Differential Mortality and the Progressivity of Social Security*

Shantanu Bagchi†

August 29, 2017

Abstract

There is a well-established negative correlation between earnings and mortality risk. Using a calibrated general-equilibrium macroeconomic model, this paper examines how this correlation interacts with Social Security's benefit-earnings rule. My findings suggest that the welfare ranking of alternative benefit-earnings rules is, in fact, sensitive to differential mortality risk. While a more progressive benefit-earnings rule provides increased insurance for households with unfavorable earnings histories, their relatively high mortality risk heavily discounts the expected utility from old-age consumption. I find that the latter effect dominates: welfare is maximized with a fully proportional (linear) benefit-earnings rule in the presence of differential mortality.

JEL Classifications: E21, E62, H55

Keywords: differential mortality; Social Security; mortality risk; labor income risk; incomplete markets; social insurance; general equilibrium

*This project received financial support from the W.E. Upjohn Foundation for Employment Research (ECRA Grant No. 15–150–01) and the College of Business & Economics Summer Support Grant Program. I would like to thank Frank Caliendo, Matt Chambers, Jim Feigenbaum, and Juergen Jung for useful comments and suggestions.

†Department of Economics, Towson University, 8000 York Road, Towson, MD 21252. Phone: +1-410-704-2191. Email: sbagchi@towson.edu

1 Introduction

Economists have traditionally viewed Social Security as a vehicle that partially insures individuals against risks that markets do not insure well, such as the risk of an uncertain lifetime, and also the risk of old-age poverty caused by unfavorable labor-market outcomes.[1] Social Security annuities are paid until death, so they insure individuals against the risk of out-living their savings. Meanwhile, Social Security benefits are a concave, increasing function of work-life earnings. The curvature of this function determines how much insurance Social Security provides against unfavorable labor-market events, such as the inability to secure a high-paying job or unemployment.[2]

While linking public pension benefits to work-life income is common within the industrialized world, the concavity of this relationship in the U.S. is unusual. Under current law, benefits replace 90% of the average work life earnings for an individual who is in the bottom 20% of the wage distribution, but only about 40-50% for individuals whose wages are higher than the average wage. Therefore, an individual at the bottom of the earnings distribution receives a higher return on every dollar of Social Security contributions paid, relative to an individual at the top of the earnings distribution. While this arrangement is intended to provide insurance against unfavorable labor income shocks, its effect on the distribution of lifetime utility across households depends on a multitude of economic and demographic factors.[3]

From the perspective of a household, an important determinant of whether the household gains or loses from Social Security is its life expectancy. Because Social Security is a retirement pension, households who expect to live longer are likely to experience a larger welfare gain, compared to households who do not. However, empirical evidence suggests that there is a significant positive correlation between income and life expectancy in the U.S. (Kitagawa and Hauser, 1987). This phenomenon, referred to as differential mortality, has important implications for Social Security, because the positive correlation between income and survivorship can potentially undo the redistribution from Social Security's current benefit-earnings rule (Coronado et al., 2002, 2011).

In this paper, I examine if differential mortality has any implications for the welfare effects of Social Security's benefit-earnings rule. To do this, I first set up a simple two-period overlapping generations model with incomplete markets, an unfunded public pension system that resembles Social Security, and rational households who experience mortality risks that are negatively correlated to their earnings. I use this simple model to theoretically demonstrate two facts. First, I show that for households with lower earnings, Social Security's internal rate of return from a progressive benefit-earnings rule may be higher in the presence of differential mortality, compared to when mortality risk is symmetric and uncorrelated to earnings. Second, I show that the welfare gain from this potentially higher internal rate of return, in terms of improved consumption smoothing, depends on the expected marginal utility of old-age consumption. Together, these results suggest that while a progressive benefit-earnings rule may provide better work-retirement consumption smoothing for households with lower earnings, their relatively high mortality risk may limit the size of welfare gain from higher old-age consumption.

Next, I evaluate the quantitative importance of these mechanisms using a rich multi-period general-equilibrium model calibrated to the U.S. economy. In the model, households experience two types of risk: an age-dependent mortality risk that is negatively correlated to their earnings,

[1] The primary justification behind the creation of Social Security in 1935 was to make "more adequate provision for aged persons". See the legislative history of the Social Security Act of 1935 at http://www.ssa.gov/history.

[2] In theory, Social Security also enables intergenerational risk sharing, but the model considered here does not have any aggregate risk to share across generations.

[3] It is well known that Social Security improves welfare in the aggregate only in the counterfactual case where the economy is dynamically inefficient (Samuelson, 1975).

and a labor income risk that consists of a permanent productivity fixed effect, and an idiosyncratic shock with both persistent and transitory components. Factor markets in the model are competitive, firms maximize profit, and the government provides public goods and Social Security, which provides partial insurance against the household-level risks. I calibrate this model to match key features of the U.S. economy, such as the institutional features of Social Security, overall capital accumulation, the pattern of labor supply over the life cycle, the earnings distribution, and the share of government expenditures in GDP.

Then, I use this model to compute the disaggregated welfare implications of alternative benefit-earnings rules, ranging from a fully proportional (or linear) function of past earnings (i.e. with zero implicit redistribution), to a fixed benefit that is completely unrelated to past earnings (i.e. with full redistribution). Finally, I examine if the welfare ranking of these rules is sensitive to the negative correlation between earnings and mortality risk. I also identify the macroeconomic effects of differential mortality on the labor market, capital accumulation, national income, and the government's budget.

My computations suggest that the welfare ranking of the alternative benefit-earnings rules is, in fact, highly sensitive to differential mortality. In the presence of differential mortality risk, more progressive benefit-earnings rules provides better work-retirement consumption smoothing for households with relatively unfavorable earnings histories. However, their relatively high mortality risk causes these households to also heavily discount the utility from old-age consumption. I find that the latter effect dominates: with differential mortality, aggregate welfare is maximized when benefits are a fully proportional (linear) function of past earnings, but when differential mortality is eliminated, the uniform-benefit rule maximizes welfare. In other words, the lower expected utility of old-age consumption limits the benefits of improved work-retirement consumption smoothing when survivorship is positively correlated with earnings.

My computations also predict that in switching from the current U.S. benefit-earnings rule to the fully proportional (linear) rule, expected Social Security benefits would decline by about 12% for households with unfavorable earnings histories, and would increase by more than 7% for households with favorable earnings histories. Finally, I find that modifying the shape of the benefit-earnings rule, given Social Security's current payroll tax rate and taxable maximum, has only a small effect on key macroeconomic aggregates: between the computational experiments, capital stock, labor, national income, and government expenditures, do not change by more than one percentage point, both with and without differential mortality.

This paper contributes to three separate strands of the literature. First, it contributes to a large literature that characterizes the optimal redistribution scheme in a heterogeneous-agent economy, accounting for distortions to consumption, saving, and labor supply.[4] Two papers in this literature that highlight the importance of earnings history-dependent tax systems are Grochulski and Kocherlakota (2010) and Michau (2014). Grochulski and Kocherlakota (2010) show that in an economy where agents have nonseparable preferences and private information about their skill levels, it is possible to implement a socially optimal allocation through a linear labor income tax during the working life, and constant payment during retirement that is conditioned on the agents' entire labor income history. Michau (2014) builds on their results and shows that an earnings history-dependent social security system can implement the optimal allocation that accounts for labor supply distortions both along the extensive and intensive margins. However, none of these studies account for differential mortality, thereby ignoring its potential redistributive consequences for such tax-and-transfer systems.

[4]Notable studies in this literature include Saez (2002), Cremer et al. (2004), Sheshinski (2008), Golosov et al. (2011), and Farhi and Werning (2013), among many others.

Second, the current paper contributes to the quantitative-macro literature on the welfare consequences of alternative social security schemes in the context of the U.S. Three studies from this literature that are closest to the current paper are Huggett and Ventura (1999), Fuster et al. (2003), and Nishiyama and Smetters (2008). Huggett and Ventura (1999) examine the distributional consequences of replacing current U.S. Social Security with a two-tier pension system with a mandatory, defined-contribution first tier, and a guaranteed second tier with a minimum retirement income.[5] In general, they do not find substantial welfare improvements in switching from the current U.S. program to the two-tier structure. Fuster et al. (2003), on the other hand, examine the welfare effects of unfunded social security in a general equilibrium model with overlapping generations of altruistic individuals that differ in lifetime expectancy and earnings ability. They find that ability shocks and uncertain lifetimes generate significant heterogeneity among households to induce different preferences for Social Security. Finally, Nishiyama and Smetters (2008) find that while the progressive linking of earnings with retirement benefits in the U.S. has beneficial insurance effects, it also introduces various marginal tax rates that distort labor supply. In fact, Nishiyama and Smetters (2008) conclude that the optimal benefit structure in the U.S. is fairly proportional. However, both Huggett and Ventura (1999) and Nishiyama and Smetters (2008) ignore differential mortality, and a crucial determinant of the results in Fuster et al. (2003) is a two-sided altruism mechanism. Also, the computational experiments in these studies do not allow for a clear interpretation of the interaction between differential mortality and the welfare effects of alternative benefit-earnings rules.

Finally, this paper also contributes to an empirical literature that measures the effect of differential mortality on the progressivity of U.S. Social Security from survey data. Studies such as Coronado et al. (2002) and Coronado et al. (2011) conclude that once the positive correlation between wealth and survivorship is accounted for, Social Security is considerably less progressive than what is defined by the benefit-earnings rule. For example, Coronado et al. (2002) compute the "net tax rate" implicit in Social Security, given by the difference between the present value of the taxes paid and benefits received over the life cycle, expressed as a fraction of potential lifetime income. They find that Social Security becomes regressive after accounting for the mortality differentials between the different income groups, or in other words, it transfers resources from poorer households with shorter lives to wealthier households with longer lives. In a separate study, Meyerson and Sabelhaus (2006) compute the "money's worth" from Social Security, given by the ratio of the present value of benefits to that of the taxes paid over the life cycle. While they conclude that Social Security remains progressive even after accounting for the mortality differences, they find that the degree of progressivity is greatly reduced. However, these studies are purely actuarial in nature, as a result of which they do not account for how households, firms, and the overall economy respond to differential mortality and the modifications to Social Security. The current paper accounts for these effects.

The rest of the paper is organized as follows: Section 2 illustrates the key mechanisms of the paper using a simple two-period model, Section 3 introduces the formal quantitative general-equilibrium model, and Sections 4 and 5 describe the baseline calibration and its results. I describe the computational experiments in Section 6, and I examine their results in Sections 7, 8, and 9. Finally, I conclude in Section 10.

[5] An example of such a reform proposal is the Boskin proposal (Boskin et al., 1987).

2 A simple two-period model

In this section, I demonstrate how differential mortality interacts with Social Security's benefit-earnings rule using a simple two-period overlapping generations model. Lifetime consists of two periods: work (1), and retirement (2). A household earns income and pays taxes in both periods. There are two types of households in every cohort: type$-l$ with lower earnings, and type$-h$ with higher earnings. We assume that both types of households experience mortality risk, but a type$-h$ household has a higher likelihood (Q_h) of surviving into retirement, compared to a type$-l$ household (Q_l), i.e. $Q_l < Q_h < 1$. Fraction a of each cohort is born as type l, and fraction $(1-a)$ is born as type h, and cohort size grows at the rate of n over time. Social Security operates as follows: it collects taxes t_l and t_h from the respective types of households during the working period, and pays benefits b_l and b_h during retirement. Finally, Social Security taxes and benefits across the household types are related as $t_h = (1+x)t_l$ and $b_h = (1+y)b_l$, with $x, y > 0$. This specification allows us to consider Social Security programs with varying degrees of implicit redistribution.

In this model environment, Social Security's budget balancing requires

$$aQ_l N_{t-1} b_l + (1-a)Q_h N_{t-1} b_h = aN_t t_l + (1-a)N_t t_h, \tag{1}$$

where N_t is the size of the cohort born on date t. Given that $b_h = (1+y)b_l$, $t_h = (1+x)t_l$, and $N_t = (1+n)N_{t-1}$, we have

$$\left[aQ_l + (1-a)Q_h(1+y)\right] b_l = \left[a + (1-a)(1+x)\right](1+n)t_l, \tag{2}$$

which gives the benefits for the type$-l$ household

$$b_l = \frac{(1+n)t_l\left[a + (1-a)(1+x)\right]}{\left[aQ_l + (1-a)Q_h(1+y)\right]}. \tag{3}$$

Therefore, the lifetime budget constraint for type$-l$ is

$$c_{l,1} + \frac{c_{l,2}}{1+r} = we_l, \tag{4}$$

where we_l is the present value of lifetime earnings, or

$$\begin{aligned} we_l &= y_{l,1} - t_{l,1} + \frac{y_{l,2} - t_{l,2}}{1+r} \\ &= y_{l,1} + \frac{y_{l,2}}{1+r} - t_{l,1} - \frac{t_{l,2}}{1+r}. \end{aligned} \tag{5}$$

Substituting for $t_{l,2} = -b_l$, we get

$$\begin{aligned} we_l(SS=1) &= y_{l,1} + \frac{y_{l,2}}{1+r} - t_{l,1} + \frac{(1+n)t_{l,1}\left[a + (1-a)(1+x)\right]}{(1+r)\left[aQ_l + (1-a)Q_h(1+y)\right]} \\ &= we_l(SS=0) + t_{l,1}\left[\frac{(1+n)\left(a + (1-a)(1+x)\right)}{(1+r)\left(aQ_l + (1-a)Q_h(1+y)\right)} - 1\right], \end{aligned} \tag{6}$$

where $we_l(SS=1)$ and $we_l(SS=0)$ represent lifetime wealth with and without Social Security, respectively. Therefore, the necessary and sufficient condition for $we_l(SS=1) > we_l(SS=0)$ is

$$\gamma_l = (1+n)\left[\frac{a + (1-a)(1+x)}{aQ_l + (1-a)Q_h(1+y)}\right] - 1 > r. \tag{7}$$

The left-hand side of (7) is the net internal rate of return of Social Security for the type-l household in this model. Therefore, Social Security has a positive wealth effect if and only if its internal rate of return is higher than the market rate of return r.

Let us define the ratio $\frac{y}{x}$ as a measure of the progressivity of Social Security. That is,

$$\frac{y}{x}\begin{cases} < & 1 \quad \text{Progressive} \\ = & 1 \quad \text{Proportional} \\ > & 1 \quad \text{Regressive} \end{cases}.$$

Then, we can first verify the internal rate of Social Security when there is no mortality risk and Social Security is proportional. Setting $\frac{y}{x} = 1$ and $Q_l = Q_h = 1$ in (7), we get

$$\gamma_l = (1+n) - 1 = n. \tag{8}$$

This is the well-known result that in a rational life-cycle framework without mortality risk and zero implicit redistribution, Social Security's internal rate of return is the population growth rate, and Social Security is welfare-improving if and only if that rate is larger than r (Samuelson, 1975).

Now let us consider the more realistic case considered in the literature, where mortality risk is symmetric (i.e. uncorrelated to earnings) and Social Security is progressive. Setting $\frac{y}{x} < 1$ and $Q_l = Q_h = Q < 1$ in (7), we get

$$\gamma_l = \frac{(1+n)}{Q}\left[\frac{a+(1-a)(1+x)}{a+(1-a)(1+y)}\right] - 1. \tag{9}$$

Given that $\frac{y}{x} < 1 \implies \frac{x}{y} > 1$, and $Q < 1 \implies \frac{1}{Q} > 1$, we can conclude that $\gamma_l > n$. Therefore, we have

$$\gamma_l(\frac{y}{x} < 1, Q_l = Q_h = Q < 1) > \gamma_l(\frac{y}{x} = 1, Q_l = Q_h = 1) = n.$$

Therefore, in a rational life-cycle framework with earnings heterogeneity and symmetric mortality risk, a progressive benefit-earnings rule yields an internal rate of return for Social Security that is higher than the population growth rate.

Finally, let us consider the question on hand: how does differential mortality affect the welfare-improving role of a progressive benefit-earnings rule? To examine this, we set $\frac{y}{x} < 1$ and $Q_h = (1+q)Q_l$, with $q > 0$ in (7). This yields

$$\gamma_l = \frac{(1+n)}{Q_l}\left[\frac{a+(1-a)(1+x)}{a+(1-a)(1+q)(1+y)}\right] - 1. \tag{10}$$

To compare (10) with (9), we set $Q = \omega Q_l + (1-\omega)Q_h$ in (9), with $0 < \omega < 1$. That is, we interpret the survival probability Q as a weighted average of the probabilities Q_l and Q_h. With this substitution, we can rewrite (9) as

$$\gamma_l = \frac{(1+n)}{(\omega Q_l + (1-\omega)Q_h)}\left[\frac{a+(1-a)(1+x)}{a+(1-a)(1+y)}\right] - 1. \tag{11}$$

Therefore, so that

$$\gamma_l(\frac{y}{x} < 1, Q_l < Q_h < 1) > \gamma_l(\frac{y}{x} < 1, Q_l = Q_h = Q < 1),$$

we need

$$Q_l\left[a+(1-a)(1+q)(1+y)\right] < (\omega Q_l + (1-\omega)Q_h)\left[a+(1-a)(1+y)\right].$$

Substituting $Q_h = (1+q)Q_l$, this condition simplifies to

$$[a + (1-a)(1+q)(1+y)] < (\omega + (1-\omega)(1+q))\,[a + (1-a)(1+y)]\,. \tag{12}$$

A necessary condition for (12) to be satisfied is

$$\begin{aligned} & \omega + (1-\omega)(1+q) > 1 \\ \implies \quad & 1 + q(1-\omega) > 1 \\ \implies \quad & q(1-\omega) > 0, \end{aligned} \tag{13}$$

which is always satisfied because $q > 0$ and $0 < \omega < 1$. Therefore, we find that Social Security's internal rate of return from a progressive benefit-earnings rule may be higher in the presence of differential mortality, compared to when mortality risk is symmetric and uncorrelated to earnings.

To examine the welfare effect of this potentially higher internal rate of return, we solve the type$-l$ household's utility maximization problem. This problem can be written as

$$\underset{c_{l,1},c_{l,2}}{\text{Max}} \quad U = u(c_{l,1}) + \beta\ Q_l\ u(c_{l,2}) \tag{14}$$

subject to

$$c_{l,1} + \frac{c_{l,2}}{1+r} = we_l, \tag{15}$$

where $we_l = y_{l,1} - t_{l,1} + \frac{y_{l,2}-t_{l,2}}{1+r}$. Substituting the budget constraint in the utility function, we can rewrite the objective function as

$$U = u(c_{l,1}) + \beta\ Q_l\ u\left((we_l - c_{l,1})(1+r)\right). \tag{16}$$

The first-order condition for this problem is

$$\begin{aligned} u'(c_{l,1}) - \beta\ Q_l(1+r)\ u'\left((we_l - c_{l,1})(1+r)\right) = 0 \\ \implies F\left(c_{l,1}; \beta, Q_l, r, we_l\right) = 0. \end{aligned} \tag{17}$$

Because

$$F_{c_{l,1}} = u''(c_{l,1}) + \beta\ Q_l(1+r)^2\ u''\left((we_l - c_{l,1})(1+r)\right) < 0, \tag{18}$$

we can use the Implicit Function Theorem to express the optimal values of the first- and second-period consumptions as

$$\begin{aligned} c^*_{l,1} &= c_{l,1}\left(\beta, Q_l, r, we_l\right) \\ c^*_{l,2} &= \left(we_l - c_{l,1}\left(\beta, Q_l, r, we_l\right)\right)(1+r), \end{aligned}$$

and the optimal value function as

$$V(\beta, Q_l, r, we_l) = u\left(c_{l,1}\left(\beta, Q_l, r, we_l\right)\right) + \beta Q_l u\left(\left(we_l - c_{l,1}\left(\beta, Q_l, r, we_l\right)\right)(1+r)\right). \tag{19}$$

Applying the Envelope Theorem on this value function, we can calculate the welfare gain from the improved consumption smoothing as

$$V_{we_l}(\beta, Q_l, r, we_l) = \beta Q_l u'\left(\left(we_l - c_{l,1}\left(\beta, Q_l, r, we_l\right)\right)(1+r)\right) > 0, \tag{20}$$

which is nothing but the expected marginal utility of old-age consumption of the type$-l$ household, weighted by its probability of surviving into retirement. Therefore, while their internal rate of return

from a progressive benefit-earnings rule may be higher in the presence of differential mortality, the higher mortality risk experienced by the type$-l$ households will also heavily discount the marginal utility of the improved work-retirement consumption smoothing.

To summarize, the simple two period model shows that it is theoretically possible for a progressive social security program to deliver higher welfare for low-income households when mortality risk is negatively correlated with earnings. However, there are two competing effects at work: while the internal rate of return from a progressive benefit-earnings rule may be higher in the presence of differential mortality, the welfare gain from this potentially higher rate of return may also be lower for these households. Next, we turn to a quantitative examination of these effects in a rich multi-period general-equilibrium environment.

3 The quantitative general-equilibrium model

The unit of our quantitative general-equilibrium model is a permanent-income household that smooths consumption and labor supply over the life cycle by accumulating a risk-free asset: physical capital. Over the course of the life cycle, this household experiences two types of risk: labor income risk, which is exogenous, and mortality risk, which is endogenous and negatively correlated to household earnings.[6] The household does not have access to markets where it can purchase insurance against these risks.

At each date, surviving households earn labor income if they work, and they also receive Social Security benefits after the full retirement age. Firms operate competitively and produce output using capital, labor and a constant returns to scale technology. The government provides public goods and Social Security; the public goods purchases are funded through general tax revenues, and Social Security is funded through a payroll tax on labor income. In this framework, Social Security provides partial insurance against labor income and mortality risks.

3.1 Preferences

Households derive utility both from consumption and leisure. A household's labor supply decision at each instant consists of two components: the extensive margin or the participation decision (P), and the intensive margin or the hours of work (h), conditional on participation. The period utility function is given by

$$u(c, 1-h, P) = \begin{cases} \frac{\left(c^{\eta}(1-h-\theta_P \cdot P)^{1-\eta}\right)^{1-\sigma}}{1-\sigma} & \text{if} \quad \sigma \neq 1 \\ \ln\left(c^{\eta}(1-h-\theta_P \cdot P)^{1-\eta}\right) & \text{if} \quad \sigma = 1 \end{cases} \tag{21}$$

where η is the share of consumption, σ is the inverse of intertemporal elasticity of substitution (IES), θ_P is the age-dependent cost of labor force participation (measured in hours), and P is the labor force participation status: $P = 1$ if the household participates, and $P = 0$ otherwise. Also, since I normalize the period time endowment to unity, $0 \leq h \leq 1$.

Expected lifetime utility from the perspective of a household is given by

$$U = E\left[\sum_{s=0}^{T} \beta^s \ \left\{\Pi_{j=0}^{s-1} Q_j\left(x_j\right)\right\} \ u\left(c_s, 1-h_s, P_s\right)\right], \tag{22}$$

[6] This approach of modeling differential mortality as a negative correlation between earnings and mortality risk follows Fuster et al. (2003) and Cristia (2009). However, some studies have also modeled health as a determinant of mortality risk (İmrohoroğlu and Kitao, 2012; Zhao, 2014). In these studies, unfavorable health shocks reduce life expectancy. In the context of current question, the exact source of differential mortality is less important; the relevant factor is how differential mortality interacts with the redistribution implicit in Social Security's benefit-earnings rule.

where β is the discount factor, and $Q_j(x_j)$ is the probability at age j of surviving to the next period, which depends on the age itself, as well as on the household's state vector that age, x_j.

3.2 Income

Conditional on labor force participation, a household earns before-tax wage income $y_s(\varphi_s) = h_s w_s e_s(\varphi_s)$ at age s, where w_s is the wage rate, and $e_s(\varphi_s)$ is a labor productivity endowment that depends on age and a stochastic productivity shock φ_s. This wage income is subject to two separate taxes: a progressive income tax $T_y(\cdot)$, and a payroll tax $T_{ss}(\cdot)$ for Social Security that is proportional up to the maximum taxable earnings of $\bar{y}$. Households also earn interest income on their asset holdings at the risk free rate r, and this interest income is subject to a capital income tax $T_k(\cdot)$. After-tax income at age s, therefore, is equal to before-tax wage plus interest income minus the income and Social Security taxes.

It is useful to note here that because insurance markets are closed, deceased households at every age leave behind accidental bequests. I assume that the government imposes a confiscatory tax on these accidental bequests, which is equivalent to assuming that the government imposes an estate tax of 100%.

3.3 Social Security

The government pays Social Security benefits to households after the full retirement age (T_r), and the amount of benefits paid to a particular household depends on its earnings history. For each household, the government calculates an average of past earnings (up to the maximum taxable earnings), referred to as the Average Indexed Monthly Earnings ($AIME$). The Social Security benefit amount, also called the Primary Insurance Amount (PIA), is calculated as

$$PIA = RR(AIME) \times AIME + \overline{PIA} \tag{23}$$

where RR is the replacement rate, which is a concave function of the $AIME$, and $\overline{PIA}$ is a fixed amount unrelated to the $AIME$.[7] Finally, the government scales the benefit amount up or down proportionally so that Social Security's budget is balanced.[8]

3.4 A household's optimization problem

A household's state vector is given by $x = \{k, \varphi, AIME\}$, where k denotes the beginning-of-period assets, φ the stochastic productivity shock, and $AIME$ the average past earnings that determine Social Security benefits. Conditional on a particular realization of the states, the household chooses consumption, assets holdings for the next period, and labor supply.

At a given age s, this optimization problem can be recursively expressed as

$$V_s(x) = \max_{c,k',P,h} \left\{ u(c, 1-h, P) + \beta\ Q_s(x)\ E\left[V_{s+1}(x')\right] \right\} \tag{24}$$

[7] For the current U.S. benefit-earnings rule, this fixed amount $\overline{PIA}$ is equal to zero.

[8] While in reality, Social Security has a trust fund and does not satisfy the definition of a Pay-As-You-Go program in the narrow sense, there is some disagreement on whether or not the trust fund assets are "real", i.e. whether or not they have increased national saving. In fact, Smetters (2003) finds that the trust funds assets have actually increased the level of debt held by the public, or reduced national saving. Because of this reason, it is a common practice in the literature to ignore the trust fund and model Social Security's budget as balanced every period (See, for example, studies such as Huggett and Ventura (1999), Conesa and Krueger (1999), İmrohoroğlu et al. (2003), Jeske (2003), Conesa and Garriga (2009), and Zhao (2014), among others).

subject to

$$c_s + k' = (1+r)k + y_s^{at} + \Theta(s - T_r)\, b(AIME) \tag{25}$$

$$y_s^{at} = h_s w_s e_s(\varphi_s) + rk - T_y\left(h_s w_s e_s(\varphi_s)\right) - T_k\left(rk\right) - T_{ss}\left(h_s w_s e_s(\varphi_s); \bar{y}\right) \tag{26}$$

$$0 \leq h_s \leq 1, \tag{27}$$

$$k' \geq 0 \tag{28}$$

$$AIME' = \begin{cases} \left[AIME \times (s-1) + \min\left\{h_s w_s e_s(\varphi_s), \bar{y}\right\}\right]/s & s < T_r \\ AIME & s \geq T_r \end{cases}, \tag{29}$$

where

$$\Theta(s - T_r) = \begin{cases} 0 & s < T_r \\ 1 & s \geq T_r \end{cases}$$

is a step function. Households are born with and die with zero assets, i.e. $k(0) = k(T+1) = 0$, and prior to age T_r, the average earnings $AIME$ evolves based on the realized labor productivity shocks and the endogenous labor supply decisions.

3.5 Technology and factor prices

Output is produced using a Cobb-Douglas production function with inputs capital and labor

$$Y \;=\; K^{\alpha} L^{1-\alpha}, \tag{30}$$

where α is the share of capital in total income. Firms face perfectly competitive factor markets, which implies

$$r = MP_K - \delta = \alpha \left[\frac{K}{L}\right]^{\alpha-1} - \delta \tag{31}$$

$$w = MP_L = (1-\alpha)\left[\frac{K}{L}\right]^{\alpha} \tag{32}$$

where δ is the depreciation rate of physical capital and w is the wage rate. There is no aggregate uncertainty.

3.6 Aggregation

The population structure in the model is as follows: at each instant a new cohort is born and the oldest cohort dies, and cohort size grows at the rate of n over time. Let us denote the measure of households at age s with state x as $\mu_s(x)$. Then, the aggregate capital stock and labor supply are given by

$$K = \sum_{s=0}^{T} N_s \sum_{x} \left\{\Pi_{j=0}^{s-1} Q_j(x)\right\} \; k_{s+1}(x)\mu_s(x) \tag{33}$$

$$L = \sum_{s=0}^{T} N_s \sum_{x} \left\{\Pi_{j=0}^{s-1} Q_j(x)\right\} \; h_s(x) e_s(x) \mu_s(x), \tag{34}$$

where N_s is the size of the age-s cohort. The budget constraints for Social Security and the general government budget are given by

$$\begin{aligned}\sum_{s=0}^{T} N_s \sum_x \left\{\Pi_{j=0}^{s-1} Q_j(x)\right\} \; T_{ss}\left(h_s(x) w_s e_s(x); \bar{y}\right) \mu_s(x) \\ = \sum_{s=0}^{T} N_s \sum_x \left\{\Pi_{j=0}^{s-1} Q_j(x)\right\} \; \Theta(s - T_r) b(x) \mu_s(x)\end{aligned} \tag{35}$$

and

$$BEQ + \sum_{s=0}^{T} N_s \sum_x \left\{\Pi_{j=0}^{s-1} Q_j(x)\right\} \; \left(T_y\left(y_s(x)\right) + T_k\left(r k_s(x)\right)\right) \mu_s(x) = G, \tag{36}$$

where BEQ is the total value of the accidental bequests from deceased households, and G is the endogenously determined level of non-Social Security government expenditures.

3.7 Competitive equilibrium

A competitive equilibrium in this environment is characterized by a collection of

1. cross-sectional consumption $\{c(s;x)\}_{s=0}^{T}$, participation $\{P(s;x)\}_{s=0}^{T}$, and labor hours allocations $\{h(s;x)\}_{s=0}^{T}$,
2. an aggregate capital stock K and labor L,
3. a rate of return r and a wage rate w,
4. Social Security benefits $b(x)$ and government expenditures G, and
5. a measure of households $\mu_s(x)\ \forall\ s$,

that

1. solves the households' optimization problems,
2. maximizes the firms' profits,
3. equilibrates the factor markets,
4. balances the government's budgets, and
5. satisfies $\mu_{s+1}(x) = R_\mu\left[\mu_s(x)\right]$, where $R_\mu(\cdot)$ is a one-period transition operator on the measure distribution.

In equilibrium, total expenditure equals consumption plus net investment plus government purchases, which is equal to the total income earned from capital and labor.

$$\begin{aligned} C + K' - (1-\delta)K + G &= C + (n+\delta)K + G \\ &= wL + (r+\delta)K \\ &= Y \end{aligned} \tag{37}$$

Finally, I consider only a steady-state equilibrium, i.e. when all aggregate quantities grow at the rate of population growth and all per-capita quantities are constant. I also normalize the initial newborn cohort size to $N(0) = 1$.

	Age groups		
Earnings quintile	35-49	50-64	65-75
Top	0.35	0.61	0.74
Fourth	0.56	0.68	0.94
Third	0.73	0.99	1.08
Second	1.13	1.10	1.14
Bottom	2.25	1.63	1.1

Table 1: Mortality ratios by lifetime earnings from Cristia (2009).

4 Calibration

4.1 Demographics

To set the demographic parameters, I first assume that households enter the model at the actual age of 25 (model age of zero), and they are alive for 75 periods (up to the actual age of 100). Second, I set the population growth rate to $n = 1\%$, which is consistent with the U.S. demographic history and also with the literature. Next, to calibrate the negative correlation between earnings and mortality risk, I use estimates of mortality ratios by earnings in Cristia (2009) to adjust the average age-specific death rates from the 2001 U.S. Life Tables (Arias, 2004).

Cristia (2009) uses data from the 1984, 1993, 1996, and 2001 panels of the Survey of Income and Program Participation (SIPP), also matched with several files administered by the Social Security Administration (SSA), to compute the mortality ratios by income groups. I report these ratios in Table 1. Each entry in Table 1 gives the likelihood of death for the respective age-income group relative to the population average for that group. For example, an individual between ages 35-49 in the bottom earnings quintile is 2.25 times more likely to die, relative to the population average in that age range (conditional on surviving up to age 35).

To use this data, I first extrapolate the mortality ratios for ages 25-35 and ages 75-100, and then I calculate the age-specific death rates for each income quintile in the model by multiplying the mortality ratios by the average age-specific death rates in the population.

4.2 Social Security

To calibrate Social Security in the model, I first set the payroll tax function to

$$T_{ss}(y;\bar{y}) = \begin{cases} \tau_{ss} y & y \leq \bar{y} \\ \tau_{ss}\bar{y} & y > \bar{y} \end{cases}$$

and then set the tax rate to $\tau_{ss} = 0.106$, which is the combined tax rate for the Old-Age and Survivors Insurance (OASI) component. The maximum taxable earnings ($\bar{y}$) is adjusted regularly relative to the average wage in the U.S. For example, the taxable maximum was set at \$76,200 in the year 2000, but was adjusted to \$106,800 in 2010 and \$113,700 in 2013. During the same period, the national average wage index increased from \$32,155 to \$41,674, and finally to \$44,888.[9] Huggett and Ventura (1999) calculate that the ratio of this taxable maximum to the average wage index has averaged at about 2.47 in the U.S., using which I set the maximum taxable earnings in the model to $\bar{y} = 2.47$.

Second, to compute the PIA amount, I incorporate the U.S. benefit-earnings rule into the model. This requires setting $\overline{PIA} = 0$ in equation (23), and then specifying the replacement rate

[9] See http://www.ssa.gov/oact/cola/awiseries.html for more details.

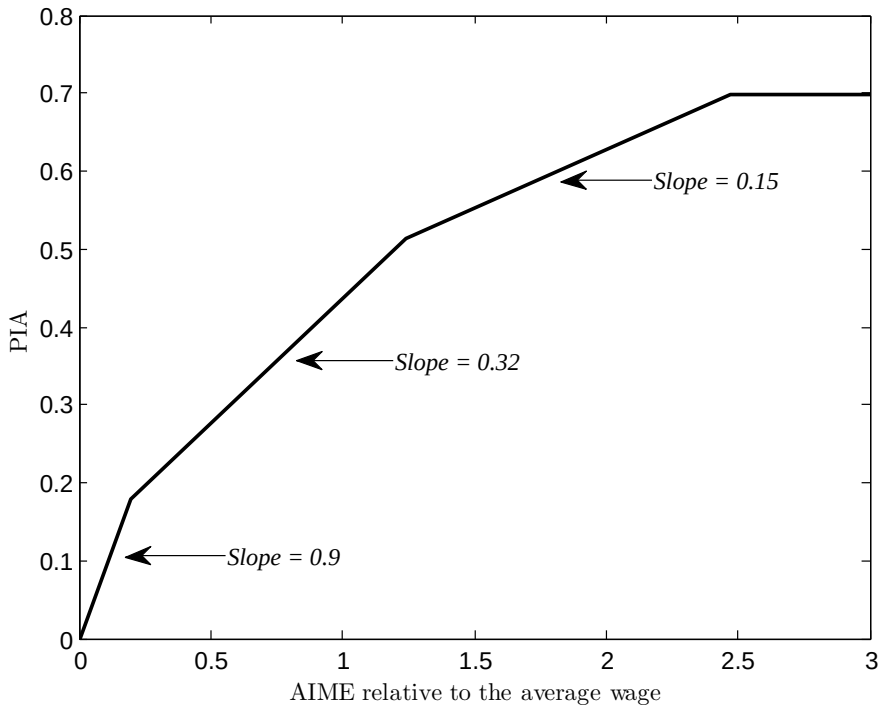

Figure 1: Benefit formula in the U.S.

(RR) as a piecewise linear function of past work-life income, the $AIME$. Depending on how large or small the $AIME$ for an individual is relative to the average wage in the economy, the SSA adjusts the replacement rate. For example, in the year 2000, the OASI benefit was 90% of the $AIME$ for the first \$531, 32% of the next \$2,671, and 15% of the remaining up to the maximum taxable earnings. As shown by Huggett and Ventura (1999), these dollar amounts come out to be roughly 20%, 124%, and 247% of the average wage in the economy. These percentage amounts are referred to as the "bend points" of the benefit rule, and I take them directly to the model. It is worth noting that the progressivity in the benefit-earnings rule is captured by the fact that the replacement rate is decreasing in the $AIME$ (see Figure 1).

Finally, I assume that households receive Social Security benefits in the model after age $T_r = 41$, which corresponds to the current full retirement age of 66 in the U.S.

4.3 Labor productivity endowment

To calibrate the labor income process, I assume that the log of labor productivity at age s can be additively decomposed as

$$\log e_s(\varphi_s) = \epsilon_s + \varphi_s, \tag{38}$$

where ϵ_s is a deterministic age-dependent component, and φ_s is a stochastic component, given by

$$\varphi_s = p + z_s + \nu_s \tag{39}$$

$$z_s = \rho z_{s-1} + \upsilon_s, \tag{40}$$

where $p \sim N(0, \sigma_p^2)$ is a permanent productivity fixed effect realized at birth, $\nu_s \sim N(0, \sigma_\nu^2)$ is a transitory shock, and z_s is a persistent shock that follows a first-order autoregressive process with $z_0 = 0$, persistence ρ, and a white-noise disturbance $\upsilon_s \sim N(0, \sigma_\upsilon^2)$.

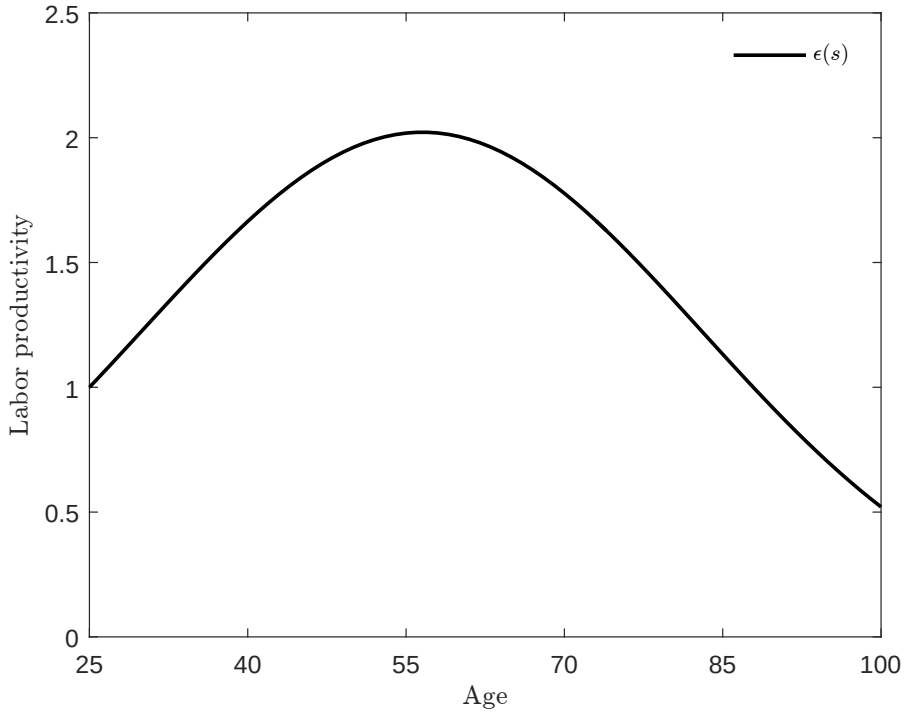

Figure 2: The age-dependent component of labor productivity from Kitao (2014).

I parameterize ϵ_s using the estimates from Kitao (2014), who uses work hour and wage data from the 2007 PSID to derive this age-dependent component of productivity as a residual of wages, after accounting for hours worked and also the part-time wage penalty. The resulting ϵ_s, normalized with respect to productivity at age 25, is plotted in Figure 2.

To calibrate the stochastic component, I use estimates from Heathcote et al. (2010) and set the persistence parameter to $\rho = 0.973$, the variances of the permanent fixed effect and the transitory shock to $\sigma_p^2 = 0.124$ and $\sigma_\nu^2 = 0.04$ respectively, and variance of the white-noise disturbance to $\sigma_v^2 = 0.005$. I use Gaussian quadrature to approximate the distribution of the permanent fixed effect using a three-point discrete distribution, and I approximate the joint distribution of the persistent and transitory shocks using a five-state first-order discrete Markov process following Tauchen and Hussey (1991).

4.4 Income tax

To calibrate the labor income tax function, I follow Storesletten et al. (2012) and Karabarbounis (2012) and assume that

$$T_y(y) = y - (1 - \tau_y)y^{1-\tau_1}, \tag{41}$$

where $\tau_y < 1$ and $\tau_1 > 0$. Note that with $\tau_1 = 0$, equation (41) reduces to a proportional tax function with a marginal rate of τ_y. With this income tax function, after-tax income is log-linear in before-tax income, and the parameter τ_1 controls the progressivity of the tax code. Following Storesletten et al. (2012), I set the value of this parameter to $\tau_1 = 0.151$. Storesletten et al. (2012) estimate this value using data from the 2000, 2002, 2004, and 2006 waves of the PSID, and also NBER's TAXSIM program, accounting for federal and state income taxes plus public transfers. I plot the average and marginal tax rates that emerge from the estimated tax function in Figure 3.

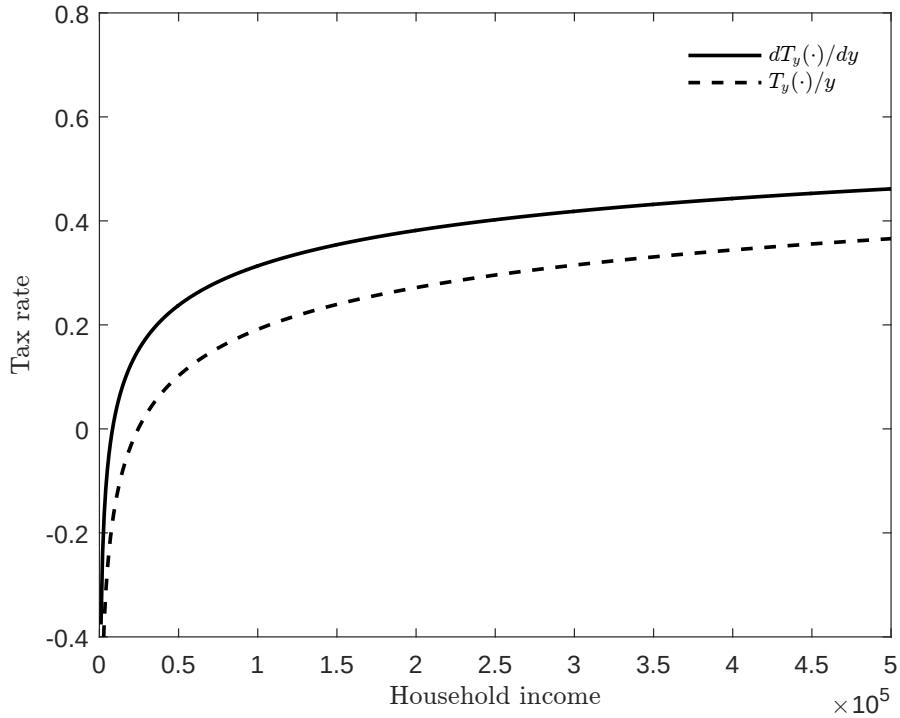

Figure 3: The average and marginal labor income tax rates from Storesletten et al. (2012).

Finally, capital income tax in the U.S. consists of taxes on interest income, and also on capital gains. However, there are no capital gains in current model because it has only one asset and there is no aggregate uncertainty. Therefore, to parameterize the capital income tax function, I follow the literature and assume that $T_k(\cdot) = \tau_k \times k$ and set $\tau_k = 30\%$ (De Nardi et al., 1999; İmrohoroğlu and Kitao, 2012).

4.5 Technology

The historically observed value of capital's share in total income in U.S. ranges between 30-40%, so I set $\alpha = 0.35$. Also, following Stokey and Rebelo (1995), I set the depreciation rate to $\delta = 0.06$.

4.6 Unobservable parameters

Once all the observable parameters have been assigned empirically reasonable values, I jointly calibrate the remaining unobservable parameters of the model, i.e. the preference parameters σ, β, and η, the age-dependent labor force participation cost $\theta_P(s)$, and also the income tax parameter τ_y, to match certain macroeconomic targets.

First, so that overall wealth accumulation in the model matches the U.S. economy, I fix the IES to $\sigma = 4$ and then calibrate the discount factor (β) to target an equilibrium capital-output ratio of 3.0. Second, two salient features of cross-sectional labor supply data in the U.S. are (i) a rapid decline in the labor force participation rate from about 90% to almost 30% between ages 55 to 70, and (ii) an average of 40 hours per week per worker spent on market work between ages 25 to 55 (Kitao, 2014). I adopt both of these empirical facts as targets.

Following Kitao (2014), I assume that the labor force participation cost increases with age based

σ	β	η	κ_1	κ_2	κ_3	τ_y
4	0.997	0.41	0.0531	0.298	5.48	0.136

Table 2: Unobservable parameter values under the baseline calibration.

	Target	Model
Capital-output ratio	3.0	2.83
Avg. hours of market work per week per worker (25-55)	40	41.5
Share of govt. expenditures in GDP	0.2–0.25	0.219
Interest rate	–	0.0636
Ratio of Social Security expenditures to GDP	–	0.051
Benefit adjustment factor	–	0.810

Table 3: Key macroeconomic variables under the baseline calibration.

on the relationship

$$\theta_P(s) = \kappa_1 + \kappa_2 s^{\kappa_3},$$

where s is model age, and then parameterize κ_1, κ_2, and κ_3 to match the observed rapid decline in labor force participation after age 55. The consumption share parameter (η) controls the fraction of time a household spends on market work (conditional of participation), so I calibrate it to match the hours per week target.[10]

Finally, I calibrate τ_y such that the model generates a ratio of government expenditures to GDP of 20–25% in equilibrium. This step ensures that the scale of tax revenues relative to GDP in the model is consistent with that in the U.S. economy.

5 Baseline economy

The unobservable parameter values under which the baseline equilibrium reasonably matches the above targets are reported in Table 2. Note that with leisure in period utility, the relevant inverse elasticity for consumption is $\sigma^c = 1 + \eta(\sigma - 1) = 2.2$, which lies within the range frequently encountered in the literature. Also, with the above values of κ_1, κ_2, and κ_3, the labor force participation cost increases at a faster rate with age (see Figure 4).

The model-generated values for key macroeconomic variables under the baseline calibration are reported in Table 3 along with their targets, and the cross-sectional labor force participation and labor hours data (conditional on participation) are reported in Figures 5 and 6. Note that the benefit adjustment factor in the baseline calibration is smaller than unity, which implies that the PIAs obtained from the U.S. benefit-earnings rule are adjusted downward to balance Social Security's budget in the baseline equilibrium.

It is clear from Figures 5 and 6 that the baseline calibration does a good job of matching observed labor supply behavior in the U.S. It replicates the rapid decline in participation at older ages quite well, and it also reasonably matches the general declining trend of weekly hours over the life cycle. However, the model fails to replicate the mild-hump shape in the hours profile observed in the data. One way to potentially improve the model's fit along this dimensions is to treat the age-dependent component of labor productivity ϵ_s as unobservable. I treat this component of labor productivity as an observable parameter in the calibration, whereas in reality it is an unobservable

[10] I set the maximum disposable time to 16 hours per day.

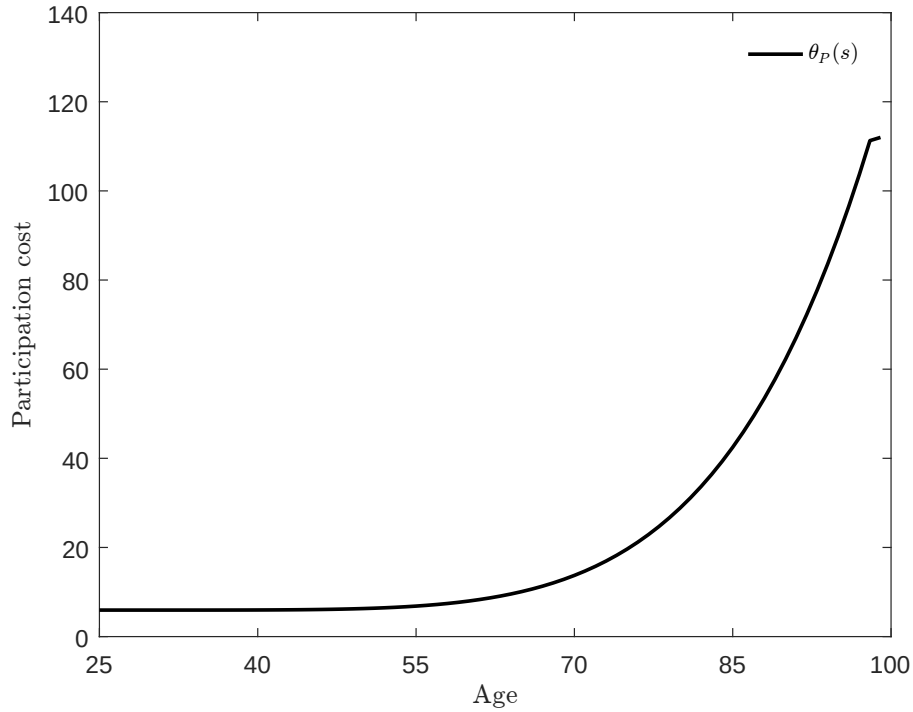

Figure 4: Age-dependent labor force participation cost $\theta_P(s)$ in weekly hours.

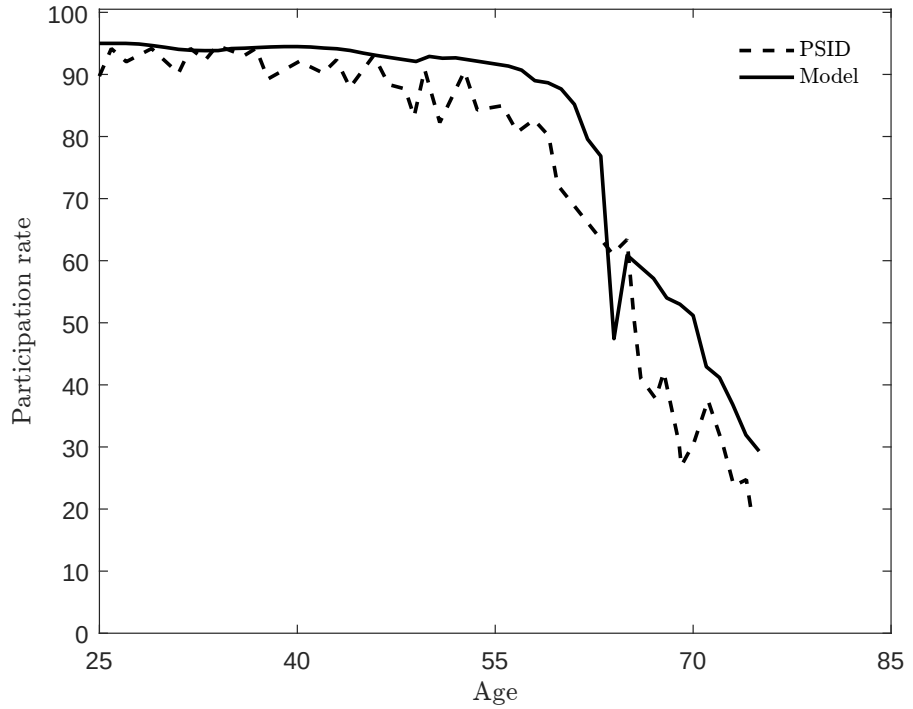

Figure 5: Cross-sectional labor force participation rates under the baseline calibration.

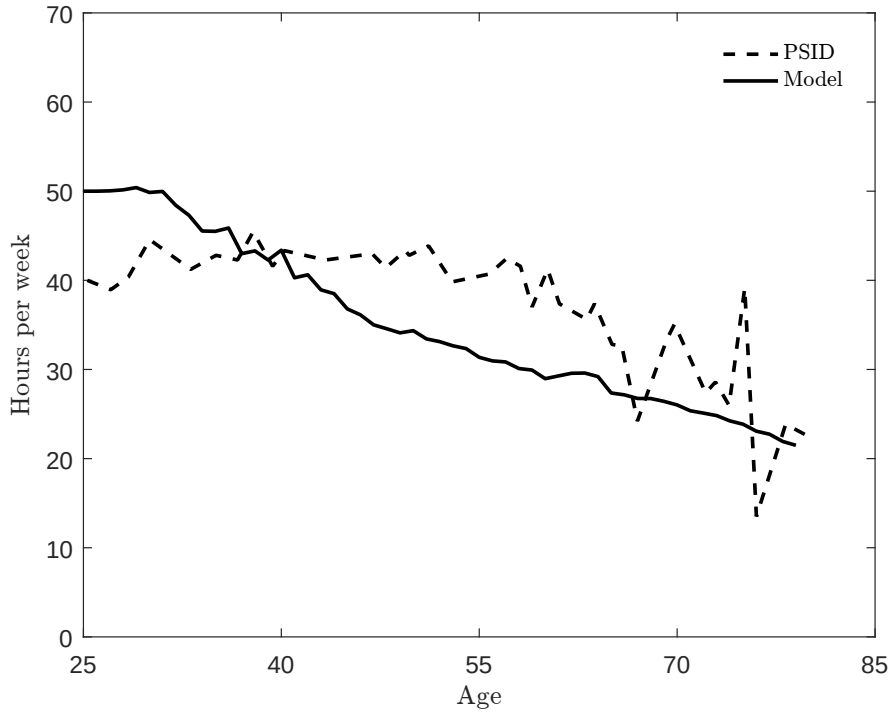

Figure 6: Cross-sectional mean of labor hours per week (conditional on participation) under the baseline calibration.

structural parameter. Treating ϵ_s as an unobservable parameter would potentially eliminate any selection bias arising from measuring it as residual wages (Bullard and Feigenbaum, 2007; Bagchi and Feigenbaum, 2014).

Next, I report the cross-sectional asset holdings by age in the baseline calibration in Figure 7. Based on the figure, it is clear that while the baseline model replicates the hump-shaped relationship between age and asset holdings, it fails to match the level of old-age asset holdings observed in the data (De Nardi et al., 2010). This is due to the fact that in the current model, households accumulate assets to smooth consumption across the work life and retirement (the life-cycle motive), and also across the stochastic realizations of the idiosyncratic productivity shock (the precautionary motive). However, both life-cycle and precautionary motives are less important at later ages, especially because the idiosyncratic productivity shock is highly persistent. Introducing a third factor that affects saving behavior, such as health risks or a bequest motive, would induce older households to increase their asset holdings and potentially improve the model's fit along this dimension.

The combined effect of the above labor supply and savings decisions, along with the calibrated labor productivity and mortality risk processes, can be seen in the baseline cross-sectional distribution of income reported in Figure 8. The figure shows a realistic amount of earnings heterogeneity in the baseline model. In particular, the baseline calibration captures the left tail of the U.S. income distribution quite well. Finally, it is worthwhile at this point to examine the correlation between earnings and mortality risk in the baseline calibration. In Figure 9, I report the death rates at ages 55, 65, and 75 as a function of earnings percentiles. Three facts are clear from the figure. First, as one would expect, mortality risk is positively related to age: death rates experienced by 75-year-old households are larger than those experienced by 65-year olds, which in turn, are almost

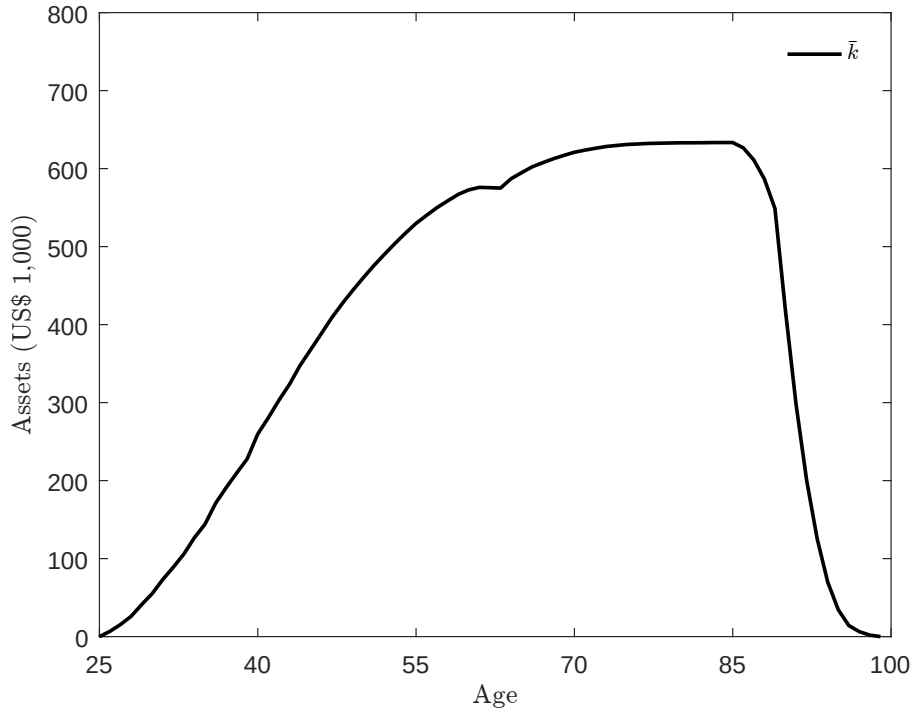

Figure 7: Cross-sectional mean of asset holdings under the baseline calibration.

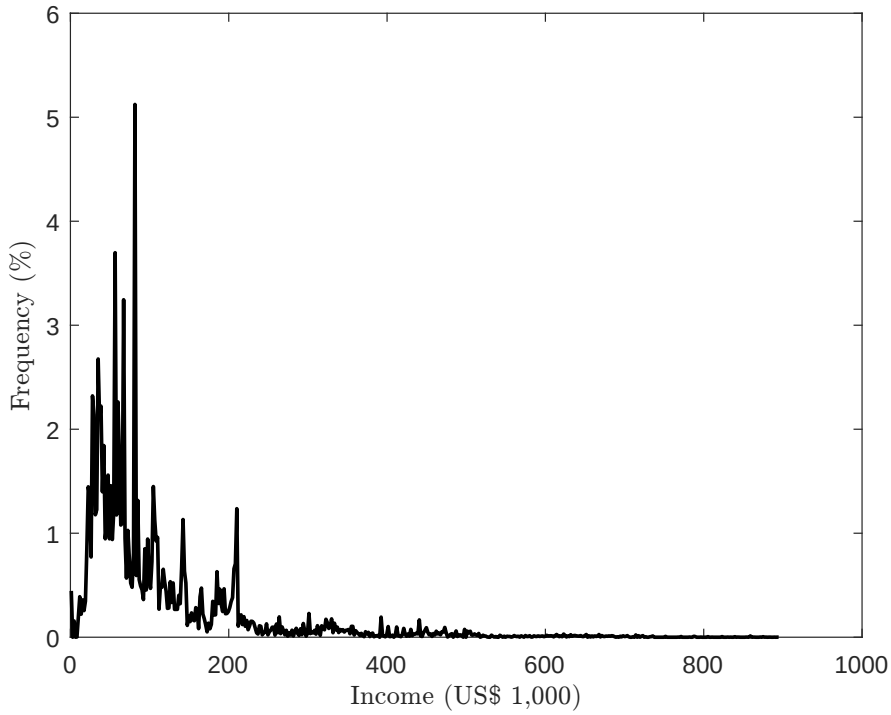

Figure 8: Cross-sectional distribution of income under the baseline calibration.

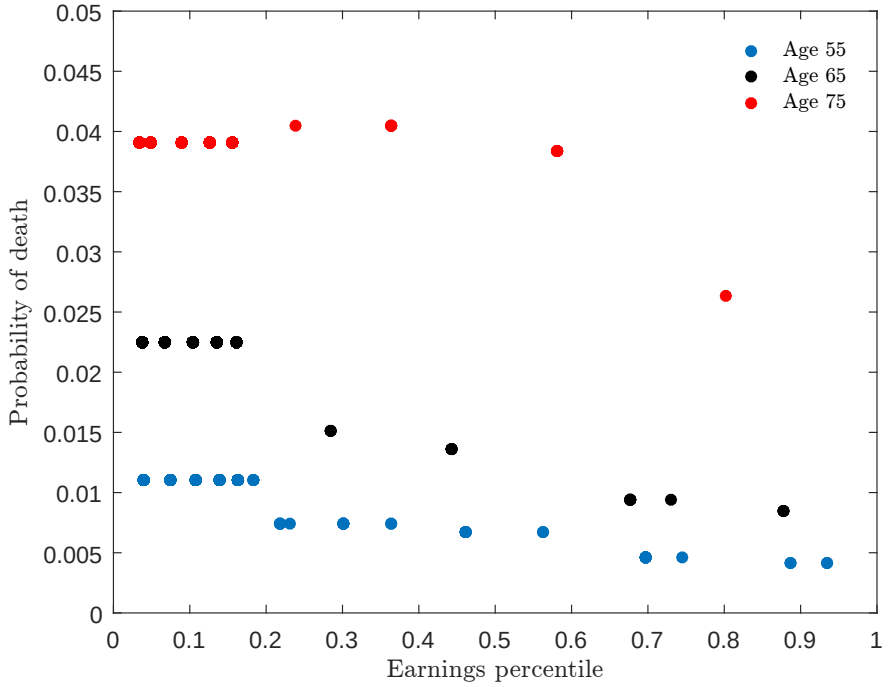

Figure 9: Age-specific death rates as a function of earnings in the baseline calibration.

always larger than those for the 55-year olds. Second, at a given age, households at the bottom of the earnings distribution experience higher death rates. For example, a 65-year-old household at the bottom of the earnings distribution is more than twice as likely to die, compared to a 65-year-old household in the 90th percentile. This pattern remains unchanged for younger as well as older households. Finally, the negative correlation between earnings and mortality risk in the baseline calibration is strong enough to cause some cross-over in survivorship, in the sense that death rates of 65-year olds in the 90th earnings percentile are lower than those experienced by 55-year olds in the 1st–20th percentiles.

6 The experiments

The goal of this paper is to examine if differential mortality has any implications for the welfare effects of Social Security's benefit-earnings rule. Essentially, this experiment involves computing new equilibria of the baseline model with alternative benefit-earnings rules, while holding all the other institutional features of Social Security fixed at their baseline level, under two scenarios: with and without earnings-dependent mortality risk. At this point, two important choices must be made. First, what kind of alternative benefit-earnings rules should be considered in the computational experiments? Second, what type of welfare measures ought to be used in evaluating these alternative benefit-earnings rules?

First, while it is certainly possible to investigate the globally optimal structure of Social Security benefits in a model such as this, I focus only on benefit-earnings rules that are structurally similar to the current U.S. rule.[11] As discussed earlier, Social Security's current benefit-earnings rule is

[11]Note that in general, the current model is not suitable for studying the globally optimal Social Security tax and benefit structure. This is because in a calibrated rational-agent model such as this, the distortions from Social

piecewise linear, with a replacement rate of 90% of the AIME for the first 20% of the average wage, 32% of the AIME for the next 104% of the average wage, and 15% of the AIME for the remaining, up to the taxable maximum of 247% of the average wage in the economy (Figure 1). I henceforth refer to these replacement rates as the primary, secondary, and the tertiary replacement rates, respectively.

With this structure, the progressivity of the benefit-earnings rule is largely determined by the primary replacement rate. Holding the secondary and the tertiary replacement rates constant, reducing the primary replacement rate makes the benefit-earnings rule more proportional (or linear). This has the effect of reducing Social Security's insurance effects, because with a linear benefit-earnings rule, households with relatively unfavorable earnings histories receive the same return on every dollar of Social Security contributions paid as households with relatively favorable earnings histories. On the other hand, increasing the primary replacement rate makes the benefit-earnings rule more progressive or concave, which has the effect of increasing Social Security's implicit insurance. This is because households with relatively unfavorable earnings histories now receive a higher return on every dollar of Social Security contributions paid, compared to those with relatively favorable earnings histories. Because of this reason, I focus only on the primary replacement rate in my computational experiments, and I consider benefit-earnings rules ranging from a linear function of past earnings (i.e. with zero implicit redistribution), to a fixed benefit that is uniform across retirees and completely unrelated to past earnings (i.e. with full redistribution).[12]

It is important to note here that even though I focus only on the primary replacement rate in the computational experiments, keeping Social Security's budget balanced with the current payroll tax rate and the taxable maximum requires the secondary and tertiary replacement rates to be adjusted as well. For example, increasing the primary replacement rate from 90%, while keeping the secondary and tertiary rates fixed at 32% and 15% respectively, leads to an overall increase in Social Security benefits. Therefore, so that Social Security can achieve Pay-As-You-Go balance with the current tax rate and taxable maximum, the secondary and tertiary replacement rates must be reduced. This is accomplished automatically in the model through the benefit adjustment factor, which adjusts the PIA obtained from every alternative benefit-earnings rule to ensure that Social Security's budget is balanced in equilibrium. In fact, as we will see, this adjustment factor even offsets some of the direct change in the primary replacement rate, in addition to the secondary and tertiary rates. This approach has the merit of allowing for the cleanest interpretation of the results, especially from a policy-making perspective, because the benefit-earnings rule fundamentally alters the progressivity in Social Security without altering the overall size of the program.

Second, to evaluate the welfare implications of the alternative benefit-earnings rules, I define the following two measures. To understand the overall welfare consequences, I follow the literature and define

$$W = \sum_{s=0}^{T} \beta^s \sum_{x} \left\{ \Pi_{j=0}^{s-1} Q_j(x) \right\} \; u\left(c_s(x), 1 - h_s(x), P_s(x)\right) \mu_s(x) \tag{42}$$

which is the ex-ante expected lifetime utility of a newborn household. Then, to understand the distributional consequences of these benefit-earnings rules, I define a consumption equivalence ψ

Security are typically larger than its insurance effects (Hubbard and Judd, 1987; İmrohoroğlu et al., 1995; Nishiyama and Smetters, 2008; Bagchi, 2015). As a result, the globally optimal Social Security tax and benefit structure in a model such as the current one would trivially warrant zero Social Security.

[12] In terms of equation (23), a fixed benefit that is completely unrelated to past earnings is obtained by setting the primary replacement rate equal to zero and allowing $\overline{PIA} > 0$.

for each realization of the permanent productivity fixed effect (p) that solves

$$E\left[\sum_{s=0}^{T}\beta^{s}\ \left\{\Pi_{j=0}^{s-1}Q_{j}(x^{C})\right\}\ u\left((1+\psi)\,c_{s}^{C}(x^{C}),1-h_{s}^{C}(x^{C}),P_{s}^{C}(x^{C})\right)\right]=$$
$$E\left[\sum_{s=0}^{T}\beta^{s}\ \left\{\Pi_{j=0}^{s-1}Q_{j}(x^{H})\right\}\ u\left(c_{s}^{H}(x^{H}),1-h_{s}^{H}(x^{H}),P_{s}^{H}(x^{H})\right)\right], \tag{43}$$

where C denotes current Social Security law, and H denotes a hypothetical Social Security law with the alternative benefit-earnings rule. Intuitively, this consumption equivalence captures the welfare gains (or losses) in units of consumption, as a function of the permanent productivity fixed effect, under each one of the computations. Taken together, these two measures provide an overall, as well as a disaggregated picture of the welfare implications of the alternative benefit-earnings rules.

7 Welfare consequences with differential mortality

A good benchmark for examining the welfare consequences of alternative Social Security benefits-earnings rules in a general-equilibrium life-cycle economy is Nishiyama and Smetters (2008). In this study, the authors examine the optimal Social Security benefit structure in an overlapping-generations macroeconomic model with labor income and mortality risk, missing annuity markets, and borrowing constraints. Calibrating the model to match some key features of the U.S. economy, Nishiyama and Smetters (2008) find that the optimal benefit-earnings rule is fairly proportional (or linear), with a strong link between benefits and past work-life income. They argue that Social Security's relatively long averaging period of 35 years already provides some insurance against negative labor income shocks, but in a manner that is more efficient than explicit redistribution through the progressive benefit-earnings rule. This is because while the progressivity in the benefit structure provides insurance against labor income risks that are difficult to insure privately, it also introduces implicit tax rates that distort labor supply. Nishiyama and Smetters (2008) find that the welfare losses from these distortions outweigh the welfare gains from the increased insurance.

I report the welfare consequences of several alternative benefit-earnings rules from the baseline model with differential mortality risk in Table 4. In the first column, I report the primary, secondary, and the tertiary replacement rates of the benefit-earnings rule being examined, and in the second column, I report the corresponding adjustment factor needed to balance Social Security's budget, given the current payroll tax rate and taxable maximum. I combine these two statistics to calculate the "effective" replacement rates in the third column, and in the last column I report overall welfare.[13]

The following two facts are clear from Table 4. First, as expected, modifying the primary replacement rate of the benefit-earnings rule requires adjusting the secondary and tertiary replacement rates, so that Social Security's budget is balanced with the current payroll tax rate and taxable maximum. This adjustment occurs through the benefit-earnings rule adjustment factor, which declines when the primary replacement rate is increased, and increases when it is reduced from its baseline level. In addition to offsetting some of the direct change in primary replacement rate, this leads to a consistent flattening of the benefit-earnings rule, as seen in the declining values of the secondary and tertiary rates in the table. Second, with differential mortality risk, ex-ante expected utility is maximized when benefits are fully proportional (or linear). Increasing the primary

[13]For the fully proportional (or linear) benefit-earnings rule, I set the primary, secondary, and the tertiary replacement rates equal to 0.6.

Replacement rates	Adjustment factor	Effective replacement rates	W
Proportional/linear (0.6/0.6/0.6)	0.383	0.23/0.23/0.23	−47.891
0.45/0.32/0.15	0.944	0.42/0.30/0.14	−47.956
Baseline (0.9/0.32/0.15)	0.810	0.73/0.26/0.12	−48.049
1.8/0.32/0.15	0.617	1.11/0.20/0.09	−48.158
Uniform benefits	–	–	−48.172

Table 4: Replacement rates and aggregate welfare under alternative benefit-earnings rules with differential mortality risk.

Permanent productivity fixed effect (p)	0.54	1.00	1.84
Proportional/linear (0.6/0.6/0.6)	−0.79	0.42	−0.67
0.45/0.32/0.15	0.13	0.17	0.02
1.8/0.32/0.15	0.42	−0.12	−0.11
Uniform benefits	1.15	−0.37	0.07

Table 5: Consumption equivalences (ψ%) under alternative benefit-earnings rules with differential mortality risk.

replacement rate from its baseline value of 90% to 180% reduces overall welfare, but reducing it consistently increases overall welfare. Therefore, similar to Nishiyama and Smetters (2008), the labor supply distortions from a more progressive benefit-earnings rule appear to outweigh the insurance effects even in the presence of differential mortality.

To assess the distribution of welfare gains and losses under the alternative benefit-earnings rules with differential mortality risk, I report in Table 5 the consumption equivalence (ψ) for each realization of the permanent productivity fixed effect (p) under each computation (in percentage terms). As expected, the table shows that increasing the degree of progressivity largely benefits households at the bottom of the income distribution, and decreasing it largely hurts them. Adopting the uniform-benefit arrangement leads to a welfare gain equivalent to an increase of more than 1% in period consumption for households with $p = 0.54$, and adopting the proportional (linear) benefit rule leads to a welfare loss equivalent to a 0.8% reduction in their period consumption. Also, note that households with $p = 1.84$ actually experience a welfare loss with the proportional (linear) benefit rule. This, as we will see, is due to the fact that their Social Security benefits decline slightly under this experiment.

To summarize, my computations suggest that in a general-equilibrium environment with uninsurable labor income and differential mortality risk, the distortionary effects of Social Security's benefit-earnings rule are large enough to warrant benefits that are strictly proportional to, or a linear function of, past work-life income. While this arrangement has negative insurance effects and leads to welfare losses for households with unfavorable earnings histories, it also reduces the implicit tax rates for households with favorable earnings histories. I find that the welfare gains from these lower implicit taxes are larger in magnitude, compared to the welfare losses from worse work-retirement consumption smoothing.

8 Eliminating differential mortality

Based on our simple two-period model outlined earlier, a progressive benefit-earnings rule may increase Social Security's rate of return for households with relatively lower earnings in the presence of differential mortality. However, due to their higher mortality risk, these households will also

Replacement rates	Adjustment factor	Effective replacement rates	W
Proportional/linear (0.6/0.6/0.6)	1.423	0.54/0.192/0.09	−48.512
0.45/0.32/0.15	0.937	0.42/0.3/0.14	−48.492
Baseline (0.9/0.32/0.15)	0.807	0.73/0.26/0.12	−48.484
1.8/0.32/0.15	0.640	1.15/0.2/0.1	−48.479
Uniform benefits	–	–	−48.460

Table 6: Replacement rates and aggregate welfare under alternative benefit-earnings rules without differential mortality risk.

heavily discount the marginal utility of better work-retirement consumption smoothing. The results in the previous section demonstrate the combined effect of these two competing mechanisms: with Social Security's current payroll tax rate and taxable maximum, a fully proportional (linear) benefit-earnings rule delivers higher welfare than Social Security's current benefit formula. In this section, I investigate the key question of this paper: how important is the role of differential mortality in the welfare ranking of these benefit-earnings rules?

To examine this, I compute a hypothetical version of the baseline model with all the observable and structural parameters held fixed at their initial values, but without differential mortality risk. Specifically, I adopt the average age-specific death rates from the 2001 U.S. Life Tables in Arias (2004) to generate the survivor functions for all households under this experiment. Then, with this modified model, I compute the welfare implications of changing Social Security'e benefit-earnings rule in Table 6. As before, I report the replacement rates of the benefit-earnings rule being examined in the first column, the corresponding adjustment factor in the second column, and the "effective" replacement rates in the third column. I report the overall welfare in the last column.

Two facts are clear from the table. First, elimination of differential mortality requires a slightly larger decline in the adjustment factor with benefit-earnings rules more progressive than the current U.S rule. This suggests that without differential mortality, the longer survival of retirees with relatively unfavorable earnings histories has a larger effect on Social Security's budget, compared to the cost savings from the reduced survival of retirees with better earnings histories, and therefore higher benefits. Second, modifying Social Security's benefit-earnings rule has vastly different welfare implications in the absence of differential mortality. In this case, overall welfare is highest when benefits are uniform and completely unrelated to past earnings, and decreases monotonically when benefit-earnings rules less progressive than the current U.S. rule are adopted. In other words, elimination of differential mortality leads to a complete reordering of the welfare ranking of the benefit-earnings rules.

In the context of our simple two-period model, the absence of differential mortality reduces Social Security's internal rate of return for households with lower earnings, but also increases their expected utility from improved work-retirement consumption smoothing. The above results suggest that the latter effect is quantitatively stronger; eliminating differential mortality from the baseline model causes benefit-earnings rules more progressive than the current U.S. to generate higher welfare.

The distributional consequences of the alternative benefit-earnings rules in the absence of differential mortality are reported in Table 7. The table shows that as expected, benefit-earnings rules more progressive than the current U.S. rule generate welfare gains for households with relatively unfavorable earnings histories, and rules less progressive than the current U.S. rule cause welfare losses. For example, adopting the uniform-benefit rule leads to a welfare improvement equivalent to almost a 1.5% increase in period consumption for households with $p = 0.54$. On the other hand, the fully proportional (linear) benefit rule causes a welfare loss equivalent to almost 2% of period

Permanent productivity fixed effect (p)	0.54	1.00	1.84
Proportional/linear (0.6/0.6/0.6)	−1.83	0.16	0.15
0.45/0.32/0.15	−0.27	0.02	0.02
1.8/0.32/0.15	0.28	−0.08	−0.04
Uniform benefits	1.47	−0.12	−0.09

Table 7: Consumption equivalences (ψ%) under alternative benefit-earnings rules without differential mortality risk.

Permanent productivity fixed effect (p)	0.54	1.00	1.84
Proportional/linear (0.6/0.6/0.6)	−0.12	0.07	−0.11
0.45/0.32/0.15	−0.02	0.01	0.0
1.8/0.32/0.15	0.02	0.02	−0.01
Uniform benefits	0.18	−0.04	0.02

Table 8: Change in expected Social Security benefits under alternative benefit-earnings rules with differential mortality risk.

consumption for these households.

9 Macroeconomic effects

So far, we have focused only on the welfare effects of the alternative benefit-earnings rules. I now turn to the macroeconomic effects of these experiments. Perhaps the most important macroeconomic effect of the modifying the benefit-earnings rule, both with and without differential mortality, is how it affects the level of Social Security benefits. I report in Table 8 the percentage change in expected Social Security benefits from the baseline with differential mortality risk, for each value of the permanent productivity fixed effect. The table shows that as expected, households with relatively unfavorable earnings histories receive lower benefits when less progressive benefit-earnings rules are adopted. Under the proportional (linear) benefits rule, benefits for households with a productivity fixed effect of $p = 0.54$ decrease by almost 12% from the baseline, and increase by 7% for those with $p = 1.0$.

It is worth noting from Table 8 that in this case, households with $p = 1.84$ also experience a reduction in their Social Security benefits when rules less progressive than the current U.S. benefit-earnings rules are adopted. This is because these households have a slightly worse earnings history than those with $p = 1.0$ in the baseline calibration, most likely due to the fact that they retire earlier. In fact, this is also the reason why these households experience a slight welfare loss when a proportional (linear) benefit-earnings rule is adopted in the presence of differential mortality (see Table 5).

The effect of modifying Social Security's benefit-earnings rule on the level of benefits is slightly different when differential mortality is eliminated from the baseline model (see Table 9). In this case, adopting the proportional (linear) benefit-earnings rule leads to a 20% reduction in the benefits for households with $p = 0.54$, and a 3.4% increase in the benefits for those with $p = 1.0$ and $p = 1.84$. On the other hand, adopting the uniform-benefit rule leads to a 20% increase in the benefits for households with $p = 0.54$, and almost a 3% reduction in the benefits of those with $p = 1.0$ and $p = 1.84$. Overall, the pattern of changes in benefits remains consistent with the fact that households with relatively worse earnings histories see an increase (reduction) in their benefits when benefit-earnings rules more (less) progressive than the current U.S. rule are adopted.

Permanent productivity fixed effect (p)	0.54	1.00	1.84
Proportional/linear (0.6/0.6/0.6)	−0.20	0.034	0.034
0.45/0.32/0.15	−0.041	−0.005	−0.005
1.8/0.32/0.15	0.046	−0.004	−0.004
Uniform benefits	0.20	−0.027	−0.027

Table 9: Change in expected Social Security benefits under alternative benefit-earnings rules without differential mortality risk.

	K	L	GDP	r	G/Y
Proportional/linear (0.6/0.6/0.6)	0.994	0.999	0.997	1.006	0.998
0.45/0.32/0.15	0.999	0.998	0.999	0.994	1.000
1.8/0.32/0.15	1.001	0.998	0.998	0.999	1.001
Uniform benefits	1.002	1.000	1.001	1.005	0.999

Table 10: Select macroeconomic variables (relative to the baseline) under alternative benefit-earnings rules with differential mortality risk.

I report the effects on other key macroeconomic variables in Table 10, such as aggregate capital, labor, national income, the interest rate, and the share of government expenditures in GDP, relative to the baseline with differential mortality risk. The table shows that in general, reducing the progressivity of the benefit-earnings rule leads to a consistent decline in capital, labor, and GDP, but the effects on the interest rate and the share of government expenditures are non-monotonic. On the other hand, increasing the degree of progressivity increases capital stock, but has non-monotonic effects on the other variables. Overall, it is clear that these changes are very small in magnitude: within a single percentage point. Therefore, modifying the benefit-earnings rule does not appear to have significant general-equilibrium effects, and this result continues to hold even without differential mortality.

To summarize, I find that the welfare ranking of the alternative benefit-earnings rules is, in fact, sensitive to differential mortality. While a more progressive benefit-earnings rule offers better consumption-smoothing benefits for households with relatively unfavorable earnings histories, their higher mortality risk also causes them to heavily discount the utility from old-age consumption. I find that this latter effect is stronger in magnitude: aggregate welfare is maximized with a fully proportional (linear) benefit rule in the presence of differential mortality. Because this effect is absent when differential mortality is eliminated, welfare is maximized with a uniform-benefit rule where benefits are completely unrelated to past earnings. Adopting benefit-earnings rules more progressive than the current U.S. rule increases Social Security benefits for households with unfavorable earnings histories, but these policy experiment do not appear to have significant effects on other macroeconomic variables. In fact, capital, labor, and national income do not change by more than a single percentage point.

10 Conclusions

While linking public pension benefits to work-life income is common within the industrialized world, U.S. Social Security is unusual in the sense that there is an explicit progressive link between average earnings over the work life, and the benefits paid out to an individual. The rationale behind this link is that it provides partial insurance against uninsurable shocks to labor income, such as unemployment or the inability to secure a high-paying job. However, recent empirical evidence

shows that there is a significant positive correlation between earnings and life expectancy, which has the potential of undoing the progressivity built into Social Security's benefit-earnings rule. Therefore, in this paper, I examine if differential mortality has any implications for the welfare effects of Social Security's benefit-earnings rule.

To do this, I first use a simple two-period overlapping generations model with incomplete markets to show that (a) for households with lower earnings, Social Security's internal rate of return from a progressive benefit-earnings rule may be higher in the presence of differential mortality, compared to when mortality risk is symmetric and uncorrelated to earnings, and (b) the welfare gain from this potentially higher internal rate of return, in terms of improved consumption smoothing, depends on households' expected marginal utility of old-age consumption.

Next, I examine this question using a rich multi-period calibrated general-equilibrium model with rational life-cycle permanent-income households, incomplete markets, and uninsurable labor income and differential mortality risks. I use this model to compute the welfare ranking of alternative benefit-earnings rules, ranging from a fully proportional (or linear) function of past earnings (i.e. with zero implicit redistribution), to a fixed benefit that is uniform across all retirees and completely unrelated to past earnings (i.e. with full redistribution), and then examine if the welfare rankings are sensitive to the positive correlation between earnings and survivorship.

My computational results suggest that the welfare ranking of these benefit-earnings rules is, in fact, highly sensitive to differential mortality. In the presence of differential mortality risk, the progressivity of the benefit-earnings rule has two competing effects on welfare. On the one hand, a more progressive benefit-earnings rule provides better work-retirement consumption smoothing for households with relatively unfavorable earnings histories. On the other hand, their relatively high mortality risk causes these households to heavily discount the utility from old-age consumption. I find that the latter effect dominates. With differential mortality, aggregate welfare is maximized when benefits are a fully proportional (linear) function of past income, but when differential mortality is eliminated, the uniform-benefit rule maximizes welfare. In other words, the lower expected utility of old-age consumption limits the benefits of improved work-retirement consumption smoothing when survivorship is positively correlated with earnings.

Both Grochulski and Kocherlakota (2010) and Michau (2014) show that an earnings history-dependent tax-and-transfer scheme can be used to implement the socially optimal allocation in a heterogeneous-agent economy with private information. The findings from this paper suggest that differential mortality may be an important determinant of the welfare implications of such a history-dependent tax system. The presence of differential mortality increases Social Security's internal rate of return for households with unfavorable earnings histories, but it limits their welfare gains from better work-retirement consumption smoothing. My findings suggest that the latter effect is quantitatively stronger: benefits that are perfectly proportional to past work-life income maximize aggregate welfare in the presence of differential mortality.

References

Arias, E., 2004. United States Life Tables, 2001. National Vital Statistics Reports 52 (14).

Bagchi, S., 2015. Labor Supply and the Optimality of Social Security. Journal of Economic Dynamics and Control 58, 167–185.

Bagchi, S., Feigenbaum, J., 2014. Is smoking a fiscal good? Review of Economic Dynamics 17 (1), 170–190.

Boskin, M. J., Kotlikoff, L. J., Puffert, D. J., Shoven, J. B., 1987. Social Security: A Financial Appraisal Across and Within Generations. Working Paper 1891, National Bureau of Economic Research.

Bullard, J., Feigenbaum, J., 2007. A Leisurely Reading of the Life-Cycle Consumption Data. Journal of Monetary Economics 54 (8), 2305–2320.

Conesa, J. C., Garriga, C., 2009. Optimal response to a transitory demographic shock in social security financing. Federal Reserve Bank of St. Louis Review 91 (1), 33–48.

Conesa, J. C., Krueger, D., 1999. Social Security Reform with Heterogeneous Agents. Review of Economic Dynamics 2, 757–795.

Coronado, J. L., Fullerton, D., Glass, T., 2002. Long-Run Effects of Social Security Reform Proposals on Lifetime Progressivity. University of Chicago Press, pp. 149–206.

Coronado, J. L., Fullerton, D., Glass, T., 2011. The Progressivity of Social Security. The B.E. Journal of Economic Analysis & Policy 11 (1), 1–45.

Cremer, H., Lozachmeur, J.-M., Pestieau, P., 2004. Social security, retirement age and optimal income taxation. Journal of Public Economics 88 (11), 2259–2281.

Cristia, J. P., 2009. Rising mortality and life expectancy differentials by lifetime earnings in the united states. Journal of Health Economics 28 (5), 984–995.

De Nardi, M., French, E., Jones, J. B., 2010. Why Do the Elderly Save? The Role of Medical Expenses. Journal of Political Economy 118 (1), 39–75.

De Nardi, M., İmrohoroğlu, S., Sargent, T. J., 1999. Projected U.S. Demographics and Social Security. Review of Economic Dynamics 2 (3), 575–615.

Farhi, E., Werning, I., 2013. Insurance and Taxation over the Life Cycle. The Review of Economic Studies 80 (2), 596–635.

Fuster, L., İmrohoroğlu, A., İmrohoroğlu, S., 2003. A Welfare Analysis of Social Security in a Dynastic Framework. International Economic Review 44 (4), 1247–1274.

Golosov, M., Troshkin, M., Tsyvinski, A., 2011. Optimal Taxation: Merging Micro and Macro Approaches. Journal of Money, Credit and Banking 43, 147–174.

Grochulski, B., Kocherlakota, N., 2010. Nonseparable preferences and optimal social security systems. Journal of Economic Theory 145 (6), 2055–2077.

Heathcote, J., Perri, F., Violante, G. L., 2010. Unequal We Stand: An Empirical Analysis of Economic Inequality in the United States: 1967-2006. Review of Economic Dynamics 13 (1), 15–51.

Hubbard, R. G., Judd, K. L., 1987. Social Security and Individual Welfare: Precautionary Saving, Borrowing Constraints, and the Payroll Tax. The American Economic Review 77 (4), 630–646.

Huggett, M., Ventura, G., 1999. On the Distributional Effects of Social Security Reform. Review of Economic Dynamics 2 (3), 498–531.

İmrohoroğlu, A., İmrohoroğlu, S., Joines, D. H., 1995. A Life Cycle Analysis of Social Security. Economic Theory 6 (1), 83–114.

İmrohoroğlu, A., İmrohoroğlu, S., Joines, D. H., 2003. Time-Inconsistent Preferences and Social Security. The Quarterly Journal of Economics 118 (2), 745–784.

İmrohoroğlu, S., Kitao, S., 2012. Social Security Reforms: Benefit Claiming, Labor Force Participation, and Long-Run Sustainability. American Economic Journal: Macroeconomics 4 (3), 96–127.

Jeske, K., 2003. Pension systems and aggregate shocks. Federal Reserve Bank of Atlanta Economic Review (Q1), 15–31.

Karabarbounis, M., 2012. Heterogeneity in labor supply elasticity and optimal taxation. In: 2012 Meeting Papers. No. 655. Society for Economic Dynamics.

Kitagawa, E., Hauser, P., 1987. Differential Mortality in the United States: A Study in Socioeconomic Epidemiology. Harvard University Press, Cambridge, MA.

Kitao, S., 2014. Sustainable social security: Four options. Review of Economic Dynamics 17 (4), 756–779.

Meyerson, N., Sabelhaus, J., 2006. Is Social Security Progressive? Economic and Budget Issue Brief, Congressional Budget Office.

Michau, J.-B., 2014. Optimal redistribution: A life-cycle perspective. Journal of Public Economics 111, 1–16.

Nishiyama, S., Smetters, K., 2008. The Optimal Design of Social Security Benefits. Working Papers wp197, University of Michigan, Michigan Retirement Research Center.

Saez, E., 2002. Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses. The Quarterly Journal of Economics 117 (3), 1039–1073.

Samuelson, P. A., 1975. Optimum social security in a life-cycle growth model. International Economic Review 16 (3), 539–544.

Sheshinski, E., 2008. Optimum Delayed Retirement Credit. The MIT Press.

Smetters, K., 2003. Is the social security trust fund worth anything? Working Paper 9845, National Bureau of Economic Research.

Stokey, N. L., Rebelo, S., 1995. Growth Effects of Flat-Rate Taxes. Journal of Political Economy 103 (3), 519–550.

Storesletten, K., Violante, G., Heathcote, J., 2012. Redistributive Taxation in a Partial Insurance Economy. 2012 Meeting Papers 588, Society for Economic Dynamics.

Tauchen, G., Hussey, R., 1991. Quadrature-based methods for obtaining approximate solutions to nonlinear asset pricing models. Econometrica 59 (2), 371–396.

Zhao, K., 2014. Social security and the rise in health spending. Journal of Monetary Economics 64 (0), 21–37.