The lost ones: the opportunities and outcomes of non-college educated Americans born in the 1960s

Margherita Borella, Mariacristina De Nardi, and Fang Yang*

March 11, 2019

Abstract

White, non-college-educated Americans born in the 1960s face shorter life expectancies, higher medical expenses, and lower wages per unit of human capital compared with those born in the 1940s, and men's wages declined more than women's. After documenting these changes, we use a life-cycle model of couples and singles to evaluate their effects. The drop in wages depressed the labor supply of men and increased that of women, especially in married couples. Their shorter life expectancy reduced their retirement savings but the increase in out-of-pocket medical expenses increased them by more. Welfare losses, measured a one-time asset compensation are 12.5%, 8%, and 7.2% of the present discounted value of earnings for single men, couples, and single women, respectively. Lower wages explain 47-58% of these losses, shorter life expectancies 25-34%, and higher medical expenses account for the rest.

*Margherita Borella: University of Torino and CeRP-Collegio Carlo Alberto, Italy. Mariacristina De Nardi: Federal Reserve Bank of Minneapolis, UCL, CEPR, and NBER. Fang Yang: Louisiana State University. De Nardi gratefully acknowledges financial support from the NORFACE Dynamics of Inequality across the Life-Course (TRISP) grant 462-16-122. We thank Jonathan Parker, who encouraged us to investigate the changes in opportunities and outcomes across cohorts, and provided us with valuable feedback. We are grateful to Marco Bassetto, John Bailey Jones, Rory McGee, Derek Neal, Gonzalo Paz-Pardo, Richard Rogerson, Rob Shimer, and seminar participants at various institutions for useful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research, the CEPR, any agency of the federal government, or the Federal Reserve Bank of Minneapolis.

1 Introduction

Much of macroeconomics either studies policies having to do with business cycle fluctuations or growth. Business cycle fluctuations are typically short-lived, do not affect a cohort's entire life cycle, and tend to have smaller welfare effects. Growth, instead, drastically improves the outcomes and welfare of successive cohorts over their entire lives compared to previous cohorts. Yet, recent evidence indicates that while we are still experiencing growth at the aggregate level, many people in recent cohorts are worse off, rather than benefiting from aggregate growth. It is important to study and better understand these cohort-level shocks and their consequences before trying to evaluate to what extent current government policies attenuate this kind of shocks and whether we should re-design some policies to reduce their impacts.

Recent research suggests that understanding these cohort-level shocks and their consequences is an important question. Guvenen et al. (2017) find that the median lifetime income of men born in the 1960s is 12-19% lower than that of men born in the 1940s, while Roys and Taber (2017) document that the wages of low-skilled men have stagnated over a similar time period. Hall and Jones (2007) highlight that the share of medical expenses to consumption has approximately doubled every 25 years since the 1950s, and Case and Deaton (2015 and 2017) have started an important debate by showing that the mortality rate of white, less-educated, middle-aged men has been increasing since 1999. In contrast with men's, the median lifetime income of women born in the 1960s is 22-33% higher than that of women born in the 1940s (Guvenen et al. 2017). The latter change, however, occurred in conjunction with much increased participation of women in the labor market.

While very suggestive, the changes in lifetime income tell us little about what happened to wages. In addition, depending on how wages, medical expenses, and mortality changed for married and single men and women, they can have weaker or stronger effects on couples, single men, and single women. Given the size of these changes and the large number of people that they affect, more investigations of their consequences is warranted.

The goal of this paper is to better measure these important changes in the lifetime opportunities of white, single and married, less-educated American men and women and to uncover their effects on the labor supply, savings, and welfare of a relatively recent birth cohort. To do so, we start by picking two cohorts of white, non-college educated Americans[1] for whom we have excellent data, those born in the 1940s and those born in the 1960s, and by using data from the Panel Study of Income Dynamics (PSID) and the Health and Retirement Study (HRS) to uncover several new facts.

First, we find that, across these two cohorts, men's average wages have decreased in real terms by 9% while women's average wages have increased by 7%, but that the increase in wages for women is due to higher human capital of women in the 1960s cohort rather than to higher wages per unit of human capital.[2] Second, we document a large increase in out-of-pocket medical expenses later in life: average out-of-pocket medical expenses after age 66 are expected to increase across cohorts by 82%. Third, we show that in middle age, the life expectancy of both female and male white, non-college-educated people is projected to go down by 1.1 to 1.7 years, respectively, from the 1940s to the 1960s cohort. All of these changes are thus large and have the potential to substantially affect behavior and welfare.

We then calibrate a life-cycle model of couples and singles to match the labor market outcomes for the 1960s cohort. Our calibrated model is a version of the life-cycle model developed by Borella, De Nardi, and Yang (2017),[3] which, in turn, builds on the literature on female labor supply (including Eckstein and Liftshitz (2011), Blundell et al. (2016a), Blundell et al (2016b), Fernandez and Wong (2017), and Eckstein et al. (2019)). Our model is well suited for our purposes for several reasons. It is a quantitative model that includes single and married people (with single people meeting partners and married people risking divorce), which matters because most people are in couples. It allows for human capital accumulation on the job, which our findings indicate is important, and includes medical expenses and life-span risk during retirement.

Our calibrated models matches key observed outcomes for the 1960s cohort very well. To evaluate the effects of the observed changes that we consider, we give the wage schedules, medical expenses, and life expectancy of the 1940s cohort to our 1960s cohort, starting at age 25, and we then study the effects of these changes on the 1960s cohort's labor supply, savings, and welfare.

[1] Because the finding of lower life expectancy is confined to less-educated whites, we focus on this group and to have a sample size that is large enough, we focus on non-college graduates.

[2] We measure human capital at a given age as average past earnings at that age (thus, our measure of human capital incorporates the effects of both years of schooling and work experience).

[3] Borella, De Nardi, and Yang (2017) develop and estimate this model to study the effects of marriage-based income taxes and Social Security benefits on the whole population, regardless of education.

We find that, of the three changes that we consider (the observed changes in the wage schedule, an increase in expected out-of-pocket medical expenses during retirement, and a decrease in life expectancy), the change in the wage schedule had by far the largest effect on the labor supply of both men and women. In particular, it depressed the labor supply of men and increased that of women. The decrease in life expectancy mainly reduced retirement savings but the expected increase in out-of-pocket medical expenses increased them by more.

We also find that the welfare costs of these changes are large. Specifically, the one-time asset compensation required at age 25 to make the 1960s households indifferent between the 1940s and 1960s health and survival dynamics, medical expenses, and wages, expressed as a fraction of their average discounted present value of earnings, are, 12.5, 8.0, and 7.2%, for single men, couples, and single women, respectively.[4] They are thus largest for single men and smallest for single women. Looking into the sources of these costs, we find that 47-58% of them are due to changes in the wage structure, 25-34% are due to changing life expectancy, and that medical expenses explain the remaining losses.

Our results thus indicate that the group of white, non-college educated people born in the 1960s cohort, which comprises about 60% of the population of the same age, experienced large negative changes in wages, large increases in medical expenses, and large decreases in life expectancy and would have been much better off if they had faced the corresponding lifetime opportunities of the 1940s birth cohort.

Our paper contributes to the previous literature along several important dimensions. First, it uncovers new facts on wages (and wages per unit of human capital), expected medical expenses during retirement, and life expectancy in middle age, for white, non-college-educated American men and women born in the 1940s and 1960s. Second, it recognizes that most people are not single, isolated individuals, but rather part of a couple and that changes in lifetime opportunities for one member of the couple could be either reinforced or weakened by the changes faced by their partner. Third, it documents these changes and introduces them in a carefully calibrated model that matches the lifetime outcomes of the 1960s cohort well. Fourth, it studies the effects of these changes in opportunities over time on the savings, labor market outcomes, and welfare of this cohort.

[4] These computations are performed for each household one at a time, keeping fixed the assets of their potential future partners in our benchmark.

The paper is organized as follows. Section 2 discusses our sample selection and the main characteristics of our resulting sample. Section 3 documents the changing opportunities for the 1940s and 1960s cohorts in terms of wages, medical expenses, and life expectancy. Section 4 describes the outcomes for our 1960s cohort in terms of labor market participation, hours worked by the workers, and savings. Section 5 discusses our structural model and thus the assumptions that we make to interpret the data. Section 6 explains our empirical strategy and documents the processes that we estimate as inputs of our structural model, including our estimated wages as a function of human capital and our estimated medical expenses and mortality as a function of age, gender, health, and marital status. Section 7 describes our results and Section 8 concludes.

2 The data and our sample

We use the PSID and the HRS to construct a sample of white, non-college educated Americans. We pick the cohort born in the 1940s (which is composed of the 1936-45 birth cohorts) as our comparison older cohort because it is the oldest cohort for which we have excellent data over most of their life cycle (first covered in the PSID and then in the HRS). We then pick our more recent cohort, the 1960s one (which is composed of the 1956-65 birth cohorts), to be as young as possible, conditional on having available data on most of their working period, which we require our structural model to match. We then compare the lifetime opportunities between these two cohorts. Appendix A reports more detail about the data and our computations.

To be explicit about the population that we are studying, we now turn to discussing our sampling choices for these cohorts and the resulting composition of our sample in terms of marital status and education level. We focus on non-college graduates for two reasons. First, we want to focus on less-educated people but we need a reasonable number of observations over the life cycle for both single and married men and women. Second, college graduates (and above) is the only group for which Case and Deaton (2017) find continued decreases in middle-age mortality over time.

Table 1 displays sample sizes before and after we apply our selection criteria. We start from 30,587 people and 893,420 observations. We keep household heads and their spouses, if present, and restrict the sample to the cohorts born between 1935 and 1965, to whites, and to include observations reporting their education. Our

Selection	Individuals	Observations
Initial sample (observed at least twice)	30,587	893,420
Heads and spouses (if present)	18,304	247,203
Born between 1935 and 1965	7,913	137,427
Age between 20 and 70	7,847	135,117
White	6,834	116,810
Non-missing education	6,675	116,619
Non-college graduates	5,039	73,944

Table 1: PSID sample selection

sample before performing the education screens comprises 6,675 people and 116,619 observations. Dropping all college graduates and those married with college graduates results in a sample of 5,039 people and 73,944 observations.[5]

Turning to our resulting PSID sample, at age 25, 90% and 77% of people in the 1940s and 1960s birth cohorts are married, respectively. To understand how education changed within our sample of interest, Table 2 reports the education distribution at age 25 for our non-college graduates in the 1940s and 1960s cohorts. It shows that the fraction of people without a high school diploma decreased by 40% for men and 43% for women from the 1940s to the 1960s cohort. Our model and empirical strategy takes into account education composition within our sample because they control for people's human capital, both at labor market entry and over the life cycle.

	Men		Women	
	1940	1960	1940	1960
Less than HS	0.29	0.17	0.23	0.13
HS	0.32	0.33	0.37	0.39
More than HS	0.39	0.50	0.40	0.48

Table 2: Fractions of individuals by education level in our two birth cohorts

One might worry about a different type of selection, that is the one coming from the fact that we drop people who completed college from our sample for all of our cohorts. If college completion rates were rising fast between 1940s and 1960s, with

[5]Thus, we also drop people with less than 16 years of education but married with someone with 16 or more years of education. Before making this selection, non-graduate husbands with a graduate wife were 5% of the sample, while non-graduate wives with a graduate husband were 9.7% of the sample.

the most able going into college, our 1960s cohort might be much more negatively selected than our 1940s cohort. Table 25 in Appendix C shows that, in the PSID, the fraction of the population having less than a college degree dropped from 83.1% in 1940s to 77.2% in 1960s. This corresponds to a 5.9 percentage points drop in non-college graduates in the population across our two cohorts (5.6 and 6.7 percentage points for men and women, respectively). Appendix C also compares the implications of our PSID and HRS samples for our model inputs with those of the corresponding samples in which we keep a constant fraction of the population for both cohorts. All of these comparisons show that our model inputs are very similar for both types of samples and that our results are thus not driven by selection out of our sample.

Because the HRS contains a large number of observations and high-quality data after age 50, we use it to compute our inputs for the retirement period. The last available HRS wave is for 2014, which implies that we do not have complete data on the life cycle of the two cohorts that we are interested in. In fact, individuals ages were, respectively, 69-78 and 49-58 in the 1936-1945 and 1956-65 cohorts as of year 2014. We use older cohorts to extrapolate outcomes for the missing periods for our cohorts of interest and we start estimation at age 50 so that the 1960s cohort is observed for a few waves in our sample.

Thus, our sample selection for the HRS is as follows. Of the 449,940 observations initially present, we delete those with missing crucial information (e.g. on marital status) and we select waves since 1996. We then select individuals in the age range 50-100. Given that we use years from 1996 to 2014, these people were born between 1906 and 1964. After keeping white and non-college-graduates and spouses, we have 19,377 individuals and 110,923 observations, as detailed in Table 3.

Selection	Individuals	Observations
Initial sample	37,495	449,940
Non-missing information	37,152	217,574
Wave 1996 or later	35,936	204,922
Age 50 to 100	34,775	197,431
White	25,693	152,688
Non-college graduates	19,377	110,923

Table 3: HRS sample selection

3 Changes in wages, medical expenses, and life expectancy across cohorts

In this section, we describe the observed changes in wages, medical expenses,[6] and life expectancy experienced by white, non-college educated Americans born in the 1960s compared with those born in the 1940s. We show that the wages of men went down by 7%, while the wages of women went up by 9%. These changes do not condition on human capital within an education group (we report wages per unit of human capital in Section 6.1, after we make explicit how we model human capital). We also show that, during retirement, out-of-pocket medical spending increased by 82%, while life expectancy decreased by 1.1 to 1.7 years.

3.1 Wages

Figure 1 displays smoothed average wage profiles for labor market participants.[7] The left-hand-side panel displays wages for married men and women in the 1940s and 1960s cohort, while the right-hand-side panel displays the corresponding wages for single people. Several features are worth noticing. First, the wages of men were much higher than those of women in the 1940s birth cohort. Second, the wages of men, both married and single, went down by 9%. Third, the wages of married and single women went up by 7% across these two cohorts.

Our model, however, requires potential wages as an input. Because the wage is missing for those who are not working, we impute missing wages (See details in Appendix B). Figure 2 shows our estimated potential wage profiles. Potential wages for men are similar to observed wages for labor market participants, except that potential wages drop faster than observed wages after age 55. Potential wages for women not only drop faster after middle age than observed wages, but also tend to be lower and grow more slowly at younger ages due to positive selection of women in

[6] All amounts in the paper are expressed in 2016 dollars.

[7] To compute these average wage profiles, we first regress log wages on fixed-effects regressions with a flexible polynomial in age, separately for men and women. We then regress the sum of the fixed effects and residuals from these regressions on cohort and marital status dummies to fix the position of the age profile. Finally, we model the variance of the shocks by fitting age polynomials to the squared residuals from each regression in logs, and use it to compute the level of average wages of each group as a function of age (by adding half the variance to the average in logs before exponentiating).

the labor market.

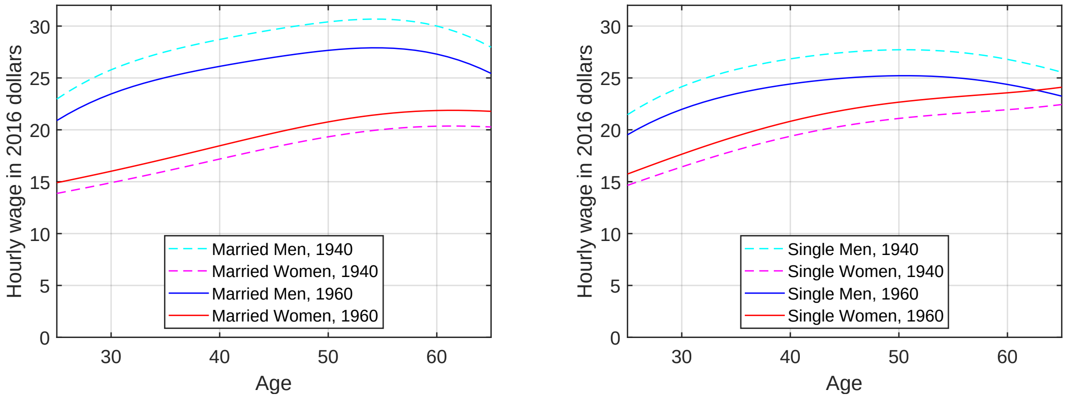


Figure 1: Wage profiles, comparing 1960s and 1940s for married people (left panel) and single people (right panel)

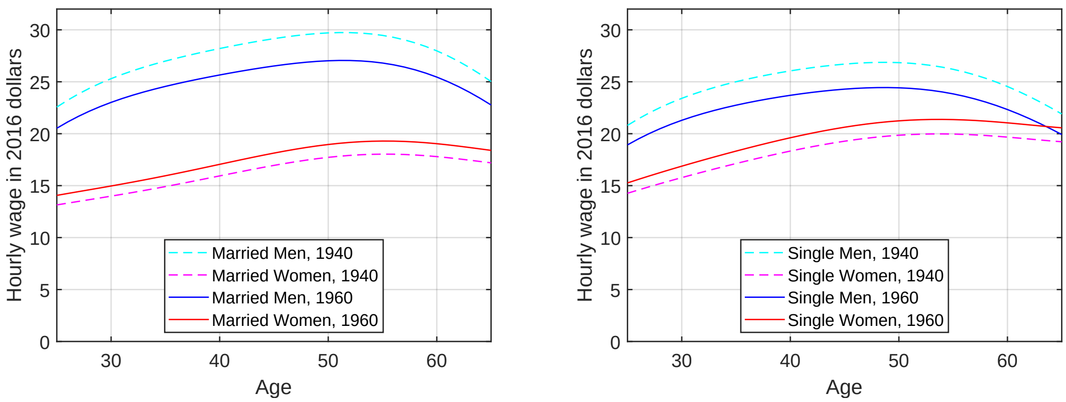


Figure 2: Potential wage profiles, comparing 1960s and 1940s for married people (left panel) and single people (right panel)

Both figures display overall similar patterns and, in particular, imply that the large wage gap between men and women in the 1940s cohort significantly decreased for the 1960s cohort because of increasing wages for women and decreasing wages for men.

3.2 Medical expenses

We use the HRS data to compute out-of-pocket medical expenses during retirement for the 1940s and 1960s cohorts.[8] Figure 3 indicates a large increase in real

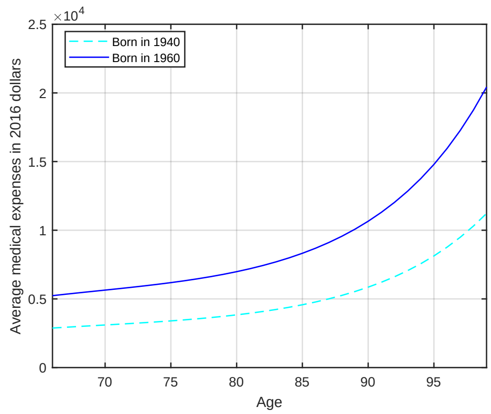


Figure 3: Average out-of-pocket medical expenses for the cohorts born in 1940s and 1960s

average expected out-of-pocket medical expenses across cohorts. For instance, at age 66, out-of-pocket medical expenses expressed in 2016 dollars are $2,878 and $5,236, respectively, for the 1940s and 1960s birth cohorts. The corresponding numbers for someone who survives to age 90 are $5,855 and $10,655. Thus, average out-of-pocket medical expenses after age 66 are expected to increase across cohorts by 82%. These are dramatic increases for two cohorts that are only twenty years apart.

3.3 Life expectancy

Case and Deaton (2015 and 2017) use data from the National Vital Statistics to study mortality by age over time and find that, interrupting a long time trend in mortality declines, the mortality of white, middle-age, and non-college educated Americans went up during the 1999 to 2015 time period. In particular, they found that individuals age 55-59 in 2015 (and thus born in 1956-1960) faced a 22% increase

[8] To generate this graph, we regress the logarithm of out-of-pocket medical expenses on a fixed effect and a third-order polynomial in age. We then regress the sum of the fixed effects and residuals from this regression on cohort dummies to compute the average effect for each cohort of interest and we add the cohort dummies into the age profile. Finally, we model the variance of the shocks fitting an age polynomial and cohort dummies to the squared residuals from the regression in logs, and use it to construct average medical expenses as a function of age.

in mortality with respect to individuals age 55-59 in 1999 (and thus born in 1940-1944). Looking at a younger group, they find that individuals age 50-54 in 2015 (thus born in 1961-1965) experienced a 28% increase in mortality compared with individuals in the same age group and born sixteen years earlier.

Using the HRS data, we find that mortality at age 50 increased by about 27 percent from the 1940s to the 1960s cohort.[9] Thus, the increases in mortality in the HRS data are in line with those found by Case and Deaton.

To further understand the HRS's data implications about mortality and their changes across our two cohorts, we also report the life expectancies that are implied by our HRS data. Table 4 shows that life expectancy at age 50 was age 77.6 and 79.8 for men and women, respectively, in the cohort born in 1940s. Conditional on being alive at age 66, men and women in this cohort expect to live until age 82.5 and 85.7, respectively. It also shows that the life expectancy of men at age 50 declined by 1.5 years across our two cohorts, which is a large decrease for cohorts that are twenty year apart and during a period of increasing life expectancy for people in other groups. The table also reveals two other interesting facts. First, the life expectancy of 50 year old women in the same group also decreased by 1.1 years. Second, life expectancy at age 66 fell slightly more than life expectancy at age 50 (by 1.6 years for men and 1.7 for women).

	Men, 1940	Men, 1960	Women, 1940	Women, 1960
At age 50	77.6	76.1	79.8	78.7
At age 66	82.5	80.9	85.7	84.0

Table 4: Life expectancy for white and non-college educated men and women born in the 1940s and 1960s cohorts. HRS data

As a comparison, for the year 2005, the life tables provided by the US Department of Health and Human Services (Arias et al., 2010) report a life expectancy at age 66 (and thus for people born in the 1940s) of 82.1 and 84.7 for white men and women respectively. Compared to the official life tables, we thus slightly overestimate life expectancy, especially for women, a result that possibly reflects that the HRS sample

[9] We obtain the results in this section by estimating the probability of being alive conditional on age and cohort and by assuming that the age profiles entering the logit regression are the same across cohorts up to a constant. We then compute the mortality rate for the cohorts of interest using the appropriate cohort dummy.

is drawn from non-institutionalized, and thus initially healthier, individuals. After the initial sampling, people ending up in nursing homes in subsequent periods stay in the HRS data set.

One might wonder whether people born in 1960s were aware that their life expectancy was shorter than that of previous generations. To evaluate this, we use the HRS question about one's subjective probability of being alive at age 75. As Table 5 shows, people born in 1960s did adjust their life expectancy downward compared to those born in 1940s. That is, men age 55 and born in 1940s report, on average, a subjective probability of being alive at age 75 of 61%, compared with 56% for those born the 1960s. For women, the drop is even larger, going from 66% for those born in 1940s to 58% for those born in the 1940s.

	Men	Women
Born in 1940s	61	66
Born in 1960s	56	58

Table 5: Average subjective probability (in percentage) reported of being alive at age 75 reported by people age 54-56 who are white and non-college educated. HRS data

4 Labor market and savings outcomes for the 1960s cohort

Figure 4 displays the smoothed life cycle profiles[10] of participation, hours worked by workers, and assets, for the 1960s cohort, by gender and marital status. Its left panel highlights several important patterns. First, married men have the highest labor market participation. Second, the participation of single men drops faster by age than that of married men. Third, single women have a participation profile that looks like a shifted down version of that of married men. Lastly, married women have

[10]The smoothed profiles of participation and hours are obtained by regressing each variable on a fourth-order polynomial in age fully interacted with marital status, and on cohort dummies, also interacted with marital status, which pick up the position of the age profiles. For assets, the profiles are obtained by fitting age polynomials separately for single men, single women and couples to the logarithm of assets plus shift parameter, also controlling for cohort. The variance of the shocks is modeled by fitting age polynomials to the squared residuals from the regression in logs and is used to obtain the average profile in levels. Our figures display the profiles for the 1960s cohort.

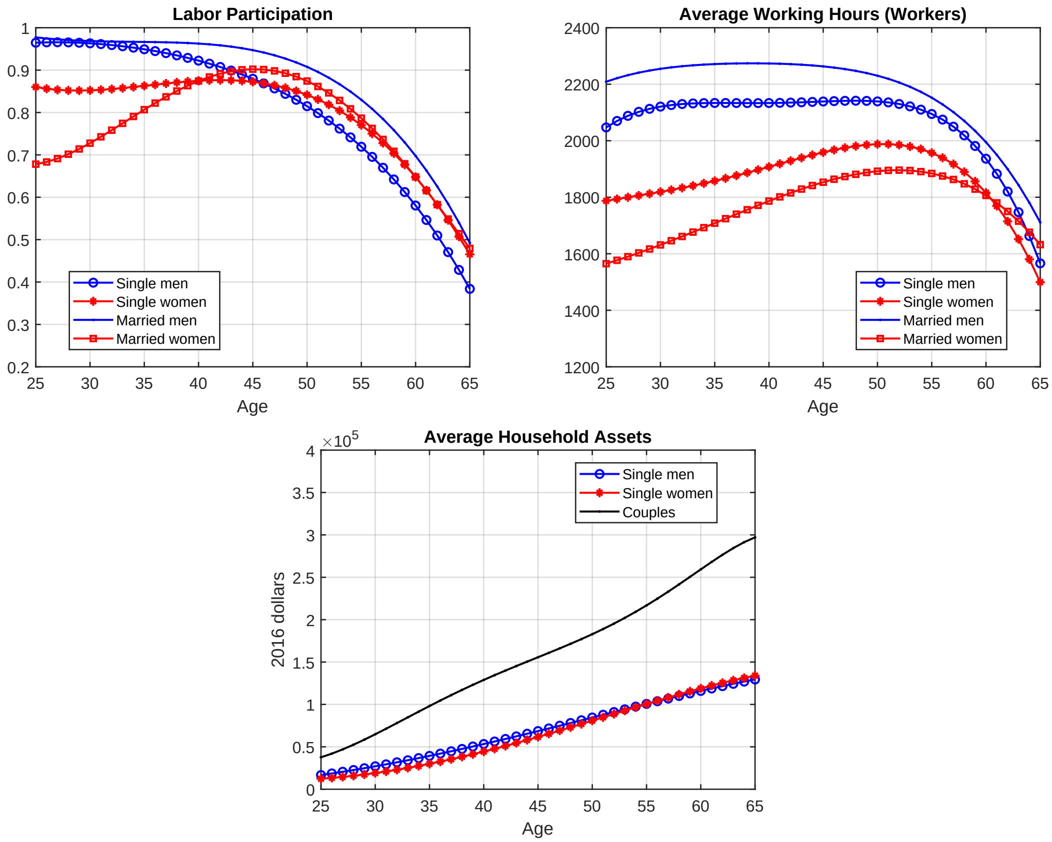


Figure 4: Participation, hours by workers, and average assets for the cohort born in 1960

the lowest participation until age 40, but it then surpasses that of single men and single women up to age 65.

The right panel displays hours worked conditional on participation, with married men working the most hours, followed by single men, single women, and married women until age 60. The bottom panel of the figure displays savings accumulation up to age 65 and shows that couples start out with more assets than singles and that this gap widens with age, to peak at about two by retirement time.

We see these outcome as important aspects of the data that we require our model to match in order to trust its implications about the effects of the changes in their lifetime opportunities that we consider.

5 The model

The model that we use is a version of that in Borella, De Nardi, and Yang (2017). Thus, we follow their exposition closely. A model period is one year long. People start their economic life at age 25, stop working at age 66 at the latest, and live up to age 99.

During the working stage, people choose how much to save and how much to work, face wage shocks and, if they are married, divorce shocks. Single people meet partners. For tractability, we make the following assumptions. People who are married to each other have the same age. Marriage, divorce, and fertility are exogenous. Women have an age-varying number of children that depends on their age and marital status. We estimate all of these processes from the data.

During the retirement stage, people face out-of-pocket medical expenses which are net of Medicare and private insurance payments, and are partly covered by Medicaid. Married retired couples also face the risk of one of the spouses dying. Single retired people face the risk of their own death. We allow mortality risk and medical expenses to depend on gender, age, health status, and marital status.

We allow for both time costs and monetary costs of raising children and running households. In terms of time costs, we allow for available time to be split between work and leisure and to depend on gender and marital status. We interpret available as net of home production, child care, and elderly care that one has to perform whether working or not (and that is not easy to out-source). In addition, all workers have to pay a fixed cost of working which depends on their age.

The monetary costs enters our model in the two ways. There is an adult-equivalent family size that affects consumption. In addition, when women work, they have to pay a child care cost that depends on the age and number of their children, and on their own earnings. We assume that child care costs are a normal good: women with higher earnings pay for more expensive child care.

We assume that households have rational expectations about all of the stochastic processes that they face. Thus, they anticipate the nature of the uncertainty in our environment starting from age 25, when they enter our model.

5.1 Preferences

Let t be age $\in \{t_0, t_1, ..., t_r, ..., t_d\}$, with $t_0 = 25$, $t_r = 66$ being retirement time and $t_d = 99$ being the maximum possible lifespan. For simplicity of notation think of the model as being written for one cohort, thus age t also indexes the passing of time for that cohort. We solve the model for our 1960s cohort and then perform our counterfactuals by changing some of its inputs to those of the 1940s cohort.

Households have time-separable preferences and discount the future at rate β. The superscript i denotes gender; with $i = 1, 2$ being a man or a woman, respectively. The superscript j denotes marital status; with $j = 1, 2$ being single or in a couple, respectively.

Each **single person** has preferences over consumption and leisure, and the period flow of utility is given by the standard CRRA utility function

$$v^i(c_t, l_t) = \frac{((c_t/\eta_t^{i,1})^{\omega} l_t^{1-\omega})^{1-\gamma} - 1}{1-\gamma} + b$$

where c_t is consumption, $\eta_t^{i,j}$ is the equivalent scale in consumption (which is a function of family size, including children) and $\eta_t^{i,1}$ corresponds to that for singles, while $b \geq 0$ is a parameter that ensures that people are happy to be alive, as in Hall and Jones (2007). The latter allows us to properly evaluate the welfare effects of changing life expectancy.

The term $l_t^{i,j}$ is leisure, which is given by

$$l_t^{i,j} = L^{i,j} - n_t - \Phi_t^{i,j} I_{n_t},$$

where $L^{i,j}$ is available time endowment, which can be different for single and married men and women and should be interpreted as available time net of home production. Leisure equals available time endowment less n_t, hours worked on the labor market, less the fixed time cost of working. That is, the term I_{n_t} is an indicator function which equals 1 when hours worked are positive and zero otherwise, while the term $\Phi_t^{i,j}$ represents the fixed time cost of working.

The fixed cost of working should be interpreted as including commuting time, time spent getting ready for work, and so on. We allow it to depend on gender, marital status and age because working at different ages might imply different time costs for married and single men and women. We assume the following functional form, whose three parameters we calibrate using our structural model,

$$\Phi_t^{i,j} = \frac{\exp(\phi_0^{i,j} + \phi_1^{i,j} t + \phi_2^{i,j} t^2)}{1 + \exp(\phi_0^{i,j} + \phi_1^{i,j} t + \phi_2^{i,j} t^2)}.$$

We assume that **couples** maximize their joint utility function

$$w(c_t, l_t^1, l_t^2) = \frac{((c_t/\eta_t^{i,2})^{\omega}(l_t^1)^{1-\omega})^{1-\gamma} - 1}{1-\gamma} + b + \frac{((c_t/\eta_t^{i,2})^{\omega}(l_t^2)^{1-\omega})^{1-\gamma} - 1}{1-\gamma} + b.$$

Note that for couples the economy of scale term $\eta_t^{i,2}$ is the same for both genders.

5.2 The environment

Households hold assets a_t, which earn rate of return r. The timing is as follows. At the beginning of each working period, each single individual observes his/her current idiosyncratic wage shock, age, assets, and accumulated earnings. Each married person also observes their partner's labor wage shock and accumulated earnings. At the beginning of each retirement period, each single individual observes his/her current age, assets, health, and accumulated earnings. Each married person also observes their partner's health and accumulated earnings. Decisions are made after everything has been observed and new shocks hit at the end of the period after decisions have been made.

5.2.1 Human capital and wages

We take education at age 25 as given but explicitly model human capital accumulation after that age. To do so, we define human capital, $\bar{y}_t^i$, as one's average past earnings at each age. Thus, our definition of human capital implies that it is a function of one's initial wages and schooling and subsequent labor market experience and wages.[11]

There are two components to wages. The first is a deterministic function of human capital: $e_t^{i,j}(\bar{y}_t^i)$. The second component is a persistent earnings shock ϵ_t^i that evolves as follows

$$\ln \epsilon_{t+1}^i = \rho_\varepsilon^i \ln \epsilon_t^i + \upsilon_t^i, \;\; \upsilon_t^i \sim N(0, (\sigma_\upsilon^i)^2).$$

[11] It also has the important benefit of allowing us to have only one state variable keeping track of human capital and Social Security contributions.

The product of $e_t^{i,j}(\cdot)$ and ϵ_t^i determines an agent's hourly wage.

5.2.2 Marriage and divorce

During the working period, a single person gets married with an exogenous probability which depends on his/her age and gender. The probability of getting married at the beginning of next period is ν_{t+1}^i.

Conditional on meeting a partner, the probability of meeting a partner p with wage shock ϵ_{t+1}^p is

$$\xi_{t+1}(\cdot) = \xi_{t+1}(\epsilon_{t+1}^p | \epsilon_{t+1}^i, i). \tag{1}$$

Allowing this probability to depend on the wage shock of both partners generates assortative mating. We assume random matching over assets a_{t+1} and average accumulated earnings of the partner $\bar{y}_{t+1}^p$, conditional on partner's wage shock. We estimate the distribution of partners over these state variables from the PSID data (see Appendix B, Marriage and divorce probabilities subsection, for details) and denote it by

$$\theta_{t+1}(\cdot) = \theta_{t+1}(a_{t+1}^p, \bar{y}_{t+1}^p | \epsilon_{t+1}^p), \tag{2}$$

where the variables $a_{t+1}^p, \bar{y}_{t+1}^p, \epsilon_{t+1}^p$ stand for partner's assets, human capital, and wage shock, respectively.

A working-age couple can be hit by a divorce shock at the end of the period that depends on age, ζ_t. If the couple divorces, they split the assets equally and each of the ex-spouses moves on with those assets and their own wage shock and Social Security contributions.

After retirement, single people don't get married anymore. People in couples no longer divorce and can lose their spouse only because of death. This is consistent with the data because in this cohort marriages and divorces after retirement are rare.

5.2.3 The costs of raising children and running a household

Consistently with the data for this cohort, we assume that single men do not have children. We keep track of the total number of children and children's age as a function of mother's age and marital status. The total number of children by one's age affects the economies of scale of single women and couples. We denote by $f^{0,5}(i,j,t)$ and $f^{6,11}(i,j,t)$ the number of children from 0 to 5 and from 6 to 11, respectively.

The term $\tau_c^{0,5}$ is the child care cost for each child age 0 to 5, while $\tau_c^{6,11}$ is the child care cost for each child age 6 to 11. Both are expressed as fraction of the earnings of the working mother.

The number of children between ages 0 to 5 and 6 to 11, together with the per-child child care costs by age of child, determine the child care costs of working mothers ($i = 2$). Because we assume that child care costs are proportional to earnings, if a woman does not work her earnings are zero and so are her child care costs. This amounts to assuming that she provides the child care herself.

5.2.4 Medical expenses and death

After retirement, surviving people face medical expenses, health, and death shocks. At age 66, we endow people with a distribution of health that depends on their marital status and gender (See Appendix B, Health status at retirement subsection).

Health status ψ_t^i can be either good or bad and evolves according to a Markov process $\pi_t^{i,j}(\psi_t^i)$ that depends on age, gender, and marital status. Medical expenses $m_t^{i,j}(\psi_t^i)$ and survival probabilities $s_t^{i,j}(\psi_t^i)$ are functions of age, gender, marital status, and health status.

5.2.5 Initial conditions

We take the fraction of single and married people at age 25 and their distribution over the relevant state variables from the PSID data. We list all of our state variables in Section 5.4.

5.3 The Government

We model taxes on total income Y as Gouveia and Strauss (1994) and we allow them to depend on marital status as follows

$$T(Y, j) = (b^j - b^j(s^j Y + 1)^{-\frac{1}{p^j}})Y.$$

The government also uses a proportional payroll tax τ_t^{SS} on labor income, up to a Social Security cap $\widetilde{y}_t$, to help finance old-age Social Security benefits. We allow both the payroll tax and the Social Security cap to change over time for the 1960 cohort, as in the data.

We use human capital $\bar{y}_t^i$ (computed as an individual's average earnings at age t) to determine both wages and old age Social Security payments. While Social Security benefits for a single person are a function of one's average lifetime earnings, Social Security benefits for a married person are the highest of one's own benefit entitlement and half of the spouse's entitlement while the other spouse is alive (spousal benefit). After one's spousal death, one's Social Security benefits are given by the highest of one's benefit entitlement and the deceased spouse's (survival benefit).

The insurance provided by Medicaid and SSI in old age is represented by a means-tested consumption floor, $\underline{c}(j)$.[12]

5.4 Recursive formulation

We define and compute six sets of value functions: the value function of working-age singles, the value function of retired singles, the value function of working-age couples, the value function of retired couples, the value function of an individual who is of working-age and in a couple, the value function of an individual who is retired and in a couple.

5.4.1 The singles: working age and retirement

The state variables for a single individual during one's working period are age t, gender i, assets a_t^i, the persistent earnings shock ϵ_t^i, and average realized earnings $\bar{y}_t^i$. The corresponding value function is

$$W^s(t,i,a_t^i,\epsilon_t^i,\bar{y}_t^i) = \max_{c_t,a_{t+1},n_t^i}\Bigg(v^i(c_t,l_t^{i,j}) + \beta(1-\nu_{t+1}(i))E_t W^s(t+1,i,a_{t+1}^i,\epsilon_{t+1}^i,\bar{y}_{t+1}^i) + \beta\nu_{t+1}(i)E_t\Big[\hat{W}^c(t+1,i,a_{t+1}^i+a_{t+1}^p,\epsilon_{t+1}^i,\epsilon_{t+1}^p,\bar{y}_{t+1}^i,\bar{y}_{t+1}^p)\Big]\Bigg) \quad (3)$$

$$l_t^{i,j} = L^{i,j} - n_t^i - \Phi_t^{i,j} I_{n_t^i}, \quad (4)$$

$$Y_t^i = e_t^{i,j}(\bar{y}_t^i)\epsilon_t^i n_t^i, \quad (5)$$

$$\tau_c(i,j,t) = \tau_c^{0,5} f^{0,5}(i,j,t) + \tau_c^{6,11} f^{6,11}(i,j,t), \quad (6)$$

[12] Borella, De Nardi, and French (2017) discuss Medicaid rules and observed outcomes after retirement.

$$T(\cdot) = T(ra_t + Y_t, j), \tag{7}$$

$$c_t + a_{t+1} = (1+r)a_t^i + Y_t^i(1-\tau_c(i,j,t)) - \tau_t^{SS}\min(Y_t^i, \widetilde{y}_t) - T(\cdot), \tag{8}$$

$$\bar{y}_{t+1}^i = (\bar{y}_t^i(t-t_0) + (\min(Y_t^i, \widetilde{y}_t)))/(t+1-t_0), \tag{9}$$

$$a_t \geq 0, \quad n_t \geq 0, \quad \forall t. \tag{10}$$

The expectation of the value function next period if one remains single integrates over one's wage shock next period. When one gets married, not only we take a similar expectation, but we also integrate over the distribution of the state variables of one's partner ($\xi_{t+1}(\epsilon_{t+1}^p|\epsilon_{t+1}^i, i)$ is the distribution of the partner's wage shock defined in Equation (1) and $\theta_{t+1}(\cdot)$ is the distribution of partner's assets and human capital defined in Equation (2)).

The value function $\hat{W}^c$ is the discounted present value of the utility for the same individual, once he or she is in a married relationship with someone with given state variables, not the value function of the married couple, which counts the utility of both individuals in the relationship. We discuss the computation of the value function of an individual in a marriage later in this section.

Equation 5 shows that the deterministic component of wages is a function of age, gender, marital status, and human capital.

Equation 9 describes the evolution of human capital, which we measure as average accumulated earnings (up to the Social Security earnings cap $\widetilde{y}_t$) and that we use as a determinant of future wages and Social Security payments after retirement.

During the last working period, a person takes the expected values of the value functions during the first period of retirement. The state variables for a retired single individual are age t, gender i, assets a_t^i, health ψ_t^i, and average realized lifetime earnings $\bar{y}_r^i$. Because we assume that the retired individual can no longer get married, his or her recursive problem can be written as

$$R^s(t,i,a_t,\psi_t^i,\bar{y}_r^i) = \max_{c_t,a_{t+1}}\left(v^i(c_t,L^{i,j}) + \beta s_t^{i,j}(\psi_t^i)E_t R^s(t+1,i,a_{t+1},\psi_{t+1}^i,\bar{y}_r^i)\right) \tag{11}$$

$$Y_t = SS(\bar{y}_r) \tag{12}$$

$$T(\cdot) = T(Y_t + ra_t, j) \tag{13}$$

$$B(a_t,Y_t,\psi_t^i,\underline{c}(j)) = \max\Big\{0, \underline{c}(j) - [(1+r)a_t + Y_t - m_t^{i,j}(\psi_t^i) - T(\cdot)]\Big\} \tag{14}$$

$$c_t + a_{t+1} = (1+r)a_t + Y_t + B(a_t, Y_t, \psi_t^i, \underline{c}(j)) - m_t^{i,j}(\psi_t^i) - T(\cdot) \tag{15}$$

$$a_{t+1} \geq 0, \quad \forall t \tag{16}$$

$$a_{t+1} = 0, \quad \text{if} \quad B(\cdot) > 0 \tag{17}$$

The term $s_t^{i,j}(\psi_t^i)$ is the survival probability as a function of age, gender, marital and health status. The expectation of the value function next period is taken with respect to the evolution of health.

The term $SS(\bar{y_r}^i)$ represents Social Security, which for the single individual is a function of the income earned during their work life, $\bar{y}_r^i$ and the fnction $B(a_t, Y_t^i, \psi_t^i, \underline{c}(j))$ represents old age means-tested government transfers such as Medicaid and SSI, which ensure a minimum consumption floor $\underline{c}(j)$.

5.4.2 The couples: working age and retirement

The state variables for a married couple in the working stage are $(t, a_t, \epsilon_t^1, \epsilon_t^2, \bar{y}_t^1, \bar{y}_t^2)$ where 1 and 2 refer to gender, and the recursive problem for the married couple ($j = 2$) before t_r can be written as:

$$\begin{aligned} W^c(t, a_t, \epsilon_t^1, \epsilon_t^2, \bar{y}_t^1, \bar{y}_t^2) = \max_{c_t, a_{t+1}, n_t^1, n_t^2} \Bigg(& w(c_t, l_t^{1,j}, l_t^{2,j}) \\ & + (1 - \zeta_{t+1})\beta E_t W^c(t+1, a_{t+1}, \epsilon_{t+1}^1, \epsilon_{t+1}^2, \bar{y}_{t+1}^1, \bar{y}_{t+1}^2) \\ & + \zeta_{t+1}\beta \sum_{i=1}^{2} \Big(E_t W^s(t+1, i, a_{t+1}/2, \epsilon_{t+1}^i, \bar{y}_{t+1}^i) \Big) \Bigg) \end{aligned} \tag{18}$$

$$l_t^{i,j} = L^{i,j} - n_t^i - \Phi_t^{i,j} I_{n_t^i}, \tag{19}$$

$$Y_t^i = e_t^{i,j}(\bar{y}_t^i)\epsilon_t^i n_t^i, \tag{20}$$

$$\tau_c(i,j,t) = \tau_c^{0,5} f^{0,5}(i,j,t) + \tau_c^{6,11} f^{6,11}(i,j,t), \tag{21}$$

$$T(\cdot) = T(ra_t + Y_t^1 + Y_t^2, j) \tag{22}$$

$$c_t + a_{t+1} = (1+r)a_t + Y_t^1 + Y_t^2(1 - \tau_c(2,2,t)) - \tau_t^{SS}(\min(Y_t^1, \widetilde{y}_t) + \min(Y_t^2, \widetilde{y}_t)) - T(\cdot) \tag{23}$$

$$\bar{y}_{t+1}^i = (\bar{y}_t^i(t - t_0) + (\min(Y_t^i, \widetilde{y}_t)))/(t + 1 - t_0), \tag{24}$$

$$a_t \geq 0, \quad n_t^1, n_t^2 \geq 0, \quad \forall t \tag{25}$$

The expected value of the couple's value function is taken with respect to the conditional probabilities of the two ϵ_{t+1}s given the current values of the ϵ_ts for each of the spouses (we assume independent draws). The expected values for the newly divorced people are taken using the appropriate conditional distribution for their own labor wage shocks.

During their last working period, couples take the expected values of the value functions for the first period of retirement. During retirement, that is from age t_r on, each of the spouses is hit with a health shock ψ_t^i and a realization of the survival shock $s_t^{i,2}(\psi_t^i)$. Symmetrically with the other shocks, $s_t^{1,2}(\psi_t^1)$ is the after retirement survival probability of husband, while $s_t^{2,2}(\psi_t^2)$ is the survival probability of the wife. We assume that the health shocks of each spouse are independent of each other and that the death shocks of each spouse are also independent of each other.

In each period, the married couple's ($j = 2$) recursive problem during retirement can be written as

$$
\begin{aligned}
R^c(t, a_t, \psi_t^1, \psi_t^2, \bar{y}_r^1, \bar{y}_r^2) = \max_{c_t, a_{t+1}} \Bigg(& w(c_t, L^{1,j}, L^{2,j}) + \\
& \beta s_t^{1,j}(\psi_t^1) s_t^{2,j}(\psi_t^2) E_t R^c(t+1, a_{t+1}, \psi_{t+1}^1, \psi_{t+1}^2, \bar{y}_r^1, \bar{y}_r^2) + \\
& \beta s_t^{1,j}(\psi_t^1)(1 - s_t^{2,j}(\psi_t^2)) E_t R^s(t+1, 1, a_{t+1}, \psi_{t+1}^1, \bar{\bar{y}}_r) + \\
& \beta s_t^{2,j}(\psi_t^2)(1 - s_t^{1,j}(\psi_t^1)) E_t R^s(t+1, 2, a_{t+1}, \psi_{t+1}^2, \bar{\bar{y}}_r) \Bigg)
\end{aligned} \tag{26}
$$

$$
Y_t = \max\left\{ (SS(\bar{y}_r^1) + SS(\bar{y}_r^2), \frac{3}{2} \max(SS(\bar{y}_r^1), SS(\bar{y}_r^2)) \right\} \tag{27}
$$

$$
\bar{\bar{y}}_r = \max(\bar{y}_r^1, \bar{y}_r^2), \tag{28}
$$

$$
T(\cdot) = T(Y_t + r a_t, j), \tag{29}
$$

$$
B(a_t, Y_t, \psi_t^1, \psi_t^2, \underline{c}(j)) = \max\left\{ 0, \underline{c}(j) - \left[(1+r)a_t + Y_t - m_t^{1,j}(\psi_t^1) - m_t^{2,j}(\psi_t^2) - T(\cdot) \right] \right\} \tag{30}
$$

$$
c_t + a_{t+1} = (1+r)a_t + Y_t + B(a_t, Y_t, \psi_t^1, \psi_t^2, \underline{c}(j)) - m_t^{1,j}(\psi_t^1) - m_t^{2,j}(\psi_t^2) - T(\cdot) \tag{31}
$$

$$
a_{t+1} \geq 0, \quad \forall t \tag{32}
$$

$$
a_{t+1} = 0, \quad \text{if} \quad B(\cdot) > 0. \tag{33}
$$

In equation (27), Y_t mimics the spousal benefit from Social Security which gives a married person the right to collect the higher of own benefit entitlement and half of the spouse's entitlement. In equation (28), $\bar{\bar{y}}_r$ represents survivorship benefits from Social Security in case of death of one of the spouses. The survivor has the right to collect the higher of own benefit entitlement and the deceased spouse's entitlement.

5.4.3 The individuals in couples: working age and retirement

We have to compute the joint value function of the couple to appropriately compute joint labor supply and savings under the married couples' available resources. However, when computing the value of getting married for a single person, the relevant object for that person is his or her the discounted present value of utility in the marriage. We thus compute this object for person of gender i who is married with a specific partner

$$\begin{aligned}\hat{W}^c(t, i, a_t, \epsilon_t^1, \epsilon_t^2, \bar{y}_t^1, \bar{y}_t^2) = v^i(\hat{c}_t(\cdot), \hat{l}_t^{i,j})+ \\ \beta(1-\zeta_{t+1})E_t\hat{W}^c(t+1, i, \hat{a}_{t+1}(\cdot), \epsilon_{t+1}^1, \epsilon_{t+1}^2, \bar{y}_{t+1}^1, \bar{y}_{t+1}^2)+ \\ \beta\zeta_{t+1}E_tW^s(t+1, i, \hat{a}_{t+1}(\cdot)/2, \epsilon_{t+1}^i, \bar{y}_{t+1}^i)\end{aligned} \tag{34}$$

where $\hat{c}_t(\cdot)$, $\hat{l}_t^{i,j}(\cdot)$, and $\hat{a}_{t+1}(\cdot)$ are, respectively, optimal consumption from the perspective of the couple, leisure, and saving for an individual of gender i in a couple with the given state variables.

During the retirement period, we have

$$\begin{aligned}\hat{R}^c(t, i, a_t, \psi_t^1, \psi_t^2, \bar{y}_r^1, \bar{y}_r^2) = v^i(\hat{c}_t(\cdot), L^{i,j}) + \beta s_t^{i,j}(\psi_t^i)s_t^{p,j}(\psi_t^p)E_t\hat{R}^c(t+1, i, \hat{a}_{t+1}(\cdot), \psi_{t+1}^1, \psi_{t+1}^2, \bar{y}_r^1, \bar{y}_r^2)+ \\ \beta s_t^{i,j}(\psi_t^i)(1-s_t^{p,j}(\psi_t^p))E_tR^s(t+1, i, \hat{a}_{t+1}(\cdot), \psi_{t+1}^i, \bar{\bar{y}}_r).\end{aligned} \tag{35}$$

where $s_t^{p,j}(\psi_t^p)$ is the survival probability of the partner of the person of gender i. This continuation utility is needed to compute Equation (34) during the last working period, when $\hat{W}^c(\cdot)$ is replaced by $\hat{R}^c(\cdot)$.

6 Estimation and calibration

We calibrate our model to match the data for the 1960s birth cohort by using a two-step strategy, as Gourinchas and Parker (2003) and De Nardi, French, and Jones

(2010 and 2016). Then, in a third step, as De Nardi, Pashchenko, and Porappakkarm (2017), we calibrate the parameter b, which affects the utility of being alive. It is important to note that this parameter does not change our decision rules and the data that we match and can thus be calibrated after the other parameters are calibrated. Nonetheless, it is necessary to calibrate it to properly evaluate welfare when life expectancy changes.

Calibrated parameters		Source
Preferences and returns		
r	Interest rate	4% De Nardi, French, and Jones (2016)
$\eta_t^{i,j}$	Equivalence scales	PSID
γ	Utility curvature parameter	2.5 see text
Government policy		
b^j, s^j, p^j	Income tax	Guner et al. (2012)
$SS(\bar{y}_r^i)$	Social Security benefit	See text
τ_t^{SS}	Social Security tax rate	See text
$\widetilde{y}_t$	Social Security cap	See text
$\underline{c}(1)$	Minimum consumption, singles	\$8,687, De Nardi et al. (2016)
$\underline{c}(2)$	Minimum consumption, couples	\$13,031, Social Security rules

Estimated processes		Source
Wages		
$e_t^{i,j}(\cdot)$	Endogenous age-efficiency profiles	PSID
ϵ_t^i	Wage shocks	PSID
Demographics		
$s_t^{i,j}(\psi_t^i)$	Survival probability	HRS
ζ_t	Divorce probability	PSID
$\nu_t(i)$	Probability of getting married	PSID
$\xi_t(\cdot)$	Matching probability	PSID
$\theta_t(\cdot)$	Partner's assets and earnings	PSID
$f^{0,5}(i,j,t)$	Number of children age 0-5	PSID
$f^{6,11}(i,j,t)$	Number of children age 6-11	PSID
Health shock		
$m_t^{i,j}(\psi_t^i)$	Medical expenses	HRS
$\pi_t^{i,j}(\psi_t^i)$	Transition matrix for health status	HRS

Table 6: First-step inputs summary

More specifically, in the third step, we choose b so that the value of statistical life (VSL) implied by our model is the middle of the range estimated by the empirical literature. The VSL is defined as the compensation that people require to bear an increase in their probability of death, expressed as "dollars per death." For example, suppose that people are willing to tolerate an additional fatality risk of $1/10,000$ during a given period for a compensation of $500 per person. Among 10,000 people there will be one death and it will cost the society 10,000 times $500 = $5 million, which is the implied VSL.

6.1 First-step calibration and estimation for the 1960s cohort

In the first step, we use the data to compute the initial distributions of our model's state variables and estimate or calibrate the parameters that can be identified outside our model. For instance, we estimate the probabilities of marriage, divorce, health transitions, and death, the number and age of children by maternal age and marital status, the wage processes, and medical expenses during retirement.

Our calibrated parameters are listed in Table 6. We set the interest rate r to 4% and the utility curvature parameter, γ, to 2.5. The equivalence scales are set to $\eta_t^{i,j} = (j + 0.7 * f_t^{i,j})^{0.7}$, as estimated by Citro and Michael (1995). The term $f_t^{i,j}$ is the average total number of children for single and married men and women by age.

We use the tax function for married and single people estimated by Guner et al. (2012). The retirement benefits at age 66 are calculated to mimic the Old Age and Survivor Insurance component of the Social Security system. The most recent paper estimating the consumption floor during retirement is the one estimated by De Nardi et al (2016) in a rich model of retirement with endogenous medical expenses. In their framework, they estimate a utility floor that corresponds to consuming $4,600 a year when healthy. However, they note that Medicaid recipients are guaranteed a minimum income of $6,670. As a compromise, we use $5,900 as our consumption floor for elderly singles, which is $8,687 in 2016 dollars, and the one for couples to be 1.5 the amount for singles, which is the statutory ratio between benefits of couples to singles.

In the subsections that follow, we describe the estimation of our wage functions, medical expenses, and survival probabilities. More details about all of our first-step inputs are in Appendix B.

6.1.1 Wage schedules

We estimate wage schedules using the PSID data and regressing the logarithm of potential wage for person k at age t,

$$\ln \overline{wag}e_{kt} = d_k + f^i(t) + \sum_{g=1}^{G} \beta_g D_g \ln(\bar{y}_{kt} + \delta_y) + u_{kt},$$

on a fixed effect d_k, a polynomial f in age t for each gender i, gender-cohort dummies D_g interacted with human capital $\bar{y}_{kt}$ and a shift parameter δ_y (to be able to take logs). Thus, we allow all coefficients to be gender-specific and for the coefficient on human capital to also depend on cohort.

We then regress the sum of the fixed effects and the residuals for each person on cohort and marital status dummies and their interactions, separately for each gender, and use the estimated effects for gender, marital status, and cohort as shifters for the wage profiles of each demographic group and cohort.

	Men	Women
Age overall	0.0015	0.0017***
Age = 30	0.0043	0.0012***
Age = 40	0.0039	0.0056***
Age = 50	-0.0018	0.0044***
Age = 60	-0.013**	-0.0025**
Married and born in 1960s vs 1940s	-0.642***	-0.395***
Single and born in 1960s vs 1940s	-0.660***	-0.381***
$ln(\bar{y}_t + \delta_y)$ and born in 1940s	0.256***	0.363***
$ln(\bar{y}_t + \delta_y)$ and born in 1960s	0.347***	0.413***

Table 7: Estimation results for potential wages, reported as percentage changes in potential wages due to 1-unit increases in the relevant variables (or changes from zero to one in case of dummy variables). In the case of $\bar{y}_t$ we report the elasticity. * p<0.10, ** p<0.05, *** p<0.01

Table 7 reports the results of our estimated equation for potential wages.[13] It shows that the effects of age on potential wages are small, especially for men.[14] The

[13]We report the percentage changes in potential wages by exponentiating the relevant marginal effect for each variable, β_x, and reporting it as $exp(\beta_x)-1$. In the case of $\bar{y}_t$, the estimated coefficient is an elasticity and we report it without any transformations.

[14]As we do not observe the complete profile for those born in 1960s, the shape of the age profile

largest age effect for men is at age 60, when their potential wage declines by 1.3%. Women's potential wages, instead, grow on average by half a percentage point until age 50 and decline only mildly around age 60.

In terms of the position of the age profile, the effect of being born in the 1960s cohort instead of the 1940s cohort is large and negative, especially for married and single men. Because these declines depend on one's human capital level, we discuss their magnitudes when illustrating the interaction between wages and human capital for the two cohorts in Figure 5. In contrast to this decline, however, returns to human capital went up for the 1960s cohort compared to the 1940s cohort, as our estimated elasticity of wages to human capital increases from 0.256 and 0.363 for the 1940s cohort to 0.347 and 0.413 for the 1960s one, respectively for men and women.

To better understand the implications of our estimates by cohort and sub-group, Figure 5 reports our estimated average wage profiles by age conditional on a fixed level of human capital during all of the working period. The human capital levels over which we condition are the 0^{th}, 25^{th}, 50^{th}, 75^{th}, and 99^{th} percentiles of the distributions of average accumulated earnings of men and women in our sample. They correspond to, respectively, \$0, \$30,100, \$41,300, \$51,600, and \$79,100 for men and to \$0, \$5,000, \$13,900, \$23,700, and \$55,900 for women (expressed in 2016 dollars). In these graphs, therefore, human capital is held fixed by age. The top graphs are for married people and the bottom ones refer to singles. The graphs on the left are for men and those on the right for women. The solid lines refer to the 1960s cohort, while the dashed ones to the 1940s cohort.

In sum, these graphs display wages as a function of age for single and married men and women in our two cohorts for five fixed levels of human capital. Hence, they illustrate the changes in the returns to human capital across cohorts and marital status for various human capital levels.

Focusing on married men with zero human capital (the lowest two lines in the top graph on the left), the effect of the lower position of the age profile for the 1960s cohort is apparent: married men entering the labor market receive an average potential hourly wage that is 3.5 dollars lower than that received by the same men in the 1940s cohort. At higher levels of human capital, the disadvantage is progressively reduced by the higher returns to human capital but is still not enough to counterbalance the drop in the level of all wages. Even at the highest level of human capital within the

is assumed to be the same across generations.

non-college graduate group the hourly wage for married men born in the 1960s is still 90 cents lower than those received by the same men in the 1960s. The bottom left panel displays the wages of single men and shows that their drops are even larger than those for married men at all human capital levels.

The right panels refer to the wages of women. The wages of married women (top panel) with zero human capital went down by about 0.9 dollars, a much smaller decrease across cohorts than that for men, both in absolute value and in percentage terms. As a consequence of the increased returns to human capital, at the median human capital level for women, their wage is 0.6 dollars lower, while it is actually higher for the high-human capital women in the 1960s than the 1940s cohort, by 0.3 dollars. The main difference between married and single women is that, from the 1940s to the 1960s, only married women in the top 1% of the human capital distribution experienced a wage increase, while single women in the top 15% of the human capital distribution experienced a wage increase.

In sum, we find that men and women in the 1960s cohort had a higher return to human capital but a lower cohort-and-gender-age wage profiles compared to those born in the 1940s. The latter drop was especially large for men. These changes imply that men and women with lower human capital had the largest drop, that wages dropped for men at all human capital levels, and that the wages of the highest human capital women increased. As a result of these changes in the wage structure and a larger increase in women's human capital (partly due to more years of education and partly to more labor market experience), average wages over the life cycle, shown in Figure 2, were higher for women and lower for men in the 1960s cohort.

6.1.2 Medical expenses

We estimate out-of-pocket medical expenses using the HRS data and regressing the logarithm of medical expenses for person k at age t,

$$\ln(m_{kt}) = X_{kt}^{m\prime}\beta^m + \alpha_k^m + u_{kt}^m$$

where the explanatory variables include a third-order polynomial in age fully interacted with gender, current health status and interactions between these variables.[15] The term α_k^m represents a fixed effect and takes into account all unmeasured fixed-

[15] We experimented adding marital status but it is not statistically different from zero.

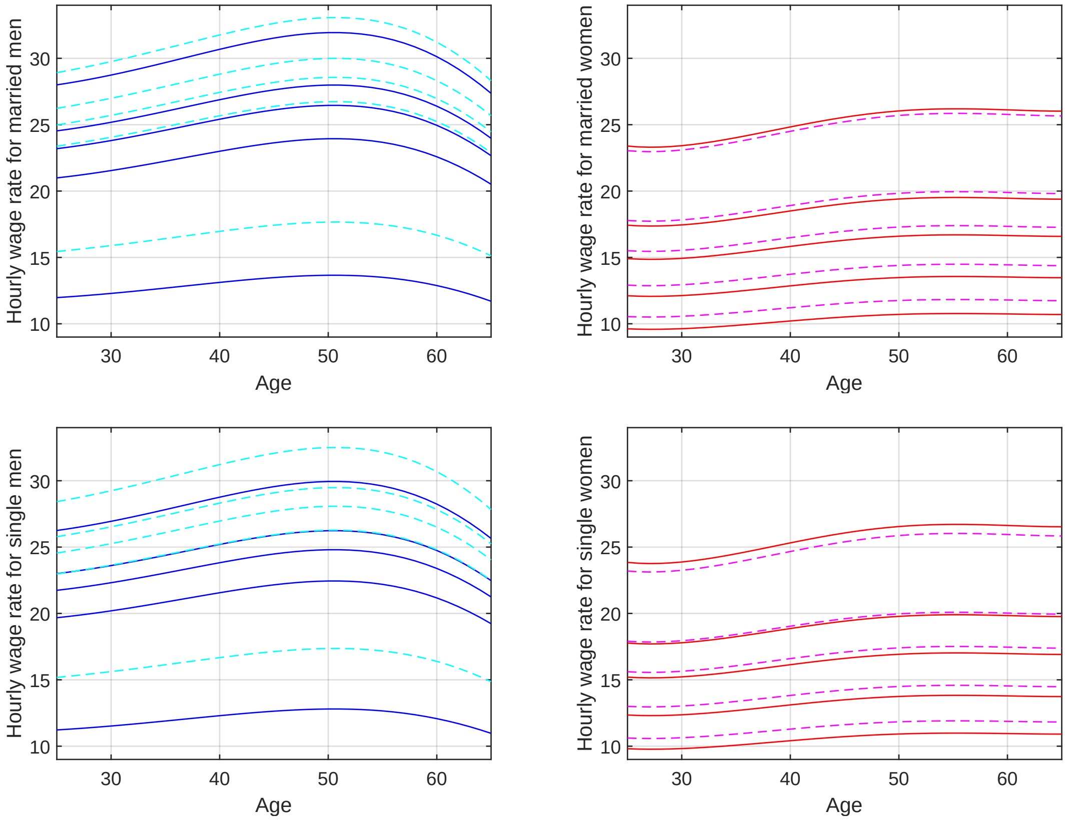


Figure 5: Wages as a function of human capital levels. Top graphs: married people. Bottom graphs: single people. Left graphs: men. Right graphs: women. The dashed lines refer to the cohort born in 1940 and the solid lines to that born in 1960, they are conditional on a fixed gender-specific level of human capital, measured at the 0^{th}, 25^{th}, 50^{th}, 75^{th} and 99^{th} percentiles of the distributions of average accumulated earnings in our sample.

over-time characteristics that may bias the age profile, such as differential mortality, as discussed in De Nardi, French and Jones (2010). We then regress the residuals from this equation on cohort, gender, and marital status dummies to compute the average effect for each group of interest. Hence, the profile of the logarithm of medical expenses is constant across cohorts up to a constant.

Table 8 reports the results from our estimates for medical expenses[16] and shows that after age 66 real medical expenses increase with age on average by 2.4 and 2.6 percent for men and women, respectively, with the growth for women being much

[16]We report the percentage changes of medical expenses by exponentiating the relevant marginal effect for each variable, β_x, and reporting it as $exp(\beta_x) - 1$.

	Men	Women
Age overall	0.024***	0.026***
At age 66	0.022***	0.019***
At age 76	0.017***	0.014***
At age 86	0.017***	0.027***
At age 96	0.023***	0.058***
Bad health	0.201***	0.209***
Married	0.327***	0.327***
Born in 1960s	0.486***	0.486***

Table 8: Estimation results for medical expenses for men and women, reported as percentage changes in medical expenses due to marginal increases in the relevant variables (or changes from zero to one in case of dummy variables). HRS data. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

faster than for men after age 76, reaching for example 5.8 percent at age 96. Finally, those born in the 1960s cohort face medical expenses that are 48.6 percent higher than those born in the 1940s cohort, even after conditioning on health status.

6.1.3 Life expectancy

As described in our model section, we allow mortality to depend on health, gender, marital status, and age, and we have health evolving over time, depending on previous health, age, gender, and marital status. We allow cohort effects to affect all of these dynamics and their initial conditions, both in our estimation of these inputs, and in our model.

More specifically, we model the probability of being alive at time t as a logit function

$$s_t = Prob(Alive_t = 1 \mid X_t^s) = \frac{\exp(X_t^{s\prime}\beta^s)}{1+\exp(X_t^{s\prime}\beta^s)}.$$

that we estimate using the HRS data. Among the explanatory variables, we include a third-order polynomial in age, gender, marital status, and health status in the previous period, as well as interactions between these variables and age, whenever they are statistically different from zero. We also include cohort dummies and use coefficients relative to the cohort of interest to adjust the constant accordingly.[17]

[17] We are thus assuming that the age profiles entering our estimated equation are the same across cohorts up to a constant. We then compute the mortality rate for the cohorts of interest using the appropriate cohort dummy.

To investigate the implications of the cohort effects that we estimate through these pathways, Table 9 reports the model-implied life expectancy at age 66 and their changes when we add, in turn, the changes in mortality, health dynamics, initial health at age 66 and initial fractions of married and single people that are driven by cohort effects on each of those components.

The first line of the table reports life expectancy using all of the inputs that we estimate for the 1960s cohort. Their implied life expectancy is very close to the one we have computed using the data and a much simpler regression for mortality and reported in Section 3.3. The second line changes the observed relationship between mortality and health and demographics from the one we estimate for the 1960s cohort to the one we estimate for the 1940s cohort. It shows that this change alone implies an increase of 0.8 and 0.7 years of life for men and women, respectively. In line three, we switch from the 1960s to the 1940s health dynamics and there is no noticeable change in life expectancy because the health dynamics are very similar. In line 4, we change the fraction of people who are in bad health at age 66, conditional on marital status, to that of the 1940s cohort. This change implies a further increase of 0.1 years of life expectancy for both men and women, indicating that a smaller part of the observed decrease in life expectancy at age 66 is captured by changing health conditions at age 66. The last line of the table not only changes initial health at age 66, but also allows for the fact that more people were married in the 1940s cohort compared to the 1960s cohort. This change in the fraction of married people at age 66 explains an additional change or 0.3 and 0.2 years of life for men and women, respectively.

	Men	Women
1960s Inputs	80.8	84.5
1940s Survival functions	81.6	85.3
1940s Survival and health dynamics	81.6	85.3
1940s Survival, health dynamics, initial health	81.7	85.4
1940s Survival, health dynamics, initial health, and marital status	82.0	85.6

Table 9: Life expectancy at age 66 for white and non-college educated men and women born in the 1940s and 1960s cohorts as we turn on various determinants of mortality. HRS data

Our decomposition thus shows that the biggest change in life expectancy in our framework comes from a change in the relationship between mortality and health dynamics after age 66, while a smaller one stems from a worsening of initial health status at age 66. Finally, the reduction in the fraction of married people also has a non-negligible effect on life expectancy of both men and women. In our experiments changing life expectancy, we do not change marital status at age 25 and we thus abstract from the effects of the small changes in life expectancy coming from that channel.

6.2 Second-step calibration

In the second step, we calibrate 19 model parameters (β, ω, $(\phi_0^{i,j}, \phi_1^{i,j}, \phi_2^{i,j})$, $(\tau_c^{0,5}, \tau_c^{6,11})$, $L^{i,j}$) so that our model mimics the observed life-cycle patterns of labor market participation, hours worked conditional on working, and savings for married and single men and women that we report in Figure 4.

Table 10 presents our calibrated preference parameters for the 1960s cohort. Our calibrated discount factor is 0.981 and our calibrated weight on consumption is 0.416.

We normalize available time for single men to 5840 hours a year (112.3 hours a week) and calibrate available time for single women and married women and men. Our calibration implies that single women have the same time endowment as single men (112 hours a week). The corresponding time endowments for married men and women are, respectively, 105 and 88 hours. This implies that people in the latter two groups spend 7 and 24 hours a week, respectively, in non-market activities such as running households, raising children, and taking care of aging parents. Our estimates of non-market work time are similar to those reported by Aguiar and Hurst (2007) and by Dotsey, Li, and Yang (2014).

Our estimates for the 1960s cohort imply that the per-child child care cost of having a child age 0-5 and 6-11 are, respectively, 35% and 3.0% of a woman's earnings. In the PSID data, child care costs are not broken down by age of the child, but per-child child care costs (for all children in the age range 0-11) of a married woman are 33% and 19% of her earnings at ages 25 and 30, respectively. Computing our model's implications, we find that per-child child care costs (for all children in the age range 0-11) of a married woman are 30% and 23% of her earnings, respectively, at ages 25 and 30. Thus, our model infers child care costs that are similar to those in the PSID data.

Calibrated parameters	1960s cohort
β: Discount factor	0.981
ω: Consumption weight	0.416
$L^{2,1}$: Time endowment (weekly hours), single women	112
$L^{1,2}$: Time endowment (weekly hours), married men	105
$L^{2,2}$: Time endowment (weekly hours), married women	88
$\tau_c^{0,5}$: Prop. child care cost for children age 0-5	35%
$\tau_c^{6,11}$: Prop. child care cost for children age 6-11	3.0%
$\Phi_t^{i,j}$: Participation cost	Fig. 6

Table 10: Second step calibrated model parameters

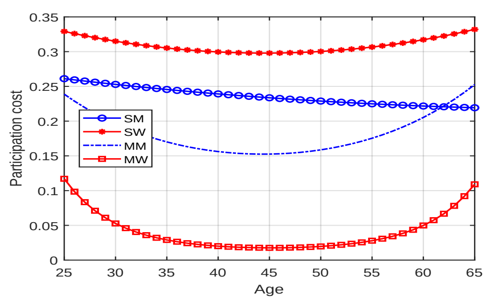


Figure 6: Calibrated labor participation costs, expressed as fraction of the time endowment of a single men. SM: single men; SW: single women; MM: married men; MW: married women. Model estimates

Figure 6 shows the calibrated profiles of labor participation costs by age, expressed as fraction of the time endowment of a single men. Participation costs are relatively high when young, decrease in middle age, and with the exception of single men, increase after 45.

6.3 Third-step calibration

To match the VSL, we proceed as follows. Because in our model we do not have mortality until age 66, we review the value of statistical life estimated for older people in previous empirical work. Within this literature, O'Brien (2013) estimates the value of statistical life by examining consumer automobile purchases by individuals up to 85 years old. He finds that the VSL is respectively \$8 million for the age 65-74 age group and $7 million for the 75-85 age group (expressed in year 2009 dollars). Alberini et al. (2004), instead, use contingent valuation surveys, which elicited respondents' willingness to pay for reductions of mortality risk of different magnitudes, and find values between $1 and $5 million for the 40 to 75 age group (expressed in year 2000 dollars). Thus, the range from these two papers, expressed in year 2016 dollars (the base year that we use in this paper) is between $1 and $9 million. Then, we choose $b = 0.009$ so that when we increase mortality after retirement and compute a compensation that makes them indifferent between this counterfactual case and our benchmark mortality, we obtain an average VSL at age 66 of $5 million.

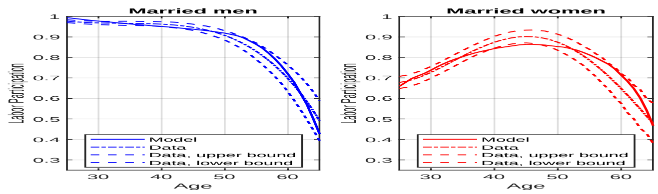
6.4 Model fit

Figures 7 and 8 report our model-implied moments, as well as the moments and 95% confidence intervals from the PSID data for our 1960s cohort. They show that our parsimoniously parameterized model (19 parameters and 448 targets) fits the data well and reproduces the important patterns of participation, hours conditional on participation, and asset accumulation for all of four demographic groups.

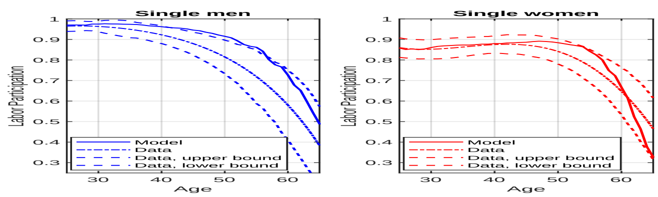
7 The effects of changing wages, medical expenses, and life expectancy

We now turn to evaluating the effects of the changes in wages, medical expenses, and life expectancy that we have documented. Because we want to isolate the effects of these changes on the 1960s cohort (while keeping everything else constant for this cohort), we only replace these three sets of inputs with those experienced by 1940s cohort, first one at a time, and then all at the same time. In doing so, we assume that, as of age 25, the 1960s cohort have rational expectations about all of the stochastic processes that they face over the rest of their lives, including when we switch some of them to their 1940s counterparts.

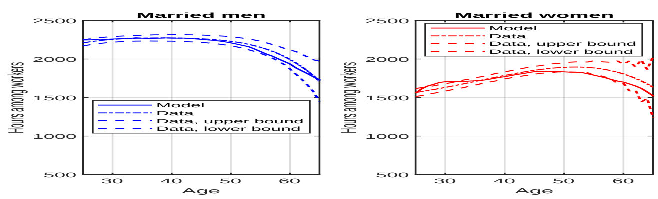
We start by studying the implications of these changes for labor participation, hours worked by workers, and savings for single and married men and women. Then, to evaluate welfare, we compute a one-time asset compensation to be given upon entering the model, that is at age 25, that makes household endowed with a given set



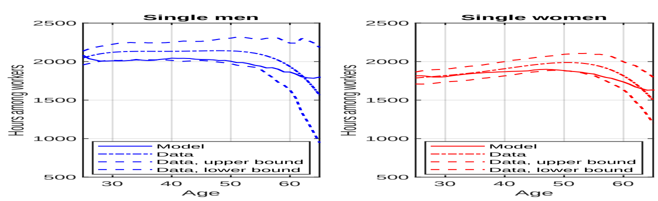
(a) Participation, couples



(b) Participation, singles



(c) Hours for workers, couples

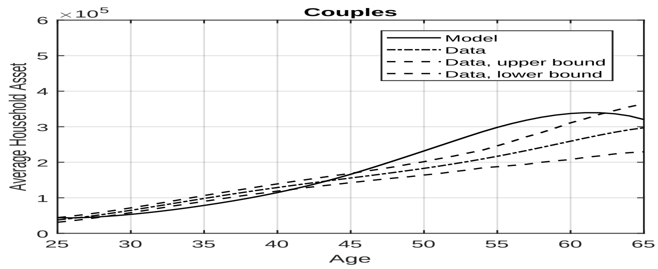


(d) Hours for workers, singles

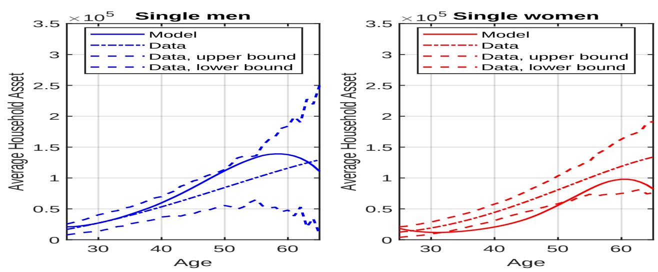
Figure 7: Model fit for participation (top graphs) and hours (bottom graphs) and 95% confidence intervals from the PSID data

of state variables, indifferent between facing the 1960s input and the 1940s input.[18] Finally, we compute the fraction of people that have lost or gained as a result of these changes and report the average welfare loss experienced by single men, single women, and married couples expressed as the average compensation that makes each of these groups indifferent between the two set of inputs.

[18]These computations are performed for each household while keeping fixed the assets of their potential future partners to those that we estimate in the data.



(a) Assets, couples



(b) Assets, singles

Figure 8: Model fit for assets and 95% confidence intervals from the PSID data

7.1 Changing wages

Figure 9 compares the participation, hours worked by workers, and savings for the 1960s cohort under their own wage schedule and under the wage schedule of the 1940s cohort. It shows that, according to our model, all of these economic outcomes would have been rather different under the 1940s wage schedule.

The largest effects occurred for married couples, with many more married women participating and working more hours under the 1960s wage schedule, while their husbands dropped out of the labor force at younger ages. At age 25, for instance, the participation of married women was 8 percentage points higher. Married men's participation started dropping faster after age 30 and was four percentage points lower than under the 1940s wage schedule at age 55. Hours worked by young married women were about 100 hours a year higher, while hours worked by young married men were only slightly higher. These changes were due to much lower wages for men, in conjunction with increasing returns to human capital. The latter, in particular, increased the returns to working when young.

Single people were affected too. They were experiencing lower wages, for the most part, and in the case of single women, they were also expecting to get married with lower-wage husbands. This negative wealth effect makes them invest more in their own human capital, work harder when young, and receive higher wages because of higher human capital accumulation. Single men reacted little to these changes by marginally reducing their participation, increasing hours worked while young, and reducing them after age 50.

As a result of the changing wage schedule and endogenous labor market decisions, average discounted lifetime income decreased by $115,000 (10%) for single men and $108,000 (9%) for married men, but increased by $28,000 (5%) for single women and $36,000 (7%) for married women. As households experienced the large negative wealth effect coming from lower wages and earnings, retirement savings were much lower. Assets at age 66 dropped by 21% for single men, 1.1% for single women, and 6.1% for couples, respectively.

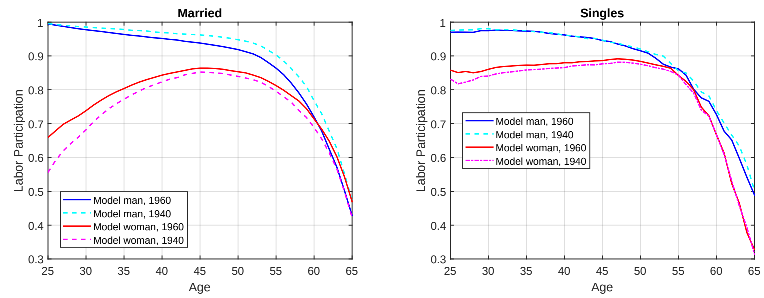
Compared with 1940 wage schedule	Single men	Single women	Couples
Men only	6.8%	2.9%	4.0%
Men and women	7.3%	3.4%	4.5%
No marriage and divorce economy			
Men only	11.1%	0.0%	4.3%
Men and women	11.1%	0.7%	4.9%

Table 11: Welfare compensations for the 1960s cohort for facing the 1960s wages schedule instead of 1940s wage schedule, computed as one-time asset compensation at age 25 and expressed as a fraction of the discounted present value of one's income. Top panel, our benchmark economy, bottom panel, an economy without marriage and divorce after age 25

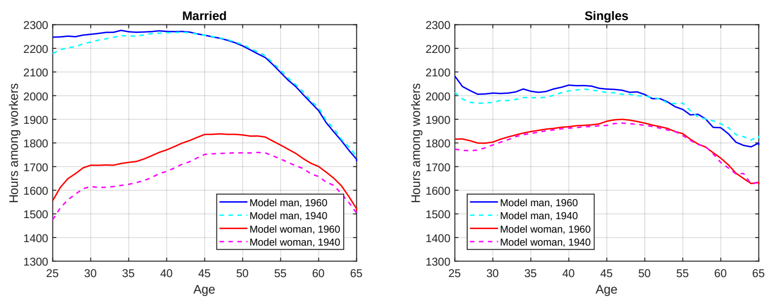
We now turn to evaluating how much worse (or better off) were people under the 1960s rather than the 1940s wage schedule.[19] We start by studying the effects of the wage changes for men only. In this case everyone loses, and the one-time asset compensation that we should give to 25 year old to make them indifferent between the two wage schedules for men, are $68,300 for single men, $17,800 for single women, and $64,800 for couples. The first line of Table 11 reports these compensations as a fraction of the present value of lifetime income for each group. They amount to 6.8%, 2.9%, and 4.0% for single men, single women, and couples, respectively.

We then turn to the welfare effects of having the wages of both men and women

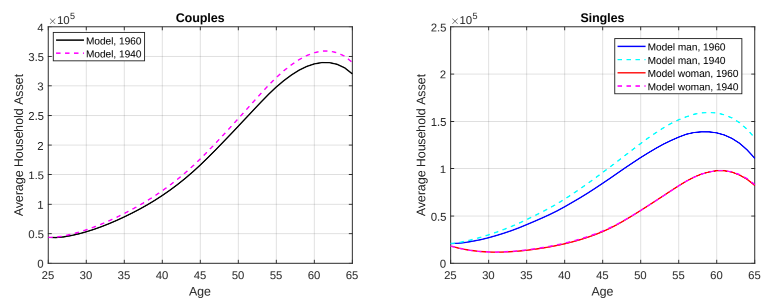
[19] When changing the wage schedule, we keep everything else (including initial conditions and prospective spouses) fixed at the levels experienced by the 1960s cohort.



(a) Participation



(b) Hours for workers



(c) Assets

Figure 9: Model outcomes with 1960s and 1940s wage schedule

set to the 1960s instead of 1940s wage schedules. Again, virtually everyone loses as a result. The one-time asset compensation that we should give to the 25 year old to make them indifferent between the 1940s and the 1960s wages, are $72,900 for single men, $20,400 for single women, and $73,600 for couples. The second line of Table 11 reports these compensations as a fraction of lifetime income for each group. They amount to 7.3%, 3.4%, and 4.5% for single men, single women, and couples, respectively. Thus, everyone loses and the welfare losses are big both in absolute value and when compared with the discounted value of lifetime income in each group.

To isolate the welfare effects coming from marriage and divorce dynamics, we also compute an economy in which there is no marriage and divorce after age 25. In it, a 25-year-old single person stays single forever, and a 25-year-old married couple stays married forever.[20] The last two lines of Table 11 report the welfare losses of the losers as a fraction of the present discounted value of lifetime income for each group.

When men's wages drop in this economy, all men and couples lose as a result. The one-time asset compensation that we should give to the 25 year old to make them indifferent between the 1940s and the 1960s men's wages when there is no marriage and divorce, are \$98,800 for single men, \$0 for single women, and \$72,000 for couples, respectively. When the wages of both men and women change, all single men and almost all couples lose, while 38% of single women gain. The single women who gain are the high-human capital one end up with higher wages. The welfare compensations for those who lose when all wages change are, respectively, \$98,800 for single men, \$5,400 for single women, and \$80,500 for couples, respectively. The average welfare gain among the 38% of single women who gain is \$6,400.

Compared with our benchmark economy, in an economy without marital dynamics after age 25 single men experience a larger welfare loss due to their much lower wages and their inability to benefit from a future working spouse. In contrast, 38% of single women gain when their wage goes up (those with high human capital), while the other single women experience a smaller welfare loss because, while their wage goes down, they no longer get married with much lower-wages husband and thus do not work as hard to help support their family.

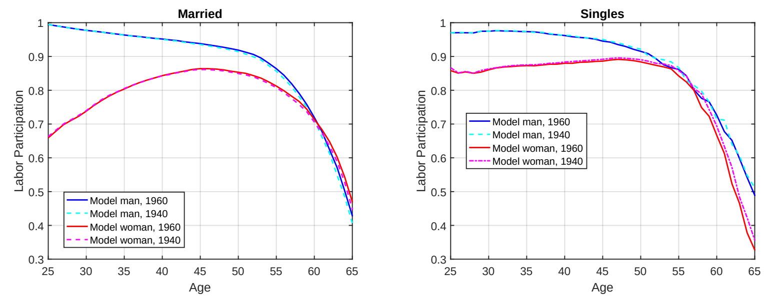
7.2 Changes in medical expenses

We now turn to studying the effects of replacing the out-of-pocket medical expenses faced by the 1960s cohort with those faced by the 1940s cohort. The present discounted value of medical expenses at age 25 for the 1960s cohort went up by \$5,000, \$7,000, and \$12,300 for single men, single women, and couples, respectively, compared with the 1940s cohort. This corresponds to a 76% increase for single men, single women, and couples.

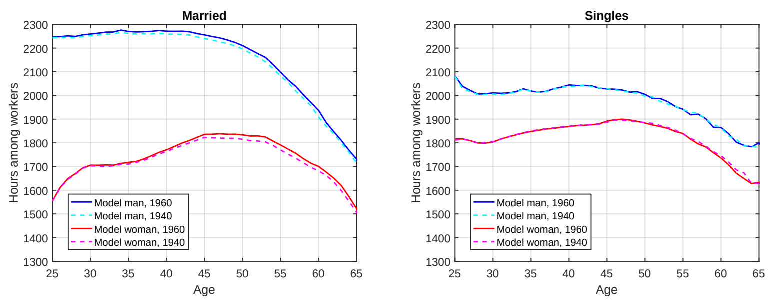
The main noticeable effects of these changes are that hours worked by married women in the 1960s cohort under the 1960s inputs are slightly higher after age 30 while those of single women, who are poorer and rely on the consumption floor more,

[20]To isolate the effect of wage changes, we eliminate marriage and divorce dynamics after age 25 both from the baseline and the counterfactual economies when performing these welfare comparisons.

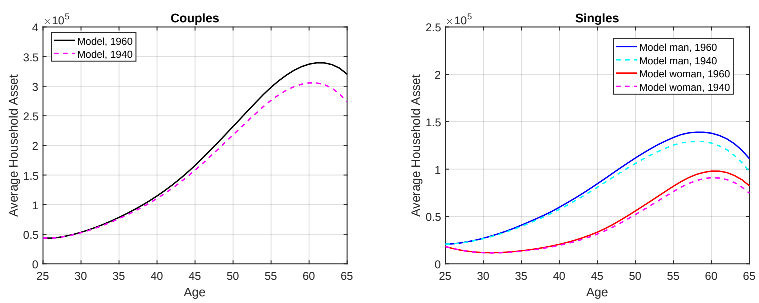
go down after age 55. Also, savings at age 66 were 14%, 11%, and 16% higher for single men, single women and couples in the 1960s cohort than they would have been under the lower medical expenses experienced by the 1940s cohort.



(a) Participation



(b) Hours for workers



(c) Assets

Figure 10: Model outcomes with 1960s and 1940s medical expenses

Turning to our welfare computations, the resulting one-time asset compensation that we should give to 25 year old to make them indifferent between the 1940s and the 1960s medical expenses, are \$14,000 for single men, \$6,000 for single women, and \$14,900 for couples. These numbers correspond, respectively, to 1.4%, 1.0%, and 0.9% of the present discounted value of their lifetime income. Despite the similar change in medical expenses for single men and women, the compensation is smaller for single women because they are poorer and rely on the consumption floor more. Thus, to the extent that they are at the consumption floor, how large their medical expenses are, is not very important to then.

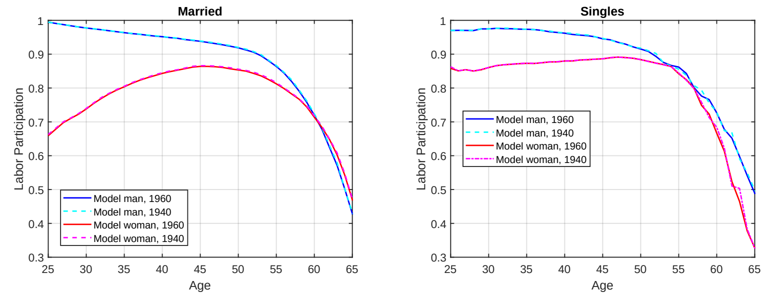
7.3 Changes in life expectancy

We endow the 1960s cohort with the mortality, that is, health initial, health transition, and survival function, and thus life expectancy, of the 1940s cohort. Because we estimate out-of-pocket medical expenses as a function of age, gender, and health, changing a cohort's health and survival dynamics also changes its medical expenses. In fact, moving from the 1940s to the 1960s health and survival dynamics not only lowers survival, but, because people die off faster, also decreases the present discounted value of medical expenses at age 25 by \$600 (4.5%) for single men, by \$670 (4.0%) for single women and \$1,300 (4.3%) for couples. Thus, both life expectancy and medical expenses go down as a result of these changes across cohorts.

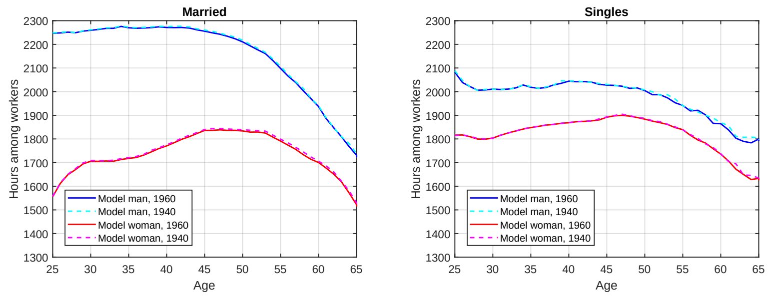
Figure 11 compares the participation and hours of married and single men and women under the two scenarios. It shows that participation and hours would have been very similar under the two scenarios but that retirement savings would have been 6.4%, 6.0%, and 4.1% higher for single men, single women and couples, respectively higher at retirement time under the 1940s health and survival dynamics. Thus savings go down, as one might expect, because of the shorter time period over which people expect to have to finance retirement consumption and decreased medical spending. Given that, in contrast, the life expectancy of college educated (and their medical expenses) went up over time, this change contributes to increasing the gap in their retirement savings and thus wealth inequality across these education groups.

Hall and Jones (2007) and De Nardi, Paschenko, and Porappakkarm (2017) find that changes in life expectancy can have large effects on welfare. One mitigating factor in our framework is that this lower life expectancy occurred together with lower medical expenses. In our model medical expenses are a shock reducing available resources; thus, reducing them increases welfare. This counters the loss in welfare due to a shorter lifespan.

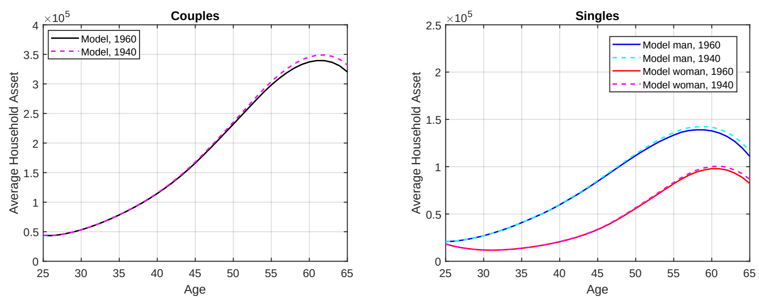
We find the welfare cost due to a shorter life expectancy dominates the welfare gain from reduced medical expenses and that all single men and women and married couples lose welfare as a result. More specifically, the one-time asset compensation



(a) Participation



(b) Hours for workers



(c) Assets

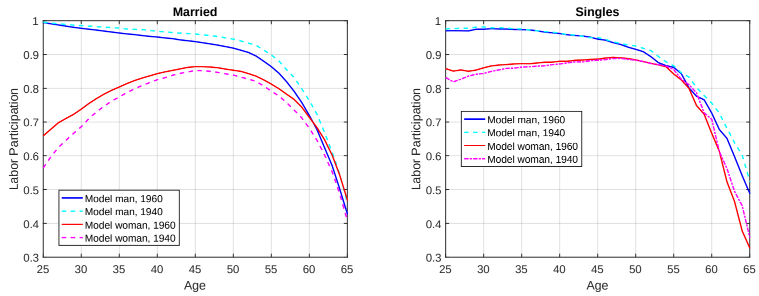
Figure 11: Model outcomes with 1960s and 1940s life expectancy

that we have to give 25 year-old households to make them indifferent between the 1940s and the 1960s health and survival dynamics are \$32,000 for single men, \$15,000 for single women, and \$36,000 for couples. These numbers correspond, respectively, to 3.2%, 2.4%, 2.2% of the present discounted value of their lifetime income.

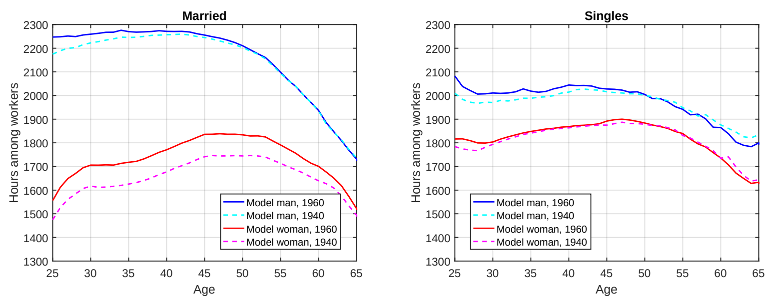
7.4 All three changes together

As we have seen from our previous three decomposition exercises, changes in the wage schedule had the largest effects on participation, hours, savings, and welfare.

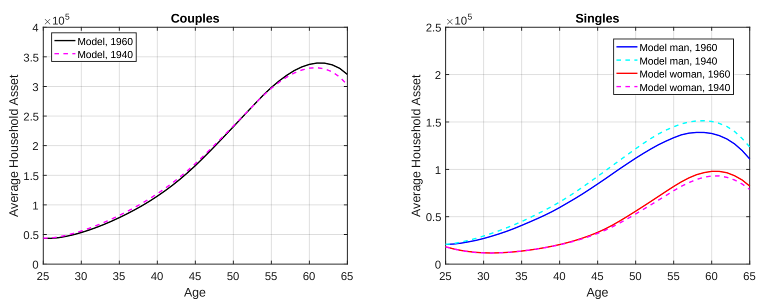
The other two changes that we consider, the decrease in life expectancy and increase in expected out-of-pocket medical costs mostly affect retirement savings and partly offset each other. They still have, however, very sizeable welfare costs.



(a) Participation



(b) Hours for workers



(c) Assets

Figure 12: Model outcomes with all changes we consider

The effects of all of these changes together on the labor market participation imply large increases in participation of both married and single women, noticeable decreases in the participation of married men after age 40, and almost no changes in the participation of single men. Hours worked by married men and women changed in opposite directions, while the hours of single men and women displayed some increases earlier on in their working period. On net, these changes depressed the retirement savings of single men, while leaving those of couples and single women roughly unchanged.

All changes considered	SM	SW	MM	MW	All
Average participation change	-1.15	-0.23	-2.42	3.12	0.12
Average hours change	0.48	0.21	0.70	4.82	1.95

Table 12: Changes in participation (in percentage points) and hours (in percentages) for the 1960s cohort when facing the 1960s inputs (wages schedules, medical expenses, and life expectancies) compared to the 1940s ones. SM = single men, SW = Single women, MM = married men, MW = married women, All=everyone

Table 12 compares outcomes for 1960s cohort. Under the 1960s inputs (wage schedules, medical expenses, and life expectancies), the participation rates of married women over their working period were 3.12 percentage points higher than under the 1940s wage schedule, while that of married men were 2.42 percentage points lower. Overall, participation was only 0.12 percentage points higher due to offsetting changes across groups. Hours worked conditional on participation, instead, were higher for all groups and especially for married women, resulting in an additional 1.95% of hours worked over the life cycle for this group.

Compared with 1940 inputs	SM	SW	Couples
Wages	7.3%	3.4%	4.5%
Medical expenses	1.4%	1.0%	0.9%
Life expectancy	3.2%	2.4%	2.2%
All changes considered	12.5%	7.2%	8.1%
All changes considered, no marriage and divorce	15.2%	2.1%	8.7%

Table 13: Welfare compensations for the 1960s cohort for facing the 1960s wages schedule, medical expenses, and life expectancies instead of the 1940s ones, computed as one-time asset compensation at age 25 and expressed as a fraction of the discounted present value of one's income. SM = single men, SW = Single women

As a result of all three changes together, the present discounted value of income went down, by 9.9%, 4.6%, and 4.0% for single men, single women, and couples and the one-time welfare loss experience by people in the 1960s cohort amounts to $126,000 for single men, $44,000 for single women, and $132,000 for couples. The fourth line of Table 13 reports that these numbers expressed as a fraction of their average discounted present value of earnings are, 12.5%, 7.2%, and 8.1%, respectively. Thus, the resulting welfare loss due to the changes between the 1940s and the 1960s birth cohort are very large.

Table 13 summarizes key information about the welfare losses and their sources. The first columns shows, for instance, that 58.4% of the total welfare loss that we consider for single men comes from wage changes and 25.6% comes from their decrease in life expectancy. The second column shows that, for single women, 47.2% of the welfare loss for single women comes from wage changes (their own and those of their prospective husbands), and that 33.3% of it comes from decreased life expectancy. The last column refers to couples and shows 55.6% of the welfare loss for couples comes from wage changes and 25.3% comes decreased life expectancy.

The last line of the table considers all changes together in an economy without marital dynamics after age 25 and finds that the welfare loss of single men is higher and that of single women lower when they have no expectations of getting married in the future. When couples no longer divorce, their welfare loss is higher because the wife works harder and no longer gets divorced.

8 Conclusions and directions for future research

Of the three changes that we consider, that is wages, out-of-pocket medical expenses during retirement, and life expectancy, we find that the observed changes in the wage schedule had by far the largest effect on the labor supply of men and women born in the 1960s cohort. Specifically, it depressed the labor supply of men and increased that of women, especially in married couples. The decrease in life expectancy mainly reduced retirement savings, while the expected increase in out-of-pocket medical expenses increased them. On net, these two changes taken together had overall modest effects on all of the outcomes that we consider, including savings.

We also find that the combined effect of the changes have large welfare costs. In fact, the one-time asset compensation required to make 25-year-old households indifferent between the 1940s and 1960s health and survival dynamics, medical expenses, and wages are \$126,000 for single men, \$44,000 for single women, and \$132,000 for couples. The corresponding numbers expressed as a fraction of their average discounted present value of earnings are, 12.5%, 7.2%, and 8.0%, respectively. Lower wages explain 47-58% of these losses, shorter life expectancies explain 26-34%, and higher medical expenses account for the rest.

Other interesting changes took place for the same cohorts, including in the number of children, marriage and divorce patterns, assortative mating, child care costs, initial conditions at age 25, and time spent in home production and raising children. Our paper suggests that studying the opportunities and outcomes of people in different cohorts and across different groups is a topic worthy of investigation, including from a macroeconomic standpoint.

We focus on the population of white and non-college educated Americans to bring to bear a large and relatively homogenous population to our structural model and study its implications. However, white non-college educated are hardly the only disadvantaged population losing ground over time in the U.S. Derek Neal (2006) extensively documents that while black-white skill gaps diminished over most of the 20th century, important measures of these gaps have not dropped since the late 1980s. A significant literature also documents a dramatic decline in employment rates and a lack of wage growth among less-skilled black men over the past four decades or more (see Neal and Rick (2016) and Bayer and Charles (2018)). However, this literature does not employ structural models that facilitate analyses of trends in aggregate welfare or overall inequality.

While employment rates for less-skilled black and white men were falling, incarceration rates were rising. However, these rising incarceration rates did not reflect rising levels of criminal activity. Neal and Rick (2014 and 2016) show that the prison boom, which began around 1980, was primarily the result of policy changes that increased the severity of punishment for all types of criminal offenders. These changes more than doubled the incarceration rates of young black and white men. As a result, a much larger fraction of the current generation of less-educated Americans have spent time in prison, and only the future can reveal the total impact of these prison experiences on their lifetime earnings and consumption (Holzer 2009). Thinking about crime and related policies and their effects in the context of structural models is an important extension to better understand the economic outcomes of disadvantaged populations.

Fella and Gallipoli (2014) estimate a rich life cycle model with endogenous education and crime choices to study the effects of two large-scale policy interventions aimed at reducing crime by the same amount: subsidizing high school education and increasing the length of prison sentences. They find that increases in high school graduation rates entails large efficiency and welfare gains which are absent if the same crime reduction is achieved by increasing the length of sentences. Intuitively, the efficiency gains of the subsidy come from its effect on the education composition of the labor force. No such effect is present in the case of a longer prison term.

Another important observation is that low income individuals are both more likely to develop a severe work-limiting disability and more likely to apply for disability insurance when they are not severely disabled. Low and Pistaferri (2015) find that by age 60, the low educated are 2.5 times more likely to be disability insurance claimants than the high educated (17% vs. 7%). In addition, a large increase in disability enrollment has been taking place over time, going from 2.2 percent in the late 1970s to 3.5 percent in the years immediately preceding the 2007-2009 recession and 4.4 percent in 2013 (Libman, 2015). Michaud and Wiczer (2018) study the increase in disability claims of men over time in the context of a structural model and evaluate the importance of changing macroeconomic conditions in driving it. They find the secular deterioration of economic conditions concentrated in populations with high health risks accounts for a third of the increase in aggregate disability claims for men. These changes occurred in conjunction with the rise of participation (and disability claiming) of women. Gallipoli and Turner (2011) show that marriage interacts with health and disability shocks in an important way and that single workers' labor supply responses to disability shocks are larger and more persistent than those of married workers. Thus, enriching our framework to allow for health shocks during the working period and disability insurance is an important area of research to better understand the changing opportunities and outcomes of the most disadvantaged groups.

References

[1] Aguiar, Mark and Erik Hurst. 2007. "Measuring Trends in Leisure: The Allocation of Time Over Five Decades." *Quarterly Journal of Economics* 122(3), 969-1006.

[2] Alberini, Anna, Maureen Cropper, and Alan Krupnick and Natalie Simon. 2004. "Does the Value of Statistical Life Vary with Age and Health Status? Evidence from the US and Canada." *Journal of Environmental Economics and Management* 48, 769-792.

[3] Arias, Elizabeth, Brian L. Rostron, and Betzaida Tejada-Vera. 2010. "United States Life Tables 2005." *National Vital Statistics Reports*, Volume 58, Number 10.

[4] Bayer, Patrick and Kerwin Kofi Charles. 2018. "Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men Since 1940s." *Quarterly Journal of Economics* 133(3), 1459-1501.

[5] Blundell, Richard, Monica Costa Dias, Costas Meghir, and Jonathan Shaw. 2016a. "Female Labor Supply, Human Capital, and Welfare Reform." *Econometrica* 84(5), 1705-1753.

[6] Blundell, Richard, Luigi Pistaferri and Itay Saporta-Eksten. 2016b. "Consumption Inequality and Family Labor Supply." *American Economic Review* 106(2), 387-435.

[7] Blundell, Richard, Jack Britton, Monica Costa Dias and Eric French. 2017. "The impact of health on labour supply near retirement." IFS Working Paper, W17/18.

[8] Borella, Margherita, Mariacristina De Nardi, and Eric French. 2018. "Who Receives Medicaid in Old Age? Rules and reality." *Fiscal Studies*, 39(1), 65-93.

[9] Borella, Margherita, Mariacristina De Nardi, and Fang Yang. 2017. "The Effects of Marriage-Related Taxes and Social Security Benefits." NBER working paper No. 23972.

[10] Case, Anne and Angus Deaton. 2015. "Rising Morbidity and Mortality in Midlife among White Non-Hispanic Americans in the 21st Century." *Proceedings of the National Academy of Sciences* 112 (49): 15078-15083.

[11] Case, Anne and Angus Deaton. 2017. "Mortality and Morbidity in the 21st Century." *Brookings Papers on Economic Activity.*

[12] Citro, Constance Forbes and Robert T. Michael, 1995, Measuring poverty: a new approach, National Academy Press.

[13] Costa Dias, Monica, Francesca Parodi, and Robert Joyce. 2018. "The Gender Pay Gap in the UK: Children and Experience in Work." IFS Working Paper 18/02.

[14] De Nardi, Mariacristina, Eric French, and John Jones. 2010. "Why do the Elderly Save? The Role of Medical Expenses." *Journal of Political Economy*, 118, 39-75.

[15] De Nardi, Mariacristina, Eric French, and John Jones. 2016. "Medicaid Insurance in Old Age." *The American Economic Review*, 106(11), 3480-3520.

[16] De Nardi, Mariacristina, Svetlana Pashchenko, and Ponpoje Porapakkarm. 2017. "The Lifetime Costs of Bad Health." NBER working paper 23963.

[17] Dotsey, Michael, Wenli Li, and Fang Yang, 2014, "Consumption and Time Use over the Life Cycle." *International Economic Review*, 55 (3), 665-692.

[18] Dustmann, Christian and María Engracia Rochina-Barrachina. 2007. "Selection Correction in Panel Data Models: An Application to the Estimation of Females' Wage Equations." *Econometrics Journal* 10(2), 263-293.

[19] Eckstein, Zvi and Osnat Liftshitz. 2011. "Dynamic Female Labor Supply." *Econometrica* 79(6), 1675-1726.

[20] Eckstein, Zvi, and Michael Keane, and Osnat Liftshitz. 2019. "Sources of Change in the Life-Cycle Decisions of American Men and Woman: 1962-2014. " *Econometrica*, 87 (1), 217-253.

[21] Fella Giulio and Giovanni Gallipoli, 2014, "Education and Crime over the Life Cycle." *The Review of Economic Studies*, 81(4), 1484-1517.

[22] Fernandez Raquel and Joyce C. Wong, 2014, "Divorce Risk, Wages and Working Wives: A Quantitative Life-Cycle Analysis of Female Labour Force Participation. " *The Economic Journal*, 124(576), 319-358.

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

[24] Gallipoli, Giovanni and Turner Laura. 2011. "Household Responses to Individual Shocks: Disability and Labor Supply."Mimeo

[25] Gouveia, Miguel and Robert P. Strauss. 1994. "Effective Federal Individual Income Tax Functions: An Exploratory Empirical Analysis." *National Tax Journal*, 47 (2), 317–39.

[26] Gourinchas, Pierre-Olivier, and Jonathan A. Parker. 2002. "Consumption over the Life Cycle." *Econometrica* 70(1), 47–89.

[27] Guner, Nezih, Remzi Kaygusuz, and Gustavo Ventura. 2012. "Income Taxation of U.S. Households: Basic Facts." CEPR Discussion Paper 9078.

[28] Guvenen, Fatih, Greg Kaplan, Jae Song, and Weidner. 2017. "Lifetime Incomes in the United States over Six Decades." NBER working paper 23371.

[29] Hall, Robert E and Jones, Charles I. 2007. "The value of life and the rise in health spending." *The Quarterly Journal of Economics*, 122, 39-72.

[30] Heckman, James. 1979. "Sample Selection Bias as a Specification Error." *Econometrica* 47(1), 153-61.

[31] Holzer, Harry J. 2009. "Collateral Costs: Effect of Incarceration on Employment and Earnings Among Young Workers." In *Do Prisons Make Us Safer?, edited by Steven Raphael and Michael A Stoll*, 239–266. New York: Russell Sage Foundation.

[32] Keane, Michael and Richard Rogerson. 2012. "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom." *Journal of Economic Literature* 50:2, 464-476.

[33] Libman, Jeffrey. 2015. "Understanding the Increase in Disability Insurance Benefit Receipt in the United States" *Journal of Economic Perspectives* 29:2, 123-150.

[34] Low, Hamish and Luigi Pistaferri. 2015. "Disability Insurance and the Dynamics of the Incentive-Insurance Tradeoff." *American Economic Review* 105:10, 2986-3029.

[35] Michaud, Amanda and David Wiczer. 2018. "The Disability Option: Labor Market Dynamics with Economic and Health Risks." Mimeo.

[36] Neal, Derek. 2011. "Why Has Black–White Skill Convergence Stopped?" *Handbook of the Economics of Education*, vol. 1, 511-576.

[37] Neal, Derek and Armin Rick. 2016. "The Prison Boom and Sentencing Policy." *Journal of Legal Studies*, 45, no. 1, 1-41.

[38] Neal, Derek and Armin Rick. 2014. "The Prison Boom and The Lack of Black Progress After Smith and Welch." NBER working paper 20283.

[39] O'Brien, James. 2013. "The Age-Adjusted Value of a Statistical Life: Evidence from Vehicle Choice." Mimeo. Georgetown.

[40] Roys, Nicolas and Christopher Taber. 2017. "Skills Prices, Occupations and Changes in the Wage Structure for Low Skilled Men." Mimeo.

[41] Semykina Anastasia and Jeffrey Wooldridge. 2005. "Estimating panel data models in the presence of endogeneity and selection." *Journal of Econometrics*, 157(2), 375-380.

[42] Wooldridge, Jeffrey. 1995. "Selection corrections for panel data models under conditional mean independence assumptions." *Journal of Econometrics*, 68(1), 115-132.

Appendix A. Data and methodology

We use the Panel Study of Income Dynamics (PSID) to estimate the wage process, the marriage and divorce probabilities, the initial distribution of couples and singles over state variables, and the sample moments that we match using our structural model.

The PSID is a longitudinal study of a representative sample of the U.S. population. The original 1968 PSID sample was drawn from a nationally representative sample of 2,930 families designed by the Survey Research Center at the University of Michigan (the "SRC sample") and from an over-sample of 1,872 low-income families from the Survey of Economic Opportunity (the "SEO sample"). Individuals and their descendants from both samples have been followed over time.

We study the two cohorts born in 1936-45 (the 1940s cohort) and in 1956-1965 (the 1960s cohort) and are not in the SEO sample.[21] More specifically, we select all individuals in the SRC sample who are interviewed at least twice in the sample years 1968-2013, select heads and their spouses, if present, and keep individuals born between 1936 and 1965 who did not graduate from college. We also include only white individuals and drop those who are married to a non-white spouse.[22] As reported in Table 1 in the main text of the paper, the resulting sample includes 5,039 individuals age 20 to 70, for a total of 73,944 observations.

The Health and Retirement Study (HRS) is a longitudinal data set collecting information on people age 50 or older, including a wide range of demographic, economic, and social characteristics, as well as physical and mental health, and cognitive functioning. We use is to compute inputs for the retirement period, because it contains a large number of observations and high-quality data for this stage of the life cycle.

Our data set is based on the RAND HRS files and the EXIT files, which include information on the wave right after death. As that data quality from the first two waves is lower, we use data from wave 3 to wave 12 (that is, from year 1996 to 2014). We select individuals in the age range 50-100 and thus born between 1906 and 1964. After keeping white, non-college graduates and their spouses, we are left with 19,377 individuals and 110,923 observations, as detailed in Table 3 in the main text of the

[21]The SEO sample includes families with income below half of the poverty line in 1968 and only 29% of them are white individuals or couples with less than a college degree. In addition, in 1997, the PSID stopped following most SEO families, that are thus no longer in the data set.

[22]Wife race is not available in the PSID until 1985. When possible, we use information gathered after that date. If still missing, we assume wife race is the same as the husband.

paper.

Appendix B. First-step estimation

This Appendix details our computations of our first-step model inputs, which are comprised of human capital, wages, health status at retirement, health dynamics after retirement, out-of-pocket medical expenses, survival probabilities, marriage and divorce probabilities, the distribution of people over state variables upon entering the model and of prospective spouses, the number of children, wealth, Social Security benefits, and taxes.

Human capital

In the model, we keep track of human capital measured as average accumulated earnings for a person ($\bar{y}_{kt}$), subject to the Social Security cap that is applied to yearly earnings and is time varying. To compute human capital, we assume that in the PSID data people start working at age 22 and we use the observed individual-level capped average earnings that they report starting from that age to compute our measure of human capital.[23]

Wages

Our framework requires that we estimate not only wage as a function of human capital, age, and gender, and the stochastic process for the wage shocks, but also the realized wage shocks for all men and women of working age in our sample (whether they are working or not). This is because we allow our initial conditions and assortative matching in marriage to depend on the realized values of these shocks.[24]

[23] For people entering the sample after age 22, we impute average accumulated earnings at the age of entry in the sample. For this purpose, we run a regression of capped earnings on cohort dummies, a polynomial in age fully interacted with gender, education dummies, and marital status and race dummies also interacted with gender. We then compute average capped earnings based on the predictions from the regression, and assign it to late entrants. Average earnings after entry in our sample is then updated for each individual following his/her observed earnings history (as done in the model).

[24] French (2005) and Blundell et al. (2016), instead, do not need the actual values of the realized wage shocks and estimate the parameters of their wage equations by using their structural models and matching moments on participants, thus relying on the model to generate the same selection patterns that are in the observed data.

To do so, we proceed as follows. First, we impute potential wages for individuals who are not working, so that we are able to construct potential wage as actual wage for participants and potential wage for non-participants. Second, we estimate potential wage as a function of age, gender, and human capital. Third, we estimate the persistence and variance of its unobserved component and the realized wage shocks using Kalman filtering, as in Borella, De Nardi, and Yang (2017).

Missing wages imputation. The observed wage rate is computed as annual earnings divided by annual hours worked. Gross annual earnings are defined as labor income during the previous year. Annual hours are given by annual hours spent working for pay during the previous year.

We impute missing wages by using coefficients from fixed effects regressions that we run separately for men and women. To avoid endpoint problems with the polynomials in age, we include individuals age 22 to 70 in the sample. Define the observed wage for labor market participants as

$$\ln wage_{kt} = I_{kt}^{n} \widetilde{\ln wage}_{kt}.$$

where k denotes an individual and t is age. The term I_{kt}^{n} is an indicator for participation (which is equal to 1 if the individual participates in the labor market and has no missing hours or earnings) and $\widetilde{\ln wage}$ is the potential wage that we wish to estimate. We estimate

$$\ln wage_{kt} = Z_{kt}'\beta_z + f_k + \varsigma_{kt},$$

where the dependent variable is the logarithm of the observed hourly wage rate, f_k is an individual-specific fixed effect and ς_{kt} is an error term. We include a rich set of explanatory variables in Z_{kt}: a fifth-order polynomial in age, a third-order polynomial in experience (measured in years of labor market participation), marital status (a dummy for being single), family size (dummies for each value), number of children (dummies for each value), age of youngest child, and an indicator of partner working if married. As an indicator of health, we use a variable recording whether bad health limits the capacity of working (this is the only health indicator available in the PSID for all years). Because this health indicator is not collected for wives, we do not include it in the regression for married women. Both regressions also include interaction terms between the explanatory variables. Variables that do not vary over time are captured by the individual effect f_k.

Using the estimated coefficients, we take the predicted value of the wage to be the potential wage for observations with missing wages. Hence, we define potential wage as

$$\ln \widetilde{wage}_{kt} = \begin{cases} \ln wage_{kt} & \text{if} \quad I^n_{kt} = 1 \\ Z'_{kt}\hat{\beta}_z + \hat{f}_k & \text{if} \quad I^n_{kt} = 0 \end{cases}$$

Wage function estimation. We model wages as a function of human capital, age, and gender, and we measure human capital as average realized earnings accrued up to the beginning of age t ($\bar{y}_t$).

To estimate the wage profiles, we proceed in two stages. First, we run the following fixed-effect regression for the logarithm of potential wages

$$\ln \widetilde{wage}_{kt} = d_k + f^i(t) + \sum_{g=1}^{G} \beta_g D_g \ln(\bar{y}_{kt} + \delta_y) + u_{kt}, \tag{36}$$

on a gender-specific fifth-order polynomial in age $f^i(t)$, gender-cohort cells g, and gender-cohort dummies D_g.[25] The shifter δ_y is set equal to \$5,000 to avoid taking the logarithm of values that are too small.[26] We also experimented by adding marital status dummies to capture the effect of changing marital status on wages, but they did not turn out to be statistically different from zero, conditional on average earnings.

Second, we regress the sum of the fixed effects and the residuals for each person on cohort and marital status dummies and their interactions, separately for each gender, and use those estimated effects for gender, marital status, and cohort as a shifter for the profiles of each group. This procedure allows for differences in average wages by marital status and cohort.

Table 14 reports the coefficients of the estimated equation from the first stage, the fixed effect regression, and Table 15 reports those the second stage, that is the regression on the residuals and fixed effects from the first stage. The marginal effects are in Table 7 in the main text of the paper.

The shock in log wage is modeled as the sum of a persistent component plus white

[25] To estimate a cohort-specific effect of human capital on wages in Equation (36), we redefine how we construct our cohorts. More specifically, we take two broader windows to define our cohorts: the 1950 cohort includes the generations born in 1935-1950, while the 1960s cohort includes those born in 1951-1965. We do so because we do not observe the complete age profile for the wages of the 1960s cohort.

[26] While we use earnings subject to the Social Security cap to compute average earnings (this is the state variable in our model), estimating this wage regression by using uncapped previous average earnings yields in very similar estimates.

	Coefficient	Standard Error
$\ln(\bar{y}_t + \delta_y)$	0.346***	(0.0191)
$\ln(\bar{y}_t + \delta_y)$*female	0.0669***	(0.0246)
$\ln(\bar{y}_t + \delta_y)$*born in 1940s	-0.0907***	(0.0246)
$\ln(\bar{y}_t + \delta_y)$*born in 1940s*female	0.0425	(0.0347)
Age	-0.0951	(0.185)
$Age^2/(10^2)$	0.497	(0.925)
$Age^3/(10^4)$	-1.222	(2.237)
$Age^4/(10^6)$	1.504	(2.626)
$Age^5/(10^8)$	-0.763	(1.198)
Age*female	0.343	(0.246)
$Age^2/(10^2)$*female	-1.897	(1.229)
$Age^3/(10^4)$*female	4.932*	(2.977)
$Age^4/(10^6)$*female	-6.101*	(3.498)
$Age^5/(10^8)$*female	2.912*	(1.598)
Constant	-1.452	(0.946)
N	62176	
R-sq	0.100	

Table 14: Coefficients from fixed effects estimates. Dependent variable: logarithm of the potential wage. PSID data. Robust standard errors in parentheses, clustered at the individual level. * p<0.10, ** p<0.05, *** p<0.01

noise, which we assume captures measurement error:

$$d_k + u_{kt+1} \equiv w_{kt+1} = \ln \epsilon_{kt+1} + \xi_{kt+1} \tag{37}$$

$$\ln \epsilon_{kt+1} = \rho_\epsilon \ln \epsilon_{kt} + v_{kt+1}, \tag{38}$$

where ξ_{kt+1} and v_{kt+1} are independent white-noise processes with zero mean and variances equal to σ^2_ξ and σ^2_v, respectively.

We use the residuals from the first stage to estimate these process separately for each gender.[27]

Because initial conditions and assortative matching in marriage are functions of one's wage shocks, we need the value of those wage shocks for each person of working age over time. To do so, we estimate the system formed by (37) and (38) by Maximum Likelihood, which can be constructed assuming that the initial state of the system and the shocks are Gaussian, and using standard Kalman Filter recursions. With this procedure, we are able to estimate both the parameters in (37) and (38) and the entire state, that is $\ln \epsilon_{kt}, t = 1, ...T$. Table 16 reports our estimates of the AR

[27]For this, we limit the age range between 25 and 65 and, because we rely on residuals also taken from imputed wages, we drop the highest 0.5% residuals both for men and women. This avoids large outliers to inflate the estimated variances (however, the effect of this drop is negligible on our estimates).

	Men	Women
Born in 1940s	1.079***	0.488***
	(0.0556)	(0.0409)
Born in 1950s	0.418***	0.278***
	(0.0503)	(0.0366)
Married	0.0730***	-0.0146
	(0.0244)	(0.0233)
Married*Born in 1940s	-0.0509	0.0137
	(0.0569)	(0.0445)
Married*Born in 1950s	0.0795	-0.0194
	(0.0527)	(0.0395)
Constant	1.110***	-1.895***
	(0.0227)	(0.0212)
N	31256	30920
R-sq	0.324	0.145

Table 15: Second stage: coefficients from OLS estimates. Dependent variable: residuals from fixed effects estimates. PSID data. Robust standard errors in parentheses, clustered at the individual level. * p<0.10, ** p<0.05, *** p<0.01

component of the wage.

Parameter	Men	Women
ρ_ϵ	0.939	0.946
$Var(v)$	0.023	0.014
$Var(\ln \epsilon_1)$	0.101	0.086

Table 16: Estimated processes for the wage shocks for men and women, PSID data.

Alternative estimation methods in the presence of sample selection.

Our results are thus obtained by running fixed-effect regressions to impute missing wages, constructing potential wages by using observed wages when available and imputing wages when missing, and running fixed effects regressions on potential wages to estimate the deterministic and stochastic components of our wage processes.

In this section, we compare the results from our procedure with those resulting from two other approaches commonly used in the literature: running fixed effects on wages for labor market participants and running fixed effects on wages for labor market participants and applying a control function approach to correct for sample selection (Dustmann and Rochina-Barrachina, 2007).

The control function approach corrects for sample selection by modeling labor market participation as a Probit, computing the Mills ratio (which is the probability that a person is working given his or her characteristics), and then using the inverse

Mills ratio as an additional regressor to the main fixed effect (or demeaned) regression for wages. This approach was pioneered by Heckman (1979) and extended by Wooldridge (1995) to panel data.

To apply the control function approach, we include the following variables to explain the participation decision: home ownership (dummy), age of the youngest child, total number of children, number of children age 0-5, and completed grades of education, all interacted with gender, cohort, and marital status. In addition, we include an age polynomial interacted with gender.

In the wage equation, we also interact the inverse of the Mills ratio with gender. As Table 17 shows, the inverse Mills ratio is not significantly different from zero for men or women, indicating no selection bias is present in the fixed effects estimates.[28] The table also shows that all of our estimated coefficients are very similar when using these three approaches, and thus robust to the specific approach used.

Health status at retirement

We use the HRS data and define health status, ψ, on the basis of self-reported health a variable that can take five possible values (excellent, very good, good, fair, poor). Bad health status is defined as a dichotomous variable equal to 1 if self-reported health is fair or poor and 0 otherwise.[29]

We cannot calculate the probability of being in bad health at the start of retirement using the observed frequencies for the 1960s cohort because we do not observe that cohort at that age. We thus resort to the following imputation procedure for health status at age 66. We estimate a logistic regression or people age 50-68 in which the dependent variable is health status (0 for good health, 1 for fair or bad health), on a third-order polynomial in age and cohort dummies, separately for single men, single women, and couples. In the case of couples we estimate a multinomial logistic regression over the four possible health states in the couple, respectively for the husband and the wife: (good, good), (good, bad), (bad, good), and (bad, bad)

[28]Because we model wages as a function of human capital and human capital is predetermined but not strictly exogenous, it should be instrumented, as suggested by Semykina and Wooldridge (2005), among others. Unfortunately it is not easy to find good instruments for human capital, in addition to those used to predict participation. We do not attempt to do so, as it is a task that goes beyond the scope of this Appendix. See Costa Dias et al. (2018) for more on this.

[29]Blundell, Britton, Costa Diaz and French (2017) study labor supply behavior around retirement time and show that self-reported health captures the effects of health well compared with a variety of health measures, including measures computed using objective health outcomes.

	(1) BASELINE	(2) FE	(3) W95	(4) W95
$\hat{\lambda}_t$			0.0560	0.140
			(0.0418)	(0.194)
$\hat{\lambda}_t * female$				-0.0902
				(0.199)
$\ln(\bar{y}_t)$	0.346***	0.379***	0.369***	0.368***
	(0.0191)	(0.0208)	(0.0207)	(0.0217)
$\ln(\bar{y}_t)$*female	0.0669***	0.148***	0.149***	0.149***
	(0.0246)	(0.0301)	(0.0306)	(0.0309)
$\ln(\bar{y}_t)$*born in 1940s	-0.0907***	-0.105***	-0.104***	-0.101***
	(0.0246)	(0.0260)	(0.0260)	(0.0271)
$\ln(\bar{y}_t)$*born in 1940s*female	0.0425	0.0501	0.0487	0.0461
	(0.0347)	(0.0444)	(0.0448)	(0.0451)
Age	-0.0951	-0.250	-0.233	-0.242
	(0.185)	(0.237)	(0.244)	(0.231)
$Age^2/10^2$	0.497	1.271	1.220	1.265
	(0.925)	(1.215)	(1.242)	(1.183)
$Age^3/10^4$	-1.222	-3.125	-3.069	-3.187
	(2.237)	(3.024)	(3.070)	(2.944)
$Age^4/10^6$	1.504	3.784	3.784	3.941
	(2.626)	(3.657)	(3.690)	(3.560)
$Age^5/10^8$	-0.763	-1.823	-1.854	-1.943
	(1.198)	(1.722)	(1.728)	(1.678)
$Age * female$	0.343	0.599*	0.562	0.574
	(0.246)	(0.358)	(0.350)	(0.354)
$Age^2 * female/10^2$	-1.897	-3.183*	-3.021*	-3.079*
	(1.229)	(1.834)	(1.787)	(1.816)
$Age^3 * female/10^4$	4.932*	8.103*	7.775*	7.921*
	(2.977)	(4.555)	(4.428)	(4.519)
$Age^4 * female/10^6$	-6.101*	-9.942*	-9.621*	-9.804*
	(3.498)	(5.501)	(5.342)	(5.470)
$Age^5 * female/10^8$	2.912*	4.731*	4.610*	4.708*
	(1.598)	(2.588)	(2.514)	(2.581)
N	62716	51662	51662	51662
R-sq	0.100	0.101	0.101	0.101

Table 17: Sample Selection. (1) Fixed effects on potential wage (our estimates in the model), (2) Fixed Effects on actual wage, (3) FE plus inverse Mills ratio λ on actual wage (Wooldridge, 1995). Robust standard errors in parentheses, clustered at the individual level. In (3), bootstrap standard errors (500 replications). * p<0.10, ** p<0.05, *** p<0.01

and we use a second-order polynomial in age because higher powers of age are not statistically different from zero for them. We then use our estimated coefficients to predict the health status at age 66 for our 1960s cohort.

Table 18 shows our estimated coefficients, while Table 19 reports our predicted probabilities of being in bad health at age 66 by demographic status, as well as the p-value of the test of equal probabilities by cohort.

Table 19 shows that single men born in 1960s are almost 8 percentage points more likely to be in bad health by the time they reach age 66 than those born in the 1940s:

	Single Men	Single Women	Husband/Wife (Good/Good)	Husband/Wife (Good/Bad)	Husband/Wife (Bad /Good)
Age	-5.749*	4.784**	-0.273*	-0.379**	-0.271
	(3.210)	(2.133)	(0.154)	(0.184)	(0.180)
$Age^2/10^2$	9.842*	-7.675**	0.187	0.290*	0.209
	(5.390)	(3.572)	(0.127)	(0.152)	(0.149)
$Age^3/10^4$	-5.564*	4.094**			
	(3.006)	(1.987)			
Born in 1930s	-0.181	-0.322***	0.484***	0.0774	0.403**
	(0.161)	(0.114)	(0.148)	(0.175)	(0.177)
Born in 1940s	-0.334**	-0.193**	0.449***	0.133	0.445***
	(0.130)	(0.0985)	(0.131)	(0.154)	(0.158)
Born in 1950s	-0.262**	-0.125	0.0943	-0.268*	0.248
	(0.121)	(0.0938)	(0.125)	(0.147)	(0.151)
Constant	110.4*	-99.60**	11.12**	12.60**	8.831
	(63.50)	(42.30)	(4.601)	(5.512)	(5.388)

Table 18: Probability of being in bad health. Logit (for singles) and multinomial logit (for couples, husbands and wives) coefficient estimates. HRS data. Robust standard errors in parentheses, clustered at the individual level. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

this difference is statistically different from zero at the 1 percent level, as illustrated by the P-value in the last row of the table. For single women the increase in the probability of being in bad health at age 66 is also substantial, almost 4.5 percentage points, and also statistically different from zero. Turning to couples, the probability that both are partners good health drops by 6.0 percentage points relatively to the cohort born in 1940s and the probability that partners are in bad health increases by 4.2 percentage points. The probability of women being in bad health and having husbands in good health increases by 3.4 percentage points with respect to the cohort born in 1940s, while the probability that husbands are in bad health and have a healthy wife decreases slightly, although the change is not statistically significant.

	Single Men	Single Women	Husband/Wife (good/good)	Husband/Wife (good/bad)	Husband/Wife (bad /good)	Husband/Wife (bad /bad)
Born in 1940s	0.349	0.347	0.576	0.148	0.173	0.104
Born in 1960s	0.428	0.392	0.516	0.182	0.156	0.146
P-value	0.013	0.055	0.021	0.025	0.229	0.008

Table 19: Predicted probabilities to be in bad health at age 66, by gender, marital status and cohort. P-value of the difference between cohorts.

Health dynamics after retirement

We model the evolution of health for people born between 1900 and 1965 as a logit function:

$$\pi_{\psi t} = Prob(\psi_t = 1 \mid X_t^{\psi}) = \frac{\exp(X_t^{\psi\prime}\beta^{\psi})}{1 + \exp(X_t^{\psi\prime}\beta^{\psi})}$$

which we then use to construct the transition matrix at each age, gender, and marital status. The set of explanatory variables X_t^{ψ} includes cohort dummies, a second-order polynomial in age, previous health status, gender, marital status, and interactions between these variables when they are statistically different from zero. We use estimated coefficients relative to the cohort of interest as input in our model. As the HRS data are collected every two years, we obtain two-year probabilities and convert them into one-year probabilities. Table 20 reports our estimated coefficients.

	Coefficient	SE
Age	-0.0188	(0.0197)
$Age^2/10^2$	0.0252*	(0.0132)
$Health_{t-1}$*age	0.105***	(0.00182)
$Health_{t-1}*age^2/10^2$	-0.0936***	(0.00240)
Male	-3.312***	(0.868)
Male*age	0.0924***	(0.0244)
Male*$age^2/10^2$	-0.0616***	(0.0169)
Married	-0.0857***	(0.0251)
Married*age	0.0666***	(0.0175)
Married*$age^2/10^2$	2.493***	(0.891)
Born in 1910s	0.308***	(0.0729)
Born in 1920s	0.132**	(0.0604)
Born in 1930s	-0.0453	(0.0531)
Born in 1940s	-0.0520	(0.0446)
Born in 1950s	-0.0524	(0.0477)
Constant	-1.539**	(0.712)

Table 20: Health dynamics over two-year periods. Logistic regression coefficients, dependent variable: health status. HRS data. Robust standard errors in parentheses, clustered at the individual level. * p<0.1, ** p<0.05, *** p<0.01

Out-of-pocket medical expenses

Out-of-pocket (oop) medical expenses are defined as the total amount that the individual spends out of pocket in hospital and nursing home stays, doctor visits, dental costs, outpatient surgery, average monthly prescription drug costs, home health care, and special facilities charges. They also include medical expenses in the last year of life, as recorded in the exit interviews. In contrast, expenses covered by public or private insurance are not included in our measure, as they are not directly incurred by the individual. The estimated equation is:

$$\ln(m_{kt}) = X_{kt}^{m\prime}\beta^m + \alpha_k^m + u_{kt}^m$$

where explanatory variables include a third-order polynomial in age fully interacted with gender and current health status, and we include these interactions whenever they are statistically different from zero. Marital status (also interacted with other variables) does not turn out to be significantly different from zero in the first step. We estimate the equation on the HRS data using a fixed effects estimator, which takes into account all unmeasured fixed-over-time characteristics that may bias the age profile, such as differential mortality (as discussed in De Nardi, French and Jones (2010)). We then regress the residuals from this equation on cohort, gender and marital status dummies to compute the average effect for each group of interest. Hence, the profile of the logarithm of medical expenses is constant across cohorts up to a constant. Table 21 reports estimated coefficients, while Table 8, in the main text, reports and discusses marginal effects.

	Coefficient	SE
Age	0.497***	(0.0477)
$Age^2/10^2$	-0.634***	(0.0675)
$Age^3/10^4$	0.277***	(0.0315)
Bad health	3.876***	(0.335)
Bad health*Age	-0.101***	(0.00947)
Bad health*$Age^2/10^2$	0.0672***	(0.00659)
Male*Age	-0.253***	(0.0842)
Male*$Age^2/10^2$	0.370***	(0.120)
Male*$Age^3/10^4$	-0.177***	(0.0560)
Constant	-3.853***	(0.929)
Second stage		
Male	5.547***	(0.00779)
Married	0.283***	(0.00808)
Born in 1910s	-0.527***	(0.0262)
Born in 1920s	-0.429***	(0.0211)
Born in 1930s	-0.396***	(0.0200)
Born in 1940s	-0.396***	(0.0199)
Born in 1950s	-0.129***	(0.0211)
Constant	-2.076***	(0.0196)
N	96098	
R-sq first stage	0.027	
R-sq second stage	0.854	

Table 21: Estimates for the logarithm of medical expenses, first stage (fixed effects) and second stage (OLS). HRS data. Robust standard errors in parentheses, clustered at the individual level. * p<0.1, ** p<0.05, *** p<0.01

Finally, we model the variance of the shocks regressing the squared residuals from the regression in logs on a third-order polynomial in age fully interacted with gender and current health status, and on cohort, gender and marital status dummies, and use it to construct average medical expenses as a function of age by adding half the variance to the average in logs before exponentiating.

Survival probabilities

We model the probability of being alive at time t as a Logit function

$$s_t = Prob(Alive_t = 1 \mid X_t^s) = \frac{\exp(X_t^{s\prime}\beta^s)}{1 + \exp(X_t^{s\prime}\beta^s)}.$$

that we estimate using the HRS data (which are biennial). Among the explanatory variables, we include a third-order polynomial in age, gender, marital status, and health status in the previous wave, as well as interactions between these variables and age, whenever they are statistically different from zero. We also include cohort dummies and use coefficients relative to the cohort of interest to adjust the constant accordingly. Table 22 reports estimated coefficients.

	Coefficient	SE
Age	-0.745***	(0.218)
$Age^2/10^2$	0.967***	(0.284)
$Age^3/10^3$	-0.476***	(0.122)
$Health_{t-1}$*age	-0.0632***	(0.00327)
$Health_{t-1}*age^2/10^2$	0.0578***	(0.00394)
Male	-0.563***	(0.0304)
Married	0.316***	(0.0323)
Born in 1910s	0.0854	(0.237)
Born in 1920s	0.106	(0.233)
Born in 1930s	0.121	(0.232)
Born in 1940s	0.116	(0.224)
Born in 1950s	0.220	(0.216)
Constant	24.94***	(5.457)

Table 22: Logistic regression coefficients, dependent variable: survival over a two-year period. HRS data. Robust standard errors in parentheses, clustered at the individual level. * p<0.1, ** p<0.05, *** p<0.01

Table 23 reports the marginal effects from our estimated equation. On average a marginal increase in age reduces the biennial probability of survival by 0.75 and 0.53 percentage points for men and women, respectively, with the effect getting larger with age: at age 96, a marginal increase in age decreases the survival probability by

4.54 and 4.27 percentage points for men and women. The effect of age also differs according to one's health status: a marginal increase in age reduces the biennial probability of survival by 0.61 percentage points if a man is in good health, and by 1.05 percentage points if he is in bad health. For women, the negative effect of age almost doubles if they are in bad health, going 0.41 to 0.77 percentage points. While being married increases the probability of survival, being born in 1960s (relative to being born in 1940s) decreases it, although the cohort effect, when conditioning on the health status, is not precisely estimated and statistically different from zero (although it would be very significant if we did not include health in the regression).

	Men	Women
Age overall	-0.0075***	-0.0053***
	(0.0002)	(0.0002)
Age = 66	-0.0031***	-0.0019***
	(0.0002)	(0.0001)
Age = 76	-0.0063***	-0.0040***
	(0.0004)	(0.0003)
Age = 86	-0.0175***	-0.0121***
	(0.0009)	(0.0006)
Age = 96	-0.0454***	-0.0427***
	(0.0028)	(0.0031)
Age overall if good health	-0.0061***	-0.0041***
	(0.0002)	(0.0001)
Age overall if bad health	-0.0105***	-0.0077***
	(0.0004)	(0.0003)
Married	0.0222***	0.0153***
	(0.0024)	(0.0015)
Born in 1960s	-0.0084	-0.0059
	(0.0169)	(0.0119)

Table 23: Average marginal effects on the two-year survival probability for men and women. HRS data. Robust standard errors in parentheses, clustered at the individual level. $^{*}p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$

We transform the biennial probability of surviving that we estimate from the HRS data into an annual probability by taking the square root of the biennial probability.

Marriage and divorce probabilities

We use the PSID to estimate the probabilities of marriage and divorce.[30] We model the probability of getting married, ν_{t+1} and separately estimate the probability

[30] Because the number of new marriages (and also of divorces) is limited in the data, we constrain the cohort effect of the 1960s cohort to be the same as the one for the 1950 cohort.

of getting married for men and women

$$\nu^i_{t+1} = Prob(Married_{t+1} = 1 | Married_t = 0, Z_t) = F(Z_t' \beta_m),$$

where Z_t include a polynomial in age, cohort dummies, and the after 1997 dummy.[31] F denotes the standard logistic distribution.

Similarly, we estimate the probability of divorce as

$$\zeta_t = Prob(Divorced_{t+1} = 1 | Married_t = 1, Z_t) = F(Z_t' \beta_d),$$

where Z_t include a polynomial in age, cohort dummies, and an indicator for biennial waves. F denotes the standard logistic distribution.

	Single Men Marriage	Single Women Marriage	Couples Divorce
Age	0.0179	0.00743	0.0224
	(0.0405)	(0.0441)	(0.0278)
$Age^2/(10^2)$	-0.0897*	-0.0988*	-0.0931***
	(0.0532)	(0.0575)	(0.0356)
$I(year > 1997)$	0.188	0.474**	0.706***
	(0.191)	(0.193)	(0.114)
Born in 1940s	0.622***	0.173	-0.119
	(0.174)	(0.179)	(0.105)
Constant	-1.672**	-1.615**	-2.886***
	(0.731)	(0.804)	(0.526)
N	4206	5410	25597
Pseudo R-sq	0.025	0.042	0.019

Table 24: Estimated coefficients from logistic regressions. Column 1: Marriage of single men; column 2: marriage of single women; column 3: divorce of couples. PSID data. Robust standard errors in parentheses, clustered at the individual level. * p<0.10, ** p<0.05, *** p<0.01

Conditional on meeting a partner, the probability of meeting with a partner p with wage shock ϵ^p_{t+1} is

$$\xi_{t+1}(\cdot) = \xi_{t+1}(\epsilon^p_{t+1} | \epsilon^i_{t+1}, i).$$

where i denotes gender. We compute the above probability using our estimates of the wage shocks, by partitioning households in age groups (25-35; 35-45; 45-60) and computing the variance-covariance matrix of newly matched partners' wage shocks in each age group. The implied correlation in the three age groups is 0.24, 0.33, and

[31]The PSID goes from a yearly to a biennial frequency in 1997. To take this into account, we include an indicator variable taking value one from 1997 on in the regression, which we then abstract from when constructing the yearly probabilities.

0.36 respectively. We then assume that the joint distribution is lognormal. As in the whole sample we observe 722 new marriages in the age range 25-60, we do not allow this probability to depend on cohort.

We assume random matching over asset and lifetime income of the partner conditional on partner's wage shock. Thus, we compute

$$\theta_{t+1}(\cdot) = \theta_{t+1}(a^p_{t+1}, \bar{y}^p_{t+1} | \epsilon^p_{t+1}).$$

using sample values of assets, average capped earnings, and wage shocks. More specifically, we assume θ_{t+1} is log-normally distributed at each age with mean and variance computed from sample values. Assets include a shifter as described for the computation of joint the distribution at age 25 (see Wealth subsection in this Appendix).

Distributions upon entering the model and for prospective spouses

For single men and women, we parameterize the joint distribution of assets, average realized earnings, and wage shocks at each age as a joint log normal distribution given by

$$\begin{pmatrix} \ln(a^i_t + \delta_a) \\ \ln(\bar{y}^i_t + \delta_y) \\ \ln\, \epsilon^i_t \end{pmatrix} \sim N \begin{pmatrix} \mu^i_{at} + \delta_a \\ \mu^i_{\bar{y}t} + \delta_y \, , \Sigma^i_{st} \\ \mu^i_{\epsilon t} \end{pmatrix}, \tag{39}$$

where Σ_s is a 3x3 covariance matrix and i denotes gender. We characterize this distribution by estimating its mean and its variance, which both depend on age t. To estimate means, we regress the logarithm of assets plus shift parameter, average earnings, and the productivity shock $\ln \hat{\epsilon}^i_t$ on a third-order polynomial in age and cohort dummies. The predicted age profile, is the age-specific estimate of the mean of the log-normal distribution. We estimate the elements of the variance-covariance matrix by taking the relevant squares or cross-products of the residuals from this regression. To obtain a a smoothed estimate of the variance-covariance matrix at each age, we regress them on a third-order polynomial in age, element by element.

For couples, we compute the initial joint distribution at age 25 of the following variables

$$\begin{pmatrix} \ln(a+\delta_a) \\ \ln(\bar{y}^1+\delta_{\bar{y}}) \\ \ln(\bar{y}^2+\delta_{\bar{y}}) \\ \ln(\epsilon^1) \\ \ln(\epsilon^2) \end{pmatrix} \sim N \begin{pmatrix} \mu_a+\delta_a \\ \mu_{\bar{y}1}+\delta_{\bar{y}} \\ \mu_{\bar{y}2}+\delta_{\bar{y}} \ , \Sigma_c \\ \mu_{\epsilon 1} \\ \mu_{\epsilon 2} \end{pmatrix}, \tag{40}$$

where Σ_c is a 5x5 covariance matrix computed on the data for couples.

Number of children

We regress the number of children on a fifth-order polynomial in maternal age, interacted with marital status and cohort dummies to construct the average age profile of children in each age group for single and married women and use the profiles relative to the cohorts of mothers born in 1960s. We runs such a regression for total number of children (used in equivalence scales), children 0-5, and children 6-11 (these two affect child care costs).

Wealth

We define wealth as total assets (defined as all assets types available in the PSID) plus home equity net. Wealth in the PSID is only recorded in 1984, 1989, 1994, and then in each (biennial) wave from 1999 onwards. We rely on an imputation procedure to compute wealth in the missing years, starting in 1968. This imputation is based on the following fixed-effect regression

$$\ln(a_{kt}+\delta_a) = Z'_{kt}\beta_z + da_k + wa_{kt}, \tag{41}$$

where k denotes the individual and t is age. The parameter δ_a is a shifter for assets to have only positive values and to be able to take logs, and the variables Z include polynomials in age, also interacted with health status, and with average earnings (uncapped), family size, and a dummy for health status. The term da_k is the individual fixed effect and wa_{kt} is a white-noise error term. Equation (41) is estimated separately for single men, single women, and couples, as wealth is measured at the household level.

We then use the imputed as well as the actual observations to estimate the wealth profiles used as target moments and to parameterize the joint distribution of initial assets, average realized earnings, and wage shocks for single men, single women, and couples.

Social Security benefits

The Social Security benefit at age 66 is calculated to mimic the Old Age and Survivor Insurance component of the Social Security system:

$$SS(\bar{y}_r) = \left\{ \begin{array}{ll} 0.9\bar{y}_r, & \bar{y}_r < 0.1115; \\ 0.1004 + 0.32(\bar{y}_r - 0.1115), & 0.1115 \leq \bar{y}_r < 0.6725; \\ 0.2799 + 0.15(\bar{y}_r - 0.6725), & 0.6725 \leq \bar{y}_r < y_t^{cap} \end{array} \right\}$$

The marginal rates and bend points, expressed as fractions of average household income, come from the Social Security Administration.[32]

The Social Security tax and Social Security cap have been changing over time. We also allow them to change over time for the households in our model.

Taxes

Guner et al. (2012b) estimate the tax function by marital status. We use their estimated parameters for married and singles unconditional of number of children. The resulting values for a married couple are $p^2 = 1.8500; b^2 = 0.2471; s^2 = 0.0006$. Those for singles are: $p^1 = 1.4150; b^1 = 0.2346; s^1 = 0.0074$. We also add a 4% state and local tax.

[32] https://www.ssa.gov/oact/cola/bendpoints.html. We use values for year 2009.

Appendix C. Robustness to sample selection

Table 25 uses the PSID data to show that the fraction of the population having less than a college degree dropped from 83.1% in 1940s to 77.2% in 1960s. This corresponds to a 6 percentage points drop in non-college graduates in the population across our two cohorts (5.6 and 6.6 percentage points for men and women respectively).

	Men		Women		All	
Grades	1940	1960	1940	1960	1940	1960
up to 11	22.9	13.5	17.6	10.5	20.4	12.0
12	50.8	38.2	51.9	38.7	51.3	38.5
13	59.6	52.0	65.2	51.0	62.2	51.5
14-15	81.7	76.1	84.8	78.1	83.1	77.2
16-17	100.0	100.0	100.0	100.0	100.0	100.0

Table 25: Cumulative distributions of grade of school completed for the sample of white people in our main cohorts, by gender and cohort. PSID data.

To check whether sample selection is an important issue for us, we compare our main model inputs for our original sample, the "education selection sample" with that in a second sample "education selection with constant fraction size." To construct this second sample, we first take white people without a college degree born in the 1940s cohort and in the 1960s cohort in both data sets. Then, we increase the size of the 1960s cohort by picking, among those who have completed college in the 1960s cohort, those that had the lowest initial human capital at age 25 (both among men and women). Thus, in this second sample we have the same fraction of men and women for both the 1960s and 1940s cohorts.

The following table and graphs compare life expectancy, wages, and medical expenses for our 1940s cohort (unchanged) and 1960s cohorts with the two selection criteria. They show that the results are very similar across the two samples.

	Men, 1940s	Men, 1960s	Women, 1940s	Women, 1960s
Original education sample				
At age 50	77.57	76.15	79.83	78.72
At age 66	82.53	80.91	85.68	84.02
Fixed fraction sample				
At age 50	77.56	76.08	79.83	78.67
At age 66	82.53	80.85	85.70	83.97

Table 26: Life expectancy for white and non-college educated men and women born in the 1940ss and 1960s cohorts. HRS data

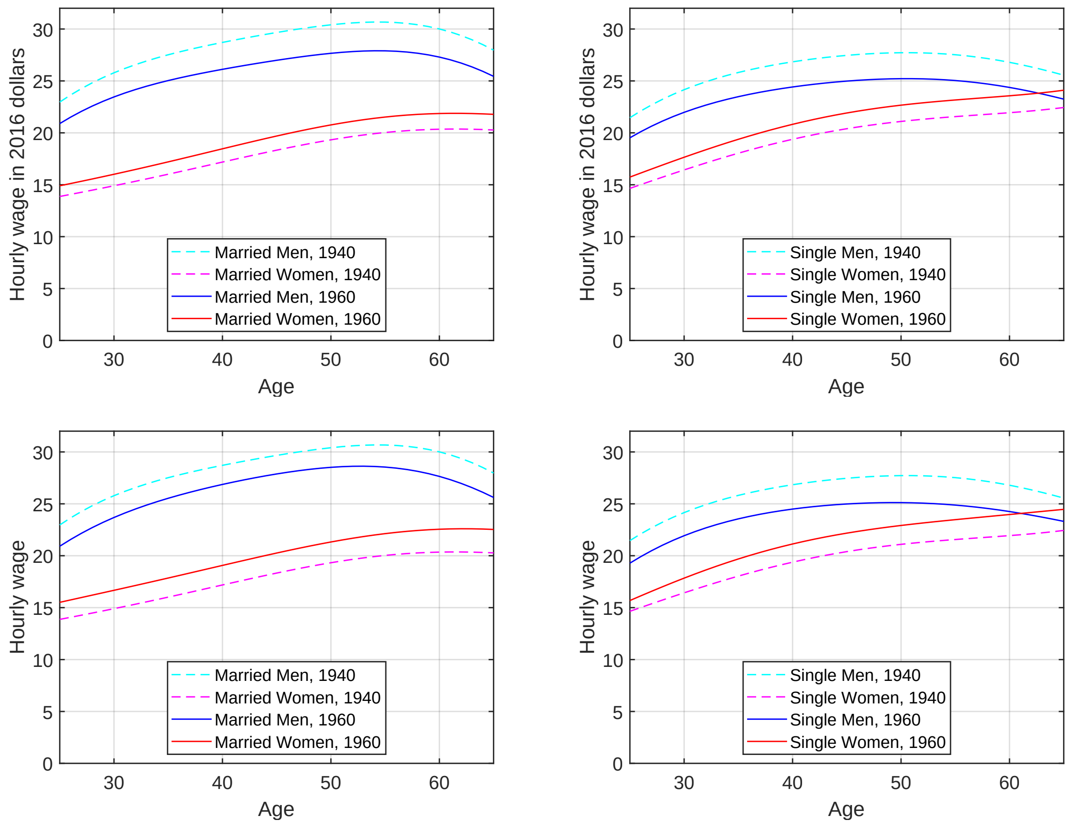


Figure 13: Wage profiles by age, comparing 1960s and 1940s for married people (left panel) and single people (right panel). Top panel: original sample. Bottom panel: same size by cohort sample.

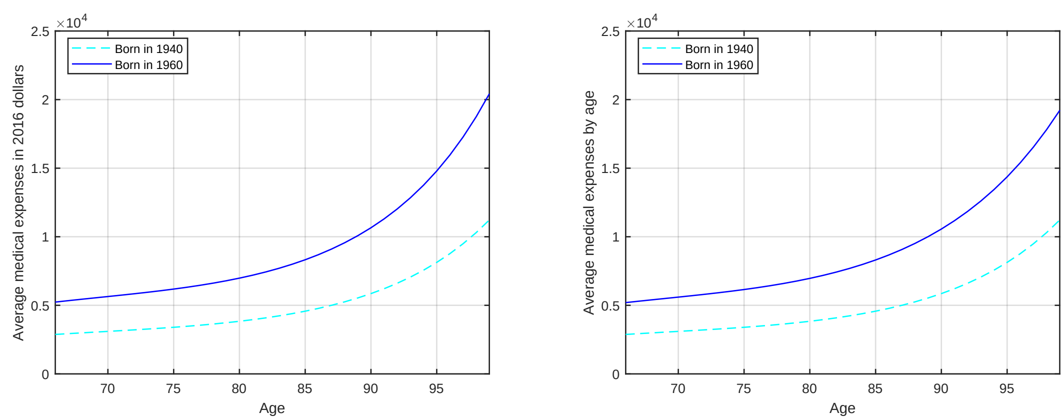


Figure 14: Average out-of-pocket medical expenses for the cohorts born in the 1940s and 1960s. Left graph: original education sample. Right graph: fixed fraction sample.