ACCOUNTING FOR SOCIAL SECURITY CLAIMING BEHAVIOR

Svetlana Pashchenko and Ponpoje Porapakkarm

CRR WP 2018-8
September 2018

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
Tel: 617-552-1762   Fax: 617-552-0191
http://crr.bc.edu

Svetlana Pashchenko is an assistant professor at the University of Georgia Terry College of Business. Ponpoje Porapakkarm is an associate professor at the National Graduate Institute for Policy Studies. The research reported herein was performed pursuant to a grant from the U.S. Social Security Administration (SSA) funded as part of the Retirement Research Consortium.

About the Steven H. Sandell Grant Program

This paper received funding from the Steven H. Sandell Grant Program for Junior Scholars in Retirement Research. Established in 1999, the Sandell program's purpose is to promote research on retirement issues by scholars in a wide variety of disciplines, including actuarial science, demography, economics, finance, gerontology, political science, psychology, public administration, public policy, sociology, social work, and statistics. The program is funded through a grant from the Social Security Administration (SSA). For more information on the Sandell program, please visit our website at: http://crr.bc.edu/?p=9570, send e-mail to crr@bc.edu, or call (617) 552-1762.

About the Center for Retirement Research

The Center for Retirement Research at Boston College, part of a consortium that includes parallel centers at the University of Michigan and the National Bureau of Economic Research, was established in 1998 through a grant from the Social Security Administration. The Center's mission is to produce first-class research and forge a strong link between the academic community and decision-makers in the public and private sectors around an issue of critical importance to the nation's future. To achieve this mission, the Center sponsors a wide variety of research projects, transmits new findings to a broad audience, trains new scholars, and broadens access to valuable data sources.

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Ave
Chestnut Hill, MA 02467
Tel: 617-552-1762   Fax: 617-552-0191
http://crr.bc.edu

Abstract

The paper examines why most individuals claim Social Security benefits before the full retirement age. Early claiming results in a substantial reduction in pension income, yet many people claim as early as possible, age 62, or soon thereafter. Since delaying claiming is equivalent to purchasing additional annuity income, this behavior is consistent with the so-called annuity puzzle. We provide a quantitative analysis of claiming decisions (or equivalently, of the demand for the Social Security annuity). Our tool is a structural lifecycle model calibrated to match many important features of the data.

The paper found that:

- One of the important factors accounting for the low demand for public annuities is a significant discrepancy between: (i) the individuals' subjective discount rate, and (ii) the discount rate implied by the implicit price of the Social Security annuity.
- Two of the commonly named impediments to private annuitization – mean-tested benefits and medical expenditures – are not important drivers of individuals' decisions for when to claim Social Security benefits.
- Pre-annuitized wealth and bequest motives play a major role in the decisions to collect Social Security benefits. Our counterfactual experiments show that if the amount of basic Social Security benefits is scaled down or if the strength of the bequest motive is diminished, significantly more people will postpone claiming.

The policy implications of the findings are:

- Given that many people consider themselves sufficiently annuitized even when they claim at age 62, late claimers should be awarded not with higher pension income but with lump-sum payments.
- We show that the policy of providing lump-sum payments instead of increasing Social Security benefits is very effective in inducing individuals to delay claiming.

1 Introduction

Why do most individuals claim Social Security benefits before the full retirement age? Individuals can claim benefits at any age between 62 and 70. Early claiming (before the full retirement age) results in a permanent reduction in the basic retirement benefits, while late claiming (after the full retirement age) results in their permanent increase. These penalties and rewards can be substantial: for example, for the cohort of individuals born in 1937, claiming at age 62 versus 65 (full retirement age for this cohort) resulted in a 20% reduction in monthly benefits, while claiming at age 70 versus 65 resulted in a more than 30% increase. Yet, among men born in 1936-1938, 67% claim benefits earlier than the full retirement age.[1]

Importantly, choosing the age of claiming benefits is equivalent to deciding how much (if any) annuity income to purchase. Every year of delay in claiming results in an increase in pension benefits, i.e., an additional lifetime annuity income, while the 'price' of this annuity is equal to one year of the foregone benefits. In this light, the reluctance of individuals to delay claiming can be interpreted as a lack of willingness to annuitize. This behavior is consistent with a well-documented annuity puzzle: a standard life-cycle model predicts that individuals should anuitize a large fraction (if not all) of their wealth, while in the data, we observe few people buying private annuities. One finding in the literature studying this puzzle is that several factors can play a role in decreasing demand for *private* annuities, among them, market frictions such as adverse selection and the existence of minimum purchase requirements (e.g., Mitchell et al, 1999; Pashchenko, 2013). However, these explanations do not apply when it comes to explaining low demand for the *public* annuity provided by Social Security, thus making it an even more difficult puzzle to explain.

Our goal in this paper is to investigate factors affecting individuals' decisions on when to claim Social Security benefits and, thus, the demand for Social Security annuity. To do this, we construct and estimate/calibrate a rich structural life-cycle model with endogenous labor supply and retirement, and detailed a representation of Social Security rules. Our model includes a number of factors previously shown to affect individuals' demand for private annuities, such as uncertain medical expenses, bequest motives, means-tested benefits, and preannuitized wealth.

Our model features the full life-cycle: a (forward-looking) individual starts his model life at age 25 and makes labor supply and saving decisions taking into account the existing Social Security rules. Every period after an individual reaches the age of 62, he decides whether or not to claim benefits (if he still hasn't done so). The key trade-off in this decision is an immediate increase in the available resources versus a higher lifetime pension income starting

[1] Own calculations based on the Health and Retirement Study.

one year later. The choice between these two alternatives depends on i) the extent of future benefits increase; ii) available resources; ii) expected medical spending and longevity, and iii) preferences: the strength of bequest motives and the rate of time preferences.

To estimate/calibrate the model we use three datasets: the Health and Retirement Study (HRS), the Medical Expenditure Panel Survey (MEPS), and the Panel Study of Income Dynamics (PSID). Apart from the claiming behavior, our estimated model can capture i) wealth accumulation profiles; ii) labor force participation over the life-cycle; iii) average labor income of workers over the life-cycle.

Our results are as follows. First, we show that one of the important factors accounting for the low demand for public annuities is a significant discrepancy between (i) the individuals' subjective discount rate and (ii) the discount rate implied by the implicit price of the Social Security annuity. We find that the number of late claimers would significantly increase if individuals had a lower rate of time preferences (were more patient) or if the Social Security annuity was "priced" based on a higher implied interest rate.

Second, we study the effects of commonly named impediments to private annuitization on claiming decisions, specifically, (i) medical and nursing home expenses, (ii) means-tested benefits, (iii) pre-annuitized wealth and (iv) bequest motives. We show that in the absence of medical spending, the demand for public annuities will decrease, thus increasing the fraction of early claimers. This happens because medical spending risk is concentrated late in life and annuities represent an effective way to partially insure it.

The means-tested benefits have only a small effect on claiming behavior. Our earlier study (Pashchenko, 2013) shows that in a context of a partial life-cycle model that includes only the retirement stage, the means-tested benefits can crowd out the demand for annuities as they also provide some degree of longevity insurance. In our framework with a full life-cycle, this effect is offset by adjustments in early life behavior: when the means-tested benefits are lower, individuals work and save more and, thus, can afford to retire (and claim benefits) earlier.

Pre-annuitized wealth and bequest motives play a major role in the decisions to collect Social Security benefits. Our counterfactual experiments show that if the amount of basic Social Security benefits is scaled down or if the strength of the bequest motive is diminished, significantly more people will postpone claiming.

Overall, our results indicate that given preferences (bequest motives and the discount rate) and the amount of basic Social Security benefits available at age 62, many people consider themselves sufficiently annuitized even when they claim as early as possible. In light of these findings, we consider an alternative policy when late claiming is rewarded not with higher pension income but with an equivalent (in present value terms) lump-sum benefit.

We show that this policy is very effective in inducing individuals to delay claiming. This result is robust to alternative interest rates used to convert pension income into lump-sum benefits.

The rest of this paper is organized as follows. Section 2 reviews the related literature. Section 3 discusses the price of Social Security annuity. Section 4 introduces the model, while Section 5 explains our estimation/calibration. The results and conclusion are presented in Sections 6 and 7, respectively.

2 Literature review

There is a growing literature examining the costs and benefits associated with claiming Social Security at different ages. A common conclusion of these studies is that in many cases, households can gain from delaying claiming, i.e., the resulting change in the expected present value of retirement income is positive (Coile et al., 2002; Meyer and Reichenstein, 2010; Shoven and Slavov, 2014a and 2014b; Sun and Webb, 2009). Despite this, many individuals claim benefits as early as possible. In order to understand this puzzle, a number of studies investigate what factors affect the claiming decisions. Hurd et al. (2004) find that individuals with a low subjective survival probability tend to claim benefits earlier. Shoven and Slavov (2014a, 2014b) find that there is no strong relationship between early claiming and factors that can potentially affect the gains from delay, e.g., gender, wealth or marital status. However, the latter study and Venti and Wise (2004) find that individuals with higher education tend to claim benefits later. Goda et al. (2015) use administrative tax data to study whether individuals who claim benefits early are financially constrained. They find that a significant fraction of early claimers have enough assets to delay claiming.

Several studies investigate claiming decisions using a structural life-cycle model. Gustman and Steinmeier (2005) construct a life-cycle model of retirement decisions allowing for heterogeneity in preferences for leisure and the discount factor. They point out that the standard life-cycle model cannot fully account for the observed claiming behavior. In their later work, Gustman and Steinmeier (2015) show that a richer version of the model with stochastic returns on assets and more flexible labor supply still falls short of capturing a large fraction of individuals claiming as early as possible; however, varying beliefs about the future of Social Security can substantially improve the fit of the model along that dimension. Hubener et al. (2016) construct a rich life-cycle model where multi-person households make decisions about labor supply, portfolio choice and Social Security claiming age. They show that family status is an important determinant of portfolio allocation and claiming decisions.

Our paper is also related to the literature that studies the annuity puzzle (Dushi and

Webb, 2004; Inkmann et al., 2011; Lockwood, 2012; Mitchell et al., 1999; Pashchenko, 2013). Based on this literature, it is difficult to explain the annuity puzzle by one single factor; the observed low demand for private annuities can be due to a combination of several factors such as market frictions, a large fraction of pre-annuitized wealth in households' portfolios, or bequest motives.

More broadly, our paper is related to the literature studying various Social Security reforms (Laitner and Silverman, 2012; Kitao, 2014, Blandin, 2015). A subset of this literature focuses on policies that can affect Social Security claiming behavior. Maurer et al. (2016) design a survey to investigate whether individuals' decisions when to claim benefits are affected by the option to substitute an increase in pension income with lump-sum transfers. They find a significant increase in the average claiming age in response to this option. Imrohoroglu and Kitao (2012) use a general equilibrium framework to compare two Social Security reforms, the first one increases the earliest claiming age by two years and the second one increases the normal retirement age by two years. They find that the second reform has much larger effect on labor supply and the Social Security budget.

3 Social Security as an annuity: a closer look

Since delaying claiming of Social Security benefits is equivalent to buying a (public) annuity an important question is how much this annuity costs. To impute the price of Social Security annuity, we use the schedule of penalties and rewards for the cohort born in 1937 as shown in the second row of Table 1.

Consider an individual who is entitled to receive annual benefits b at the full retirement age of 65 and who decides whether to claim at 62 or at 63. If he claims at 63 he will receive additional lifetime income equal to $0.067b$. This will cost him $0.8b$ in terms of forgone benefits at age 62. Thus, the price of 1$ stream of this additional annuity income is equal to $0.8b/0.067b = 12$. In the same way, an individual who still did not claim by age 63 faces a trade-off of further increasing his annuity income by additional $0.067b$ or claiming right away and receiving $0.867\bar{b}$ in benefits. In this case, he can increase his annuity income at a price of $0.067b/0.867b = 13$ per 1$ of the extra income.

Age	62	63	64	65	66	67	68	69	70
% of benefits	80	86.7	93.3	100	106.5	113	119.5	126	132.5
Imputed price	12	13	14	15.38	16.38	17.38	18.38	19.38	-

Table 1: Reduction (increase) in benefits for early (late) claiming, in % of the benefits received at the full retirement age (1937 cohort); imputed price of SS anuity

The third row of Table 1 reports imputed price for each age between 62 and 69. One observation form this table is that the imputed price of Social Security annuity increases with age. This is not surprising because the increment of increase in annuity income with each additional year of delay is constant ($0.067b$ before the full retirement age and $0.065b$ after that), while the amount of the forgone benefits increases.

Next, we compare the imputed price with the actuarially fair annuity price. We compute the actuarially fair price q_m^{AF} of annuity purchased at age m as follows:

$$q_m^{AF} = \sum_{t=m}^{T-1} \frac{\zeta_{t+1|m}}{(1+r)^{t+1-m}},$$

where $\zeta_{t|m}$ is the probability to survive to age t conditional on being alive at age m, T is the last period of life set to 100 years old, and r is the interest rate. We use $\zeta_{t|m}$ from the Social Security cohort life tables for individuals born in 1940 (the closest cohort to 1937 for which life tables are available). Figure (1) plots the actuarially fair price for the three values of the interest rate (1%, 2%, and 3%) alongside the imputed price of Social Security annuity.

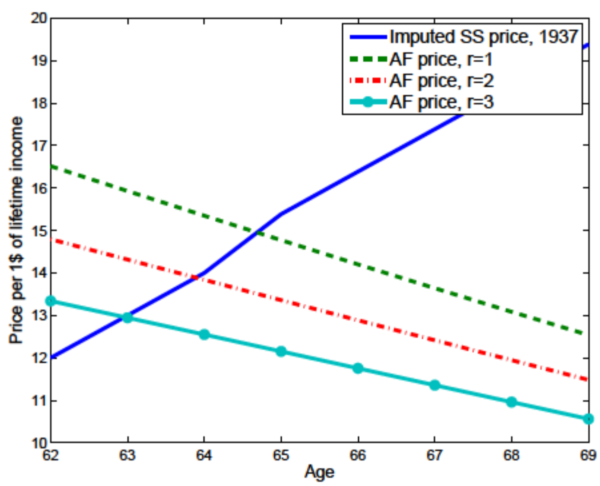

Figure 1: The imputed price of the Social Security annuity vs the actuarially fair annuity prices computed for different levels of interest rates

As can be seen from this figure, for most ages, the prices for Social Security annuity and the actuarially fair annuity are quite different. For ages 62 and 63, the former is significantly cheaper than the latter even when it is priced based on the interest rate of 3%.[2] However, for ages above 65, private annuity is significantly cheaper than public annuity. This happens because the price of private annuities decreases with age as older individuals get entitled to

[2] It is important to point out, that private annuity market in the US has a load of around 10% for a person with average mortality as estimated by Mitchell et al. (1999), which increases the gap between the prices of Social Security and private annuities even further.

shorter stream of income, while public annuity demonstrates the opposite pattern.

This comparison reveals that Social Security annuity is relatively badly priced for individuals above the full retirement age, which can possibly explain why very few individuals claim after that age. However, Social Security annuity is very attractive for individuals with average mortality and younger than the full retirement age; thus making this a puzzle why so many individuals claim as early as possible.[3]

4 Baseline Model

4.1 Demographics and preferences

A model period is one year. Individuals enter the model at age $t = 25$. Until age R^E individuals make only labor supply and consumption/saving decisions, between ages R^E and R^D individuals also decide whether to start collecting Social Security pension benefits, after age R^D individuals cannot work and only make consumption/saving decisions.

Individuals face health uncertainty: at age t, an agent's health condition h_t can be either good ($h_t = 1$) or bad ($h_t = 0$), where h_t evolves according to an age-dependent Markov process, $\mathcal{H}_t^j(h_t|h_{t-1})$. Health affects productivity, medical expenses, and survival probability. We denote the probability to survive from period t to $t+1$ as ζ_t^h. Each period an agent faces a stochastic out-of-pocket medical expenditure shock x_t^h which depends on his age and health; we denote the probability distibution of medical shock as $\mathcal{G}_t(x_t^h)$. Individuals after a certain age are also exposed to the risk of needing a long-term care; these shocks arrive with age- and health-dependent probability pn_t^h. An agent who needs to move to a nursing home has to pay for it xn_t out-of-pocket.

An individual is endowed with one unit of time that can be used for either leisure or work. Labor supply (l_t) is indivisible: $l_t \in \left\{0, \bar{l}\right\}$. Work brings disutility modeled as a fixed costs of leisure ϕ_w. The leisure of an individual can be represented as $\widetilde{l_t}$ where :

$$\widetilde{l_t} = 1 - l_t - \phi_w \mathbf{1}_{\{l_t>0\}}.$$

Here $\mathbf{1}_{\{.\}}$ is an indicator function equal to one if its argument is true. In addition to consumption and leisure individuals derive utility from leaving bequests.

To be able to separate the risk aversion from the intertemporal elasticity of substitution

[3] Consistent with our finding in this section, Bronshtein et al. (2016) show that individuals who claim benefits early and buy private annuities or opt for defined benefits annuity are making a financial mistake that can cost them up to $250,000.

(IES), we incorporate Epstein-Zin preferences in our model (Epstein and Zin, 1989).[4] More specifically, we assume than an individual's utility function over streams of consumption (c_t), leisure ($\widetilde{l_t}$), and bequeathed in case of not surviving to the next period assets (k_{t+1}) can be represented in recursive form:

$$U_t = \left[\left(c_t^{\chi} \widetilde{l_t}^{1-\chi} \right)^{1-\gamma} + \beta \left\{ \zeta_t^h E_t U_{t+1}^{1-\psi} + (1-\zeta_t^h) \eta \left(k_{t+1} + \phi \right)^{1-\psi} \right\}^{\frac{1-\gamma}{1-\psi}} \right]^{\frac{1}{1-\gamma}}$$

where χ is a parameter determining the relative weight of consumption in the consumption-leisure composite, ψ is the risk-aversion, $1/\gamma$ is the IES, β is the discount factor, η is the strength of the bequest motive and ϕ is a shift parameter that controls to what extent bequest is a luxury good.[5]

4.1.1 Labor income, taxation, transfers and Social Security

The earnings of an individual are equal to $wz_t^h l_t$, where w is wage and z_t^h is the idiosyncratic productivity that depends on age (t) and health (h_t). All individuals pay an income tax $\mathcal{T}(y_t)$, where taxable income y_t is based on both labor and capital income. Working households also pay Medicare (τ_{MCR}) payroll tax.

Individuals impoverished because of low earnings or high medical spending get means-tested transfers T_t^{SI} that guarantee each household a minimum consumption level $\underline{c}$. This safety net is a reduced form representation of the existing public transfer programs such as food stamps, Supplemental Security Income, disability insurance, and uncompensated care.

Working individuals pay Social Security payroll tax (τ_{ss}). The Social Security tax rate for earnings above $\overline{y}_{ss}$ is zero. Social Security benefits $ss(AE, j^R)$ is a concave function of the average lifetime earnings (AE) and the age when the benefits were claimed (j^R). Average earnings evolve as follows:

$$AE_{t+1} = \begin{cases} AE_t + \dfrac{y_t}{35} & ; \text{ if } t < 60 \\ AE_t + \dfrac{1}{35} \max\{0, y_t - AE_t\} & ; \text{ otherwise} \end{cases}$$

where

$$y_t = \max \quad wz_t^h l_t, \overline{y}_{ss}$$

Note that over the 35-year period from age 25 to 60, AE_t is updated every period, while after

[4] We separate the risk aversion from the inverse of the intertemporal elasticity of substitution because it allows us to better macth wealth profiles over the working stage of the life cycle.

[5] In this formulation of bequest motive we follow De Nardi (2004) and De Nardi et al (2010). Note that when $\phi = 0$ bequests become a necessity.

age 60 it is updated only if the current earnings exceed the average of previous earnings.[6]

The basic level of Social Security benefits ss^b corresponding to the full retirement age R^F, $ss(AE_t, j^R = R^F)$, is calculated as follows:

$$ss^b = \begin{cases} 0.9AE_t & ; \text{ if } AE_t < B_1 \\ 0.9B_1 + 0.32(AE_t - B_1) & ; \text{ if } B_1 \leq AE_t < B_2 \\ 0.9B_1 + 0.32(B_2 - B_1) + 0.15(AE_t - B_2) & ; \text{ if } AE_t \geq B_2, \end{cases} \quad (1)$$

where B_1 and B_2 are the bend points, i.e., the levels of AE_t when the replacement rate changes first from 0.9 to 0.32, then from 0.32 to 0.15. Social Security rules regarding benefits calculations change for each cohort; we use individuals born in 1936-1938 as our base cohort. The full retirement age (R^F) for our base cohort is 65 years, so we set $R^F = 65$.[7] We set the bend points B_1 to \$6,372 and B_2 to \$38,424 based on the Social Security benefits formula for 2000.[8]

The earliest age an individual can start receiving benefits (R^E) is 62 and the latest age the benefits can be claimed (R^D) is 70. The benefits of early takers are reduced by 6.7% for ages between 62 and 65. Individuals who claim benefits after the full retirement age get the basic benefits increased by 6.5% for every year up to age 70. The full schedule of benefits/rewards for early/late claiming is shown in the second row of Table 1.

Individuals who are younger than the full retirement age, who receive Social Security benefits but continue to work are subject to Social Security earning tax t^{earn}.[9] This tax rate is determined as follows:

$$t^{earn} = \begin{cases} 0 & ; \text{ if } wz_t^h l_t < \$10,080 \\ \min\left\{ ss(AE_t, j^R), \frac{wz_t^h l_t - \$10,080}{2} \right\} & ; \text{ otherwise} \end{cases},$$

i.e., for individuals whose earnings exceed an exempt amount (\$10,080 in 2000), \$1 of benefits is withheld for every \$2 of earnings in excess of the exempt amount. It is important to note that benefits withheld this way are not lost but go towards increasing the future benefits. More specifically, Social Security earning tax allows to partially offset the penalty for early

[6] The Social Security benefits are a function of the average earnings of the 35 years with the highest earnings. We use a simplified version of this rule because otherwise we have to keep track over the entire previous earnings history as additional state variables which makes our computation infeasible.

[7] For individuals born in 1936 and 1937 the full retirement age is 65 years, for individuals born in 1938 it is 65 years and 2 months.

[8] These numbers correspond to the annual benefits, they are derived by multiplying the bend points corresponding to monthly benefits by 12.

[9] Starting 2000, the Social Security earning tax for individuals who reach full retirement age was abolished.

claiming. For example, if an individual has all benefits withheld for the entire year his benefits will be adjusted as if he claimed them one year later. To avoid keeping track of withheld benefits as an additional state variables, we approximate these rules as follows. If more than 50% of benefits of an individual are withheld due to the earning tax we increase j^R by one year and we do not do any adjustments otherwise.[10]

4.1.2 Timing of the model

The timing of the model is as follows. In the beginning of the period, individuals learn their productivity and health status. Based on this information, an individual decides his labor supply (l_t). An individual who is older than age R^E also decides whether to claim Social Security benefits. We denote the decision regarding claiming as i_t^C; $i_t^C = 1$ if an individual claims benefits and $i_t^C = 0$ otherwise. After that, the out-of-pocket medical shock (x_t^h) is realized; for individuals older than age R^D the nursing home shock (xn_t) is realized. In the very end of the period, consumption/saving decisions are made. An individual who reaches age R^D and who still did not claim benefits, must claim benefits. Individuals after age R^D only make consumption/saving decisions.

4.1.3 Optimization problem

Individuals younger than the earliest claiming age $\left(t < R^E\right)$. The state variables for individuals younger than age R^E at the beginning of each period are capital ($k_t \in \mathbb{K} = R^+ \cup \{0\}$), health ($h_t \in \mathbb{H} = \{0,1\}$), idiosyncratic labor productivity $\left(z_t^h \in \mathbb{Z} = R^+\right)$, average lifetime earnings ($AE_t \in \mathbb{A} = R^+$), and age $\left(t \in \mathbb{T} = \ 1, 2, ..., R^E - 1\ \right)$. We denote the vector of state variables of an individual of age t as $\mathbb{S}_t$: $\mathbb{S}_t = (k_t, h_t, z_t^h, AE_t)$.

The value function of an individual in this age range can be written as follows:

$$V_t(\mathbb{S}_t) = \max_{l_t} \left\{ \sum_{x_t^h} \mathcal{G}_t\left(x_t^h\right) W_t(\mathbb{S}_t; l_t, x_t^h)^{1-\psi} \right\}^{\frac{1}{1-\psi}} \tag{2}$$

where

$$W_t(\mathbb{S}_t; l_t, x_t^h) = \max_{c_t, k_{t+1}} \left\{ \begin{array}{c} c_t^{\chi} \widetilde{l}_t^{1-\chi} \ ^{1-\gamma} + \\ \beta \left[\zeta_t^h E_t \left(V_{t+1}(\mathbb{S}_{t+1})\right)^{1-\psi} + (1-\zeta_t^h)\eta\left(k_{t+1} + \phi\right)^{1-\psi}\right]^{\frac{1-\gamma}{1-\psi}} \end{array} \right\}^{\frac{1}{1-\gamma}} \tag{3}$$

[10] Our results do not significantly change if we assume that the claiming age is rest only if 100% of benefits are withheld.

subject to

$$k_t\,(1+r) + wz_t^h l_t + T^{SI} = k_{t+1} + c_t + x_t^h + Tax \tag{4}$$

$$T_t^{SI} = \max\left(0, \underline{c} + x_t^h + Tax - k_t\,(1+r) - wz_t^h l_t\right) \tag{5}$$

$$Tax = \mathcal{T}\left(y_t^{tax}\right) + \tau_{ss}\min\left(wz_t^h l_t, \overline{y}_{ss}\right) + \tau_{MCR} wz_t^h l_t \tag{6}$$

$$y_t^{tax} = k_t r + wz_t^h l_t \tag{7}$$

The conditional expectation on the right-hand side of Eq (3) is over z_{t+1}^h and h_{t+1}. Eq (4) is the budget constraint. Eq (5) describes the means-tested transfers that provide the minimum consumption guarantee $\underline{c}$. In Eq (6), the first term is income tax and the last two terms are payroll taxes. Eq (7) describes the taxable income.

Individuals older than the earliest claiming age but younger that the latest claiming age $\left(R^E \leq t < R^D\right)$ and who still did not claim benefits. Individuals in this age range have to decide whether to claim Social Security benefits or not. Their value function can be written as follows:

$$V_t(\mathbb{S}_t) = \max_{l_t, i_t^C} \left\{ {}_{x_t^h} \mathcal{G}_t\left(x_t^h\right) W_t^E(\mathbb{S}_t; l_t, i_t^C, x_t^h)^{1-\psi} \right\}^{\frac{1}{1-\psi}} \tag{8}$$

$$W_t^E(\mathbb{S}_t; l_t, i_t^C = 0, x_t^h) = \max_{c_t, k_{t+1}} \left\{ \begin{array}{c} \left(c_t^{\chi} \widetilde{l}_t^{1-\chi}\right)^{1-\gamma} + \\ \beta\left(\zeta_t^h E_t\left(V_{t+1}(\mathbb{S}_{t+1})\right)^{1-\psi} + (1-\zeta_t^h)\eta\left(k_{t+1}+\phi\right)^{1-\psi}\right)^{\frac{1-\gamma}{1-\psi}} \end{array} \right\}^{\frac{1}{1-\gamma}}$$

$$W_t^E(\mathbb{S}_t; l_t, i_t^C = 1, x_t^h) = \max_{c_t, k_{t+1}} \left\{ \begin{array}{c} \left(c_t^{\chi} \widetilde{l}_t^{1-\chi}\right)^{1-\gamma} + \\ \beta\left(\zeta_t^h E_t\left(V_{t+1}^C(\mathbb{S}_{t+1}, j^R)\right)^{1-\psi} + (1-\zeta_t^h)\eta\left(k_{t+1}+\phi\right)^{1-\psi}\right)^{\frac{1-\gamma}{1-\psi}} \end{array} \right\}^{\frac{1}{1-\gamma}}$$

subject to

$$k_t\,(1+r) + wz_t^h l_t + ss(AE_t, t)\mathbf{1}_{\{i_t^C=1\}} + T^{SI} = k_{t+1} + c_t + x_t^h + Tax \tag{9}$$

$$T_t^{SI} = \max \quad 0, \underline{c} + x_t^h + Tax - k_t\,(1+r) - wz_t^h l_t - ss(AE_t, t)\mathbf{1}_{\{i_t^C=1\}}$$

$$Tax = \mathcal{T} \quad y_t^{tax} \quad + \tau_{ss} \min \; wz_t^h l_t, \overline{y}_{ss} \quad + \tau_{MCR} wz_t^h l_t + t^{earn}\mathbf{1}_{\{t<R^F, i_t^C=1, l_t=\bar{l}\}} \tag{10}$$

$$y_t^{tax} = k_t r + wz_t^h l_t + ss(AE_t, t)\mathbf{1}_{\{i_t^C=1\}}$$

$$j^R = \begin{cases} t & ;\ \text{if } t^{earn} < 0.5ss(AE_t, t) \\ t+1 & ;\ \text{otherwise} \end{cases} \tag{11}$$

Note that the interim value function W_t^E takes different form depending on whether an individual claims benefit or not; in the latter case, there will be another state variable next period: age at which he starts collecting benefits. Eq (9) includes Social Security benefits $ss(AE_t, t)$ for individuals who choose to collect benefits in the current period (i.e., $i_t^C = 1$). Eq (10) includes Social Security earning tax for individuals who are younger than the full retirement age and who claimed benefits but continue working. Eq (11) shows that the claiming age can be increased by one year for working individuals who claimed in the current period if most of the benefits are taxed away.

Individuals older than the earliest claiming age but younger than the latest claiming age $R^E \le t < R^D$ and who already claimed benefits. This category of individuals have an additional state variable j^R, the age at which they started collecting benefits. The value function of these individuals can be written as follows:

$$V_t^C(\mathbb{S}_t, j^R) = \max_{l_t} \left\{ \quad_{x_t^h} \mathcal{G}_t \quad x_t^h \quad W_t^C(\mathbb{S}_t, j^R; l_t, x_t^h)^{1-\psi} \right\}^{\frac{1}{1-\psi}} \tag{12}$$

$$W_t^C(\mathbb{S}_t, j^R; l_t, x_t^h) = \max_{c_t, k_{t+1}} \left\{ \begin{array}{c} c_t^{\chi} \widetilde{l}_t^{1-\chi} \quad {}^{1-\gamma} + \\ \beta \quad \zeta_t^h E_t \quad V_{t+1}^C(\mathbb{S}_{t+1}, \widetilde{j}^R) \quad {}^{1-\psi} + (1-\zeta_t^h)\eta\,(k_{t+1}+\phi)^{1-\psi} \quad {}^{\frac{1-\gamma}{1-\psi}} \end{array} \right\}^{\frac{1}{1-\gamma}}$$

subject to

$$k_t\,(1+r) + wz_t^h l_t + ss(AE_t, j^R) + T^{SI} = k_{t+1} + c_t + x_t^h + Tax \tag{13}$$

$$T_t^{SI} = \max \ 0, \underline{c} + x_t^h + Tax - k_t\,(1+r) - wz_t^h l_t - ss(AE_t, j^R)$$

$$Tax = \mathcal{T} \ \ y_t^{tax} \ \ + \tau_{ss} \min \ wz_t^h l_t, \overline{y}_{ss} \ \ + \tau_{MCR} wz_t^h l_t + t^{earn} \mathbf{1}_{\{t<R^F, i_t^C=1, l_t=\bar{l}\}}$$

$$y_t^{tax} = k_t r + wz_t^h l_t + ss(AE_t, j^R) \tag{14}$$

$$\widetilde{j^R} = \begin{cases} j^R & ; \text{ if } t^{earn} < 0.5ss(AE_t, j^R) \\ j^R + 1 & ; \text{ otherwise} \end{cases} \tag{15}$$

Note that the age at which an individual first claimed benefits j^R affects his pension income $ss(AE_t, j^R)$ but this age can be increased if he is subject to Social Security earning tax and most of his benefits are taxed away (Eq 15).

Individuals after age R^D Individuals after age R^D make only consumption-saving decisions and their state variables are capital (k_t), health (h_t), average lifetime earnings ($AE \in \mathbb{A} = R^+$), the age when they claimed benefits ($j^R \in \mathbb{J} = \ \ R^E, ..., R^D \ \ $), and age ($t$). Denote the vector of the state variables as $\mathbb{S}_t^R = (k_t, h_t, x_t^h, AE, j^R)$. The value function of these individuals is:

$$V_t^R(\mathbb{S}_t^R) = \left\{ \quad \mathcal{G}_t \ \ x_t^h \ \ pn_t^h W_t^R(\mathbb{S}_t^R; x_t^h, xn_t)^{1-\psi} \right\}^{\frac{1}{1-\psi}}$$

$x_t \quad xn_t$

where

$$W_t^R(\mathbb{S}_t^R; x_t^h, xn_t) = \max_{c_t, k_{t+1}} \left\{ \begin{array}{c} c_t^{\chi} \widetilde{l}_t^{1-\chi} \ \ {}^{1-\gamma} + \\ \beta \ \ \zeta_t^h E_t \ \ V_{t+1}^R(\mathbb{S}_{t+1}^R) \ \ {}^{1-\psi} + (1-\zeta_t^h)\eta\,(k_{t+1}+\phi)^{1-\psi} \ \ {}^{\frac{1-\gamma}{1-\psi}} \end{array} \right\}^{\frac{1}{1-\gamma}} \tag{16}$$

subject to:

$$k_t\,(1+r) + ss(AE, j^R) + T^{SI} = k_{t+1} + c_t + \mathcal{T} \ \ y_t^{tax} \ \ + x_t^h + xn_t$$

$$T_t^{SI} = \max \ 0, \underline{c} + \mathcal{T} \ \ y_t^{tax} \ \ + x_t^h + xn_t - k_t\,(1+r) - ss(AE, j^R)$$

$$y_t^{tax} = k_t r + ss(AE, j^R)$$

Note that the interim value function W_t^R is conditional on the realization of the out-of-pocket medical spending shock x_t^h and the nursing home shock xn_t.

5 Data and calibration

5.1 Data and sample selection

An ideal dataset for our study would be a representative panel that tracks individuals over the entire life-cycle and includes information on labor supply, labor income, savings, medical spending, and Social Security claiming behavior. Since a dataset like this does not exist for the US, we combine information from the three datasets: the Medical Expenditure Panel Survey (MEPS), the Health and Retirement Study (HRS) and the Panel Study of Income Dynamics (PSID). In all three datasets, we select a sample of male individuals. We use 2002 as the base year, and all level variables were normalized to the base year using the Consumer Price Index (CPI).

The MEPS is a nationally representative survey of households with a particular focus on medical usage and health insurance variables. It contains individuals of all ages but age is top-coded at 85. The MEPS has a short panel dimension: each individual is observed for at most two years. Medical spending reported in the MEPS are cross-checked with insurers and providers which improves their accuracy.[11] We use the MEPS to construct data moments related to medical spending (except for nursing home spending), health, labor income, and employment.[12] We use fourteen waves of the MEPS from 1999 to 2012. We construct a sample of male individuals who are at least 20 years old. There are 80,984 individuals (or 152,308 individual-year observations).

The HRS is a nationally representative sample of individuals over the age of 50. We use the RAND Version P of this dataset to construct moments related to claiming behavior and nursing home costs. When constructing claiming behavior, we use males born around 1937 as our base cohort. We choose this cohort because we need to consider individuals that i) face similar rules regarding early/late claiming benefit adjustments, ii) are entirely retired (or are older than 70) by the last wave of the HRS we considered. To increase the number of observations, we use the window of 3 years, i.e., we consider all males born in

[11] Pashchenko and Porapakkarm (2016) provide more details on the MEPS dataset.

[12] The MEPS dose not contain information on nursing home spending because it only contains non-institutionalized population and thus excludes nursing home residents.

years 1936-1938, which leaves us with 864 individuals. To construct moments related to nursing home costs, we use a larger sample by pooling together waves 2002-2012 of HRS. We use a sample of individuals older than 70 that do not have missing information on nursing home use, health or age. This leaves us with 8,546 individuals or 35,487 individual-year observations.

The PSID is a national representative panel survey of individuals and their families. It started in 1968 on an annual basis and from 1997 it is administered bi-annually. We use the PSID to construct data moments related to wealth accumulation.

5.2 Demographics, preferences and technology

In the model, agents are born at age 25 and can live to a maximum age of 99. For survival probabilities, we use the cohort life table for men born in 1940 provided by the Social Security Administration.[13] To adjust conditional survival probabilities ζ_t^h for the difference in health, we follow Attanasio et al. (2011). In particular, we use the HRS to estimate the difference in survival probabilities for people in different health categories. Specifically, we use the HRS data to estimate the survival probability as a function of a cubic polynomial of age, using a probit model for each health status. Then, we compute the *survival premium* - the difference between the estimated survival probabilities of healthy and unhealthy males for each age. From the Social Security Administration cohort life table, we know the av age survival probability of males. From the MEPS, we can construct the fraction of people in each health category for each age. Using this information, we can recover the survival probabilities of healthy and unhealthy people for each age.

We set the consumption share in the utility function χ to 0.5 to facilitate matching the employment profile. This number is in the range estimated by French (2005).[14] We set the labor supply of those who choose to work ($\bar{l}$) to 0.4. We define a person as employed if he earns at least $2,678 per year in base year dollars (this corresponds to working at least 10 hours per week and earning a minimum wage of $5.15 per hour). Fixed leisure costs of work ϕ_w is calibrated to match the age profile of the employment.

A common approach in structural life-cycle and macroeconomic models is to set the discount factor β to match the wealth accumulation over the life-cycle or the aggregate wealth to income ratio. This approach is justified by the fact that saving decisions of individuals are very responsive to the value of the discount factor. However, the studies that use wealth

[13] The Social Security Administration publishes cohort life tables with ten-year intervals, i.e., for individuals born in 1930,1940, etc. We use the cohort born in 1940 since it is the closest to our base cohort's birth year (1937).

[14] Given that we have an indivisible labor supply, we cannot pin down this parameter using a moment in the data.

accumulation decisions to identify the discount factor usually abstract from decisions to claim Social Security benefits.[15] In our model, the decisions to claim Social Security benefits are endogenous and we show that they are strongly affected by the value of the discount factor. Because of this, we adjust the discount factor to match the percentage of people claiming benefits at the earliest possible age (62). The resulting β is 0.962. In Section 6.1, we discuss our identification of the discount factor in more details.

Because in our calibration strategy the discount factor is used to match the claiming behavior, we are left with four parameters to match the wealth profiles over the life-cycle: the intertemporal elasticity of substitution (IES, $\frac{1}{\gamma}$), the risk aversion (ψ), the strength of the bequest motive (η) and the degree to which bequest is a luxury good (ϕ). The risk aversion and the IES affect the wealth accumulation over the working stage of the life-cycle; we set the risk aversion to 4 and 1/IES equal to 1.5. It is important to note that to match wealth accumulation profile we need to set the risk aversion to relatively high number and to make it significantly different from 1/IES. The risk aversion by itself have limited power to affect wealth accumulation if it is tied to be equal to the inverse of the IES because even though higher risk aversion results in stronger precautionary motive it also implies lower IES. Low IES increases preferences for flatter consumption profile and thus flattens wealth accumulation profile. To break this relationship, we need to spread the risk aversion and the inverse of the IES apart.

After the middle age bequest motives have a strong impact on wealth accumulation/ deccumulation. The bequest function that we use implies that bequests is a luxury good, i.e., the bequest motives become operational only when individuals' assets are above a certain threshold, in which case the amount of assets they allocate to bequests is controlled by the marginal propensity to bequeath (MPB). The threshold and the MPB can be expressed as functions of the parameters η and ϕ in a simple two-period consumption-savings model (more on this see De Nardi et al. (2010) and Pashchenko (2013)). We adjust the threshold to match the wealth profiles of individuals in the bottom 25 percentile of the wealth distribution and we adjust the MPB to match the profiles for the median and the 75th percentile. The resulting numbers are \$3,605 for the threshold and 0.969 for the MPB.[16]

[15] Gustman and Steinmeier (2005 and 2015) represent an exception: they allow for endogenous claiming decisions but still use wealth profiles to identify the discount factor. They show that the model where the discount factor is identified this way falls short of replicating the claiming decisions as in the data.

[16] The corresponding values of η and ϕ are 2.4^{11} and $115,000$, respectively.

5.3 Health, medical expense and nursing home shocks

To construct our health measure, we use self-reported health status reported in the MEPS. In the MEPS an individual's self-reported health status is coded as 1 for excellent, 2 for very good, 3 for good, 4 for fair and 5 for poor. Individuals in MEPS are interviewed five times over two-year period and the question about health is asked in every interview round. We classify a person as being in bad health if his average health score over that year is greater than 3.

To construct the age-dependent health transition matrix, we first compute the transition matrices for ages 30, 40,...70. In each case, we use a sample in a 10-year age bracket. For example, to construct the transition matrix for age 40, we pool individuals between ages 35 and 44. Then we construct the health transition matrix for all the remaining ages by using the polynomial degree two approximation.

Medical expenses in our model correspond to the out-of-pocket medical expenditures in the MEPS dataset. In our calibration, medical expense shock is approximated by a 3-state discrete health- and age-dependent stochastic process.[17] For each age and health status, these three states correspond to the average out-of-pocket medical expenses of the three groups: those with out-of-pocket medical spending below the 50th, 50th to 95th, and above 95th percentiles, respectively.[18] To construct the transition matrix, we measure the fraction of people who move from one group to another between two consecutive years separately for those between ages 25 and 64 and for those who are 65 and older.

We estimate the risk of incurring nursing home shock (pn_t^h) from the HRS as follows. First, we compute the probabilities to enter a nursing home for selected ages: 67, 72, 77, 82, 87, and 95. In each case, we use a sample in a 5-year age bracket. To do this, we compute the percentage of individuals reporting staying in a nursing home in each interview round for the following age groups: 65-69, 70-74, 75-79, 80-84, 85-89, and older than 90. Since the HRS is a bi-annual survey, we convert these numbers into the annual probabilities under the assumption that the probability to stay in a nursing home over the two-year interval is equal to the product of the annual probabilities. Then we extrapolate the probability to

[17] Given short panel structure of MEPS (2 years), we cannot estimate the medical shock process parametrically.

[18] The MEPS tends to underestimate aggregate medical expenditures (Pashchenko and Porapakkarm, 2016). The ratio of aggregate medical spending in the National Health Expenditure Account (NHEA) divided by aggregate medical spending in MEPS for people younger and older than 65 years old constitute 1.6 and 1.9, respectively. These numbers were computed by averaging over the years 2002, 2004, 2006, 2008, and 2010 (the years when NHEA provides the aggregate statistics by age). The larger discrepancy for the older group is due to the fact that MEPS does not include nursing home expenditures. To bring aggregate medical expenses computed from the MEPS in line with the corresponding statistics in the NHEA, we multiply our estimated medical expenses by 1.60. We use this number because we explicitly account for nursing home spending in our model.

stay in a nursing home for other ages using polynomial degree three approximation.We do this separately for healthy and unhealthy males. The HRS also reports the number of nights over all nursing home stays. To compute the average nursing home costs, we multiply the number of nights by the average daily rate for a semiprivate room in a nursing home, which was $158.26 in 2003 Metlife (2003).[19]

5.4 Taxes and government transfers

We parameterize the tax function $\mathcal{T}(y)$ following Gouveia and Strauss (1994):

$$\mathcal{T}(y) = a_0 \left[y - (y^{-a_1} + a_2)^{-1/a_1} \right]$$

As in Gouveia and Strauss (1994), we set a_0 and a_1 to 0.258 and 0.768, respectively. We set the parameter a_2 to 0.616 following Pashchenko and Porapakkarm (2013).

The Medicare, Social Security and consumption tax rates were set to 2.9 percent, 12.4 percent and 5.67 percent, respectively. The maximum taxable income for Social Security ($\overline{y}_{ss}$) is set to $76,200 (corresponding to year 2000).

When calibrating the consumption minimum floor $\underline{c}$, we use the fact that this safety net has a significant effect on labor supply of individuals with low assets, such as the young. We set the minimum consumption floor to $3,500 to match the employment rate among individuals in the age group 25-29 years old. Our estimate of the consumption floor is in line with with other models with medical expense shocks that consider the entire life-cycle (e.g. Capatina, 2015).

5.5 Labor productivity process

We specify the individual productivity as following:

$$z_t^h = \lambda_t^h \Upsilon_t = \lambda_t^h \exp(v_t) \exp(\xi) \tag{17}$$

where λ_t^h is the deterministic component that depends on age and health; the stochastic component of productivity Υ_t consists of the persistent shock v_t and a fixed productivity type ξ:

$$v_t = \rho v_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N(0, \sigma_\varepsilon^2) \tag{18}$$

$$\xi \sim N(0, \sigma_\xi^2)$$

[19] The MetLife Market Survey of Nursing Home and Home Care Costs, August 2003 available at http://www.lifestyleinsurance.com/media/2003%20NHHC%20Market%20survey.pdf

For the persistent shock v_t, we set ρ to 0.98 and σ_ε^2 to 0.02 following the incomplete market literature (Storesletten et al., 2004; Hubbard et al., 1994; French, 2005). We set the variance of the fixed productivity type (σ_ξ^2) to 0.242 as in Storesletten et al. (2004). In our computation, we discretize the shock processes using 9 gridpoints for v_t and 2 gridpoints for ξ. To construct the distribution of individuals just entering the model, we draw v_1 in Eq (18) from the $N(0, 0.352^2)$ distribution following Heathcote et al. (2010).

To estimate the deterministic part of productivity λ_t^h, we need to take into account the fact that we only observe labor income of workers and we do not know the potential labor income of non-workers, which are not necessarily the same because there can be a selection into employment. To avoid the selection bias, we adapt the method developed by French (2005). We start by estimating the labor income profiles from the MEPS dataset for all workers. Then, given other parameters of the model, we guess λ_t^h in Eq. (17). Next, we feed the resulting productivity into our model. After solving and simulating the model, we compute the average labor income profile of workers in our model and compare it with the data. We update our guess and reiterate until the labor income and the employment profiles in the model are the same as in the data.

We set the wage rate w so that the level of the average earnings in our model is the same as in the data. The model parametrization is summarized in Table 5 in Appendix A.

5.6 Baseline model performance

Figure (2) compares the employment profile (left panel) and the average labor income of workers (right panel) in the data and in the model. The model closely tracks the data. The average labor income profiles and employment profiles were targeted in our calibration by adjusting the exogenous productivity, the disutility from work parameter and the consumption floor.

The left panel of Figure (3) shows that our calibration strategy of adjusting the risk aversion, IES, and the bequest function parameters allows us to capture the wealth profiles for the bottom 25th percentile, median and top 25th percentile constructed from the data.

The right panel of Figure (3) compares the claiming behavior in our model with the cohort born in 1936-1938 in the data. In our calibration, we target the percentage of individuals who start collecting Social Security benefits as early as possible (at age 62) but the model is able to capture the overall pattern of claiming for other ages as well.

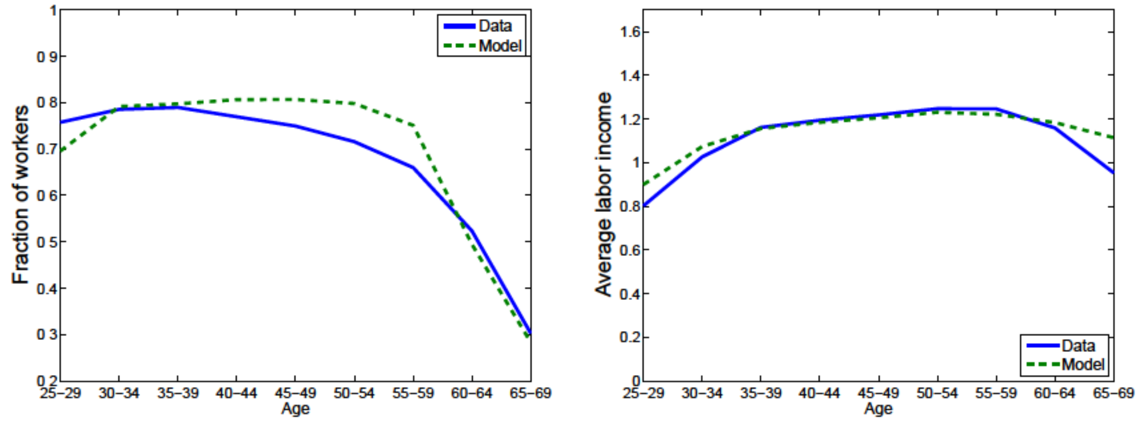

Figure 2: Left panel: fraction of workers by age. Right panel: average income among workers by age

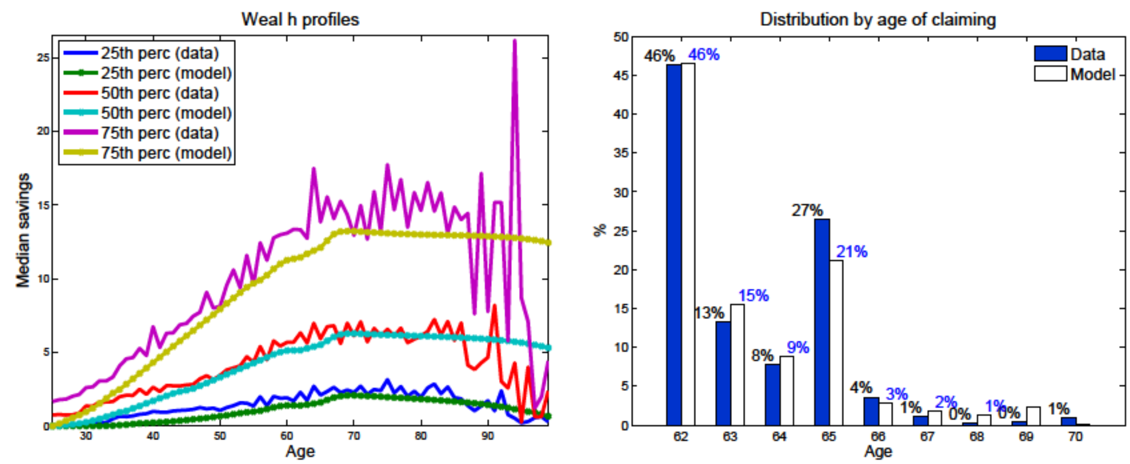

Figure 3: Left panel: wealth profiles by age. Right panel: distribution by claiming age

6 Results

This section is organized as follows. We start by illustrating the role of the discount factor in individuals' decisions at what age to claim Social Security benefits. Then we investigate the role of the price of Social Security annuity in claiming decisions. Next, we consider how factors previously shown to matter for the decisions to annuitize through the private market affect claiming behavior. Finally, we consider a policy experiment where we allow individuals who delay claiming to receive the resulting increase in their pension income as lump-sum transfers.

6.1 The role of the discount factor

To illustrate the role of the discount factor in claiming decisions, in this section we consider two alternative versions of the model. In the first version, the discount factor is set to 0.95 which is lower than the baseline value of 0.962. In the second version, we set the discount factor to a higher value of 0.97. Then, we recalibrate our model until we match the wealth and employment profiles as in the data.

Figure (4) shows that for each of the alternative discount factors, the model parameters can be adjusted to match the employment and wealth profiles as in the data. However, as shown in Figure (5), these two versions of the model fail to account for the Social Security claiming behavior, especially for the percentage of individuals claiming at age 62. In particular, the model with the low discount factor produces too many people claiming at age 62 (56% in the alternative model versus 46% in the baseline), while in the model with the high discount factor too few people claim at age 62 (31% in the alternative model versus 46% in the baseline).

The intuition for why the discount factor plays a key role in accounting for claiming behavior is as follows. Individuals who delay claiming are 'purchasing' Social Security annuity, i.e., they forgo current benefits to increase the future stream of income. The subjective valuation of this extra stream of income crucially depends on the discount factor used. If the discount factor is low, this stream of income is valued less and individuals choose to forgo this annuity and claim as early as possible.

Thus, given how much annuity income individuals are entitled to at the earliest claiming age and given the price of buying additional annuity they prefer to forgo the option to increase their annuity-like income. We illustrate this point further in the next section when we show how people respond to the change in the in est rate used to compute the price of social security annuity.

6.2 The role of the annuity price

In Section 3 we show that the price of the Social Security annuity differs from the actuarially fair price for individuals with average mortality (see Figure 1). To understand how this affects the demand for this annuity, in this section we consider a counterfactual experiment where the schedule of penalties/rewards for early/late claiming is changed in a way that the Social Security annuity is actuarially fair for individuals with average mortality.

The third row of Table 2 displays the resulting benefits' adjustments when we use the interest rate of 2% to compute the underlying actuarially fair price (the interest rate of 2% is used in our baseline calibration). We explain the details of this computation in Appendix B.

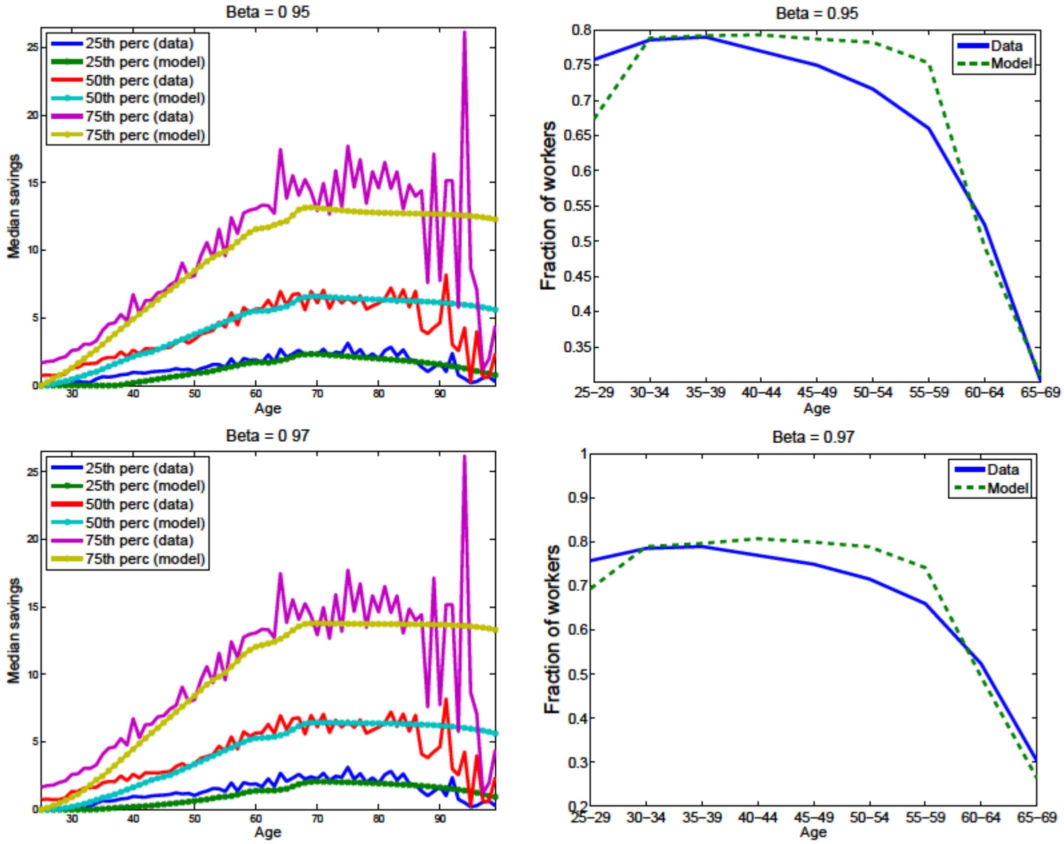

Figure 4: Wealth and employment profiles in the two versions of the model. Top panel: the discount factor is lower than in the baseline (0.95). Bottom panel: the discount factor is higher than in the baseline (0.97)

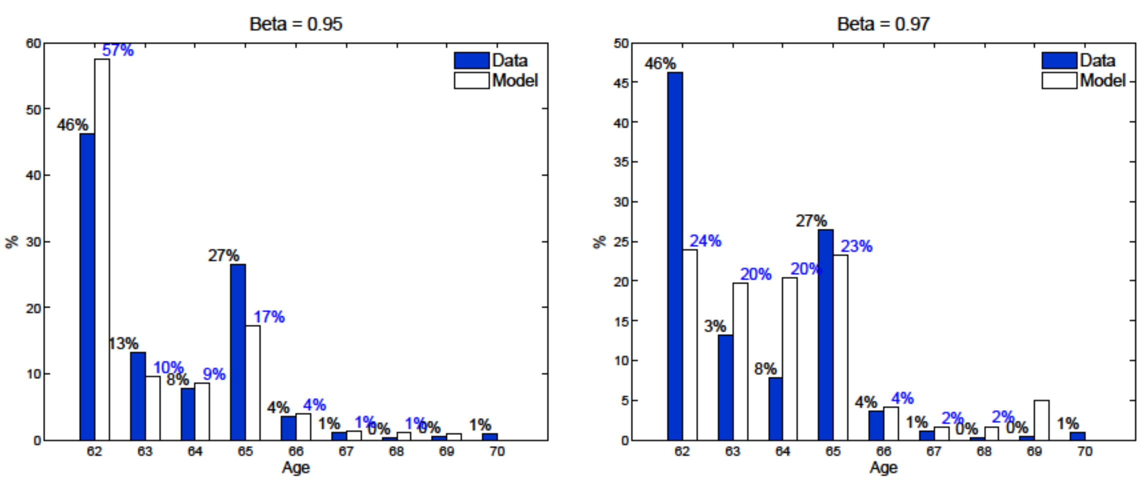

Figure 5: Distribution by claiming age in the two versions of the model. Left panel: the discount factor is lower than in the baseline (0.95). Right panel: the discount factor is higher than in the baseline (0.97)

Note, that compared with the actual schedule of penalties/rewards, the actuarially fair one implies smaller penalty for early claiming and significantly higher reward for late claiming.

Age	62	63	64	65	66	67	68	69	70
Actual	0.8	0.867	0.933	1.0	1.065	1.13	1.195	1.26	1.325
Counterf ($r = 2$)	0.816	0.872	0.933	1.0	1.075	1.158	1.252	1.356	1.475
Counterf ($r = 5$)	0.766	0.835	0.913	1.0	1.098	1.209	1.334	1.477	1.640

Table 2: The counterfactual adjustments of benefits for early/late claiming that results in the actuarial fairness of the Social Security annuity

The left panel of Figure (6) shows the distribution by claiming age of individuals who face the schedule of penalties/rewards corresponding to this actuarially fair annuity price. An important observation from this figure is that at age 62, individuals are more price-elastic than after the full retirement age. A small increase in the annuity price at age 62 leads to a large increase in the percentage of individuals claiming as early as possible: from 46% (baseline) to 56%. At the same time, a decline in the annuity price for individuals after the full retirement age leads to a relatively small increase in the number of late claimers: in total, the percentage of individuals claiming after 65 increase from 8% (baseline) to 10%.

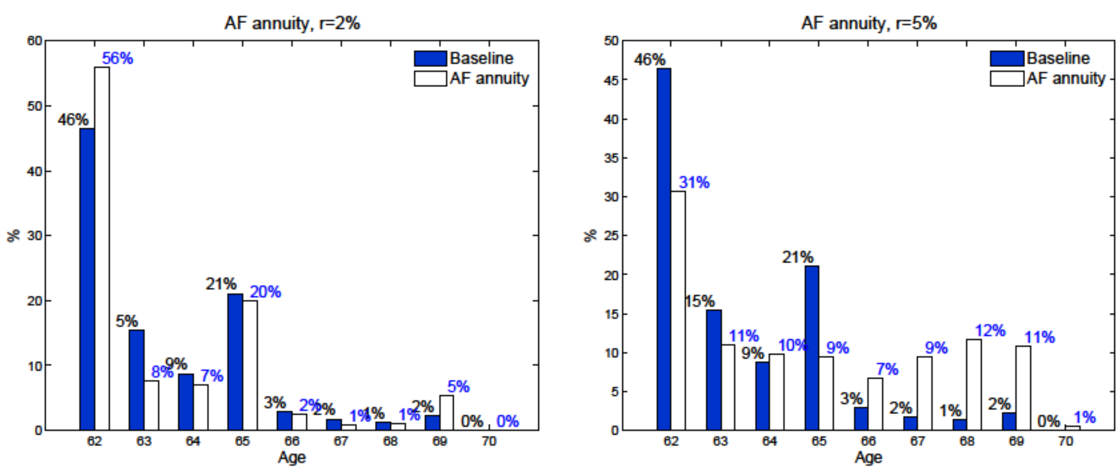

Figure 6: Distribution by claiming ages in two versions of the model when SS annuity is priced actuarially fair. Left panel: the price is based on the interest rate of 2%. Right panel: the price is based on the interest rate of 5%

This experiment shows that if Social Security annuity is priced actuarially fair more people claim at age 62 without a significant increase in the percentage of late claimers. This to some extent can be due to the fact that the actuarially fair price is computed based on the interest rate of 2% meaning that the future stream of annuity income is discounted based on the factor $1/(1 + 0.02) = 0.98$. In our model, individuals discount the future at a lower rate of 0.962. In other words, the annuity is priced based on a lower rate (2%) than the rate of time preferences which is equal to $1/\beta - 1 = 3.95\%$.

To understand what role this plays for annuity demand, we consider an experiment where Social Security annuity is priced actuarially fair but based on the interest rate that is above the rate of time preferences of individuals. The fourth row of Table 2 displays the penalties/rewards schedule that corresponds to actuarially fair annuity based on the interest rate of 5% and the right panel of Figure (6) shows the corresponding claiming behavior. Note that when the interest rate used to price annuity is above the rate of time preferences, demand for this annuity significantly increases. The percentage of people claiming as early as possible decreases from 46% to 31%, and the percentage of individuals claiming after the full retirement age increases from 8% to 39%. Thus, changing the way Social Security annuity is priced can lead to a significant delay in claiming but for this the interest rate use for pricing this annuity should be above the (relatively high) individuals' rate of time preferences.

6.3 The role of various impediments to annuitization

In this section, we consider several factors that were previously shown to affect annuity demand in order to understand their quantitative importance for claiming decisions. Specifically, we consider the following potential impediments to annuitization: medical spending (Turra and Mitchell, 2008), means-tested benefits (Pashchenko, 2013), pre-annuitized wealth (Dushi and Webb, 2004), and bequest motives (Lockwood, 2012).

Medical spending Davidoff et al. (2005) show theoretically that uncertain medical expenses can affect demand for annuities and the direction of this effect depends on the timing of the risk: a medical spending risk early in life can decrease demand for annuities while late in life it can produce the opposite effect. To understand the effect of uncertain medical expenses on demand for Social Security annuity, we consider the experiment when both medical and nursing home shocks are set to zero.

Panel (a) in Figure (7) and the first column in Table 3 illustrate the results of this experiment. Notice that the percentage of individuals claiming as early as possible increases (from 46% in the baseline to 53%), in other words, without medical expenses there is even less demand for annuities provided by Social Security.

This happens because midical spending increases quickly with age, i.e., medical risk is concentrated late in life. Thus, when an individual survives until very old age (an insurable event for annuities) this likely coincides with a situation when an individual faces high medical or nursing home spending. This complementarity makes Social Security annuity more valuable in presence of medical spending.

Means-tested benefits Pashchenko (2013) points out that means-tested benefits can decrease the demand for private annuities because they can de facto represent annuity-like income for individuals who outlive their assets and thus crowd-out demand for private longevity insurance. In our model, means-tested benefits are modeled as the consumption minimum floor that is set to \$3,500 in the baseline calibration. To understand the importance of this public program in determining claiming behavior, we consider an experiment where the consumption floor is decreased to \$2,000, and its results are displayed in the Panel (b) of Figure (7) and the second column of Table 3.

Lowering the consumption floor does not have effect on the percentage of individuals claiming as early as possible but more people start claiming at age 63 and less after the full retirement age. Note that since we are considering a full life-cycle model, the means-tested programs matter not only for decisions to annuitize but also for labor supply and savings decisions early in life. Individuals facing less generous consumption floor work and save more and arrive at the retirement stage with more assets, they can also afford to retire earlier which explains a small shift towards early claiming. This experiment shows the importance of taking the early stage of life into account when considering the effect of public insurance programs on late-life decisions. Individuals who face less generous public support adjust their behavior over working stage of the life-cycle and this can have more impact on their demand for public annuities that the change in the insurance arrangements per se.

% claiming	Baseline	No medical shocks	Low $\overline{c}$	Low SS benefits	High bequest threshold	Low bequest strength
Early (62-64)	71	76	76	58	66	43
FRA (65)	21	18	20	33	23	21
Late (66-70)	8	6	4	9	12	35

Table 3: Percentage of individuals claiming early, late and at the full retirement age (FRA) in the baseline vs counterfactuals

Pre-annuitized wealth Even individuals who claim Social Security benefits as early as possible get entitled to a stream of life-time income, i.e., they already have part of their lifetime wealth annuitized. One reason behind the reluctance to delay claiming can be that the fraction of this pre-annuitized wealth is already high and individuals do not want to increase it any further.

To understand the role of this factor in claiming decisions, we consider an experiment where we scale down the Social Security program. Specifically, we assume that individuals pay twice lower payroll tax (6.2% as opposed to 12.4%) and receive twice lower benefits, i.e.,

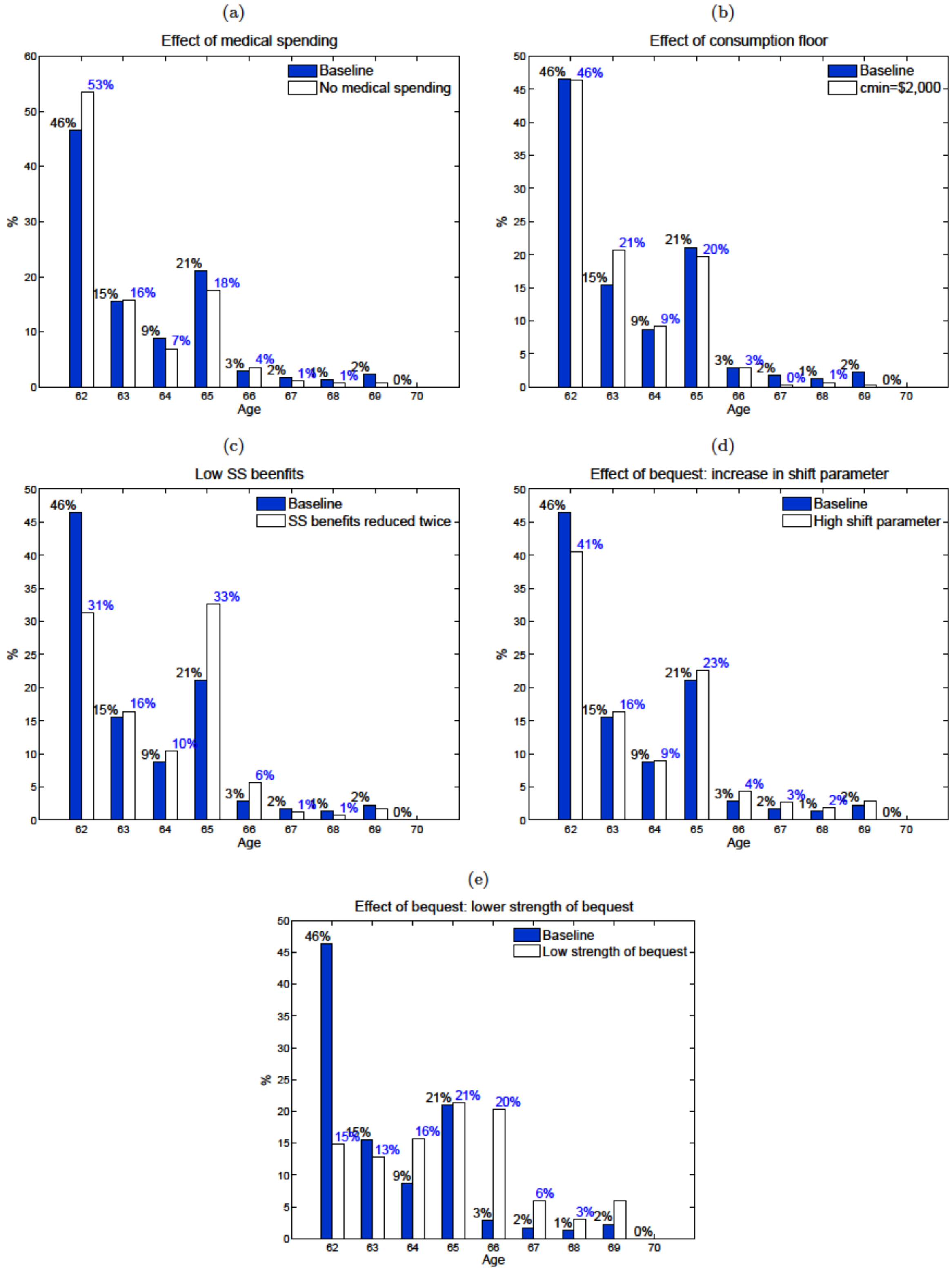

Figure 7: Distribution by claiming age in the baseline vs counterfactuals

the basic benefits ss^b in Eq (1) are multiplied by 0.5.

Panel (c) in Figure (7) and the third column in Table 3 illustrate the results of this experiment. There is a noticeable decline in the number of early claimers: the percentage of people claiming before the full retirement age decreases from 71% (baseline) to 58%. Thus, individuals who are entitled to lower annuity income from the beginning are more interested in acquiring additional lifelong income by delaying claiming.

Bequest motives A key property of an annuity is that it only pays out in the state when an individual is alive; this can be a serious drawback for an individual who cares about the state when he is not alive because he has bequest motives. To understand how this mechanism affects claiming decisions, we consider two experiments with weaker bequest motives. In the first experiment, we increase the bequest threshold (i.e., the level of assets above which bequest motive is operational) from the baseline level of $3,600 to $6,000 while keeping the MPB unchanged. In the second experiment, we decrease the MPB from the baseline level of 0.97 to 0.95 while keeping the threshold unchanged. Note that in the first experiment, bequest motive af smaller group of people who are relatively rich; while in the second experiment, bequest motive is less strong but it is operational at the similar level of assets as in the baseline.

The results of these experiments are presented in Panels (d) and (e) of Figure (7) and the fourth and fifth columns of Table 3. Note that in both cases, the demand for public annuity increases but the effect is significantly more pronounced for the case with lower MPB: the percentage of people claiming before the full retirement age goes down from 71% (baseline) to 66% in the first experiment and to 43% in the second one. Moreover, in the second case the percentage of individuals claiming as early as possible declines almost three times to 15%. Thus, bequest motive represents a quantitatively important factor in explaining low demand for Social Security annuity.

6.4 Policy implications: lump-sum option

Our results in the previous subsections show that individuals are not willing to acquire public annuity by delaying claiming because they have relatively high discount factor, strong bequest motives, and are already well-annuitized, i.e., they have substantial annuity income even if they claim at the earliest possible age. One policy implication of these findings is that to incentivize individuals to delay claiming it is important to take into account this unwillingness to annuitize and offer alternative rewards for delaying claiming. One way to do it is to substitute the increase in future pension with lump-sum transfers. Specifically, in the current environment individuals who delay claiming are offered a lifetime annuity;

instead, they can be offered a lump-sum transfer equivalent to the present discounted value of this annuity.

To understand the quantitative implications of this policy, we consider an experiment where an individual who is entitled to the basic retirement benefits of b at age 65 and who claims at age m is offered a lump-sum transfer LS_m that is determined as follows:

$$LS_m = \begin{cases} \sum_{t=m}^{T-1} \frac{\zeta_{t+1|m} 0.067b}{(1+r)^{t+1-m}} & ; \text{ if } m = 63, 64 \\ \sum_{t=m}^{T-1} \frac{\zeta_{t+1|m} 0.065b}{(1+r)^{t+1-m}} & ; \text{ if } m = 65, .., 70 \end{cases}$$

Note that the difference in transfers for individuals below and above the full retirement age arises because the accrual in extra pension income for each year of delay is higher for the former group than for the latter one ($0.067b$ vs $0.065b$).

The left panel of Figure (8) and the second column of Table 4 display the results of this experiment when the interest rate used to compute the present discounted value of additional pension income is set to 2%, which is the same value as in our baseline calibration. This policy option results in a large change in the pattern of claiming behavior: the percentage of individuals claiming as early as possible drops from the baseline 46% to only 3%, at the same time the percentage claiming as late as possible (age 70) increases from almost zero to 25%. Overall, the percentage of individuals claiming before the full retirement age decreases from 71% in the baseline to 11%, and the percentage claiming after the full retirement age increases from 8% in the baseline to 83%. This illustrates that individuals value the lump-sum option significantly more than increase in future pension benefits and are willing to delay claiming if this option is offered. Note that this result is consistent with the findings of Maurer et al. (2016) who using survey responses to a specifically designed set of questions show that individuals would be willing to delay claiming if they are offered a lump-sum option.

	Baseline	$r = 2$	$r = 5$
Early claim (62-64)	71%	11%	22%
Full retirement age (65)	21%	7%	9%
Late claim (66-70)	8	83%	69%

Table 4: The effects of offering lump-sum transfers for delaying claiming

It is worth stressing that when the increase in pension benefits is converted to lump-sum transfers the interest rate plays an important role, i.e., the higher is the interest rate the

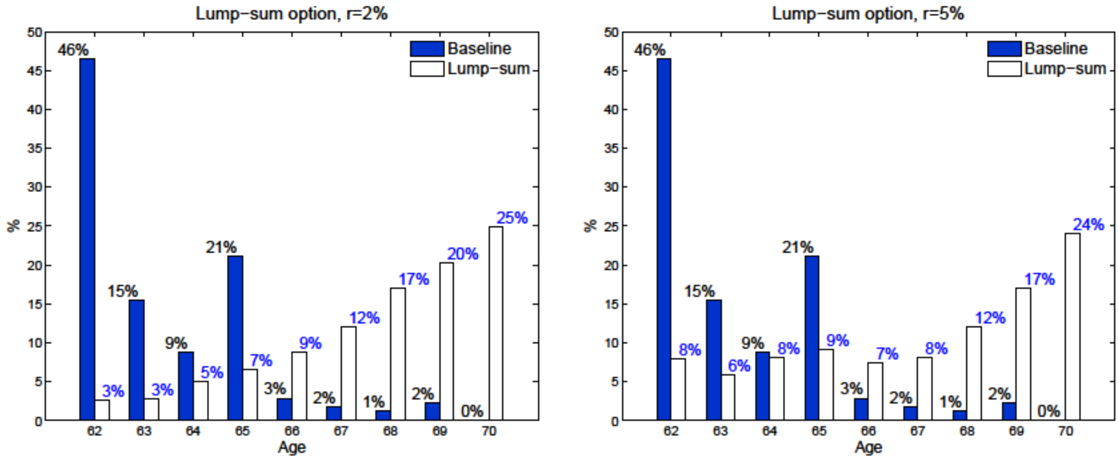

Figure 8: Distribution by claiming age when the lump-sum option is offered. Left panel: the lump-sum transfers are based on the interest rate of 2%. Right panel: the lump-sum transfers are based on the interest rate of 5%

lower is the computed present discounted value of pension income and thus the lower are the lump-sum benefits. To check the sensitivity of this policy to the interest rate used, we consider a case with a higher interest rate. As discussed in Section 6.2, the interest rate of 2% is below the rate of time preferences in our model (which is around 4%) and this matters for the valuation of annuity income. We next compute the lump-sum benefits using the discount rate of 5%.

The right panel of Figure (8) and the third column of Table 4 show the resulting claiming behavior. Compared with the case of a 2% interest rate, more people are claiming before the full retirement age (22% vs 11%) and less after that age (69% vs 83%). However, the pattern of claiming is still remarkably different from the baseline case because most people prefer to collect benefits after the full retirement age. This suggests that even if annuity income is converted to lump-sum transfers at a high interest rate, this policy provides much stronger incentives to delay claiming than an increase in pension income.

7 Conclusion

In this paper, we construct a rich structural model with heterogeneous agents in order to analyze individuals' decisions about when to claim Social Security benefits. An important fact from the data is that despite relatively large increases in pension benefits for those who delay claiming, most individuals claim before the full retirement age. Moreover, delay in claiming is equivalent to purchasing additional annuity income because individuals forgo benefits for one year so they can be entitled to a higher lifetime stream of benefits in the future; thus, the prevalence of early claiming is equivalent to low demand for the Social Security annuity.

We show that this Social Security annuity is priced below the actuarially fair level for individuals younger than the full retirement age and with the average mortality, making it even more puzzling why the demand for it is so low. We show that several versions of our calibrated model can be consistent with the observed labor supply and savings decisions over the life-cycle, but to simultaneously account for the claiming behavior, a model has to feature a relatively low discount factor. Put differently, the unwillingness of individuals to annuitize in the current settings can be rationalized by the fact that the present value of the extra annuity income obtained by delaying claiming is too low when their rate of time preferences is taken into account. We also show, that if individuals are offered actuarially fair annuity prices based on the discount rate that exceeds their rate of time preferences, substantially more people will delay claiming.

We use our model to investigate how commonly named impediments to annuitization such as means-tested benefits, medical spending, pre-annuitized wealth and bequest motives affect claiming decisions. We show that the latter two impediments have the highest quantitative impact, i.e., even individuals who claim as early as possible (age 62) have substantial annuity income and they may prefer not to annuitize any further, especially given their desire to leave bequests. We show that medical spending uncertainty increases the demand for public annuities since this risk is especially high late in life and annuities can be a better financial instrument to transfer resources to this stage of the life-cycle. Means-tested transfers have limited impact on claiming behavior: less generous public transfers make individuals more vulnerable to the risk of outliving their assets but they prepare for this starting from younger ages by accumulating more resources rather than acquiring more public annuities.

Our policy analysis shows that rewarding individuals for delaying claiming with lump-sum benefits is substantially more effective than offering them higher pension income. Individuals value additional resources they can obtain by postponing claiming but they strongly prefer them to be paid immediately as opposed to spreading them over their remaining lifetimes.

Overall, our results suggest that a well-documented unwillingness of individuals to purchase private annuities (the so-called annuity puzzle) can be interpreted more broadly in the context of the observed demand for the Social Security annuity. The low interest in any type of annuities suggests that to incentivize individuals to claim later, offering them extra annuity income may not be the best option, and a more flexible structure of rewards such as lump-sum benefits can be used.

References

[1] Blandin, A., 2015. Reforming the Social Security Earnings Cap: The Role of Endogenous Human Capital. Mimeo, Arizona State University

[2] Bronshtein, G., Scott, J., Shoven, J., Slavov, S., 2016. Leaving Big Money on the Table: Arbitrage Opportunities in Delaying Social Security. NBER Working Paper

[3] Coile, C, Diamond, P., Gruber, J., Jousten, A., 2002. Delays in Claiming Social Security Benefits. Journal of Public Economics, 84(3), pp 357-385

[4] Davidoff, T., Brown, J., Diamond, P., 2005. Annuities and Individual Welfare. American Economic Review, 95(5), pp 1573-1590

[5] Dushi, I., Webb, A., 2004. Household Annuitization Decisions: Simulations and Empirical Analysis. Journal of Pension Economics and Finance 3(2), 109-143

[6] French, E., Jones, J., 2011. The Effects of Health Insurance and Self-Insurance on Retirement Behavior. Econometrica, 79(3), 693-732

[7] Goda, G.S., Ramnath, S., Shoven, J., Slavov, S., 2015. The Financial Feasibility of Delaying Social Security: Evidence from Administrative Tax Data. NBER Working Paper

[8] Gustman, A., Steinmeier, T., 2005. The Social Security Early Retirement Age In a Structural Model of Retirement and Wealth. Journal of Public Economics, 89(23), pp. 441-463

[9] Gustman, A., Steinmeier, T., 2015. Effects of Social Security Policies on Benefit Claiming, Retirement and Saving. Journal of Public Economics, 129, pp.1-62

[10] Heathcote, J., Storesletten, K., Violante, G., 2010. The Macroeconomic Implications of Rising Wage Inequality in the United States. Journal of Political Economy, vol. 118(4), pp.681-722

[11] Hubbard, G., Skinner, J., Zeldes, S., 1994. The Importance of Precautionary Motives in Explaining Individual and Aggregate Saving. Carnegie-Rochester Conference Series on Public Policy, 40(1), pp.59-125

[12] Hubener, A., Maurer, R., Mitchell, O., 2016. How Family Status and Social Security Claiming Options Shape Optimal Life Cycle Portfolios. The Review of Financial Studies, 29(4), pp.937-978

[13] Hurd, M., Smith , J., Zissimopoulos, J., 2004. The Effects of Subjective Survival on Retirement and Social Security Claiming. Journal of Applied Econometrics, 19(6), pp.761-775

[14] Imrohoroglu, S., Kitao, S., 2012. Social Security Reforms: Benefit Claiming, Labor Force Participation and Long-run Sustainability. American Economic Journal: Macroeconomics, 4(3), pp.96-127

[15] Inkman J., Lopes, P., Michaelides, A., 2011. How Deep is the Annuity Market Participation Puzzle? Review of Financial Studies, 24(1), pp.279-319

[16] Kitao, S., 2014. Sustainable Social Security : Four Options. Review of Economic Dynamics, Vol17(4), pp.756-779

[17] Laitner, J. Silverman, D., 2012. Consumption, Retirement and Social Security: Evaluating the Efficiency of Reform that Encourages Longer Careers. Journal of Public Economics, 96, pp.615-634

[18] Lockwood, L., 2012. Bequest Motives and the Annuity Puzzle. Review of Economic Dynamics, 15(2), pp. 226-243

[19] Maurer, R., Mitchell, O., Rogalla, R., Schimetschek, T., 2016. Will They Take the Money and Work? People's Willingness to Delay Claiming Social Security Benefits for a Lump Sum. Journal of Risk and Insurance, 9999, pp.1-33

[20] Meyer, W., Reichenstein, W., 2010. Social Security: When to Start Benefits and How to Minimize Longevity Risk. Journal of Financial Planning, 23(3), pp.49-59

[21] Mitchell, O. S., Poterba, J.M., Warshawsky, M.J., Brown, J.R., 1999. New Evidence on the Money's Worth of Individual Annuities. American Economic Review 89(5), pp.1299-1318

[22] Pashchenko, S., 2013. Accounting for Non-Annuitization. Journal of Public Economics, 98, pp.53-67

[23] Pashchenko, S., Porapakkarm, P., 2013. Quantitative Analysis of Health Insurance Reform: Separating Regulation from Redistribution. Review of Economic Dynamics 16, pp.383-404.

[24] Pashchenko, S., Porapakkarm, P., 2016. Medical Spending in the U.S.: Facts from the Medical Expenditure Panel Survey Database. Fiscal Studies. 37(3-4), pp.689-716

[25] Pashchenko, S., Porapakkarm, P., 2016. Work Incentives of Medicaid Beneficiaries and the Role of Asset Testing. International Economic Review, 58(4), pp.1117-1154

[26] Shoven, J., Slavov, S., 2014a. Does It Pay to Delay Social Security? Journal of Pension Economics and Finance, 13(2), pp.121-144

[27] Shoven, J., Slavov, S., 2014b. Recent Changes in the Gains from Delaying Social Security. Journal of Financial Planning, 27(3), pp.32-41

[28] Storesletten, K., Telmer, C., Yaron, Y., 2004. Consumption and Risk Sharing Over the Life Cycle. Journal of Monetary Economics 51(3), pp.609-633

[29] Sun, W., Webb, A., 2009. How Much Do Households Really Lose by Claiming Social Security at Age 62? Center of Retirement Research at Boston College Working Paper

[30] Turra, C., Mitchell, O., 2008. The Impact of Health Status and Out-of-Pocket Medical Expenditures on Annuity Valuation. I Ameriks, J., and Mitchell, O., editors, Recalibrating Retirement Spending and Saving, pages 227-250. Oxford University Press

[31] Venti, S., Wise, D., 2004. The Long Reach of Education: Early Retirement. NBER Working Paper

Appendix

A Summary of the parametrization of the baseline model

Parameter name	Notation	Value	Source
Parameters set outside the model			
Consumption share	$\varkappa$	0.5	French (2005)
Labor supply	$\bar{l}$	0.4	
Tax function parameters	a_0	0.258	Gouveia and Strauss (1994)
	a_1	0.768	"
	a_2	0.616	Pashchenko and Porapakkarm (2013)
Labor productivity			
- Persistence parameter	ρ	0.98	Storesletten, et al (2000)
- Variance of innovations	σ_ε^2	0.02	"
- Fixed effect	σ_ξ^2	0.24	"
Parameters used to match some targets			
Discount factor	β	0.962	% claiming at age 62
Risk aversion		4	Wealth accumulation before 60
1/IES	γ	1.5	- " -
Bequest parameters			
- MPB	-	0.97	Wealth profile before 60 for p50 and p75
- Bequest threshold	-	$3,600	Wealth profile before 60 for p25
Consumption floor	$\underline{c}$	$3,500	% employment among 25-29
Wage rate	w	1.55	average earnings
Fixed costs of work	ϕ_w	0.255	employment profiles (healthy)

Table 5: Parameters of the model

B Computing actuarially fair adjustments to Social Security benefits

In this section, we explain how the adjustments to Social Security benefits for early/late claiming reported in Table 2 are computed. Denote the adjustments for age 62 as x_{62}, for age 63 as x_{63}, etc. As in the actual schedule of benefits and rewards, we set x_{65} to 1, i.e., individuals who claim at age 65 get full benefits. In order for the underlying price of the Social Security annuity to be actuarially fair, these adjustments have to satisfy the following:

$$q_t^{AF} = \frac{x_t}{x_{t+1} - x_t}, \qquad t = 62, ..., 69$$

where q_t^{AF} is actuarially fair price for the annuity at age t. This represents a system of 8 equations which can be solved for x_t because $x_{65} = 1$.

Recent Working Papers from the Center for Retirement Research at Boston College

The Minimum Wage and Annual Earnings Inequality
Gary V. Engelhardt and Patrick J. Purcell, August 2018

Exploring the Consequences of Discrimination and Health for Retirement by Race and Ethnicity: Results from the *Health and Retirement Study*
Ernest Gonzales, Yeonjung Jane Lee, William V. Padula, and Lindsey Subin Jung, July 2018

Financial Management Support for SSA Beneficiaries: Looking Beyond the Payee
Annie Harper, May 2018

What Factors Explain the Decline in Widows' Poverty?
Alicia H. Munnell, Geoffrey T. Sanzenbacher, and Alice Zulkarnain, May 2018

Exploring the Rise of Mortgage Borrowing among Older Americans
J. Michael Collins, Erik Hembre, and Carly Urban, May 2018

How Might Earnings Patterns and Interactions Among Certain Provisions in OASDI Solvency Packages Affect Financing and Distributional Goals?
Melissa M. Favreault, March 2018

Distributional Effects of Alternative Strategies for Financing Long-Term Services and Supports and Assisting Family Caregivers
Melissa M. Favreault and Richard W. Johnson, March 2018

How to Pay for Social Security's Missing Trust Fund?
Alicia H. Munnell, Wenliang Hou, and Geoffrey T. Sanzenbacher, December 2017

Retirement Prospects for the Millennials: What is the Early Prognosis?
Richard W. Johnson, Karen E. Smith, Damir Cosic, and Claire Xiaozhi Wang, November 2017

Mom and Dad We're Broke, Can You Help? A Comparative Study of Financial Transfers Within Families Before and After the Great Recession
Mary K. Hamman, Daniela Hochfellner, and Pia Homrighausen, November 2017

Homeownership, Social Insurance, and Old-Age Security in the United States and Europe
Stipica Mudrazija and Barbara A. Butrica, October 2017

How Much Does Motherhood Cost Women in Social Security Benefits?
Matthew S. Rutledge, Alice Zulkarnain, and Sara Ellen King, October 2017

All working papers are available on the Center for Retirement Research website (http://crr.bc.edu) and can be requested by e-mail (crr@bc.edu) or phone (617-552-1762).