the number of observations in each cell. We exclude cells with less than 5 observations.



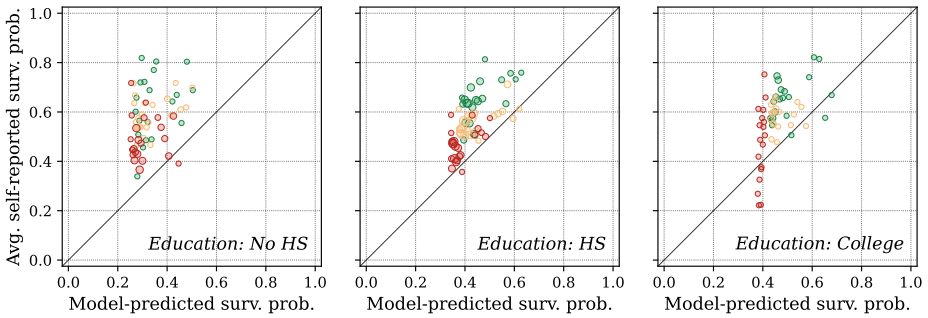
(a) Nonblack men



(b) Nonblack women

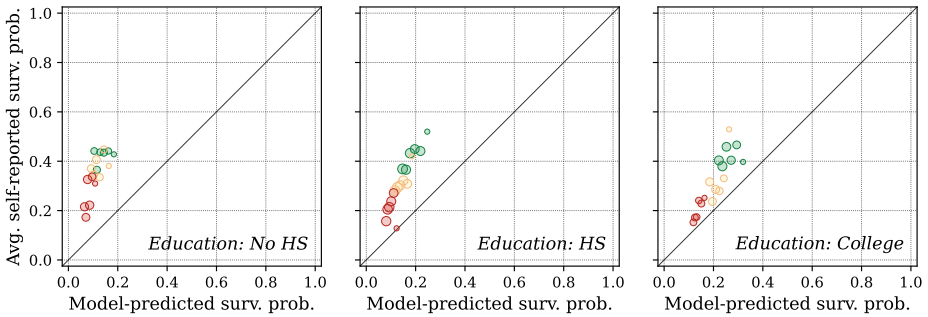


(c) Black men

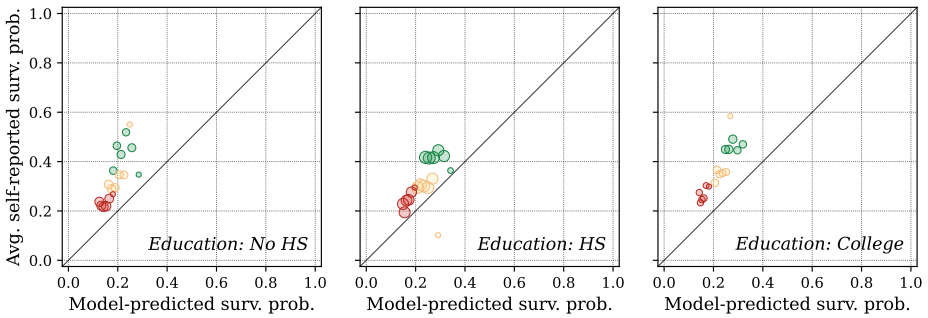


(d) Black women

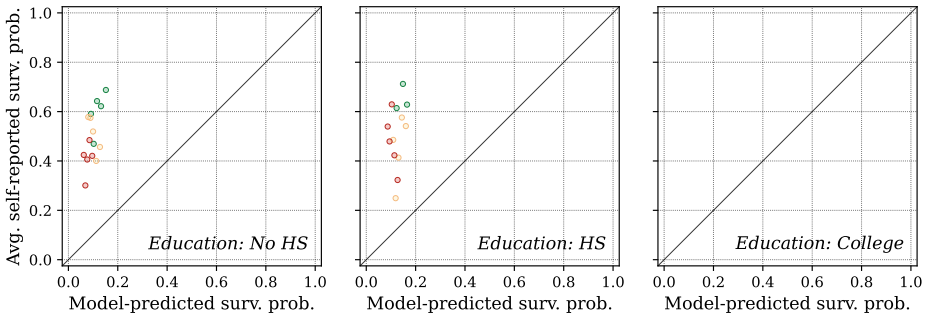
Figure C.4: Elicited beliefs about survival vs. estimated objective (statistical) survival probabilities for *target age 85*. Each bubble represents the average for a race/sex/age/-education/health group. The size indicates the number of observations in each cell. We exclude cells with less than 5 observations.



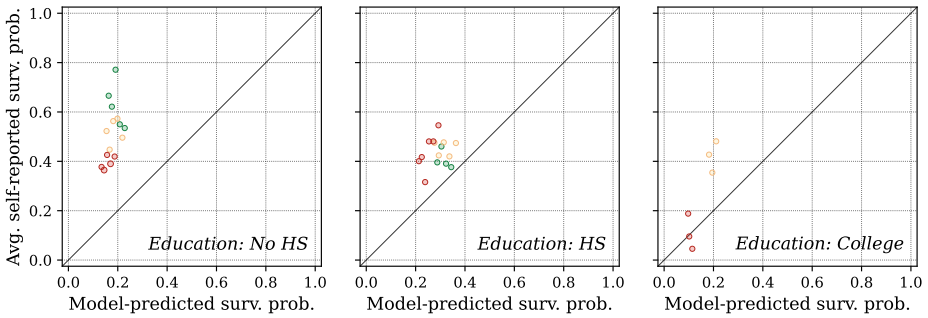
(a) Nonblack men



(b) Nonblack women



(c) Black men



(d) Black women

Figure C.5: Elicited beliefs about survival vs. estimated objective (statistical) survival probabilities for *target age 95*. Each bubble represents the average for a race/sex/age/-education/health group. The size indicates the number of observations in each cell. We exclude cells with less than 5 observations.

C.4 Survival bias by sex, race, age and self-reported health

Figure 5 in the main text shows the survival bias by sex, age and self-reported health for the nonblack population. Figure C.6 shows the corresponding graphs for the black population. As the figure shows, the patterns are the same, but standard errors are larger due to smaller sample size.

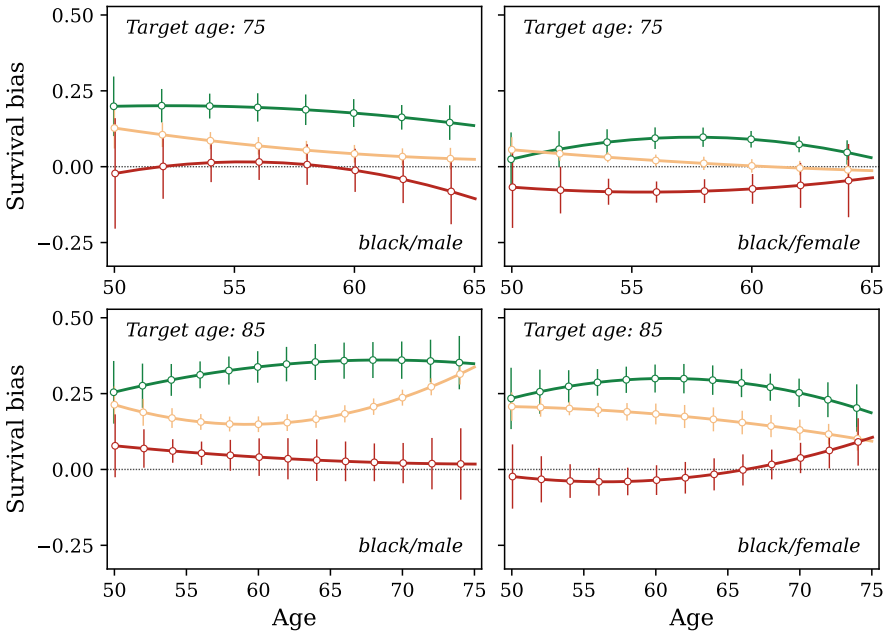


Figure C.6: Survival bias by sex, age and self-reported health. Black population. A positive bias of 0.1 indicates that the group is overestimating their survival probability by 10 percentage points. Colors indicate the health state: dark green is excellent while red is poor health. Error bars indicate 95% confidence intervals.

C.5 Comparison to Grevenbrock et al. (2021)

In a related paper, Grevenbrock et al. (2021) (henceforth GGLZ) estimate survival based on several additional characteristics besides self-reported health, age, and sex, such as smoking and drinking behavior, and chronic diseases. Grouping individuals based on their estimated objective survival probability, they find that individuals with low objective survival probability in general overestimate, while individuals with high objective probability underestimate their survival probabilities, in other words, at a first glance the reverse pattern compared to what we find.

There are three main differences between our approaches:

1. Different estimation strategies for objective survival probabilities (OSPs)
2. Different samples (both for objective survival and subjective beliefs)
3. Different grouping of individual data

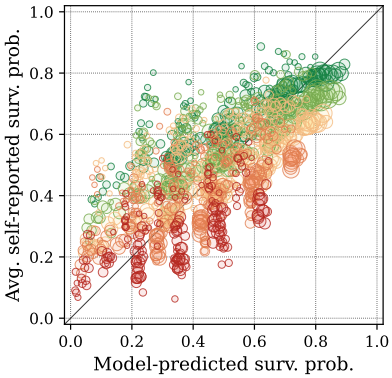
As we show below, point 3 is the most important. If we apply the same sample restrictions in terms of age and survey years as GGLZ (which leaves us with about a fifth of our original sample) and group our data in the same way as GGLZ do, we more or less recover the same pattern.

Figure C.7 shows the importance of the way the observations are grouped in practice. First, Figure C.7(a) shows the "raw data", merged into the same graph. In other words: all sex/race/age/health cells presented separately in Figure 3(a), Figure 3(b), and Figure C.2 are here placed together in the same graph.

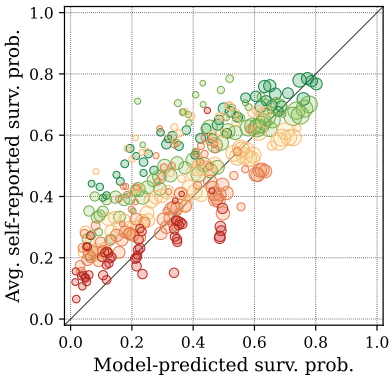
Next, Figure C.7(b) presents a restricted sample: only observations where a respondent is 65 or older from wave 8–12 are included. This corresponds to the sample used in GGLZ.

Finally, Figure C.7(c) shows the resulting pattern if we group the observations in the previous panel into 5% objective probability bins. As this last figure shows, grouped in this manner there is a flatness pattern, as in the GGLZ paper.

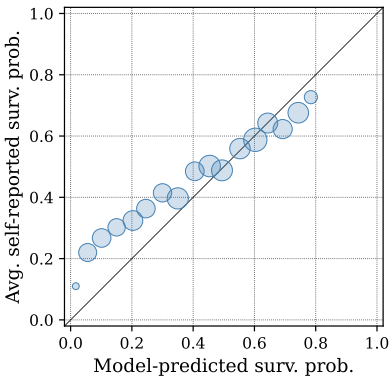
It is interesting to note that GGLZ regress observed biases on psychological variables and show that more optimistic respondents overestimate their survival chances. This could help shed some light on why the pattern we observe emerges, if individuals in good health are generally more optimistic.



(a) Pooled sample across race and sex. Each bubble represents the average for a sex/race/age/health group. Colors indicate the health state: dark green is excellent while red is poor health.



(b) Same plot as subfigure (a), but restricting the sample to age 65+ and only waves 8–12 (year 2006–2012). This corresponds to the sample used by GGLZ.



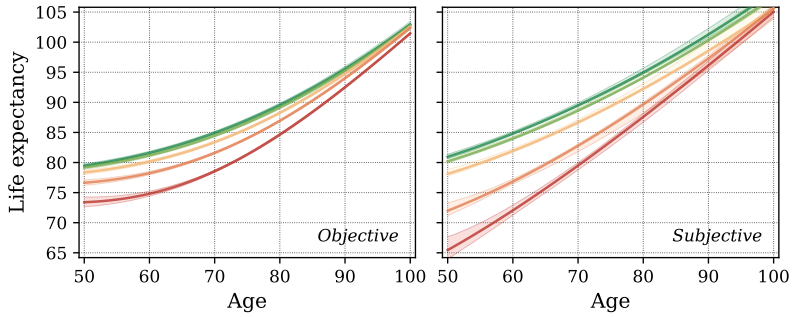
(c) Same data as subfigure (b), grouped into 5% objective probability bins.

Figure C.7: Illustration of the effect of regrouping the data used in our paper. All graphs show beliefs about survival vs. estimated objective (statistical) survival probabilities. The x-axis shows the model-predicted (objective) survival probability, the y-axis the average self-reported survival probability for that group. The size indicates the number of observations in each cell. We exclude cells with less than 20 observations.

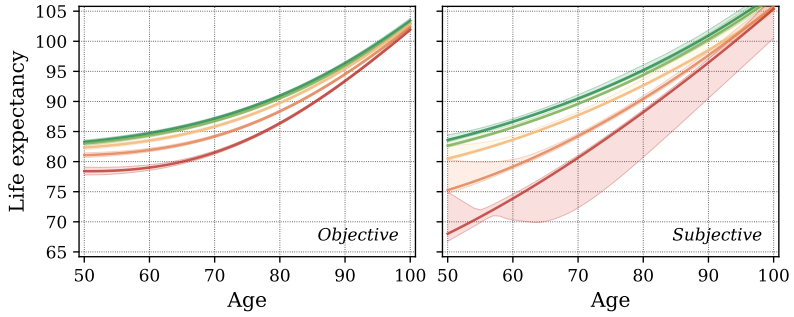
C.6 Objective vs. subjective life expectancy by age, sex, and health

This section contains additional graphs contrasting the objective and subjective expected longevity for all demographic subsamples. In Figure C.8, we plot the objective life expectancy on the left and the life expectancy implied by the subjective survival beliefs on the right (panel Figure (a) is the same as Figure 7 in the main text).

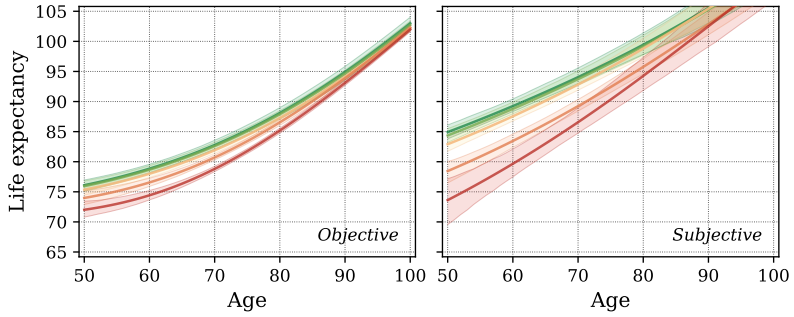
The main take-away from these graphs is that the survival bias exhibits a similar pattern across all groups, especially among the nonblack population. However, due to the small black sample size, in particular among the very old (see Tables B.1 and 3), the confidence intervals for these estimates are very wide.



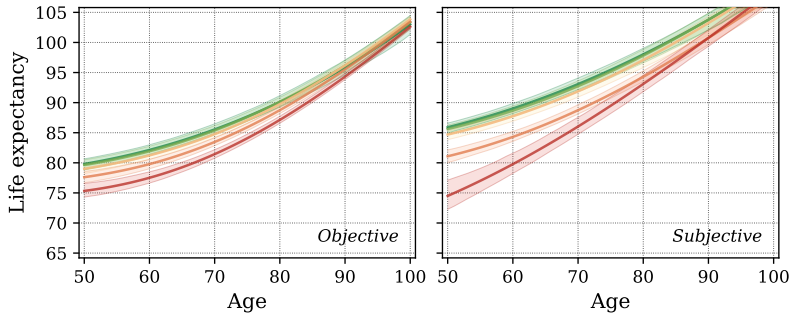
(a) Nonblack men



(b) Nonblack women



(c) Black men

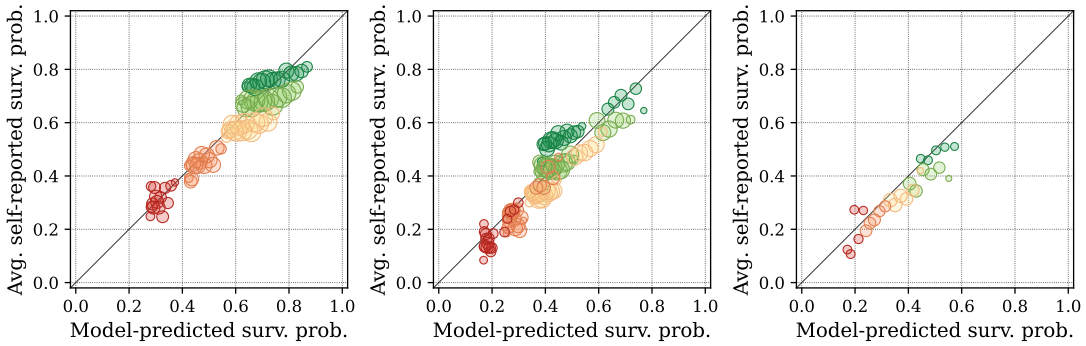


(d) Black women

Figure C.8: Life expectancy by age and health. Colors indicate the health state: dark green is excellent while red is poor health. Shaded areas indicate 95% confidence intervals. For each bootstrapped sample we re-estimate the objective health process.

C.7 Estimated subjective survival probability vs. elicited beliefs

In Figure C.9, we plot the model-predicted subjective survival against elicited beliefs. As can be seen, the estimated model for subjective beliefs captures the main picture very well, since the dots, each representing an age/health/target-age group, line up closely along the 45-degree line.



(a) Target age 75 **(b)** Target age 85 **(c)** Target age 95

Figure C.9: Elicited beliefs about survival vs. estimated *subjective* survival probabilities. Each bubble represents the average for an age/health group. The x-axis shows predicted survival probability according to the subjective model and the y-axis the average self-reported survival probability for that group. Colors indicate the health state: dark green is excellent while red is poor health.

D Model details

D.1 Intergenerational linkages and bequests

This section provides a detailed discussion of the stochastic intergenerational links used to distribute bequests left by parents to children in a way that creates positive intergenerational sorting similar to what is observed in US data.

D.1.1 Household problem

Consider the retired agent's problem from the main text which we restate here for convenience:

$$
\begin{aligned}
V_r\left(a,p,h,\eta,\mathbb{1}_b,t\right) = \max_{c,\,k}\Bigg\{ & u(c) + \beta\pi^s_{ht}\mathbb{E}\left[\, V_r\left(\boldsymbol{x}'\right) \,\middle|\, p,h,\eta,\mathbb{1}_b,t \,\right] \\
& + \beta\left(1-\pi^s_{ht}\right)\mathbb{E}\left[\, V_b\left(b'\right) \,\middle|\, h,\eta,\mathbb{1}_b,t \,\right]\Bigg\} \qquad \text{(D.1)}
\end{aligned}
$$

$$
\begin{aligned}
\text{s.t.}\quad c+k &= a \\
a' &= Rk + b'_* + \iota' - m\left(h',\eta',\nu',t+1\right) + \xi' \\
\xi' &= \max\left\{0,\ \underline{c} + m\left(h',\eta',\nu',t+1\right) - Rk - b'_* - \iota'\right\}
\end{aligned}
$$

$$
\left(b'_*,\ \mathbb{1}'_b\right) \,\middle|\, \left\{\mathbb{1}_b = 1\right\} \sim \begin{cases} \left(b_*(p,h,t),0\right) & \text{with prob. } \left(1-\overline{\pi}^s_{t_*}\right)\pi^b_{pht} \\ (0,0) & \text{with prob. } \left(1-\overline{\pi}^s_{t_*}\right)\left(1-\pi^b_{pht}\right) \\ (0,1) & \text{with prob. } \overline{\pi}^s_{t_*} \end{cases} \qquad \text{(D.2)}
$$

$$
\left(b'_*,\ \mathbb{1}'_b\right) \,\middle|\, \left\{\mathbb{1}_b = 0\right\} = (0,0)
$$

$$
\begin{aligned}
b' = \max\left\{0, Rk + b'_* - m\left(\eta',\nu',t+1\right)\right\} & \\
- T_b\left(\max\left\{0, Rk + b'_* - m\left(\eta',\nu',t+1\right)\right\}\right) & \qquad \text{(D.3)}
\end{aligned}
$$

Throughout this section, we adopt the convention that quantities designated with a $*$ subscript pertain to parents, e.g., b_* being bequests left by an agent's parents, while b denotes the bequests left by the agent.

Instead of directly modeling the intergenerational joint distribution of the children's and parents' state vectors, we proxy such links via an intergenerational correlation of lifetime income quintiles which we take from Chetty et al. (2014). However, since lifetime income quintiles are not part of an agent's state, we approximate these using the states (p,h,t) in a probabilistic fashion. This implies that the bequests b_* received by

children are fully characterized by (p, h, t), and thus, conditional on receiving a positive bequest, deterministic.

The probabilistic mapping between children's and parents' states has three components and works as follows:

1. $Q_1 : (p, h, t) \mapsto \Delta^4(q)$ defines a mapping of children's states into lifetime income quintiles q where $\Delta^4(q)$ is a 4-dimensional simplex which represents the PMF over these quintiles. For example, a child who was in the highest productivity state and in excellent health at the age of 50 is more likely to be in the top lifetime income quintile than someone who was at the bottom of the labor productivity distribution or in poor health.

2. $Q_2 : q \mapsto \Delta^4(q_*)$ represents the intergenerational linkages by mapping each child's lifetime income quintile into a distribution over the parents' lifetime income quintiles. This mapping is taken directly from Chetty et al. (2014).

3. $Q_3 : q_* \mapsto \Delta^{|\mathcal{P}|\cdot|\mathcal{H}|-1}(p_*, h_* \,|\, t_* = t + T_*)$ maps the parental lifetime income quintile back into a distribution over the parental (p_*, h_*) conditional on the parents' age $t_* = t + T_*$. We impose that parents are always $T_* = 30$ years older than their children.

For simplicity, the probability that an agent's parents survive (and thus no bequest is passed on),

$$\overline{\pi}^s_{t_*} \equiv \mathbb{E}\left[\, \pi^s_{h_*,t_*} \,\middle|\, t_* \,\right], \tag{D.4}$$

is independent of both the parents' health (which is averaged out) as well as the child's health. However, the intergenerational correlation in lifetime income quintiles and the interdependency of lifetime income and health implies that healthier children are more likely to have healthier parents, who it turn are more likely to live longer. We ignore this additional intergenerational link since it is computationally infeasible to keep track of, and we have no data on self-reported health across generations.

D.1.2 Intergenerational income links

We use the 5×5 double stochastic matrix from Chetty et al. (2014, Table II) shown in Table D.1 which summarizes the joint distribution of parent and child incomes in the US:

A few notes on how this transition matrix was obtained:

	Parent quintile				
Child quintile	1	2	3	4	5
1	33.7%	24.2%	17.8%	13.4%	10.9%
2	28.0%	24.2%	19.8%	16.0%	11.9%
3	18.4%	21.7%	22.1%	20.9%	17.0%
4	12.3%	17.6%	22.0%	24.4%	23.6%
5	7.5%	12.3%	18.3%	25.4%	36.5%

Table D.1: Intergerational quintile transition matrix for the US. Source: Chetty et al. (2014), Table II

1. The intergenerational links are based on US families who had children born in 1980–1982.

2. Parental income is measured as the average gross family income including taxable capital returns between 1996–2000.

3. Child income is measured in the same way for the years 2011–2012, i.e., when the child generation was about 29–32.

4. Chetty et al. (2014) do various robustness checks, for example looking at earlier child cohorts (from another, smaller, data source). The main conclusion is that the rank-rank correlation is mostly unaffected because the income ranks have stabilized by the age of 30.

Relying on the last point, we abstract from the fact that these rank-rank correlations are computed at a particular age. Instead, we interpret the matrix in Table D.1 as the joint distribution of parent and child *lifetime income* quintiles.

Lifetime income in the model. In the model, we don't keep track of lifetime income (we don't keep track of individuals during the simulation when solving for the equilibrium). However, for each point on the state space, we need to compute the probabilistic distribution over the population lifetime income quintiles as a function of states (p, h, t). To this end, we run an auxiliary simulation of the exogenous earnings process for a panel of one million households where we keep track of individuals until retirement (this simulation is completely independent of the model solution and needs to be run only once). Using this simulated panel, we establish the mapping from (p, h, t) as follows:

1. Simulate household $i = 1, \ldots, N$ starting at model age 1 until model age 45. This involves drawing the exogenous shocks for health h and the persistent and

transitory income shocks p and ϵ.

2. Compute the average working-age gross income for individual i as

$$I_i = \frac{1}{T_r - 1} \sum_{t=1}^{T_r - 1} y_{it}$$

In what follows, we use the simpler term *lifetime income* whenever we refer to this working-age income in the model.

3. Rank individuals by their lifetime income I_i and bin them into quintiles.

4. Using the simulated data, for any tuple of values (p, h, t) take all simulated individuals who "passed through" this particular point on the state space and compute the probability distribution over the population lifetime income quintiles. For example, assuming that 100 individuals were observed in the cell (p, h, t), we compute the fraction of these individuals who ended up in each of the lifetime income quintiles. Because the income component p and health h are somewhat persistent, this distribution is not uniform and thus informative about an individual's likely position in the lifetime income distribution.

This creates the mapping

$$Q_I : \mathcal{P} \times \mathcal{H} \times \{1, 2, \ldots, T_r - 1\} \mapsto \Delta^4 \tag{D.5}$$

where Δ^4 is the 4-dimensional simplex that defines the PMF over lifetime income quintiles.

5. For retired agents, income no longer fluctuates and is no longer health dependent, so for them we need to keep track of their last health realization h just before retirement to be able to map them into lifetime income quintiles. During the model simulation (the one used to find the equilibrium, not the auxiliary one outlined above), we therefore need to store agents' health in their last year of working life, $h_{T_r - 1}$. Note that this does not add an additional state to the agent's problem, but only needs to be stored during the simulation to be able to apply the mapping Q_I.

D.1.3 Bequests left by parents

Since our intergenerational mapping works via the joint distribution over income quintiles from Chetty et al. (2014), it is most convenient to characterize the equilibrium in

terms of the aggregate bequests left by parents of age $t_* = t + T_*$ who are in income quintile q_*,

$$B(q_*, t_*) = \mathbb{E}\left[b \,\middle|\, q_*, t_* \right] \Pr\left(q_* \mid t_* \right) \mu_t(t_*) \tag{D.6}$$

where $\mathbb{E}\left[b \,\middle|\, q_*, t_* \right]$ is the average bequest left by an individual of age t_* in income quintile q_*, $\Pr\left(q_* \mid t_* \right)$ is the probability of being in that income quintile and $\mu_t(t_*)$ is the mass of individuals aged t_*.

First, note that the probability of being in lifetime income quintile q_* conditional on age can be computed as

$$\Pr\left(q_* \mid t_* \right) = \sum_{p_*} \sum_{h_*} \Pr\left(q_* \mid p_*, h_*, t_* \right) \mu_p(p_*) \mu_h(h_* \mid t_*) \tag{D.7}$$

where $\Pr\left(q_* \mid p_*, h_*, t_* \right)$ is given by the mapping (D.5).

Next we derive the average bequests left by parents conditional on their income quintile and age. Because an agent's lifetime income quintile does not enter the state space, this expression can only be approximated using the probabilitistic mapping from parental lifetime income quintiles to (p_*, h_*, t_*) as follows:

$$\begin{aligned}
\mathbb{E}\left[b \,\middle|\, q_*, t_* \right] =& \mathbb{E}\left[\mathbb{E}\left[b \mid p_*, h_*, q_*, t_* \right] \,\middle|\, q_*, t_* \right] \\
\approx& \mathbb{E}\left[\mathbb{E}\left[b \mid p_*, h_*, t_* \right] \,\middle|\, q_*, t_* \right] && \text{(D.8)} \\
=& \sum_{p_*} \sum_{h_*} \Pr\left(p_* \wedge h_* \mid q_*, t_* \right) \mathbb{E}\left[b \,\middle|\, p_*, h_*, t_* \right] && \text{(D.9)}
\end{aligned}$$

The probability of observing (p_*, h_*) conditional on income quintile and age is not directly part of our model, but can be computed from Bayes' rule:

$$\Pr\left(p_* \wedge h_* \mid q_*, t_* \right) = \frac{\Pr\left(q_* \mid p_*, h_*, t_* \right) \Pr\left(p_* \wedge h_* \mid t_* \right)}{\Pr\left(q_* \mid t_* \right)} \tag{D.10}$$

where

$$\Pr\left(p_* \wedge h_* \mid t_* \right) = \mu_p(p_*) \, \mu_h(h_* \mid t_*)$$

since persistent income is independent of health and age, and $\mu_p(\bullet)$ is the PMF over persistent income while $\mu_h(\bullet \mid t_*)$ is the age-dependent PMF over health. In the absence of heterogeneous mortality risk, the numerator $\Pr\left(q_* \mid t_* \right)$ in (D.10) would always be 20% since at each age, a fifth of the surviving cohort of households must be in any given lifetime income quintile. However, because both mortality and lifetime income are correlated with health, this need not hold here, as shown in (D.7).

Note that the approximation in (D.8) arises because from the law of iterated expectations we have

$$\mathbb{E}\left[\, b \,\middle|\, q_*, t_* \,\right] = \mathbb{E}\left[\, \mathbb{E}\left[\, b \,\middle|\, p_*, h_*, q_*, t_* \,\right] \,\middle|\, q_*, t_* \,\right] \neq \mathbb{E}\left[\, \mathbb{E}\left[\, b \,\middle|\, p_*, h_*, t_* \,\right] \,\middle|\, q_*, t_* \,\right]$$

where the last two expressions need not be the same in general since (p_*, h_*, t_*) are only an imperfect measure of q_*. We assume that they are sufficiently close since it is computationally infeasible to include the lifetime income quintile in the state space of an already very rich model, and thus the exact expression cannot be computed.

Finally, the average bequest left by the parent generation conditional on their state (p_*, h_*, t_*) which enters (D.9) is given by

$$\begin{aligned}\mathbb{E}\left[\, b \,\middle|\, p_*, h_*, t_* \,\right] = (1 - \pi^s_{h_* t_*}) \sum_{\mathbb{1}_{b*} \in \{0,1\}} \sum_{\eta_*} \mu_\eta(\eta_*) \Big\{ &\Pr\left(\, \mathbb{1}_{b*} \,|\, p_*, h_*, t_* \,\right) \\ &\times \int_a b'\left(a, p_*, h_*, \eta_*, \mathbb{1}_{b*}, t_*\right) d\Gamma(a \,|\, p_*, h_*, \eta_*, \mathbb{1}_{b*}, t_*) \Big\} \end{aligned} \tag{D.11}$$

where b' is the bequest left by a parent who dies at the end of the period, defined in (D.3).

Note that the expression in (D.11) is not directly needed to find an equilibrium and is only used to compute aggregate bequests in (D.6). For the equilibrium itself, we only need to find a fixed point for $B(q_*, t_*)$ which is a matrix of $5 \times 60 = 300$ elements, whereas the expectation in (D.11) can take on $5 \times 5 \times 60 = 1500$ distinct values. See section H for details.

D.1.4 Bequests received by children

The central quantities entering the equilibrium definition of the model are the bequests left by parents $B(q_*, t_*)$ from (D.6). In this section, we characterize how these are distributed among children which in turn defines the conditional bequests received by children $b_*(p, h, t)$ which enters the household problem (D.2).

Consider the bequests left by parents in income quintile q_* which need to be distributed among children of age $t = t_* - T_*$ who can potentially have any combination of permanent income and health (p, h). The probability that a child of such a parent is in a particular state is

$$\Pr\left(\, p \wedge h \,|\, q_*, t_* \,\right) = \sum_q \Pr\left(\, p \wedge h \,|\, q, t \,\right) \Pr\left(\, q \,|\, q^* \,\right)$$

where the probabilistic intergenerational income quintile linkages are given by Table D.1. We again use Bayes' rule to express $\Pr(\, p \wedge h \,|\, q, t\,)$ in terms of quantities we know,

$$\Pr(\, p \wedge h \,|\, q, t\,) = \frac{\Pr(\, q \,|\, p, h, t\,) \Pr(\, p \wedge h \,|\, t\,)}{\Pr(\, q \,|\, t\,)}$$

Because the probability to receive a bequest conditional on the parents' death π^b_{pht} also depends on the child's income quintile, bequests from $B(q_*, t_*)$ are ultimately distributed proportionally to

$$\Pr(\, p \wedge h \wedge b_* > 0 \,|\, q_*, t_*\,) = \sum_q \Pr(\, p \wedge h \wedge b_* > 0 \,|\, q, t\,) \Pr(\, q \,|\, q^*\,)$$

Note that the probability to receive bequests conditional on the parents' death is independent of (p, h) *conditional* on q, so that

$$\Pr(\, p \wedge h \wedge b_* > 0 \,|\, q, t\,) = \Pr(\, p \wedge h \,|\, q, t\,) \Pr(\, b_* > 0 \,|\, q\,)$$

where we also use the assumption that, conditional on parental death, $\Pr(\, b_* > 0 \,|\, q\,)$ is independent of the parents' age. As described in section 4 in the main text, we estimate the probabilites $\Pr(\, b_* > 0 \,|\, q\,)$ for the five income quintiles using the Survey of Consumer Finances.

It can be shown that the conditional bequest received by children is then given by the expression

$$b_*(p, h, t) = \varsigma(p, h, t) \cdot \sum_{q_*} B(q_*, t_*) \frac{\Pr(\, p \wedge h \wedge b_* > 0 \,|\, q_*, t_*\,)}{\overline{\pi}^p_{q_*}} \tag{D.12}$$

with

$$\overline{\pi}^p_{q_*} = \Pr(\, b_* > 0 \,|\, q_*\,) = \sum_q \Pr(\, b_* > 0 \,|\, q\,) \Pr(\, q \,|\, q_*\,)$$

being the average probability to receive a positive bequest conditional on parental death among all children linked to parents in income quintile q_*. The scaling factor $\varsigma(p, h, t)$ ensures that the bequests received by children are equal to the bequests left by parents and is given by

$$\varsigma(p, h, t) = \left[\left(1 - \overline{\pi}^s_{t_*}\right) \mu_t(t) \mu_b(\mathbb{1}_b = 0 \,|\, t)\ \mu_p(p)\ \mu_h(h \,|\, t)\ \pi^b_{pht} \right]^{-1} \tag{D.13}$$

which is the mass of children at (p, h) whose parents were alive at the beginning of the

period and died at the end, and who receive a positive bequest. Since children can die before their parents, there is a subtle distinction between the mass of parents alive and the mass of children whose parents are still alive, where the latter is (weakly) smaller. Denote by $\mu_b(\mathbb{1}_b = 1 \mid t)$ the fraction of a cohort with state $\mathbb{1}_b = 1$. Since the mass of children with $\mathbb{1}_b = 1$ must be equal to the mass of parents whose children are still alive, this fraction is implicitly defined by

$$\mu_t(t)\mu_b(\mathbb{1}_b = 0 \mid t) = \mu_t(t_*)\frac{\mu_t(t)}{\mu_t(1)} \implies \mu_b(\mathbb{1}_b = 0 \mid t) = \frac{\mu_t(t_*)}{\mu_t(1)} \tag{D.14}$$

where fraction of children alive is expressed as the cohort size aged t relative to the cohort size at birth, $\mu_t(1)$.

As mentioned earlier, it is computationally infeasible to link children's and parents' health directly, hence we use the average parental survival probability at age t_* in (D.13) which was defined in (D.4).

D.1.5 Equilibrium

It is straightforward to show that if the bequests received by children are given by the expression in (D.12), then the aggregate bequests received by cohort t, $B_*(t)$, are equal to the bequests left behind by their parent cohort, $B(t_*)$:

$$\begin{aligned}
B_*(t) &\equiv \mu_t(t)\frac{\mu_t(t_*)}{\mu_t(1)}\left(1 - \overline{\pi}^s_{t_*}\right)\sum_p\sum_h \mu_p(p)\mu_h(h \mid t)\pi^b_{pht}b_*(p,h,t) \\
&= \sum_p\sum_h\sum_{q_*} B(q_*,t_*)\frac{1}{\overline{\pi}^p_{q_*}}\sum_q \frac{\Pr\left(q \mid p,h,t\right)\mu_p(p)\mu_h(h \mid t)}{\Pr\left(q \mid t\right)}\Pr\left(b_* > 0 \mid q\right)\Pr\left(q \mid q^*\right) \\
&= \sum_{q_*} B(q_*,t_*)\frac{1}{\overline{\pi}^p_{q_*}}\sum_q \frac{\Pr\left(b_* > 0 \mid q\right)\Pr\left(q \mid q^*\right)}{\Pr\left(q \mid t\right)}\underbrace{\sum_p\sum_h \Pr\left(q \mid p,h,t\right)\mu_p(p)\mu_h(h \mid t)}_{=\Pr(q \mid t)} \\
&= \sum_{q_*} B(q_*,t_*)\frac{1}{\overline{\pi}^p_{q_*}}\sum_q \frac{\Pr\left(b_* > 0 \mid q\right)\Pr\left(q \mid q^*\right)}{\Pr\left(q \mid t\right)}\Pr\left(q \mid t\right) \\
&= \sum_{q_*} B(q_*,t_*)\frac{1}{\overline{\pi}^p_{q_*}}\underbrace{\sum_q \Pr\left(b_* > 0 \mid q\right)\Pr\left(q \mid q^*\right)}_{=\Pr(b_* > 0 \mid q_*)} \\
&= \sum_{q_*} B(q_*,t_*) \\
&= B(t_*)
\end{aligned}$$

D.2 Social Security system

D.2.1 Retirement benefits

First, consider the following stylized version of the actual retirement income formula used in the US social security system, where $\bar{e}$ is an (annualized) measure of historical average monthly earnings, $b_1^\$$ and $b_2^\$$ are bend points in USD for some reference year, and $e_{max}^\$$ is the contribution and benefit base (CBB), i.e., the maximum earnings subject to payroll taxes. Retirement income $\iota^\$$ measured in USD is approximately given by

$$\iota^\$(\bar{e}) = \begin{cases} \rho_1 \bar{e} & \text{if } \bar{e} \leq b_1^\$ \\ \rho_1 b_1^\$ + \rho_2 \left(\bar{e} - b_1^\$\right) & \text{if } b_1^\$ < \bar{e} \leq b_2^\$ \\ \rho_1 b_1^\$ + \rho_2 \left(b_2^\$ - b_1^\$\right) + \rho_3 \left(\min\left\{e_{\max}^\$, \bar{e}\right\} - b_2^\$\right) & \text{else} \end{cases} \tag{D.15}$$

where ρ_1, ρ_2 and ρ_3 are decreasing replacement rates applied to earnings ranges bracketed by the bend points $b_i^\$$.

In the model, we define retirement income ι as the product of the following components:

$$\iota_r(p) = w \cdot y_r(p) = w \cdot \overline{\omega}_r \cdot \mathcal{R}_{ss}(p) \tag{D.16}$$

where we construct the function $\mathcal{R}_{ss}(\bullet)$ to mimic (D.15) but define it in terms of persistent labor productivity p.

To this end, denote by $e_{med}^\$$ the median earnings in dollars in the reference year 2000. To express the bend points in terms of the persistent labor productivity, we implicitly define a model bend point p_i^* corresponding to $b_i^\$$ by the relationship

$$\frac{b_i^\$}{e_{med}^\$} = \frac{w\,\overline{\omega}_r\,p_i^*}{w\,y_{med}} \qquad i \in \{1, 2\}$$

That is, we normalize the dollar bend point $b_i^\$$ by dollar median earnings $e_{med}^\$$ and its model counterpart by the median earnings in the model. Solving for p_i^*, we obtain

$$p_i^* = \frac{\left(b_i^\$ / e_{med}^\$\right) y_{med}}{\overline{\omega}_r}.$$

Analogously, the CBB in terms of persistent labor productivity is

$$p_{max}^* = \frac{\left(e_{max}^\$ / e_{med}^\$\right) y_{med}}{\overline{\omega}_r}.$$

By factoring out the common term $w \cdot \overline{\omega}_r$ that is independent of a retired individual's idiosyncratic state vector, we can write the replacement formula purely in terms of the permanent labor state p as follows:

$$\mathcal{R}_{ss}(p) = \begin{cases} \rho_1 p & \text{if } p \leq p_1^* \\ \rho_1 p_1^* + \rho_2 \left(p - p_1^*\right) & \text{if } p_1^* < p \leq p_2^* \\ \rho_1 p_1^* + \rho_2 \left(p_2^* - p_1^*\right) + \rho_3 \Big(\min\Big\{ p_{max}^*, p \Big\} - p_2^* \Big) & \text{else} \end{cases}$$

Retirement income can then be computed as a function of p and the equilibrium wage w according to (D.16).

D.2.2 Social security budget balance

Government expenditures on retirement benefits in each period are given by

$$G_{ss} = \sum_{t=T_r}^{N_t} \sum_{p} \mu_t(t) \mu_p(p) \iota_r(p) = \Pi_r \cdot w \cdot \overline{\omega}_r \cdot \overline{p}_r \tag{D.17}$$

which is a weighted sum over the retirement incomes received by all retired cohorts, with weights μ_t and μ_p denoting the PMFs of the ergodic distribution over age and persistent labor productivity, respectively. We denote the mass of retired individuals by

$$\Pi_r = \sum_{t=T_r}^{N_t} \mu_t(t)$$

and the average permanent component of retirement income after applying the replacement rate function as

$$\overline{p}_r = \sum_{p} \mu_p(p) \mathcal{R}_{ss}(p).$$

Retirement income is fully financed by payroll taxes each period which are given by

$$\overline{T}_{ss} = \sum_{t=1}^{T_r - 1} \sum_{p} \sum_{h} \sum_{\epsilon} \mu_t(t) \mu_p(p) \mu_h(h,t) \mu_\epsilon(\epsilon) T_{ss}\left(y\right) w \tag{D.18}$$

where labor productivity of workers $y = y(p, h, \epsilon, t)$ is defined in (12), $\mu_h(h,t)$ is the age-dependent distribution over health states and $\mu_\epsilon(\epsilon)$ is the distribution over transitory labor shocks. The payroll tax function is defined as

$$T_{ss}\left(y\right) = \tau_{ss} \cdot \min\{y_{max}, y\}$$

where y_{max} are maximum taxable earnings, which we obtain from the dollar values using

$$y_{max} = \left(e^{\$}_{max} / e^{\$}_{med}\right) y_{med}.$$

To balance the social security system, we need to find τ_{ss} such that $G_{ss} = \overline{T}_{ss}$. Equating $G_{ss} = \overline{T}_{ss}$ implies that

$$\tau_{ss} = \frac{\Pi_r \, \overline{\omega}_r \, \overline{p}_r}{\sum_{t=1}^{T_r-1} \sum_p \sum_h \sum_\epsilon \mu_t(t) \mu_p(p) \mu_h(h,t) \mu_\epsilon(\epsilon) \min\{y_{max}, \omega_{ht} p \, \epsilon\}}. \tag{D.19}$$

D.3 Government budget

D.3.1 Expenditures

As described in the main text, the government makes means-tested lump-sum transfers to individuals hit by substantial medical expenditures shocks to ensure a minimum consumption level $\underline{c}$ as defined in (10). First, consider agents of age $t \geq T_r - 1$, i.e., retired individuals or those in the terminal period of their working life who will retire next period, conditional on survival. The aggregate transfers to this group are given by

$$\begin{aligned}\Xi_r = \sum_{t=T_r-1}^{N_t} \sum_p \sum_h \sum_\eta \sum_{\mathbb{1}_b \in \{0,1\}} & \mu_t(t) \mu_p(p) \mu_h(h,t) \mu_\eta(\eta) \mu_b(\mathbb{1}_b \mid t) \pi^s_{ht} \\ \times \int_a \mathbb{E}\Big[\max\Big\{0, \; & \underline{c} + m\left(h', \eta', \nu', t+1\right) - R \cdot \mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right) - \iota(w,p) \\ & - b_*(p', h', t+1)\Big\} \,\Big|\, h, \eta, \mathbb{1}_b, t\Big] d\Gamma(a | p, h, \eta, \mathbb{1}_b, t)\end{aligned} \tag{D.20}$$

where $\mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right)$ is the savings policy function and the after-tax retirement income ι is defined in (7). Note that these government transfers are only paid to surviving individuals with π^s_{ht} denoting the survival probability for a given age and health state.

Second, the aggregate transfers to workers of age $t < T_r - 1$ are given by

$$\begin{aligned}\Xi_w = \sum_{t=1}^{T_r-2} \sum_p \sum_h \sum_\eta \sum_{\mathbb{1}_b \in \{0,1\}} & \mu_t(t) \mu_p(p) \mu_h(h,t) \mu_\eta(\eta) \mu_b(\mathbb{1}_b \mid t) \pi^s_{ht} \\ \times \int_a \mathbb{E}\Big[\max\Big\{0, \; & \underline{c} + m\left(h', \eta', \nu', t+1\right) - R \cdot \mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right) \\ & - \iota(w, \omega_{th+1}, p', \epsilon') - b_*(p', h', t+1)\Big\} \,\Big|\, p, h, \eta, \mathbb{1}_b, t\Big] d\Gamma(a | p, h, \eta, \mathbb{1}_b, t)\end{aligned} \tag{D.21}$$

where ι is the after-tax income of workers defined in (13). Compared to the retired

individuals, we additionally need to take expectations over the next-period realizations of persistent and transitory labor income shocks.

D.3.2 Estate tax revenues

The government taxes estates using the tax schedule $T_b(\bullet)$, defined as

$$T_b(x) = \begin{cases} 0 & \text{if } x \leq \chi_b \\ \tau_b(x - \chi_b) & \text{else} \end{cases} \tag{D.22}$$

where χ_b is the maximum inheritance that is exempt from estate taxes and τ_b is the marginal tax are on taxable estates. As described in the main text, individuals who die at the beginning of the next period leave their savings Rk. Moreover, with a continuum of individuals, there is positive mass of agents who die in the same period as their parents. In this case we adopt the convention that older cohorts "die first" so that their estates are passed on to their children and become part of the child's estate. Finally, any out-of-pocket medical expenditures incurred in their terminal period of life are deducted (but children are not liable to pay any medical bills that exceed the inherited amount), so that estates before taxes are given by

$$\max\Big\{0,\ Rk + b_*(p', h', t+1) - m\left(\eta', \nu', t+1\right)\Big\}.$$

Estate tax revenues are thus determined as

$$\begin{aligned} T_{estate} = \sum_{t=1}^{N_t} \sum_{p} \sum_{h} \sum_{\eta} \sum_{\mathbb{1}_b \in \{0,1\}} & \mu_t(t)\mu_p(p)\mu_h(h,t)\mu_\eta(\eta)\mu_b(\mathbb{1}_b \mid t)\left(1 - \pi^s_{ht}\right) \\ & \times \int_a \mathbb{E}\Big[\max\Big\{0,\ R \cdot \mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right) + b_*(p', h', t+1) \\ & \qquad - m\left(\eta', \nu', t+1\right)\Big\} \ \Big|\ p, \eta, \mathbb{1}_b, t\Big] d\Gamma(a|p, h, \eta, \mathbb{1}_b, t) \end{aligned} \tag{D.23}$$

D.3.3 Income tax revenues

In this section, we derive an expression for the total amount of income taxes raised by the government. Before proceeding, we state the following useful definitions: We denote by $\overline{p}$ and $\overline{\epsilon}$ the average persistent and transitory labor components in the economy,

$$\overline{p} = \sum_{p} \mu_p(p)p \qquad\qquad \overline{\epsilon} = \sum_{\epsilon} \mu_\epsilon(\epsilon)\epsilon \tag{D.24}$$

where $\mu_p(p)$ and $\mu_\epsilon(\epsilon)$ are their respective PMFs,[1] and by Π_w the size of the labor force,

$$\Pi_w = \sum_{t=1}^{T_r-1} \mu_t(t) = 1 - \Pi_r$$

Additionally, average labor productivity can be defined as

$$\begin{aligned}\overline{y} &= \Pi_w^{-1}\left[\sum_{t}^{T_r-1}\sum_{p}\sum_{h}\sum_{\epsilon}\mu_t(t)\mu_p(p)\mu_h(h,t)\mu_\epsilon(\epsilon)\omega_{ht}p\,\epsilon\right] \\ &= \Pi_w^{-1}\left[\overline{p}\,\overline{\epsilon}\sum_{t}^{T_r-1}\sum_{h}\mu_t(t)\mu_h(h,t)\omega_{ht}\right].\end{aligned}$$

Now, consider the aggregate tax revenues raised from working individuals, which are given by

$$\begin{aligned}T_w &= \sum_{t=1}^{T_r-1}\sum_{p}\sum_{h}\sum_{\epsilon}\mu_t(t)\mu_p(p)\mu_h(h,t)\mu_\epsilon(\epsilon)\Big\{\big(\omega_{ht}p\,\epsilon - T_{ss}\left(\omega_{ht}p\,\epsilon\right)\big)w \\ &\qquad - \lambda\Big[\big(\omega_{ht}p\,\epsilon - T_{ss}\left(\omega_{ht}p\,\epsilon\right)\big)w\Big]^{1-\tau}\Big\} \\ &= \Big[w\Pi_w\overline{y} - \lambda w^{1-\tau}\overline{y}_{-ss,\tau}\Big] - \overline{T}_{ss}\end{aligned}$$

where we define

$$\overline{y}_{-ss,\tau} = \sum_{t=1}^{T_r-1}\sum_{p}\sum_{h}\sum_{\epsilon}\mu_t(t)\mu_p(p)\mu_h(h,t)\mu_\epsilon(\epsilon)\Big(\omega_{ht}p\,\epsilon - T_{ss}\left(\omega_{ht}p\,\epsilon\right)\Big)^{1-\tau}$$

to simplify the notation. Income taxes raised from retired individuals amount to

$$\begin{aligned}T_r &= \sum_{t=T_r}^{N_t}\sum_{p}\mu_t(t)\mu_p(p)\left[w\,y_r(p) - \lambda\Big(w\,y_r(p)\Big)^{1-\tau}\right] \\ &= \sum_{t=T_r}^{N_t}\sum_{p}\mu_t(t)\mu_p(p)\left[w\,\overline{\omega}_r\mathcal{R}_{ss}(p) - \lambda\Big(w\,\overline{\omega}_r\mathcal{R}_{ss}(p)\Big)^{1-\tau}\right] \\ &= \Pi_r\left[w\overline{\omega}_r\overline{p}_r - \lambda w^{1-\tau}\overline{\omega}_r^{1-\tau}\overline{p}_{r,\tau}\right] \\ &= \overline{T}_{ss} - \lambda\Pi_r w^{1-\tau}\overline{\omega}_r^{1-\tau}\overline{p}_{r,\tau}\end{aligned}$$

[1] In the numerical implementation these shocks are rescaled so that $\overline{p} = \overline{\epsilon} = 1$.

with

$$\overline{p}_{r,\tau} = \sum_p \mu_p(p)\mathcal{R}_{ss}(p)^{1-\tau}$$

Thus, the total revenue from income taxes is

$$T_{inc} = T_w + T_r = w\Pi_w\overline{y} - \lambda w^{1-\tau}\left[\overline{y}_{-ss,\tau} + \Pi_r\,\overline{\omega}_r^{1-\tau}\,\overline{p}_{r,\tau}\right] \tag{D.25}$$

D.3.4 Government budget balance

Using the expenditure and tax components derived in the previous sections, government budget balance requires that

$$G + \Xi = T_{estate} + T_{inc}(\lambda) \tag{D.26}$$

where non-discretionary government spending is

$$G = gY = g\frac{w\,L}{1-\alpha_k},$$

which follows from the constant returns to scale production function and perfectly competitive markets. The aggregate lump-sum transfers to households $\Xi = \Xi_r + \Xi_w$ are defined in (D.20) and (D.21), T_{estate} is defined in (D.23) and T_{inc} is given in (D.25). This equation implicitly defines the equilibrium λ which has to satisfy

$$\lambda = \frac{w\Pi_w\overline{y} - G + T_{estate} - \Xi}{w^{1-\tau}\left[\overline{y}_{-ss,\tau} + \Pi_r\,\overline{\omega}_r^{1-\tau}\,\overline{p}_{r,\tau}\right]} = \frac{wL\left(1 - \frac{g}{1-\alpha_k}\right) + T_{estate} - \Xi}{w^{1-\tau}\left[\overline{y}_{-ss,\tau} + \Pi_r\,\overline{\omega}_r^{1-\tau}\,\overline{p}_{r,\tau}\right]}$$

The second equality uses the fact that aggregate labor supply is $L = \Pi_w\overline{y}$, i.e., the size of the labor force times average productivity.

D.4 Equilibrium definition

A recursive competitive equilibrium is given by a set of prices $\{R, w\}$, tax rates $\{\tau_{ss}, \lambda\}$, decision rules $\mathcal{C}\left(a, p, h, \eta, \mathbb{1}_b, t\right)$ for consumption and $\mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right)$ for savings, the aggregate bequest function $B(q_*, t_*)$, and a stationary distribution Γ such that:

1. The decision rules solve the agents' problem for all $(a, p, h, \eta, \mathbb{1}_b, t)$.

2. Factor prices are given by:

$$r = F_K(A, K, L) - \delta_k \qquad \text{and} \qquad w = F_L(A, K, L)$$

and gross returns on savings are $R = 1 + r$.

3. Capital and labor markets clear:

$$K = \int \mathcal{K}\left(a, p, h, \eta, \mathbb{1}_b, t\right) \mathrm{d}\Gamma \tag{D.27}$$
$$L = \sum_{t=1}^{T_r-1} \sum_{p} \sum_{h} \sum_{\epsilon} \mu_t(t)\mu_p(p)\mu_h(h,t)\mu_\epsilon(\epsilon)\omega_{ht} p\, \epsilon$$

where $\mu_t(t)$, $\mu_h(h,t)$, $\mu_p(p)$ and $\mu_\epsilon(\epsilon)$ are the ergodic distributions over age, health, the persistent and the transitory labor shocks, respectively.

4. Payroll taxes τ_{ss} are set according to (D.19) so that the Social Security budget is balanced, $G_{ss} = \overline{T}_{ss}$.

5. The equilibrium value of λ is implicitly defined by (D.25) so that the government budget is balanced, $gY + \Xi = T_{estate} + T_{inc}(\lambda)$.

6. The distribution Γ over $(a, p, h, \eta, \mathbb{1}_b, t)$ is stationary.

7. The function $B(q_*, t_*)$ is consistent with the after-tax estates left by agents aged t_* in line with (D.6), (D.7) (D.9) and (D.11).

We discuss the numerical implementation of how to find this equilibrium in section H.

E Calibration details

E.1 Health process for ages 20–50

We estimate the health transitions for nonblack men aged 20–50 using PSID data from 1984 (when the question about self-reported health was asked for the first time) up until 2019. During this period, the PSID was annual up until 1997, and thereafter biennial. As explained in section 2.3, our estimator can handle arbitrary and varying transition lengths and can thus estimate annual transition probabilities from such data. We assume that survival is certain below the age of 50, thus we do not have to worry about the PSID not recording deaths with the same precision as the HRS.

The resulting annual health transition probabilities for the ages 50–70, in other words, those ages that overlap between HRS and PSID, are shown in Figure E.1. The results indicate that the health transition probabilities are very similar across data sets.

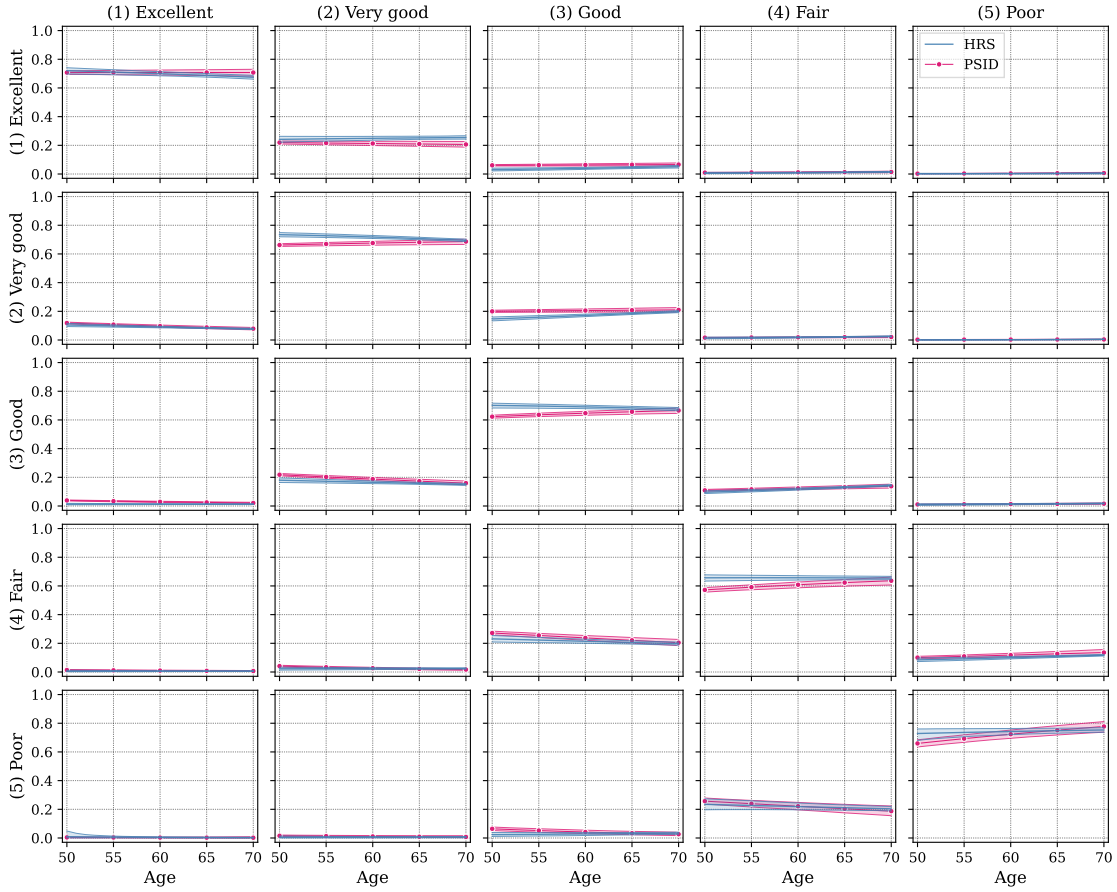


Figure E.1: One-year health transition probabilities for nonblack men. Shaded areas indicate bootstrapped 95% confidence intervals.

E.2 Ergodic distribution over age and health

Figure E.2 shows the relative cohort sizes and distribution of health states in the model, where we rescale the cohort size of newborns to unity for the purpose of illustration.

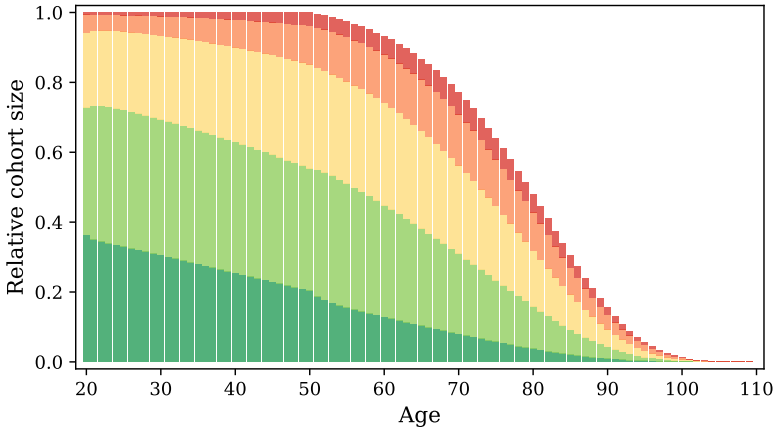


Figure E.2: Relative cohort size and health state distribution. Colors indicate the health state: dark green is excellent while red is poor health.

E.3 Life cycle profiles of labor earnings by health

To estimate labor income life cycle profiles by health we use PSID data, and restrict the sample to nonblack men. We use survey years from 1992 (to be consistent with our estimates from the HRS) up until 2019 (to increase the number of observations).

We use a total measure of labor earnings, including several components: wages and salaries, separate reports of bonuses, overtime compensation, tips, commissions, income from professional practice or trade, market gardening, additional job income, and miscellaneous labor income.

Figure E.3 shows the resulting labor earnings profiles. Note that the profiles are *not* conditional on having positive labor income. Thus, the lower profile for individuals in poor health is partly due to lower income conditional on working, but also due to lower participation. Among the individuals in poor health, 61 percent report zero labor income. Among individuals in excellent health the corresponding figure is only 8 percent.

Since the quantitative model does not have an endogenous labor choice, this measure of labor income captures the health effects on labor income on both the extensive and the intensive margin. In the model we use a third-order polynomial approximation, as shown in Figure E.3.

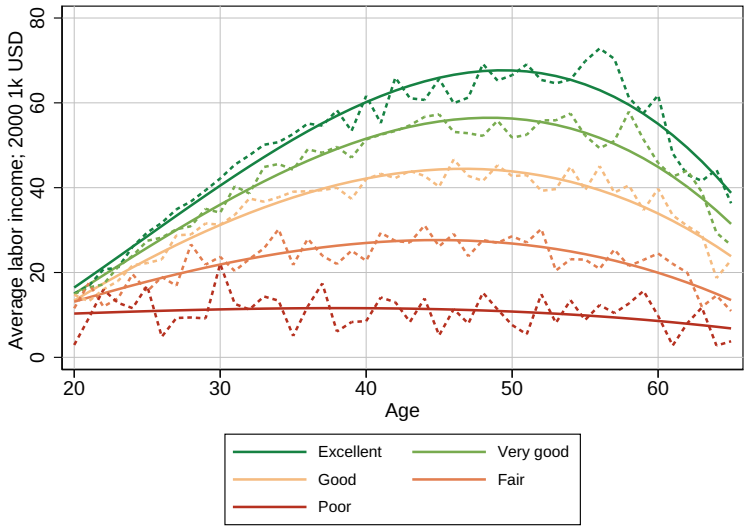


Figure E.3: Labor earnings by health state, nonblack males. Dashed lines indicate raw (weighted) averages by health and age. Solid lines indicate a fitted third-order polynomial, which is what is used in the quantitative model. Colors indicate the health state: dark green is excellent while red is poor health.

E.4 Medical expenditure process

We follow French and Jones (2004) and De Nardi, French, and Jones (2010) and estimate medical expenditures from the HRS which surveys all out-of-pocket expenses that accrued since the previous wave (or over the two years preceding an interview if a respondent was not surveyed in the previous wave). For deceased participants, medical expenditures are collected as part of the exit interview.

Medical expenditure sub-components include out-of-pocket costs for hospital stays, outpatient surgery, doctor and dentist visits, prescription drugs, home health services, as well as special facilities and services. We exclude expenses for nursing homes since we are interested in modeling medical expenditure risk, whereas nursing home expenses are at least in part voluntary consumption choices. We convert all values to 2000 USD (using U.S. Bureau of Labor Statistics (2024b)).

The RAND version of the HRS contains imputations for missing medical expenditure categories. However, these imputed values can be unreasonably large, so we drop any imputed values from the analysis. Additionally, we winsorize each sub-component of medical expenditures at the 99.5% level within five-year age bins to eliminate outliers.

We include individuals aged 50 and above for whom we have at least 5 observations on medical expenditures in our estimation sample. No consistently measured data is available prior to wave 3, so the initial two waves are excluded from the analysis. Moreover, since we calibrate the economic model to nonblack men, we restrict the sample to this group for a total of 5,754 individuals with 41,376 observations.

We impose that log medical expenditures follow a heteroskedastic process given by

$$\log m_{it} = \alpha_i + \boldsymbol{x}_{it}'\boldsymbol{\beta} + \mathbf{z}_{it}'\boldsymbol{\gamma} + \sigma(\boldsymbol{x}_{it})\left(\eta_{it} + \nu_{it}\right) \tag{E.1}$$

$$\eta_{it} = \rho_m \eta_{it-1} + \zeta_{it} \tag{E.2}$$

$$\zeta_{it} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\zeta^2) \tag{E.3}$$

$$\nu_{it} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma_\nu^2) \tag{E.4}$$

where $\boldsymbol{x}_{it}$ contains a third-order polynomial in age, health, as well as health interacted with age. Additionally, $\mathbf{z}_{it}$ includes controls not present in the OLG model such as marital status, education level, 5-year cohort dummies, and time fixed effects, as well as interactions of these terms. We moreover control for the length of the reference period since the distance between two consecutive interviews need not be two years (and can be substantially shorter if a respondent dies shortly after an interview).

The variance of the error term $u_{it} \equiv \sigma(\boldsymbol{x}_{it})\left(\eta_{it} + \nu_{it}\right)$ is modeled via the skedastic

function $\sigma(\bullet)$. For identification, we need to impose the normalization

$$\text{Var}\left(\eta_{it} + \nu_{it}\right) = 1\,,$$

and we define

$$\text{Var}\left(\eta_{it}\right) = \phi \tag{E.5}$$
$$\text{Var}\left(\nu_{it}\right) = 1 - \phi \tag{E.6}$$
$$\text{Var}\left(\zeta_{it}\right) = \sigma_\zeta^2 = (1-\rho^2)\sigma_\eta^2 = (1-\rho^2)\phi$$

where ϕ is a parameter governing the share of variance attributed to the persistent component η. The term $\eta_{it} + \nu_{it}$ is scaled by $\sigma(\boldsymbol{x}_{it};\boldsymbol{\delta})$ which is a function of covariates $\boldsymbol{x}_{it}$, parametrized by the vector $\boldsymbol{\delta}$, so that the variance and the lag-k auto-covariance are given by

$$\text{Var}\left(u_{it} \,|\, \boldsymbol{x}_{it};\boldsymbol{\theta}\right) = \sigma^2(\boldsymbol{x}_{it}) = e^{\boldsymbol{x}_{it}'\boldsymbol{\delta}} \tag{E.7}$$
$$\text{Cov}\left(u_{it}, u_{it-k} \,|\, \boldsymbol{x}_{it}, \boldsymbol{x}_{it-k};\boldsymbol{\theta}\right) = \sqrt{\sigma^2(\boldsymbol{x}_{it})}\sqrt{\sigma^2(\boldsymbol{x}_{it-k})}\rho_m^k\phi \tag{E.8}$$

where $\boldsymbol{\theta} = \left(\rho_m, \phi, \boldsymbol{\delta}\right) \in \mathbb{R}^{K_\theta}$ is the parameter vector to be estimated.

We estimate the coefficient vector $\boldsymbol{\beta}$ of the mean equation with a standard fixed-effects estimator and use the residuals to estimate the parameters $\boldsymbol{\theta}$ via GMM where we include lags of up to 10 years in the covariance moment conditions. We find that medical expenditures shocks are quite persistent with $\rho_m = 0.836$, while the share of variance attributed to the persistent component is $\phi = 0.609$ at a two-year frequency. The details are shown in Table E.1.

To get some intuition for the moments matched by the GMM estimator, Figure E.4 contrasts the sample means of log medical expenditures by age with the predicted values for nonblack men. The gray bubbles represent the (weighted) raw data means whereas the pink bubbles correct for covariates $\mathbf{z}$ not present in the OLG model (we assume that

Parameter	Estimate	95% Confidence interval
ρ_m	0.835	[0.820, 0.850]
ϕ	0.609	[0.583, 0.632]
σ_ζ^2	0.184	[0.164, 0.205]
σ_ν^2	0.391	[0.368, 0.417]

Table E.1: Parameter estimates for the medical expenditure shock process at two-year frequency.

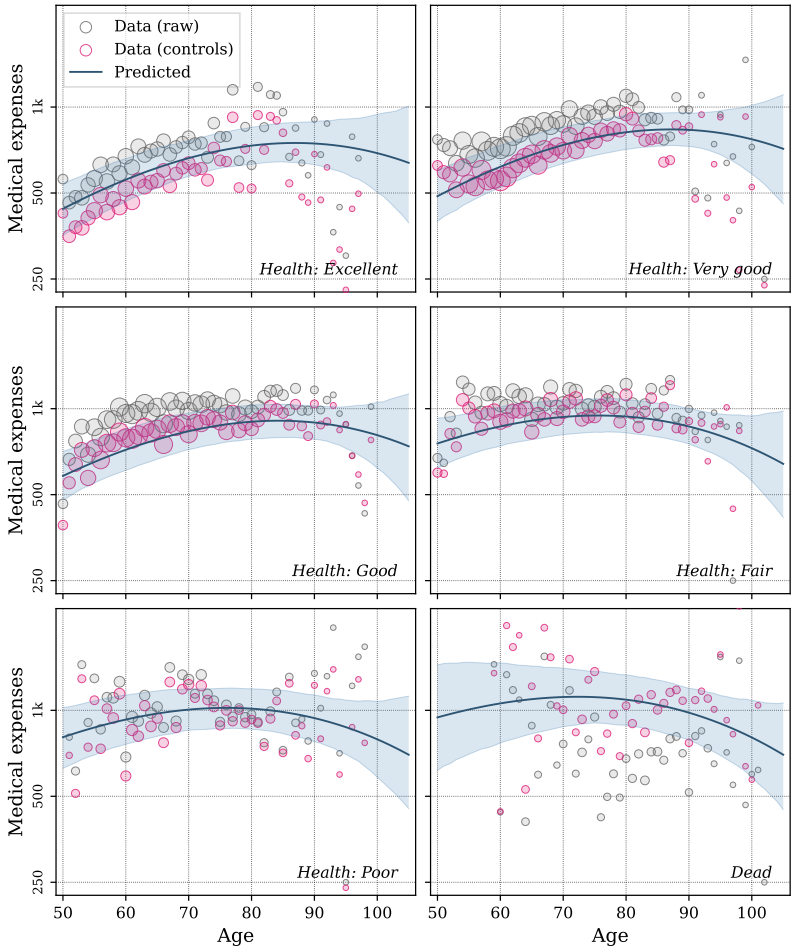


Figure E.4: Log medical expenditures vs. predicted values for nonblack men. Grey data points show raw average log medical expenditures by age, while the pink data points control for marital status, cohort, education, and year. Shaded areas indicate 95% confidence intervals.

the reference category is an unmarried man with high school education who was born in 1920–1924). Figure E.5 plots the analogous graphs for the variance of log medical expenditures, where the data moments are given by the average squared residuals by age. Finally, Figure E.6 shows the covariances for lags of up to 20 years observed in the data and their model counterparts.[2] In general, the model predicts that both the mean and variance of log medical expenditures are increasing in age, with the exception of very high age where the log mean is decreasing.

[2]To reduce clutter, the graph shows the covariance at the age of 70. As can be seen from (E.8), the lag-k covariance is a function of both $\boldsymbol{x}_{it}$ and $\boldsymbol{x}_{it-k}$, so the model predictions in the plot need to be weighted by the sample distribution of $\boldsymbol{x}_{it-k}$ for any fixed $\boldsymbol{x}_{it}$.

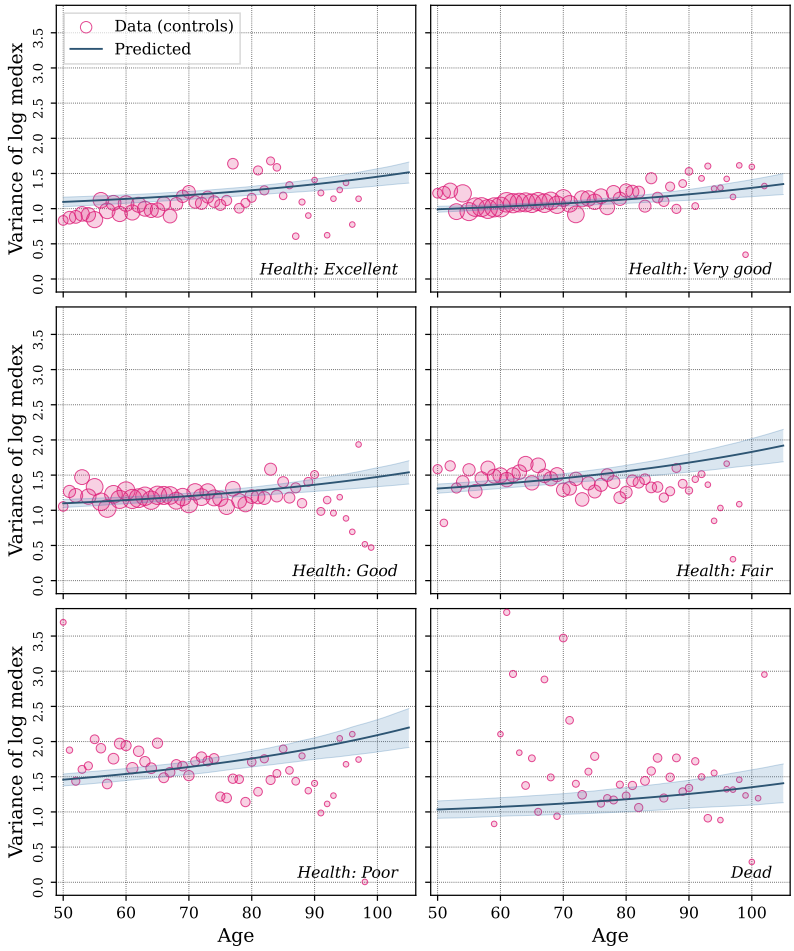


Figure E.5: Squared residuals of log medical expenditures vs. predicted variance for nonblack men. Data points show average squared residuals by age. Shaded areas indicate 95% confidence intervals.

Since the HRS data is mostly biennial but the OLG model is calibrated to annual frequency, we use a simulated method of moments (SMM) algorithm to convert the biennial parameter estimates to their annual counterparts. To do this, we assume that annual log medical expenditures follow exactly the same process as laid out in (E.1) to (E.8), but we allow the parameters to differ. For a given guess of the annual parameter vector, we simulate a cross-section of 100,000 individuals for up to 20 years, using health transitions and survival probabilities from the objective health/survival model discussed in the main text. We then aggregate annual medical expenditures to two-year observations, compute the implied two-year moments and compare these to our biennial estimates. Our SMM algorithm minimizes the distance between these two sets of moments. We obtain $\rho_m = 0.920$ and $\phi = 0.543$ as our annual estimates. The shock

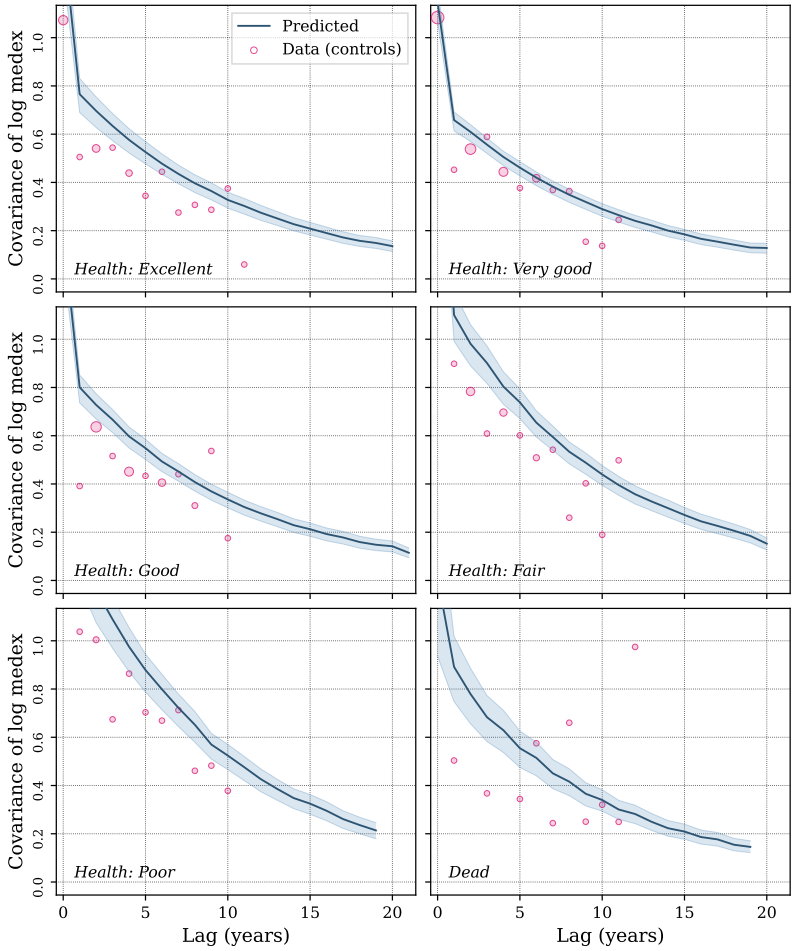


Figure E.6: Covariances of log medical expenditures vs. predicted covariance at age 70. Model predictions are weighted by the empirical distribution of covariates at each lag. Shaded areas indicate 95% confidence intervals.

variances at annual frequency are therefore given by $\sigma_\zeta^2 = 0.083$ and $\sigma_v^2 = 0.457$ which we use to discretize the persistent and transitory medical expenditure shocks in the OLG model.

F Additional model results

F.1 Discount rates

De Nardi, Pashchenko, and Porapakkarm (2017) (henceforth DPP) estimate that individuals in bad health on average have a higher discount rate (i.e., are more impatient) than individuals in good health. The subjective beliefs about life expectancy we document in this paper provides one plausible explanation behind this pattern. In this section we, compare the magnitudes of our respective estimates.

In DPP, individuals differ ex-ante in their patience and health type. The discount factor can take two values, $\beta_{low} = 0.877$ or $\beta_{high} = 0.992$. The health type can take three values, $\eta_i \in \{\eta_1, \eta_2, \eta_3\}$. Individuals with "lower" health type have a higher probability of being in bad health (due to both higher probability of transitioning to bad health, and higher probability of staying in bad health once there). DPP estimate a strong correlation between discount factor and health type: the fraction of impatient people (with discount factor β_{low}) among those with health type η_1 or η_2 is about 80%, while the fraction among those with the best health type (η_3) is slightly less than 40%. The average discount factors among the η_1- and η_2-types is 0.90 and the average among the η_3-type is 0.95.

It is difficult to make a fair comparison between DPP's estimates and our effective discount factor heterogeneity due to subjective survival beliefs. In DPP, the discount factor is an ex-ante fixed characteristic, while in our model, the effective discount rate of an individual is changing over time as the health state evolves. Moreover, the health types in DPP do not match one-to-one with the observed health state. Nevertheless, to give an idea, we do the following: we define the *"subjective discount rate addition"* as the difference between the subjective discount rate and the objective discount rate. This is a measure of the discount factor heterogeneity due to subjective survival beliefs.

In Figure F.1 we plot this *"subjective discount rate addition"* for ages 50 and 70 for different time horizons. As can be seen, at the age of 50 the difference between individuals in excellent and poor health is substantial, more than 10 percent on a one or two year horizon. At a 10-year horizon, the difference is 5 percent, and at all horizons, individuals in bad health have a higher difference between their subjective and objective discount rate. Thus, our results are well in line with the magnitudes estimated by DPP.

F.2 Policy functions

Figure F.2 shows the average savings policy functions for selected ages for the model with objective survival heterogeneity (OSH). Since we cannot plot the full six dimensions

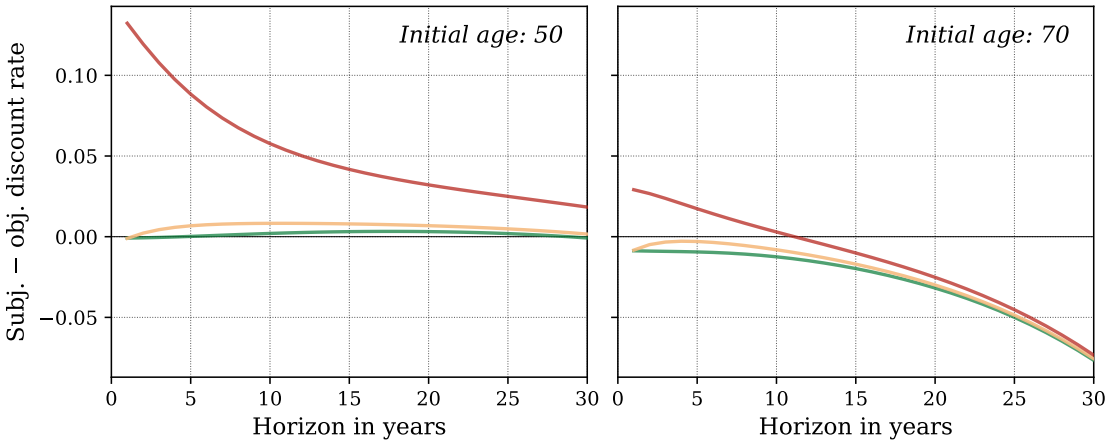
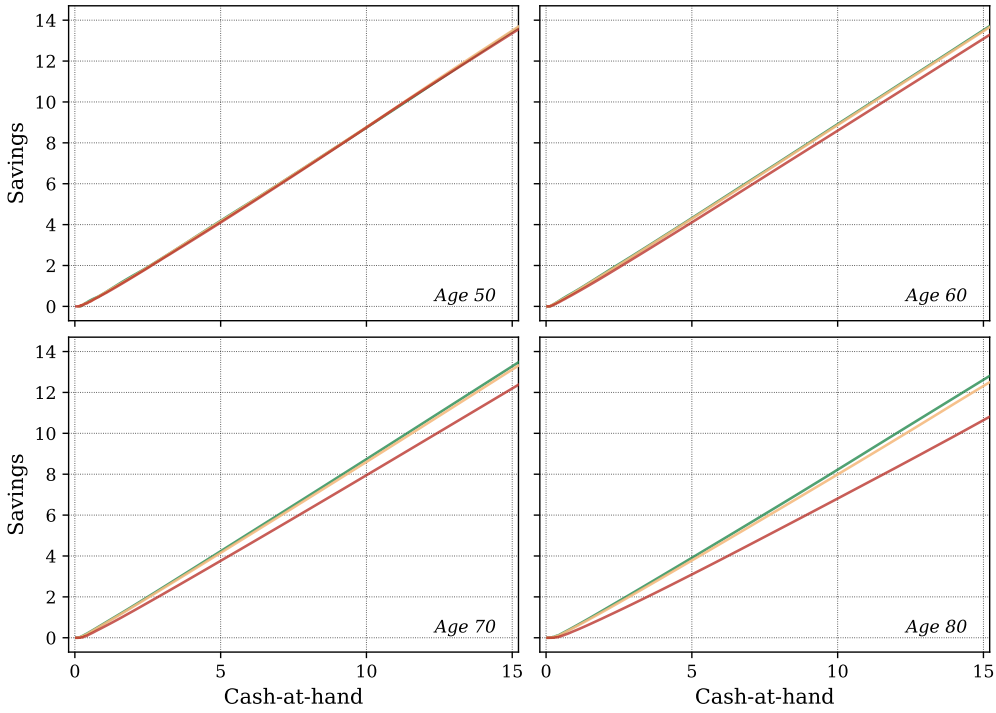


Figure F.1: Subjective discount rate addition. Difference between subjective and objective discount rate as a function of current age (50 or 70) and forecast horizon (on the x-axis). Colors indicate the health state: dark green is excellent while red is poor health.

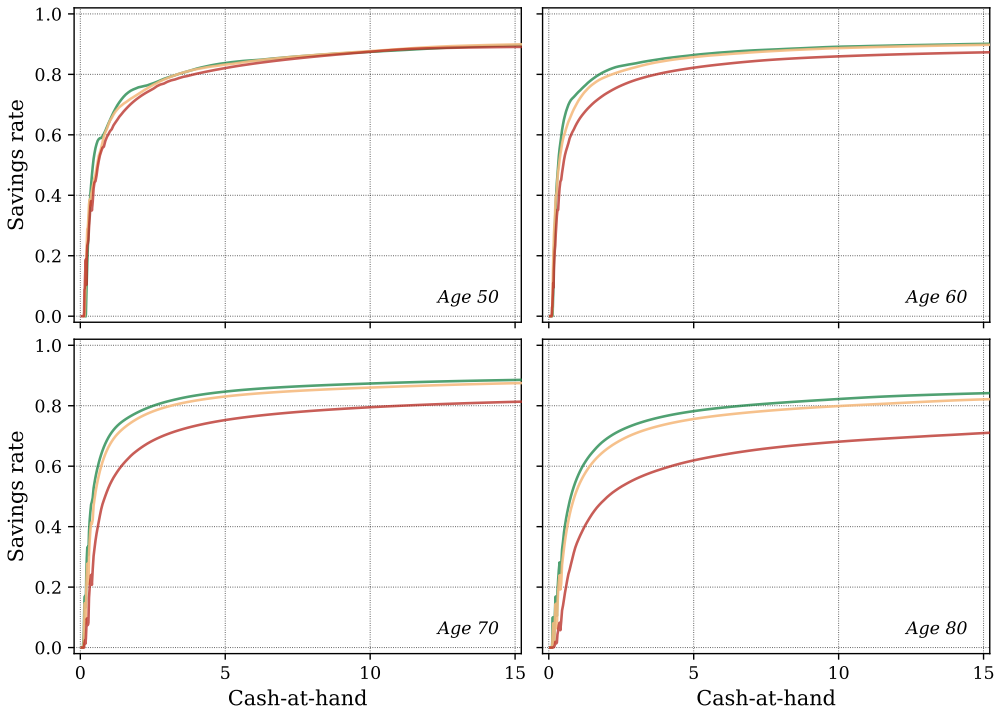
of the idiosyncratic state space, we average out the states of lesser interest, namely labor productivity, persistent medical expenditures, and the bequest indicator, using the corresponding stationary distribution in equilibrium. Some of the non-smoothness in these policy functions is due to this averaging, while the remainder arises because even the disaggregated policy functions are discountinuous, as discussed in section H.

Panel (a) depicts how much an agent saves as a function of current cash-at-hand, while panel (b) shows how much the agent saves as a fraction of current cash-at-hand.

Figure F.3 shows the same information from the model with subjective survival heterogeneity (SSH), where we average out labor and medical expenditure heterogeneity using the stationary distribution of the SSH economy.

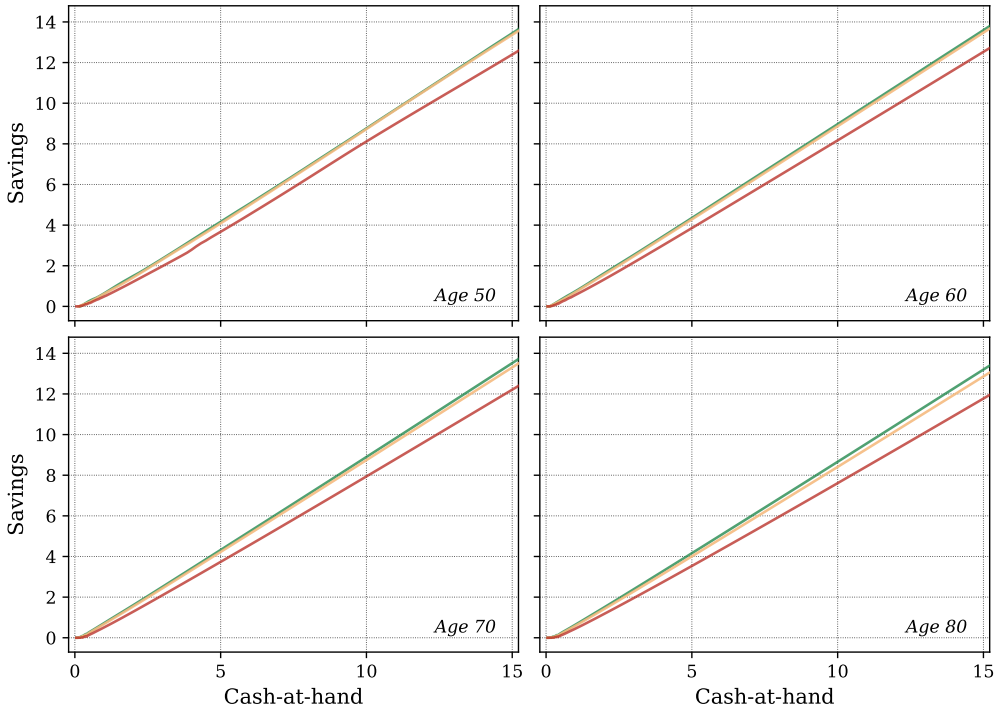


(a) Savings policy functions

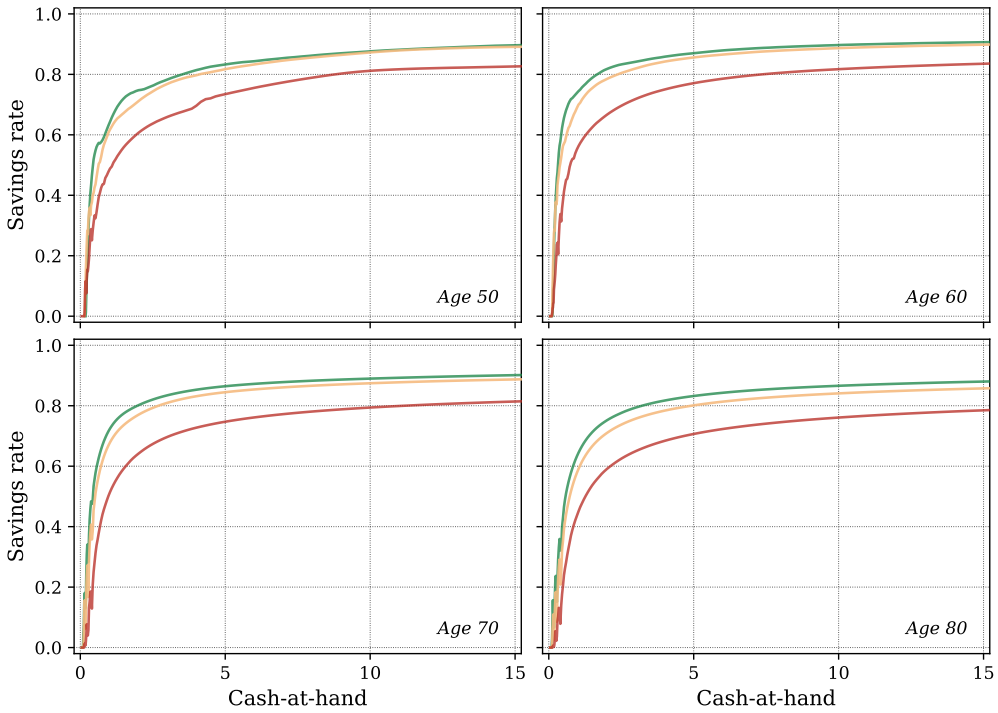


(b) Total savings rate (defined as fraction of total cash-at-hand saved to next period).

Figure F.2: Policy functions for the model with *objective* survival heterogeneity (OSH) by age, health and cash-at-hand, with remaining state variables averaged out. Colors indicate the health state: dark green is excellent while red is poor health.



(a) Savings policy functions



(b) Total savings rate (defined as fraction of total cash-at-hand saved to next period)

Figure F.3: Policy functions for the model with *subjective* survival heterogeneity (SSH) by age, health and cash-at-hand, with remaining state variables averaged out. Colors indicate the health state: dark green is excellent while red is poor health.

G Model with an active bequest motive

G.1 A two-period model

In this section, we illustrate how the introduction of a warm-glow bequest motive can have ambiguous affects on savings. To build intuition, we consider a simple two-period model with a bequest motive where the only uncertainty is survival risk. The survival probability between the first and second period is denoted by π. After the second period, the agent dies with certainty. We assume that the agent has some initial assets a_1, the discount factor is $\beta < 1$, and the gross interest rate is R. There is no income in the second period, thus it is a simple cake-eating problem. The agent solves

$$V(a_1) = \max_{c_1, c_2, a_2, a_3} \left\{ u(c_1) + \beta \left(\pi \Big[u(c_2) + \beta V_b(a_3) \Big] + (1 - \pi) V_b(a_2) \right) \right\} \tag{G.1}$$

subject to the usual non-negativity restrictions and the budget constraints

$$a_2 = a_1 - c_1, \qquad a_3 = R a_2 - c_2.$$

The first-order conditions are given by

$$\begin{aligned} u'(c_1) &= \beta \pi R u'(c_2) + \beta (1 - \pi) V_b'(a_2) \\ u'(c_2) &= \beta V_b'(a_3). \end{aligned}$$

Let the utility function $u(\bullet)$ take the usual CRRA functional form,

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma}$$

and let $V_b(\bullet)$ take the standard warm-glow bequest motive form as in De Nardi (2004),

$$V_b(a) = \phi_1 \frac{(a + \phi_2)^{1-\sigma} - 1}{1 - \sigma} \tag{G.2}$$

where ϕ_1 determines the strength of the bequest motive, and ϕ_2 determines to what extent bequests are a luxury good. For simplicity, we assume $\phi_2 = 0$ and solve for the agent's problem as given in (G.1). Some algebra shows that whether an increase in the survival probability makes the agent consume more or less in the first period (the sign

of $\frac{\partial c_0^*}{\partial \pi}$) depends on the strength of the bequest motive, more precisely by the relationship

$$\phi_1 \lesseqgtr \left(\frac{R^{\frac{1-\sigma}{\sigma}}}{1 - R^{\frac{1-\sigma}{\sigma}} \beta^{1/\sigma}} \right)^{\sigma} \equiv \widehat{\phi}_1. \tag{G.3}$$

If $\phi_1 < \widehat{\phi}_1$, an increase in the survival probability makes the agent consume less in the current period and save more given that it is optimal to transfer more resources to the next period since the probability to still be alive increases.

On the other hand, if the weight on the bequest motive is large, $\phi_1 > \widehat{\phi}_1$, an increase in the survival probability leads to *decreased* savings and more consumption in the first period. There are two mechanisms behind this. First, when the probability of surviving increases, the effective discounting of the next-period bequest utility increases, and hence the incentive to save decreases. We call this the expected-date-of-handover channel. Second, an increase in the survival probability leads to a higher expected interest rate income over the remaining life, and therefore the agent can afford more consumption also in the first period. We call this channel the income effect. If the weight on the bequest motive is high, these two effects dominate the effect of wanting to save for a longer expected life.

In the example above, we assumed $\phi_2 = 0$. It can be shown that if $R = 1$, the sign of $\frac{\partial c_0^*}{\partial \pi}$ is independent of ϕ_2 (although it still affects the level of savings). However, if we allow for a positive interest rate, the extent to which the bequest is a luxury good in combination with the level of initial assets a_0 matters. Moreover, we assumed no second-period income in our example above. If we allow for income in the second period, the level of income compared to the initial assets affects the incentives. For a given a_0, the higher the second-period income, the lower is the bequest weight needed to get $\frac{\partial c_0^*}{\partial \pi} > 0$ since then the bequest becomes relatively more important as a reason to save.

G.2 Calibration of the bequest motive

As discussed in the main text, for the model with a warm-glow bequest motive we target a capital-output ratio of 3 and a fraction of taxable estates of 2% as in the benchmark model. In addition, we attempt to match the relative decumulation of assets at higher ages relative to median asset holdings at ages 55. The latter moments and their data counterparts are shown in Table G.1. Some aspects of our model are too simplistic to match the data moments exactly. For example, because we impose an exogenous retirement age of 65, the life cycle profile of assets peaks exactly at this age, whereas this

	Median asset holdings					
Age	**60**	**65**	**70**	**75**	**80**	**85**
Data	1.05	1.09	1.13	1.16	1.18	1.17
Model	1.44	1.77	1.52	1.32	1.17	1.07

Table G.1: Life cycle asset profile in the model with a bequest motive, relative to the median asset holdings at age 55.

is not the case in the data.

To assess the magnitude of bequests in the model, it is informative to look at non-targeted moments. The bequest-to-wealth ratio in the model is 2.3%. It is difficult to precisely measure this figure in the data, but according to Gale and Scholz (1994), using SCF data, it should be closer to 0.9%. However, according to others, 2% is "a conservative estimate".[3] Another measure is estate tax revenue as a fraction of GDP. In the model, this figure is 0.02%. In the US, this figure varied during the last few years, from a high of 0.17% in 2007 to a low of 0.07% in 2011.[4] The model cannot replicate these magnitudes since it is unable to generate the right tail of the wealth distribution and hence also lacks the estates left behind by the richest. In sum, the magnitude of bequests left in the model seems to be roughly in line with the data.

[3]See Larry Summer's opinion at https://www.reuters.com/article/column-summers/column-how-to-target-untaxed-wealth-lawrence-summers-idUSL1E8NG2MC20121216?irpc=932.

[4]Calculated as the ratio of net estate taxes paid, as reported by the IRS, and official GDP figures.

H Computational appendix

H.1 The agent's problem

Solving the agent's problem is complicated by two factors:

1. The number of dimensions of the idiosyncratic state space requires us to solve the agent's problem on about 3.5 million grid points (see Table H.1 for the individual grid sizes).

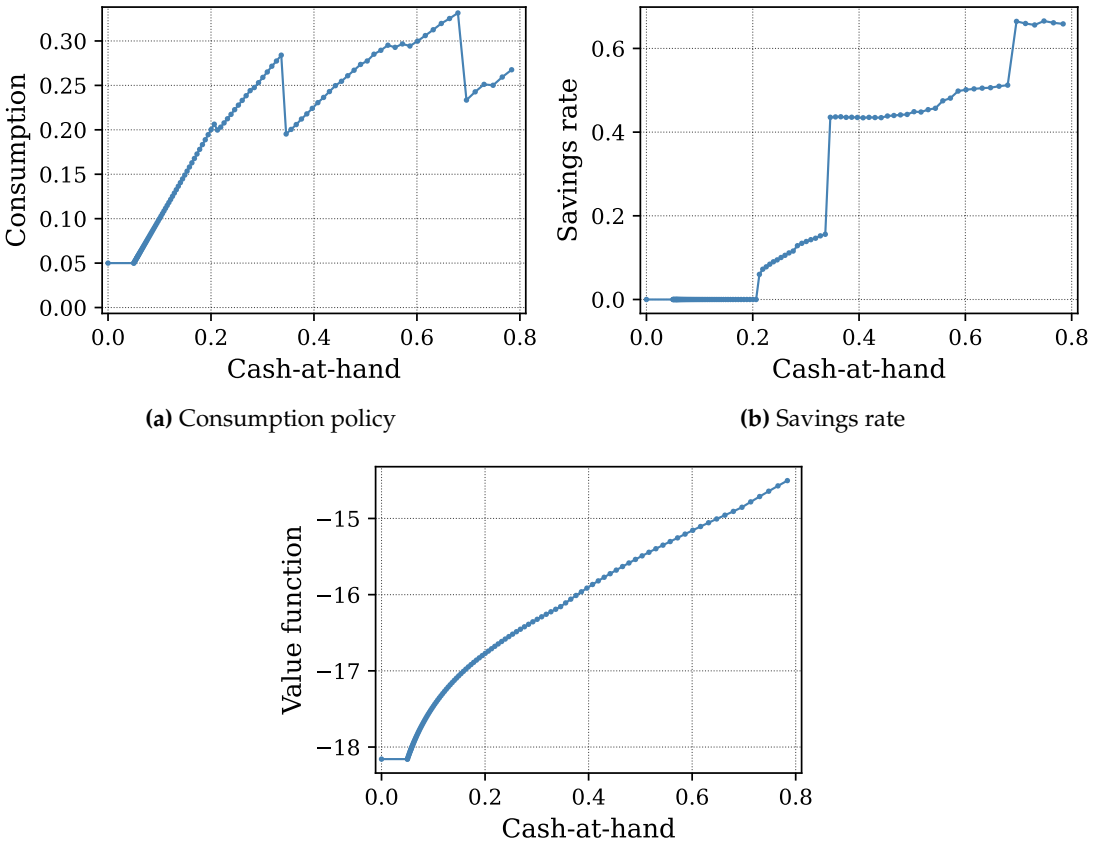
Dimension	Size
Individual state space	
Cash-at-hand	200
Age	90
Health	5
Bequest indicator	2
Persistent labor productivity	5
Persistent medical expenditures	7
Total number of grid points	3,425,000
Transitory shocks	
Labor productivity	3
Medical expenditures	5

Table H.1: Grid dimensions for the agent's problem

   Note that the total number of grid points is not the product of individual grid dimensions because the state space is not rectangular. For example, agents below the age of 50 do not face medical expenditure shocks, while agents above the age of 89 cannot receive bequests since their parents are no longer alive.

2. Introducing a consumption floor and/or a bequest motive combined with potentially large medical expenditure shocks leads to consumption policy functions with downward jumps and value functions that are not concave. Standard methods that either rely on the concavity and differentiability of the value function or on the monotonicity of the consumption policy function (such as the endogenous gridpoint method, EGM) are therefore not applicable.

To illustrate, in Figure H.1 we plot the consumption policy, the savings rate and value function for the model with subjective beliefs, solved on a larger cash-at-hand grid of 500 points to more clearly bring out any discontinuities. Consider an individual at age 80 who is in the worst health and productivity states and faces the most severe persistent medical expenditure risk, a point on the state space where these problems are particularly pronounced.



(a) Consumption policy

(b) Savings rate

(c) Value function

Figure H.1: Solution to the agent's problem for an individual aged 80 in the worst health and productivity state who is facing the highest medical expenditure risk. Markers indicate the cash-at-hand grid points.

Panel Figure (a) shows the optimal consumption choice. Below the consumption floor of $\underline{c} = 0.05$, the agent consumes the consumption floor amount and does not save by assumption. However, even above the consumption floor, this individual does not save until a cash-at-hand level a of about 0.2 because any savings $k \in (0, a)$ will be wiped out by even the smallest realization of medical expenditures next period. Thus the agent would be back at the consumption floor next period, the same outcome as if nothing had been saved. Hence positive savings are suboptimal in that region.

For higher cash-at-hand levels a, the agent can choose a sufficiently high savings level $k \leq a$ so that some savings remain as long as the medical expenditure shock realization is small. This introduces a downward jump in the consumption policy because saving only makes sense if some minimum amount is saved, as otherwise everything will be wiped out by medical expenditure shocks. A similar logic generates the downward

jumps at higher levels of a shown in Figure (a): an even higher minimum level needs so be saved so that some wealth is left when the two lowest possible medical expenditure shocks realize, and so on.

The consumption floor and downward jumps in consumption moveover break the concavity of the value function, as shown in panel (c). In the region below the consumption floor, the value function is in fact flat because the agent will consume $\underline{c}$ in each period and will never be able to save out of that region. The value function is thus constant for sufficiently low levels of cash-at-hand. Additionally, as is well known, downward jumps in consumption correspond to downward kinks in the value function, as is barely visible at the cash-at-hand of around 0.35. Consequently, the value function is neither globally concave nor differentiable.

Note that this phenomenon is even more pronounced in the model with a warm-glow bequest motive because the same logic applies to bequests: estates are passed on net of medical expenditures incurred in the last year of life but are bounded by zero from below. Therefore, the incentive to leave bequest depends on whether the estate is wholly or partially wiped out by medical expenses.

While there are extensions of the EGM algorithm to deal with downward jumps in consumption policy functions, these are usually designed to address discontinuities introduced by an additional binary discrete choice (such as retirement) and often accompanied by extreme-value distributed taste shocks to smooth out the resulting kinks in the value function. In our setting, these jumps are much more frequent (because of the discrete nature of medical expenditure shocks, with as many as $7 \times 5 = 35$ possible realizations), and adding taste shocks makes no sense in our setting (since we are not dealing with discrete choices). We therefore opt to do the most conservative but robust algorithm and solve the agent's problem using grid search.

Grid search is know to create very jumpy consumption rules even for high values of cash-at-hand, introducing artificial discontinuities that are solely an artifact of the solution algorithm. We alleviate this issue by constructing the savings grid as follows:

1. For each point $(a, p, h, \eta, \mathbb{1}_b, t)$, we create a candidate grid of savings levels that is centered around the optimal savings level k_{-1} found in the previous iteration and covers an interval of approximately $\pm$ 0.15 $\frac{k_{-1}}{a}$. Since the agent's problem is computed many times when solving for the equilibrium, such a point always exists except for the very first iteration, where we simply create a candidate grid covering the feasible space $[0, a]$.

2. The savings grid we create is much denser close to the previous savings level k_{-1}

and more sparse towards the boundaries of the interval $\pm\, 0.15\, \frac{k_{-1}}{a}$.

Constructing the candidate savings grid this way creates remarkably smooth savings and consumption policy functions for higher cash-at-hand levels where the purely artificial downward jumps are not present, as can be seen in Figure F.3.

H.2 Equilibrium

To find a steady-state equilibrium defined in section D.4, we need to find fixed points in terms of

1. The aggregate capital stock K (or the equilibrium interest rate r);
2. The tax level λ.
3. The equilibrium wage rate which we pin to $w = 1$ by finding the corresponding TFP level A. This is only relevant for the benchmark models with subjective beliefs as we keep the TFP level unchanged when solving the models with objective survival beliefs (OSH) or no health heterogeneity (NHH).
4. The function $B(q_*, t_*)$ characterizing after-tax estates left by parents in lifetime income quintile q_* who die at age t_*.

To find the fixed point in terms of 1–3, we use the multidimensional root-finder known as Broyden's "good method" (Broyden 1965) to locate a root in terms of

1. Excess demand for capital (determining the equilibrium r);
2. The government budget balance (determining λ); and
3. "Excess" wages $w - 1$ (fixing $w = 1$).

Having to additionally find a fixed point in terms of $B(q_*, t_*)$ introduces a substantial complication as $B(q_*, t_*) \in \mathbb{R}_+^{5\times 60}$ is a matrix of $5 \times 60 = 300$ elements for five income quintiles and 60 years of generational overlap between parents and children. We found it infeasible to add these additional 300 elements to the root-finder, and instead we devised the following algorithm:

▷ *Initial guess:* ◁
$(K_1, \lambda_1, A_1) \leftarrow (K_{guess}, \lambda_{guess}, A_{guess})$
$B_1(q_*, t_*) \leftarrow \mathbf{0}_{[5\times 60]}$
▷ *Outer loop: Broyden's root-finding algorithm* ◁

5: **repeat**

6: $r_i \leftarrow \alpha F_K(A_i, K_i, L) - \delta_k, \quad w_i \leftarrow (1-\alpha) F_L(A_i, K_i, L)$

7: ▷ *Initial guess for inner loop* ◁

8: $B_{i,1}(q_*, t_*) \leftarrow B_i(q_*, t_*)$

9: ▷ *Inner loop: Fixed point of* $B(q_*, t_*)$ *for given* (K_i, λ_i, A_i) ◁

10: **repeat**

11: $b_{i,j}(p, h, t) \leftarrow$ from $B_{i,j}$ using (D.12)

12: $\left\{ \mathcal{C}_{i,j}(a, p, h, \eta, \mathbb{1}_b, t), \ \mathcal{K}_{i,j}(a, p, h, \eta, \mathbb{1}_b, t) \right\} \leftarrow$ solve agent's problem

13: $\Gamma_{i,j} \leftarrow$ simulate economy

14: $B_{i,j+1}(q_*, t_*) \leftarrow$ (D.6), (D.7), (D.9), (D.11)

15: **until** $\left\| B_{i,j} - B_{i,j+1} \right\|_\infty < \epsilon_B$

16: $K_{i+1} \leftarrow$ (D.27)

17: ▷ *Government budget deficit* ◁

18: $\Delta \leftarrow G + \Xi - T_{estate} - T_{inc}(\lambda)$ from (D.26)

19: **until** $\left\| \left(\frac{K_i - K_{i+1}}{K_i}, \Delta, w_i - 1 \right) \right\|_2 < \epsilon$

The outer loop on lines 5–19 represents Broyden's root-finder which iterates over candidate values of (K_i, λ_i, A_i), starting from some initial guess. It is sufficient to initialize aggregate bequests to a zero matrix $\mathbf{0}_{[5\times 60]}$ since it would be hard to come up with a good guess anyhow.

For given (K_i, λ_i, A_i), the inner loop on lines 10–15 finds a fixed point in terms of the bequest function $B(q_*, t_*)$. To this end, we compute the inheritances $b(p, h, t)$ expected by agents, solve the agent's problem and simulate the economy. We use the simulated distribution to update the amount of after-tax estates $B(q_*, t_*)$ and iterate on this inner loop until these are numerically close to the input value. Note that the agent's problem needs to be solved in each iteration of the inner loop (since the expected bequests potentially change) which is computationally very costly. However, it is sufficient to re-solve the problem only for agents in state $\mathbb{1}_b = 1$, as expected bequests have no effect on the decisions of agents whose parents are no longer alive.

References

Broyden, C. G. 1965. A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation* 19 (92): 577–593.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. Where is the land of opportunity? the geography of intergenerational mobility in the united states. *The Quarterly Journal of Economics* 129 (4): 1553–1623.

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

De Nardi, Mariacristina, Eric French, and John B Jones. 2010. Why do the elderly save? The role of medical expenses. *Journal of Political Economy* 118 (1): 39–75.

De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. *The lifetime costs of bad health.* Technical report. National Bureau of Economic Research.

Foltyn, Richard, and Jonna Olsson. 2021. *Health dynamics and heterogeneous life expectancies.* Working Papers 2021-17. Business School - Economics, University of Glasgow.

French, Eric, and John Bailey Jones. 2004. On the distribution and dynamics of health care costs. *Journal of Applied Econometrics* 19 (6): 705–721.

Gale, William G, and John Karl Scholz. 1994. Intergenerational transfers and the accumulation of wealth. *Journal of Economic Perspectives* 8 (4): 145–160.

Grevenbrock, Nils, Max Groneck, Alexander Ludwig, and Alexander Zimper. 2021. Cognition, optimism, and the formation of age-dependent survival beliefs. *International Economic Review* 62 (2): 887–918.

Heeringa, Steven G, Brady T West, and Patricia A Berglund. 2017. *Applied survey data analysis.* 2nd ed. Chapman / Hall/CRC.

HRS Staff. 2008. *HRS sample evolution: 1992-1998.* Survey Research Center, Institute for Social Research, University of Michigan, Ann Arbor, MI.

HRS Staff and Co-Investigators. 2006. *Getting started with the Health and Retirement Study.* Edited by Cindi Leacock.

Rao, J. N. K., and C. F. J. Wu. 1988. Resampling inference with complex survey data. *Journal of the American Statistical Association* 83 (401): 231–241.

Rust, Keith F, and J. N. K. Rao. 1996. Variance estimation for complex surveys using replication techniques. PMID: 8931197, *Statistical Methods in Medical Research* 5 (3): 283–310.

U.S. Bureau of Labor Statistics. 2024a. *Consumer Price Index for All Urban Consumers: all items in U.S. city average [CPIAUCSL].* Retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/CPIAUCSL.

———. 2024b. *Consumer Price Index for All Urban Consumers: medical care services in U.S. city average [CUUR0000SAM].* Retrieved from BLS; https://beta.bls.gov/dataViewer/view/timeseries/CUUR0000SAM.