Social Security: Progressive Benefits but Regressive Outcome?

By

Monisankar Bishnu
Nick L. Guo
Cagri S Kumru

ANU Working Papers in Economics and Econometrics
656

December 2017 JEL: E21, E43, G11

ISBN: 0 86831 656 3

Social Security: Progressive Benefits but Regressive Outcome?*

Monisankar Bishnu,† Nick L. Guo,‡ and Cagri S. Kumru§

December 30, 2017

Abstract

In this paper, we study under what conditions a Pay As You Go (PAYG) type social security program can have regressive outcomes even though the benefits of this program are designed to be progressive. Since a PAYG social security program collects payroll taxes whenever agents are working, and it pays retirement benefits as long as retirees are alive, each individual's well being depends on how *long* they contribute to and receive payments from this program as well as how *much*. Empirical evidence suggests that agents who have low income tend to start working earlier and have shorter longevity than those with middle or high income. Implications of the low income groups' shorter mortality are examined both analytically and quantitatively in this paper. We find the conditions under which a PAYG social security program may have a regressive outcome in a simple two period partial equilibrium model. Afterwards, we created a large scale quantitative OLG model calibrated to the US economy to compare aggregate and welfare implications of the US type PAYG, a no progressive PAYG, and a means tested pension program. Our results indicate that incorporating differential mortality into account change the welfare implications.

Key Words: Social Security, Inequality, Progressiveness.

JEL code: E21, E43, G11

*We thank to conference participants at the Finance and Economic Growth in the Aftermath of the Crisis Conference in Milan, in September 2017 and Sixth Delhi Macroeconomics Workshop in New Delhi in October 2017 for their valuable comments.

†Economics and Planning Unit, Indian Statistical Institute - Delhi Centre. Email: mbishnu@isid.ac.in

‡Department of Economics, The University of Wisconsin - Whitewater. Email: nicklguo@gmail.com

§Research School of Economics, The Australian National University. Email: cagri.kumru@anu.edu.au

1 Introduction

In this paper, we show that when the differential mortality rates across income groups taken into account, Pay As You Go (PAYG) social security systems can have regressive outcomes even though the benefits of these programs are designed to be progressive. Since a PAYG social security program collects payroll taxes whenever agents are working, and it pays retirement benefits as long as retirees are alive, each individual's well being depends on how *long* they contribute to and receive payments from this program as well as how *much*. Empirical evidence suggests that agents who have low income tend to start working earlier and have shorter longevity than those with middle or high income. The implications of the differences in mortality rates in the context of social security are examined both analytically and quantitatively in this paper.

Social security programs are large expenditure items and play important insurance and redistribution roles. Hence, aggregate and welfare implications of various programs are well analyzed both analytically and computationally starting with Diamond (1965) and Auerbach and Kotlikoff (1987), respectively. Imrohoroglu et al. (1995), Huggett and Ventura (1999), Nishiyama and Smetters (2007), Kitao (2014), Fehr and Uhde (2013) quantify the aggregate and welfare effects of PAYG, fully funded, and means-tested social security programs. The overall conclusion is redistributive and insurance benefits of PAYG and means-tested programs are exceeded by behavioral distortions generated by those programs. Fehr et al. (2013) quantitatively characterizes the consequences of rising pension progressivity in an incomplete market OLG model and show that more redistributive pension system will improve welfare.

Most developed countries have nominally progressive PAYG social security programs as their benefits are. The US Social Security program has a highly progressive benefit formula to determine monthly payments. Hence, the people with low lifetime earnings get a much higher replacement rate than those with high lifetime earnings. For instance, Social Security might replace 70 percent of earnings for someone with a full-length career in the bottom quantile of the earnings distribution (see Goda et al. (2011) for a detailed discussion). Since benefits are paid as annuities, the total amount of benefits an individual receives depends on the that individual's longevity. If individuals from high income groups can relatively live long enough, the progressive structure of the PAYG system would disappear. Starting with Kitagawa and Hauser (1973), the extent, causes, and trends of differential mortality in the US has been well analyzed empirically. Meara et al. (2008) and Hadden and Rockswold (2008) found increases in indices of mortality inequality by education groups. Waldron (2007) found evidence of a significant increase in differential mortality by lifetime earnings for the 60 and older male in the 1972–2001 period. Cristia (2009) extends Waldron (2007), Meara et al. (2008), and Hadden and Rockswold (2008), which investigated whether higher earners or the better educated enjoyed larger advantages in mortality reductions, and explores increases in life expectancy by individuals in different quintiles of the lifetime earnings distribution. Cristia finds large differentials in age-adjusted mortality rates across individuals in different quintiles of the individual lifetime earnings distribution (for instance men ages 35–49 in the bottom quintile have age-adjusted mortality rates 6.4 times larger than those in the top quintile). The existence of strong empirical evidence regarding mortality differentials across different earning quintiles requires evaluating social security programs once again. The aforementioned earlier studies assume away differences in mortality rates. In this paper, we aim to analyze the implications of social security programs taking differential mortality rates across different earnings quintiles into account.

There is a limited number of studies that analyze the implications of differential moralities across different income groups in the context of PAYG social security. Bommier et al. (2011) study the normative problem of redistribution between individuals who differ in their lifespans. They show the social optimum is obtained when long-lived individuals retire later and consume less per period than short-lived individuals. Le Garrec and Lhuissier (2017) study macroeconomic and distributional consequences of global gain in life expectancy. By considering a framework where individuals decide to acquire skills depending on economic incentives and differential mortality, they show that introducing a 'long career' exception cannot be to the advantage of future unskilled workers unless education yields no spillover effects. Goda et al. (2011) calculate internal rates of return and net present values for the US PAYG program under the assumption of differential mortality without providing any formal model. They show that under the assumption of constant mortality across lifetime income subgroups, the Social Security system is progressive but a good deal of the progressivity is undone or even reversed when differential mortality is taken into account. [1]

In this paper, we first generate a simple two period partial-equilibrium OLG model with differential mortality to lay out the conditions under which a PAYG program can be regressive despite its progressive benefits design. Then, we generate a large scale general equilibrium incomplete market OLG model that is calibrated to the US economy. The model mimics the features of the US income tax system and PAYG Social Security program. We then generate models in which a means-tested pension program and a non-progressive PAYG program replaces the current US PAYG program. We show that once we take into account differential mortality risks, welfare rankings of the PAYG and means-tested programs do not change.

[1] Tan (2015) and Bagchi (2017) also show that differential mortality matters in welfare rankings of various pension programs.

In both non-differential and differential mortality cases, the fixed tax means-tested pension programs dominate the PAYG. Among the means-tested pension programs, the least redistributive one in which benefit reduction rate is equal to zero generates the highest welfare. When we fixed the maximum benefits instead, the most progressive means-tested program in which benefit reduction rate is 100% generates the highest welfare since it comes with the least tax burden. Yet, the welfare ranking of the PAYG and non-redistributional programs depend on whether mortality differentials are taken into account or not. More precisely, when we ignore mortality differentials, progressive PAYG dominates non-progressive-non-redistributional PAYG program. This result changes when take differential mortality into our account and non-redistributional PAYG dominates the progressive PAYG which is in line with our analytical results.

In sum, both analytical and computational models imply that the existence of mortality differences have important aggregate and behavioral implications and should have been taken into account seriously. This is because low income individuals receive pension benefits for relatively shorter period of time. As a result, the progressive benefits would be outweighed by differential mortality risks, and hence the social security becomes regressive in terms of welfare.

The paper is organized as follows. In section 2, we use an analytical model to show that the regressive outcome is possible as a result of a social security program, even though its benefits are designed to be progressive. The only driving force behind this qualitative result is the differential mortality risks. In section 3, we introduce the quantitative model. Section 4 introduces parameter values. In section 5, we calibrate the overlapping generations model to data and provide the results implied by the model. In section 6, we conclude.

2 An Analytical Model

In this section we use a two period partial equilibrium OLG model to analyze the implications of the differential mortalities with the existence of a PAYG type social security system exists in the economy.

2.1 Homogeneous Agents

Assume that there is only one representative agent in each cohort and each agent can live up to two periods indexed by 1 and 2. The survival probability is s. The agent works and receives labor income w in the first period. The income is subject to a social security tax at rate τ. In return, the agent receives a social security benefit b if she survives to the second period. If the agent dies early, his saving, a will be collected by the government and redistributed to the young generation as a bequest income η. This accidental bequest and transfer program is also managed by the government. For simplicity, we assume that there is no population growth and the net return to capital is zero. Individuals preferences are model by a CRRA utility function, where c represents consumption and σ stands for the relative risk aversion coefficient. In this environment, PAYG and fully funded social security problems are equivalent. The representative individual solves life cycle maximization problem

$$\max_{c_1,c_2,a} \frac{c_1^{1-\sigma}}{1-\sigma} + s\frac{c_2^{1-\sigma}}{1-\sigma}, \tag{1}$$

subject to,

$$\begin{aligned} c_1 + a = (1-\tau)w + \eta \quad , \\ c_2 = a + b \qquad\qquad . \end{aligned} \tag{2}$$

The optimal saving, a, and consumption levels are

$$c_1 = \frac{1}{1+s^{\frac{1}{\sigma}}}\left[(1-\tau)w + \eta + b\right], \tag{3}$$

$$c_2 = \frac{s^{\frac{1}{\sigma}}}{1+s^{\frac{1}{\sigma}}}\left[(1-\tau)w + \eta + b\right], \tag{4}$$

$$a = \frac{s^{\frac{1}{\sigma}}}{1+s^{\frac{1}{\sigma}}}\left[(1-\tau)w + \eta\right] - \frac{1}{1+s^{\frac{1}{\sigma}}}b. \tag{5}$$

The government runs a social security program with the budget constraint

$$sb = \tau w. \tag{6}$$

The government also runs a transfer program. It collects accidental bequests and transfer them to the young:

$$(1-s)a = \eta. \tag{7}$$

Given the balanced budget conditions of social security and bequest-transfer programs, the agent's optimal saving is:

$$a = \frac{s^{\frac{1}{\sigma}}}{1+s^{1+\frac{1}{\sigma}}}w - \frac{\tau}{s}w. \tag{8}$$

This illustrates that the private saving is lower after the introduction of the social security program ($\tau > 0$). The equilibrium bequest income is:

$$\eta = (1-s)a = \frac{(1-s)s^{\frac{1}{\sigma}}}{1+s^{1+\frac{1}{\sigma}}}w - \frac{(1-s)\tau}{s}w. \tag{9}$$

The life-time income can thus be written as:

$$\begin{aligned}(1-\tau)w+\eta+b &= (1-\tau)w + \frac{(1-s)s^{\frac{1}{\sigma}}}{1+s^{1+\frac{1}{\sigma}}}w - \frac{(1-s)\tau}{s}w + \frac{\tau}{s}w \\ &= \frac{1+s^{\frac{1}{\sigma}}}{1+s^{1+\frac{1}{\sigma}}}w. \end{aligned}\tag{10}$$

Now we can restate the optimal consumption in each periods are as follows:

$$c_1 = \frac{1}{1+s^{\frac{1}{\sigma}}}\left[(1-\tau)w+\eta+b\right] = \frac{1}{1+s^{1+\frac{1}{\sigma}}}w, \tag{11}$$

$$c_2 = \frac{s^{\frac{1}{\sigma}}}{1+s^{\frac{1}{\sigma}}}\left[(1-\tau)w+\eta+b\right] = \frac{s^{\frac{1}{\sigma}}}{1+s^{1+\frac{1}{\sigma}}}w. \tag{12}$$

As one can notice that the consumption levels are not affected by the social security system. Hence, the welfare, measured by expected life time utility is not affected by social security program either.

$$\begin{aligned}\frac{c_1^{1-\sigma}}{1-\sigma} + s\frac{c_2^{1-\sigma}}{1-\sigma} &= \frac{c_1^{1-\sigma}}{1-\sigma} + s\frac{\left(c_1 s^{\frac{1}{\sigma}}\right)^{1-\sigma}}{1-\sigma} = \frac{c_1^{1-\sigma}}{1-\sigma}\left(1+s^{\frac{1}{\sigma}}\right) \\ &= \frac{1+s^{\frac{1}{\sigma}}}{1-\sigma}\left(\frac{1}{1+s^{1+\frac{1}{\sigma}}}\right)^{1-\sigma} w^{1-\sigma}. \end{aligned}\tag{13}$$

This result is known from Caliendo et. al. (2014). The introduction of social security program pools the contributions and gives benefits only to the survivors. However, social security, on the other hand, decreases private saving and thus reduces bequest income. Since it does not alter the intertemporal choice (the Euler Equation), the social security only has a wealth effect. Hence, in this environment (no private annuity markets, no population growth, interest rate is zero, no other uncertainty), social security does not change welfare.

2.2 Heterogeneous Agents

In this section, we extend the above model by incorporating differences in survival rates. In this model, we assume there are two types of agents, denoted by h and l, representing those who have either high or low wages. Each agent can live up to two periods. The mass of all the young agents is normalized as 1. Type h young agents, have mass α, and thus type l young agents $1-\alpha$. The survival probability of the i type is s^i, where $i=\{h,l\}$. Hence the total population in this economy has mass $1+\alpha s^h+(1-\alpha)s^l$.

For an agent of type i, his problem is

$$\max_{c_1^i, c_2^i, a^i} \frac{(c_1^i)^{1-\sigma}}{1-\sigma} + s^i \frac{(c_2^i)^{1-\sigma}}{1-\sigma}, \tag{14}$$

subject to

$$c_1^i + a^i = (1-\tau)w^i + \eta, \tag{15}$$

$$c_2^i = a^i + b^i. \tag{16}$$

The government runs a balanced budget social security program:

$$\tau\left[\alpha w^h + (1-\alpha)w^l\right] = s^h \alpha b^h + s^l(1-\alpha)b^l. \tag{17}$$

Finally, the bequest-transfer program requires:

$$\alpha(1-s^h)a^h + (1-\alpha)(1-s^l)a^l = \eta. \tag{18}$$

Let's define the maximized utility of each type, $U^h(\tau)$ and $U^l(\tau)$, when the social security is in place and its tax rate at τ. We would like to show in this section that there exist parameters $(\alpha, s^h, s^l, w^h, w^l)$ that, when we set social security policy as $(\tau > 0, b^h, b^l)$, we have progressive benefits and regressive welfare.

The social security program benefits are progressive, in the sense that

$$\frac{b^h}{b^l} < \frac{w^h}{w^l}.$$

When $\sigma < 1$, the welfare outcome is regressive if

$$\frac{U^h(\tau)}{U^l(\tau)} > \frac{U^h(0)}{U^l(0)}.$$

When $\sigma > 1$, the condition is reversed:

$$\frac{U^h(\tau)}{U^l(\tau)} < \frac{U^h(0)}{U^l(0)}.$$

From the previous section, we learned that the indirect utility of each type can be defined:

$$\begin{aligned}\frac{(c_1^i)^{1-\sigma}}{1-\sigma}+s^i\frac{(c_2^i)^{1-\sigma}}{1-\sigma} &= \frac{1+(s^i)^{\frac{1}{\sigma}}}{1-\sigma}(c_1^i)^{1-\sigma}=\frac{1+(s^i)^{\frac{1}{\sigma}}}{1-\sigma}\left[\frac{(1-\tau)w^i+\eta+b^i}{1+(s^i)^{\frac{1}{\sigma}}}\right]^{1-\sigma} \\ &= \frac{\left(1+(s^i)^{\frac{1}{\sigma}}\right)^{\sigma}}{1-\sigma}\left[(1-\tau)w^i+\eta+b^i\right]^{1-\sigma}\equiv P(s^i)(W^i)^{1-\sigma}, \quad (19)\end{aligned}$$

where $P(s^i)=\frac{\left(1+(s^i)^{\frac{1}{\sigma}}\right)^{\sigma}}{1-\sigma}$ is determined by preference parameter and survival probability, and $W^i=(1-\tau)w^i+\eta+b^i$ is the life time wealth of a type i agent.

In order to show that the social security is regressive in outcome, we need to show, when $\sigma<(>)1$, $\frac{U^h(\tau)}{U^l(\tau)}>(<)\frac{U^h(0)}{U^l(0)}$. When $\sigma<1$,

$$\begin{aligned}\frac{U^h(\tau)}{U^l(\tau)} &> \frac{U^h(0)}{U^l(0)}\Longleftrightarrow\frac{P(s^h)(W^h(\tau))^{1-\sigma}}{P(s^l)(W^l(\tau))^{1-\sigma}}>\frac{P(s^h)(W^h(0))^{1-\sigma}}{P(s^l)(W^l(0))^{1-\sigma}}\Longleftrightarrow\left[\frac{W^h(\tau)}{W^l(\tau)}\right]^{1-\sigma}>\left[\frac{W^h(0)}{W^l(0)}\right]^{1-\sigma} \\ &\Longleftrightarrow \frac{W^h(\tau)}{W^l(\tau)}>\frac{W^h(0)}{W^l(0)}\Longleftrightarrow\frac{(1-\tau)w^h+\eta(\tau)+b^h}{(1-\tau)w^l+\eta(\tau)+b^l}>\frac{w^h+\eta(0)}{w^l+\eta(0)} \\ &\Longleftrightarrow \frac{w^h+\eta(\tau)+(b^h-\tau w^h)}{w^l+\eta(\tau)+(b^l-\tau w^l)}>\frac{w^h+\eta(0)}{w^l+\eta(0)}. \quad (20)\end{aligned}$$

Similarly, when $\sigma>1$,

$$\frac{U^h(\tau)}{U^l(\tau)}<\frac{U^h(0)}{U^l(0)}\Longleftrightarrow\left[\frac{W^h(\tau)}{W^l(\tau)}\right]^{1-\sigma}<\left[\frac{W^h(0)}{W^l(0)}\right]^{1-\sigma}\Longleftrightarrow\frac{W^h(\tau)}{W^l(\tau)}>\frac{W^h(0)}{W^l(0)},$$

leading to the same condition as in (20). That is to say, the social security is regressive whenever the life time wealth is regressive. This is the case since the social security does not alter the Euler equation, or intertemporal choices. Note that the following equations showing the equilibrium levels of bequests when there is and isn't social security in place hold:

$$\begin{aligned}\eta(\tau) &= \alpha(1-s^h)a^h(\tau)+(1-\alpha)(1-s^l)a^l(\tau), \quad (21)\\ \eta(0) &= \alpha(1-s^h)a^h(0)+(1-\alpha)(1-s^l)a^l(0), \quad (22)\end{aligned}$$

Note that

$$a^i(\tau)=\frac{(s^i)^{\frac{1}{\sigma}}\left[(1-\tau)w^i+\eta(\tau)\right]-b^i}{1+(s^i)^{\frac{1}{\sigma}}},$$

and therefore incorporating this expression into the expression for $\eta(\tau)$ above gives

$$\begin{aligned}\eta(\tau) &= \alpha(1-s^h)\frac{(s^h)^{\frac{1}{\sigma}}\left[(1-\tau)w^h+\eta(\tau)\right]-b^h}{1+(s^h)^{\frac{1}{\sigma}}} \\ &\quad +(1-\alpha)(1-s^l)\frac{(s^l)^{\frac{1}{\sigma}}\left[(1-\tau)w^l+\eta(\tau)\right]-b^l}{1+(s^l)^{\frac{1}{\sigma}}}.\end{aligned}$$

We now derive the following expression:

$$\eta(\tau)=\frac{\alpha(1-s^h)\frac{(s^h)^{\frac{1}{\sigma}}(1-\tau)w^h-b^h}{1+(s^h)^{\frac{1}{\sigma}}}+(1-\alpha)(1-s^l)\frac{(s^l)^{\frac{1}{\sigma}}(1-\tau)w^l-b^l}{1+(s^l)^{\frac{1}{\sigma}}}}{\Pi},$$

where $\Pi\equiv 1-\alpha(1-s^h)(s^h)^{\frac{1}{\sigma}}/(1+(s^h)^{\frac{1}{\sigma}})-(1-\alpha)(1-s^l)(s^l)^{\frac{1}{\sigma}}/(1+(s^l)^{\frac{1}{\sigma}})$. It is straightforward to show that $\Pi>0$ (please see the Appendix for the derivation). Further, using the above expression, we get

$$\eta(0)=\frac{\alpha(1-s^h)\frac{(s^h)^{\frac{1}{\sigma}}w^h}{1+(s^h)^{\frac{1}{\sigma}}}+(1-\alpha)(1-s^l)\frac{(s^l)^{\frac{1}{\sigma}}w^l}{1+(s^l)^{\frac{1}{\sigma}}}}{\Pi}.$$

Hence,

$$\eta(\tau)-\eta(0)=-\left[\frac{\alpha(1-s^h)(s^h)^{\frac{1}{\sigma}}\tau w^h}{\Pi\left(1+(s^h)^{\frac{1}{\sigma}}\right)}+\frac{(1-\alpha)(1-s^l)(s^l)^{\frac{1}{\sigma}}\tau w^l}{\Pi\left(1+(s^l)^{\frac{1}{\sigma}}\right)}\right]\equiv-\Theta$$

where $\Theta>0$. This implies

$$\begin{aligned}\frac{U^h(\tau)}{U^l(\tau)} &> \frac{U^h(0)}{U^l(0)}\Longleftrightarrow\frac{w^h+\eta(\tau)+(b^h-\tau w^h)}{w^l+\eta(\tau)+(b^l-\tau w^l)}>\frac{w^h+\eta(0)}{w^l+\eta(0)} \\ &\Longleftrightarrow\frac{w^h+\eta(0)-\Theta+(b^h-\tau w^h)}{w^l+\eta(0)-\Theta+(b^l-\tau w^l)}>\frac{w^h+\eta(0)}{w^l+\eta(0)}.\end{aligned}$$

Thus this is the condition needed for regressivity. It simply says that the if the relative gain from the social security program for the rich is higher than that of the poor, we can very well end up with regressivity in utility. Note that that the probability of survival appears in the above inequality via the expression of τ. In the next step, we simplify the above inequality by assuming $w^l=\beta w^h$ where $\beta\in(0,1)$ and $b^l=\delta b^h$ where we do not restrict the value of δ. With these specifications,

$$\frac{w^h}{w^l}>\frac{b^h}{b^l}\Leftrightarrow\frac{\beta}{\delta}<1. \tag{23}$$

Given these assumptions, from (17) we get

$$\tau\left[\alpha w^h+(1-\alpha)\beta w^h\right]=s^h\alpha b^h+s^l(1-\alpha)\delta b^h$$

which ensures the following expression for the tax rate τ

$$\tau=\frac{s^h\alpha+s^l(1-\alpha)\delta}{\alpha+(1-\alpha)\beta}\frac{b^h}{w^h}=\Phi\frac{b^l\beta}{w^l\delta}$$

where $\Phi\equiv s^h\alpha+s^l(1-\alpha)\delta/\alpha+(1-\alpha)\beta>0$. With this expression of τ, the regressivity condition

$$\frac{U^h(\tau)}{U^l(\tau)}>\frac{U^h(0)}{U^l(0)}\Longleftrightarrow\frac{w^h+\eta(0)-\Theta+\left(1-\Phi\frac{\beta}{\delta}\right)b^h}{w^l+\eta(0)-\Theta+(1-\Phi)b^l}>\frac{w^h+\eta(0)}{w^l+\eta(0)}$$

which implies that the parametric condition needed for regressiveness is

$$\frac{(1-\Phi\frac{\beta}{\delta})b^h-\Theta}{(1-\Phi)b^l-\Theta}>\frac{w^h+\eta(0)}{w^l+\eta(0)}. \tag{24}$$

Thus we have the following proposition.

Proposition 1. *If the condition (24) holds, welfare outcome can be regressive even though the social security benefit program is progressive.*

In this section we showed that the regressive outcome is possible when we take the mortality differentials into account even though the PAYG program is progressive in benefits. Now we extend this model and generate a multi-period incomplete market general equilibrium model that mimics the stylized facts in the US economy to investigate the aggregate and welfare implications of various progressive and relatively non-progressive pension programs.

3 The Model Economy

We use a general equilibrium OLG model economy with uninsured idiosyncratic shocks to labor productivity and mortality. The main features of our model follow those of Conesa et al. (2009).

3.1 Demographics

Time is discrete. In each period a new generation is born. Individuals live a maximum of J periods. The population grows at a constant rate n. All individuals face a probability (s_j) of surviving from age j to $j+1$ conditional on surviving up to age j. Individuals retire at exogenously determined retirement age $j^*$and receive relevant pension benefits.

3.2 Endowments

Let $j \in \hat{J} = \{1, 2, ...J\}$ denote age. An individual's labor productivity in a given period depends on age, permanent differences in productivity due to differences in education or abilities, and an idiosyncratic productivity shock to the individual's labor productivity. In other words, agents are heterogeneous in terms of labor productivity. Age-dependent labor productivity is denoted by $\bar{e}_j$. Each individual is born with a permanent ability type $\hat{e}_i \in \hat{E} = \{\hat{e}_1, \hat{e}_2, ..., \hat{e}_m\}$ with probability $p_i > 0$. An individual's average income up to age j is given by $y_j^{avg} \in Y^{avg} A \subset R^+$. Individuals face an idiosyncratic shock $\psi \in \Psi = \{\psi_1, \psi_2, ..., \psi_n\}$ to labor productivity. The stochastic process for ψ is identical and independent across individuals and follows a finite-state Markov process with a stationary distribution over time: $Q(\psi, \Psi) = \Pr(\psi' \in \Psi | \psi)$. We assume that Q consists of only strictly positive entries and, hence, Π is the unique, strictly positive, invariant distribution associated with Q. Initially each individual has the same average stochastic productivity given by $\overline{\psi} = \sum\limits_{\psi} \psi \Pi(\psi)$, where $\Pi(\psi)$ is the probability of ψ. Hence, an ability type $\hat{e}_i$ individual's labor supply at age j in terms of efficiency units is written as $\bar{e}_j \hat{e}_i \psi l_j$, where l_j is hours of work. Let $a \in A \subset R^+$, where a denotes asset holdings. A is a compact set. Its upper bound never binds and its lower bound is equal to zero. We define the space of individuals' state variables as follows: $X = \hat{J} \times A \times \hat{E} \times Y^{avg} \times \Psi$. Note that at any time t, an individual is characterized by the state set $x = (j, a, \hat{e}_i, y^{avg}, \psi) \in X$. Let M be the Borel σ-algebra generated by X and let $B \in M$. Define μ as the probability measure over M. Hence, we can represent individuals' type distribution by the probability space (X, M, μ).

3.3 Preferences

Individuals have preferences over consumption and leisure sequence $\{c_j, (1 - l_j)\}_{j=1}^{J}$ represented by a standard time separable utility function:

$$E\left[\sum_{j=1}^{J} \beta^{j-1} u(c_j, 1 - l_j)\right], \tag{25}$$

where E is the expectation operator and β is the time-discount factor. Expectations are taken over the stochastic processes that govern idiosyncratic labor productivity risk and longevity.

3.4 Technology

A representative firm produces output Y at time t by using aggregate labor input measured in efficiency units (L) and aggregate capital stock (K). The technology is represented by a Cobb-Douglas constant returns to scale production function:

$$Y_t = A_t K_t^{\alpha} L_t^{1-\alpha}. \tag{26}$$

A_t is the level of total factor productivity. Output shares of capital stock and labor input are given by α and $(1-\alpha)$, respectively. The capital stock depreciates at a constant rate $\delta \in (0,1)$. The representative firm maximizes its profit by setting wage and rental rates equal to the marginal products of labor and capital, respectively:

$$w_t = A_t(1-\alpha)(\frac{K_t}{L_t})^{\alpha}, \tag{27}$$

$$r_t = A_t\alpha(\frac{K_t}{L_t})^{\alpha-1}. \tag{28}$$

3.5 The Public Sector

A j year old individual's labor income, capital income, and gross taxable income in year t are given as follows:

$$\begin{aligned} y_{l,t} &= w_t \bar{e}_j \hat{e}_i \psi l_j, \\ y_{k,t} &= r_t(a_t + \eta_t), \\ y_t &= y_{l,t} + y_{k,t}. \end{aligned}$$

The state variable y_j^{avg} denotes an individual's average earnings up to age j.

In the benchmark economy, the government runs an earnings-dependent PAYG pension program. This program taxes an individual's labor income before the retirement age $j^*$and pays old age pension. Payroll taxes are proportional to labor earnings up to the maximum taxable level $y_{l,t}^{max}$. Earnings more than the maximum taxable level are not taxed. Hence, the payroll tax paid at age j in year t can be equal to the following:

$$\tau_p min\{y_{l,t} l_t, y_{l,t}^{max}\},$$

where τ_p is the payroll tax rate. Starting with the retirement age j^*, a PAYG benefit $b_t(y_j^{avg})$, which is a fixed function of an accounting variable y_j^{avg} is transferred:

$$b_t(y_j^{avg}) = \begin{cases} 0.9y_j^{avg} & if \quad y_j^{avg} \leq 0.21\bar{y} \\ 0.189\bar{y} + 0.32(y_j^{avg} - \bar{y}) & if \quad 0.21\bar{y} < y_j^{avg} \leq 1.29\bar{y} \\ 0.5346\bar{y} + 0.15(y_j^{avg} - \bar{y}) & if \quad y_j^{avg} \geq 1.29\bar{y} \end{cases}$$

where $\bar{y}$ represents average yearly earnings in the economy. Following Huggett and Parra (2010), we set the bend points, the maximum earnings $y_{l,t}^{max}$ and the slopes of the benefit function equal to the actual values used in the US social security system.

We run two experiments. In the first experiment, we replace the PAYG system with a means tested benefit system similar to ones in the UK and Australia. Means-tested benefits are determined as follows:

$$b_t^*(x) = \max[\overline{b_t} - \phi y_t, 0], \tag{29}$$

where $b_t^*(y_t)$ is the means-tested benefit received by a retired individual at time t; $\overline{b_t}$ is the maximum amount of means-tested pension benefits that can be received at time t; and ϕ is the taper (benefit reduction) rate.[2] As in the PAYG case, this system is also financed through payroll taxes.

In the second experiment, we simply impose non-progressive PAYG by making benefits proportional to an individual's average earnings as follows:

$$\widehat{b_t}(y_j^{avg})) = \mathit{\Gamma} y_j^{avg},$$

where Γis the replacement rate.

Since individuals face stochastic life-span and private annuity markets are closed by assumption, a fraction of the population will leave accidental bequests. The government confiscates all accidental bequests and delivers them to the remaining population in a lump-sum manner. We denote these transfers by η_t.

Finally, the government faces a sequence of exogenously given consumption expenditures $\{G_t\}_{t=1}^{\infty}$. To finance its consumption, the government levies taxes on capital income, labor income, and consumption. Pension programs in the model are self-financing and benefits are financed through payroll tax collections.

As in Huggett and Parra (2010), we determine income taxes in the model by applying

[2] In our model individuals can receive the means-tested benefits only after they reach the exogenously determined retirement age and benefits are income tested only. In countries that run means-tested pension programs such as the UK and Australia, individuals might be entitled to means-tested benefits before they reach the pension age and the means-tested benefits are also subject to asset tests. In our model, since individuals do not work after the retirement, retirement income comes from asset holdings only and hence, two tests are equivalent. In addition, in our model, means-tested pension program is self-financed as the PAYG program. In the UK and Australia, programs are financed from the general budget.

an income tax function to an individual's income. More precisely, we choose income taxes $T_t(y_t, j^*)$ before and after the retirement age j^* to approximate the average tax rates in the US.

In addition to taxes on capital and labor incomes, the government taxes consumption expenditures at an exogenously given proportional rate τ_c, which does not change in all experiments.

3.6 An Individual's Decision Problem

In the benchmark economy, individuals face the following budget constraint:

$$\left\{ \begin{array}{c} (1+\tau_c)c_t + a_{t+1} \leq y_t - T_t(y_t) - \tau_p y_{l,t} \text{ when } j < j^* \\ (1+\tau_c)c_t + a_{t+1} \leq y_t - T_t(y_t) + b_t(x) + b_t^*(x) \text{ when } j \geq j^* \\ (1+\tau_c)c_t = y_t - T_t(y_t) + b_t(x) + b_t^*(x) \text{ when } j = J. \end{array} \right\} \tag{30}$$

Individuals also face the following borrowing constraint:

$$a_{t+1} \geq 0. \tag{31}$$

The decision problem of an individual in our model economy can be written as a dynamic programming problem. Denoting the value function of the individual at time t by V_t, the decision problem is represented by the following problem:

$$V_t(x_t) = \max_{c,l} \{u(c_t, 1-l_t) + \beta s_j \int V_{t+1}(x_{t+1}) Q(\psi, d\psi')\} \tag{32}$$

subject to the aforementioned budget and borrowing constraints.

3.7 Equilibrium

Our competitive and stationary competitive equilibrium definition are as follows. Given sequences of government expenditures $\{G_t\}_{t=1}^{\infty}$, consumption tax rates $\{t_c\}_{t=1}^{\infty}$, payroll tax rate $\{\tau_p\}_{t=1}^{\infty}$, the PAYG benefit formula given by the function $b(y_j^{avg})$ and initial conditions K_1 and Φ_1, a competitive equilibrium is a sequence of value functions $\{V_t\}_{t=1}^{\infty}$ and optimal decision rules $\{c_t, a_{t+1}, l_t\}_{t=1}^{\infty}$, measures $\{\Phi_t\}_{t=1}^{\infty}$, aggregate stock of capital and aggregate labor supply $\{K_t, L_t\}_{t=1}^{\infty}$, prices $\{r_t, w_t\}_{t=1}^{\infty}$, transfers $\{\eta_t\}_{t=1}^{\infty}$, and tax policies $\{T_t(.)\}_{t=1}^{\infty}$ such that [3]

[3]If the means tested pension program is in place, replace the above sentence with the following "Given sequences of government expenditures $\{G_t\}_{t=1}^{\infty}$, consumption tax rates $\{t_c\}_{t=1}^{\infty}$, payroll tax rate $\{\tau_p\}_{t=1}^{\infty}$, the maximum amount of means-tested benefits can be received $\{\overline{b_t}\}_{t=1}^{\infty}$, benefit reduction rate $\{\phi\}_{t=1}^{\infty}$ and initial conditions K_1 and Φ_1, a competitive equilibrium is a sequence of value functions $\{V_t\}_{t=1}^{\infty}$ and optimal decision rules $\{c_t, a_{t+1}, l_t\}_{t=1}^{\infty}$, measures $\{\Phi_t\}_{t=1}^{\infty}$, aggregate stock of capital and aggregate labor supply $\{K_t, L_t\}_{t=1}^{\infty}$,

1. $\{V_t\}_{t=1}^{\infty}$ is a solution to the maximization problem defined above by 32. Associated optimal decision rules are given by the sequence $\{c_t, a_{t+1}, l_t\}_{t=1}^{\infty}$.

2. The representative firm maximizes its profit according to the equations 27 and 28.

3. All markets clear:

   (a) $K_t = \int a\Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$,

   (b) $L_t = \int \bar{e}_j \hat{e}_i \psi l_j(j, a, \hat{e}_i, y^{avg}, \psi)\Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$,

   (c) $C_t = \int c_t(j, a, \hat{e}, y^{avg}, \psi)\Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$,

   (d) $C_t + K_{t+1} + G_t = Y_t + (1-\delta)K_t$.

4. Law of motion

   (a) for all $\hat{J}$ such that $1 \notin \hat{J}$ is given by $\Phi_{t+1}(\hat{J}\times A\times \hat{E}\times Y^{avg}\times\Psi) = \int P_t((j, a, \hat{e}_i, y^{avg}, \psi); \hat{J}\times A \times \hat{E} \times Y^{avg} \times \Psi)\Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$, where

   (b) $P_t((j, a, \hat{e}_i, y^{avg}, \psi); \hat{J}\times A\times\hat{E}\times Y^{avg}\times\Psi) = \begin{cases} Q(\psi, \Psi)s_j \text{ if } j+1 \in J, a_{t+1}(j, a, \hat{e}_i, y^{avg}, \psi) \in A, \hat{e}_i \\ \text{0 else} \end{cases}$

   (c) for $\hat{J} = \{1\}$: $\Phi_{t+1}(\{1\}\times A\times\hat{E}\times Y^{avg}\times\Psi) = (1+n)^t \begin{cases} \sum_{\hat{e}_i\in\hat{E}} p_{\hat{e}_i} \text{ if } 0 \in A, \overline{\psi} \in \Psi \\ \text{0 else} \end{cases}$

5. Transfers are given by $\eta_{t+1}\int \Phi_{t+1}(dj\times da\times d\hat{e}_i\times dy^{avg}\times d\psi) = \int(1-s_j)a_{t+1}(j, a, \hat{e}_i, y^{avg}, \psi)\Phi_t(dj\times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$.

6. PAYG pension program is self financing: $\tau_{p,t}\int y_{l,t}\Phi_t(\{1, ..., j^*-1\} \times da \times d\hat{e}_i \times dy^{avg} \times d\psi) = \int b_t(j, a, \hat{e}_i, y^{avg}, \psi)\Phi_t(\{j^*, ..., J\}\times da\times d\hat{e}_i\times dy^{avg}\times d\psi)$. If it is non-redistributive PAYG, replace b_t by $\widehat{b_t}$.

---

prices $\{r_t, w_t\}_{t=1}^{\infty}$, transfers $\{\eta_t\}_{t=1}^{\infty}$, and tax policies $\{T_t(.)\}_{t=1}^{\infty}$ such that." If the non-redistributional PAYG is in place replace the above sentence with the following "Given sequences of government expenditures $\{G_t\}_{t=1}^{\infty}$, consumption tax rates $\{t_c\}_{t=1}^{\infty}$, payroll tax rate $\{\tau_p\}_{t=1}^{\infty}$, the non redistributive PAYG benefit formula given by the function $\widehat{b_t}(y_j^{avg}))$ and initial conditions K_1 and Φ_1, a competitive equilibrium is a sequence of value functions $\{V_t\}_{t=1}^{\infty}$ and optimal decision rules $\{c_t, a_{t+1}, l_t\}_{t=1}^{\infty}$, measures $\{\Phi_t\}_{t=1}^{\infty}$, aggregate stock of capital and aggregate labor supply $\{K_t, L_t\}_{t=1}^{\infty}$, prices $\{r_t, w_t\}_{t=1}^{\infty}$, transfers $\{\eta_t\}_{t=1}^{\infty}$, and tax policies $\{T_t(.)\}_{t=1}^{\infty}$ such that."

7. Means-tested pension program is self-financing: $\tau_{p,t} \int y_{l,t} \Phi_t(\{1, ..., j^* - 1\} \times da \times d\hat{e}_i \times dy^{avg} \times d\psi) = \int b_t^*(j, a, \hat{e}_i, y^{avg}, \psi) \Phi_t((dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$.

8. Government runs a balanced budget: $G_t = \int T_t[y_{l,t}] \Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi) + \int \tau_k r_t(a + \eta_t) \Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi) + \tau_c \int c_t \Phi_t(dj \times da \times d\hat{e}_i \times dy^{avg} \times d\psi)$.

Definition. A stationary equilibrium is a competitive equilibrium in which per capital variables and functions, prices, and policies are constant. Aggregate variables grow at the constant rate n.

4 Calibration

This section defines the parameter values of our model. The values of calibrated parameters for the benchmark economy are presented in Table 2.

Demographics

Each model period corresponds to a year. Individuals are born at a real age of 25 (model age of 1) and they can live up to a maximum real life age of 85 (model age of 61). The population growth rate is assumed to be equal to the long-term average US population growth rate between 1960 and 2009, i.e. $n = 1, 1\%$.[4]

In calculating survival probabilities for different income groups we benefited from Cristia (2009) and Bell and Miller (2002). Table 1 reports the differential mortality rates calculated by Cristia for three different age groups and five different income groups. The mortality ratios present the likelihood of death of a respective income group relative to the population average at that same age.

	Ages		
Income Groups	**35 - 49**	**50 - 64**	**65 - 75**
Top	0.35	0.61	0.74
4th	0.56	0.68	0.94
3rd	0.73	0.99	1.08
2nd	1.13	1.10	1.14
Bottom	2.25	1.63	1.10

Table 1: Mortality ratios by income levels

[4]See the Statistical Abstract of the US (2012).

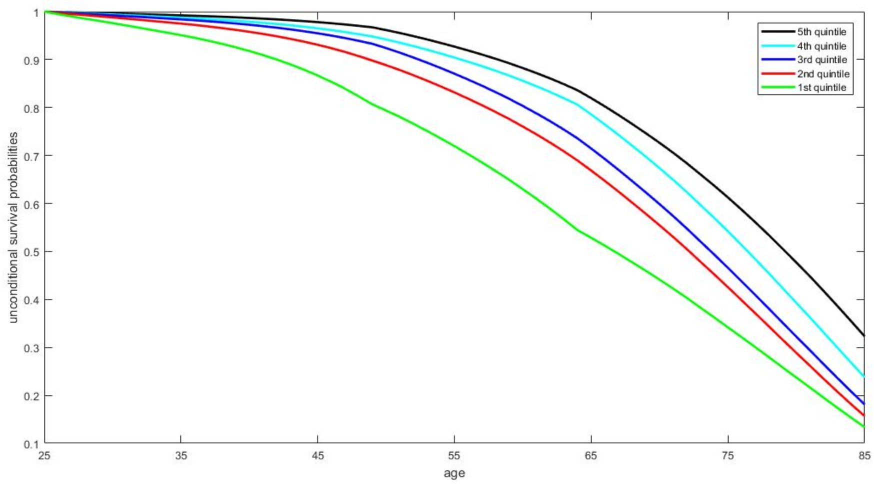

Figure 1: Unconditional survival probabilities for different income groups

We first set the average conditional survival probabilities in accordance with Bell and Miller (2002) estimates by adjusting each cohort's share of population by taking the population growth rate into our account. To find income group specific conditional survival probabilities, we then took the differential mortality rates into our account. The unconditional survival probabilities for five income groups are given in Figure 1.

Finally, we set the mandatory retirement age to 65 (model age of 41).

Endowments

An individual's wage income at time t, expressed in logarithms, is given by $\log(w_t) + \log(\bar{e}_j) + \log(\hat{e}_i) + \log(\psi)$. The age-dependent efficiency index,$\bar{e}_j$ taken from Peterman (2016). Permanent and persistent idiosyncratic shocks to individuals' productivity are normally distributed with a mean zero and the values of the shock parameters are set equal to Kaplan (2012)'s estimates: $\rho = 0.958$, $\sigma^2_{\hat{e}} = 0.065$, $\sigma_\psi = 0.017$.

Preferences

Individuals have time-separable preferences over consumption and leisure. We use the following additively separable utility function:

$$u(c, 1-l) = \frac{c^{1-\upsilon}}{1-\nu} + \vartheta \frac{(1-l)^{1-\sigma}}{1-\sigma}. \tag{33}$$

We set the utility function parameters are equal to Kaplan (2012)'s estimates i.e. the coefficient of relative risk aversion $\nu = 1.66$; the coefficient that governs the Frisch elasticity,

σ=5.55; the parameter that captures the relative importance of leisure, $\vartheta = 0.13$. We set time-discount factor $\beta = 0.965$ in the benchmark model to generate the capital-output ratio of approximately 2.7.

Technology

We set the value of capital's income share to 0.36. We set the value of δ in such a way that we can generate investment-output ratio of 25.5%. The technology level, A, can be chosen freely and we set it to 1.

Government Policy

In the benchmark economy, we use the PAYG benefit function we introduced earlier in calculation of PAYG benefits. The respective payroll tax rate is endogenously determined. In the no redistributional PAYG program we use the same replacement rate for all individuals. We find the replacement rate keeping the payroll tax at the same rate as in the benchmark. When means-tested pension program is in use, we find the value of the maximum amount of means-tested benefits that can be received, $\bar{b}$, by keeping the payroll tax rate the same as in the benchmark model and setting the benefit reduction rate to 100%. We set government expenditure G to 17% of GDP and set the consumption tax rate τ_c to 5%.

5 Results

In section 2, employing a simple model, we showed that a PAYG program can be regressive if we take the differential mortality into our account. In this section, we provide the results of our large scale quantitative model. More precisely, we compare the aggregate and welfare implications of the current earnings dependent PAYG program with various means-tested programs and earnings dependent non redistributive PAYG programs. In the model, both PAYG and means-tested pension programs are self-financing and financed through the payroll taxes. The only difference between the programs are the way benefits are calculated. In the earnings dependent PAYG programs, benefits depend on average past earnings. In the means-tested programs, benefits depend private income after the retirement. The aggregate and welfare implications of the meas-tested programs and comparisons between PAYG and means tested programs are already well analyzed (see Kitao (2014) and Fehr and Uhde (2013)). Yet, the earlier studies often overlooked the mortality differentials across different income groups. Our experiments will offer an answer regarding the role of differential mortality in comparing the PAYG program with means-tested programs. In addition, we analyze the welfare and aggregate implications of replacing the current PAYG program with a non redistributive PAYG program.

Parameter	Value	Source/Target
Maximum possible life span J	61 (real age of 85)	By assumption
Obligatory retirement age j^*	41 (real age of 65)	By assumption
Growth rate of population n	1.1%	Data
Conditional survival probabilities $\{s_j\}_{j=1}^{J}$	See text	Data
Endowments		
Age efficiency profile $\{\bar{e}_j\}_{j=1}^{j^*-1}$	Peterman (2016)	Data
Variance types $\sigma^2_{\hat{e}}$	0.065	Kaplan (2012)
Variance shocks σ^2_{ψ}	0.017	Kaplan (2012)
Persistence ρ	0.958	Kaplan (2012)
Preferences		
Annual discount factor of utility β	0.995	K/Y=2.7
Risk aversion υ	1.66	Kaplan (2012)
Frisch elasticity σ	5.55	0.27
Value of leisure ϑ	0.13	Kaplan (2012)
Production		
Capital share of the GDP α	0.36	Data
Annual depreciation of capital stock δ	8.33%	Peterman (2016)
Scale parameter A	1	Normalization
Government		
Consumption tax rate τ_c	5%	Conesa et al. (2009)
Government expenditures G	17%	Peterman (2016)

Table 2: Calibration parameters

In order to compare welfare across economies with different pension programs we compute the consumption equivalent variation (CEV), which is simply the uniform percentage decrease in consumption required to make an agent indifferent between being born under the new pension program (comparison case) relative to being born under the benchmark economy. A positive CEV reflects a welfare increase due to the new program compared to the baseline case.

In sum, we compare welfare and aggregate implications of PAYG, means-tested, and non-redistributive PAYG programs under two different economies. In the fist economy, we assume that all individuals face the same age dependent survival probabilities. In the second economy, we assume that individuals who differ from each other due to the permanent differences in abilities also differ in terms of mortality rates they face.

5.1 No Differential Mortality

In the benchmark economy, the tax-transfer system mimics the US tax system and PAYG social security program. We calibrated the model economy to the US economy by hitting the

	τ_p	L	K	Y	CEV(%)
PAYG	0.22	100.000	100.000	100.000	
MT 100%	0.22	98.674	91.362	95.689	0.59
MT 80%	0.22	98.768	93.169	96.619	0.95
MT 60%	0.22	98.863	94.446	97.210	1.20
MT 40%	0.22	98.943	95.895	97.689	1.29
MT 20%	0.22	98.975	98.259	98.724	1.78
MT 0%	0.22	99.048	100.4873	99.613	2.07

Table 3: No differential mortality - PAYG vs Means-tested pensions with the fixed tax rate

aforementioned targets. In Table 3, we normalized the values of the benchmark economy at 100 to make the comparison easier. After calibrating the benchmark economy, we replaced the PAYG pension program with a means-tested program by keeping the payroll tax rates across the economies same. Our means-tested pension programs differ from each other by two dimensions: benefit reduction rate and the maximum benefit. In order to make a meaningful comparison across economies, we needed to keep the tax burden same. Since higher benefit reduction rates (Θ) reduce revenue requirements of the means-tested system, we increased the maximum pension benefits to keep the tax burden same across economies. Higher benefit reduction rates with higher maximum pension benefits imply more redistributive means-tested programs. L is aggregate level of labor supply; K is the aggregate capital stock; and Y is the output.

When benefit reduction rate is 100%, individuals with low accumulated wealth receive very generous pension benefits. In contrast, some individuals will end up receiving no pension benefits if their accumulated wealth is large enough. This pension program is quite progressive as the current PAYG program. The only difference is while in the current PAYG program the average past earnings determine the pension benefits, in the means tested program, individuals' private wealth at retirement determines their pension benefits. When we replace the current PAYG with the means tested program with 100% replacement rate, we see that both aggregate labor supply and capital stock decrease substantially. Since relatively low income groups face large pension benefits, leisure becomes relatively cheap and hence, we see a huge drop in labor supply. Similarly, ex-ante more productive types might choose to reduce their labor supply to be eligible for generous pension programs. No surprisingly the capital stock decreases at a very large level. There are two reasons. First, all types of individuals prefer to save less in order to maximize the amount of pension benefits they will receive. Second, relatively rich individuals prefer to decumulate their private wealth as early as possible to receive the generous pension benefits. Since the aggregate capital stock and labor supply decrease substantially, the aggregate output decreases substantially as well. Although the economic aggregates decrease at higher margins, the replacement of the PAYG with a means tested program with 100% benefit reduction rate improves welfare moderately. The increase in leisure one of the factors that contributes to welfare gain.

Zero percent benefit reduction rate implies that all individuals in the economy receive the same level of means-tested pension benefit. In Table 3, the payroll taxes are the same across experiments. As a result, when we increase the benefit reduction rate, the maximum possible pension benefit increases as well. This in turn implies that means-tested pension programs with higher benefit reduction rates are more progressive i.e. they provide generous benefits to relatively low income groups. As we mentioned earlier, in means-tested programs individual's past earning histories are irrelevant but their their private retirement incomes from their own savings are relevant. Only exception to this is the case when the benefit reduction rate is 0%. In this case neither past earnings nor private retirement savings are relevant for pension benefits. When benefit reduction rate is 0%, we see that aggregate labor supply decreases but aggregate capital stock increases slightly. Since low income individuals now face relatively less generous pension program, leisure becomes relatively more expensive and hence, labor supply is larger than those of other means-tested pension programs. Compared to the PAYG program, there is a slight decrease in the labor supply. One possible explanation is as follows. Since high income groups now receive more generous pensions compared to the PAYG they prefer taking more leisure and hence, labor supply decreases slightly. The capital stock increases. The intuition is as follows. Compared to the PAYG and other means tested programs, relatively poor individuals now receive less pension benefits. Hence, they would increase their savings to compensate the decrease in their pension benefit entitlements. Higher income groups now no need to decrease their saving and/or decumulate their savings when they are retired to receive the pension benefits. A combination of these two effects imply an increase in the capital stock. Output slightly decreases since the decrease in labor supply more pronounced than the increase in the capital stock. In this case, welfare increases substantially. This is due to increase in leisure and increase in consumption as a result on an increase in output. Means tested pension programs with benefit reduction rate higher than 0% and lower than 100% decrease aggregate capital, labor supply, and output. Yet, the drops are less pronounced than that of means-tested pension program with 100% benefit reduction rate. Welfare increase linearly with an increase in benefit reduction rate.

The upper panel of Figure 2 demonstrates the average life-cycle asset holdings and consumption profiles for PAYG and two extreme means-tested programs i.e. means-tested program with 0% benefit reduction rate and means-tested program with 100% benefit reduction rate. One can easily say that means-tested program with 100% benefit reduction rate affects life-cycle asset holdings quite negatively. In terms of life-cycle consumption profile, it looks like means-tested program with 0% benefit reduction rate provides better consumption smoothing. This figure supports our explanations above regarding possible causes of welfare differences among programs.

	τ_p	L	K	Y	CEV (%)
PAYG	0.224	100.000	100.000	100.000	
MT 0%	0.221	99.048	100.487	99.613	2.07
MT 20%	0214	98.947	100.491	99.505	2.37
MT 40%	0.207	98.853	100.468	99.381	2.60
MT 60%	0.200	98.746	100.186	99.181	2.73
MT 80%	0.193	98.660	100.045	98.777	2.90
MT 100%	0.190	98.618	99.723	98.814	2.86

Table 4: No differential mortality - PAYG vs Means-tested pensions with the variable tax rates

In Table 4, we fixed the maximum pension benefits at the level of the means-tested pension program with 0% benefit reduction rate's maximum benefit level. Hence, in the subsequent mans-tested programs, the payroll tax rate decreases implying less tax burden on earnings. Now an increase in benefit reduction rate, decreases the tax burden of the program. Hence, individuals have higher net income to allocate between consumption and savings. This positively contribute to the aggregate saving. As we explained earlier, with an increase in benefit reduction rate, relatively rich individuals save less and decumulate their wealth as quick as possible to receive pension benefits. Although this is the case in here, the positive impact of low tax rate dominates the negative impact of having higher pension benefit reduction rate and hence, K increases. With a increase in the benefit reduction rate, labor supply decreases since leisure become relatively cheap for low income individuals. When we kept the maximum pension benefit constant, means-tested pension benefits with higher benefit reduction rates generate substantial welfare improvement. This is due to increase in labor and relatively less reduction in output compared to the earlier case we considered.

	τ_p	L	K	Y	CEV (%)
PAYG	0.22	100.000	100.000	100.000	
No Red 40%	0.22	100.222	101.824	100.456	-0.34
No Red 45%	0.23	100.436	100.264	100.314	-1.01

Table 5: No differential mortality - PAYG vs No redistributional PAYG programs

Now, we conduct our main analysis and check out what would happen if we replace the current PAYG pension program with a non-redistributive PAYG program. Table 5 presents results. In this new program, benefits are earnings dependent but the benefit formula is not progressive. In other words, there is no redistribution across various income groups and all income groups receive benefits that is proportional to their past earnings histories. In order to make a meaningful comparison, we kept the pension tax rate same as the PAYG program and look for the flat replacement rate. It turns out that 40% replacement rate generate the same pension tax rate. When we ignore mortality differential across various income groups, replacing the current PAYG with a non-redistributive PAYG implies a slight welfare loss. No redistributive PAYG program leads to an increase in aggregate labor supply since pension benefits are proportional to past earnings. Hence, making the program not progressive generates positive labor supply incentives. In a similar fashion, aggregate capital stock increases. In comparison to the benchmark case, high income individuals now receive higher pension benefits. Low income individuals on the other hand, receive substantially less pension income. In the benchmark, an individual with middle income receives around 41.5% of his past earnings as pension benefits. In contrast, higher income individuals receive 29.1% of their past earnings as pension benefits. When we replaced the PAYG with non-redistributive PAYG, all income groups receive 40% of their average past earnings as pension benefits. Since low income groups now receive relatively less pension income, they need to save more for retirement. In contrast high income groups do not need to save as much as in the benchmark case. It looks like, low income groups' increase in savings substantial enough to generate a sizable increase in overall capital stock. Since aggregate capital stock and labor supply increase, aggregate output increases as well. This, in turn positively affects aggregate consumption. Slight welfare reduction should be consequence of a decrease in leisure and a negative impact on low income individuals' life-cycle consumption due to substantial decrease in their pension benefits. Notice that when we increase the replacement rate from 40% to 45%, the non-resinstributional PAYG reduces welfare substantially due to an increase in the tax burden.

The lower panel of Figure 2 compares the life-cycle asset holdings and consumption profiles between PAYG and non-redistributional PAYG with 40% replacement rate. It shows that asset holdings and consumption do not vary much. The figure provides another support to our explanation regarding welfare differences.

5.2 Differential Mortality

Now we re-calibrate the benchmark economy by using type dependent unconditional survival probabilities. As in the previous case, we use same targets and same parameter values except β, which is re-calibrated to generate the same capital-output ratio. In this

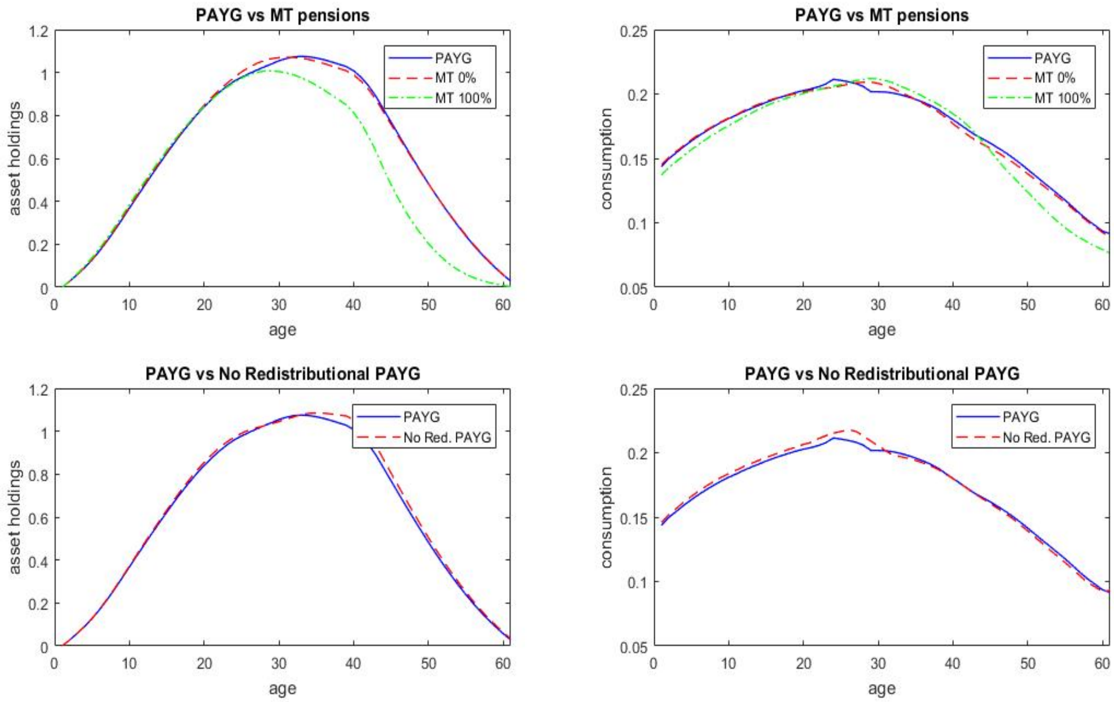

Figure 2: No differential mortality - Life-cycle profiles

section, we repeat the exact same set of exercises as in the previous section.

In the first set of exercises, we replaced the PAYG program by various mans-tested pension programs. Means-tested pension programs imposed same level of tax burden but benefits varied. Our results are in the same direction as in the previous case. When benefit reduction rate is 100%, labor supply and capital stock is the lowest. Welfare gain is the lowest as well. When we increase benefit reduction rate i.e. when we replace the most progressive means-tested program with less progressive ones, labor supply and capital stock increases. We also see more pronounced welfare gains. The intuition we provided earlier applies here as well. Adding differential mortality to our model did not change our conclusion regarding the fixed tax rate means-tested pension programs.

The upper panel of Figure 3 provides life-cycle provides showing that means-tested program with 100% benefit reduction rate affect life-cycle asset holdings negatively and the means-tested program with 0% benefit reduction rate provides slightly better consumption smoothing. Once again, the figure supports our explanation regarding welfare differentials.

Now we fix the maximum possible means-tested benefits and vary pension tax accordingly. In this case, more progressive means-tested pension programs generate higher welfare since they come up with lower payroll tax. Results are at the same direction as in no differential mortality case. Yet, in this case, welfare gains larger.

	τ_p	L	K	Y	CEV
PAYG	0.20	100.000	100.000	100.000	
MT 100%	0.20	98.990	94.008	95.052	1.01
MT 80%	0.20	99.052	95.538	95.998	1.22
MT 60%	0.20	99.065	96.133	96.265	1.45
MT 40%	0.20	99.072	96.898	96.548	1.66
MT 20%	0.20	99.046	98.441	96.989	2.15
MT 0%	0.20	99.088	101.187	98.214	2.51

Table 6: Differential mortality - PAYG vs Means-tested pensions with the fixed tax rate

	τ_p	L	K	Y	CEV
PAYG	0.197	100.000	100.000	100.000	
MT 0%	0.194	99.088	101.187	98.214	2.51
MT 20%	0.189	99.004	101.088	97.625	3.29
MT 40%	0.181	98.888	101.541	97.653	3.46
MT 60%	0.170	98.753	99.473	99.398	4.24
MT 80%	0.150	98.551	105.108	101.826	4.90
MT 100%	0.154	98.574	103.539	101.225	4.98

Table 7: Differential mortality - PAYG vs Means-tested pensions with variable pension tax rates

In our last experiment, we replace the current PAYG with a non-redistributive PAYG program. As in the previous section, we look for a flat benefit replacement rate that generates the same pension tax rate. We ended up with 40% replacement rate. As it is clear from Table 8, replacing the current PAYG with a no redistributive PAYG program generates slight welfare gain. This result is in line with the analytical result we presented in Section 2: when we take the differential mortality into our account, PAYG program with progressive benefits can generate regressive outcomes. Since low income groups now receive less generous benefits, they are inclined to increase their labor supply and savings. On the other hand, relatively high income groups would prefer taking more leisure and saving less due to more generous pensions. Hence, we see overall increase in labor supply and a slight decrease in aggregate capital stock, which generates a drop in output. In this case, low income individuals face shorter life spans compared to higher income individuals and hence, higher income individuals' population shares increase especially after retirement. Hence, the policy change now affects relatively small amount of individuals negatively and benefits relatively larger amount of individuals compared to the no differential mortality case. This in turn leads to the modest welfare gain. When we increase the replacement rate, the positive welfare gained is reversed due to a jump in the tax burden. The lower panel of Figure 3 shows that average life-cycle profiles do not change much which is reflected in relatively smaller welfare

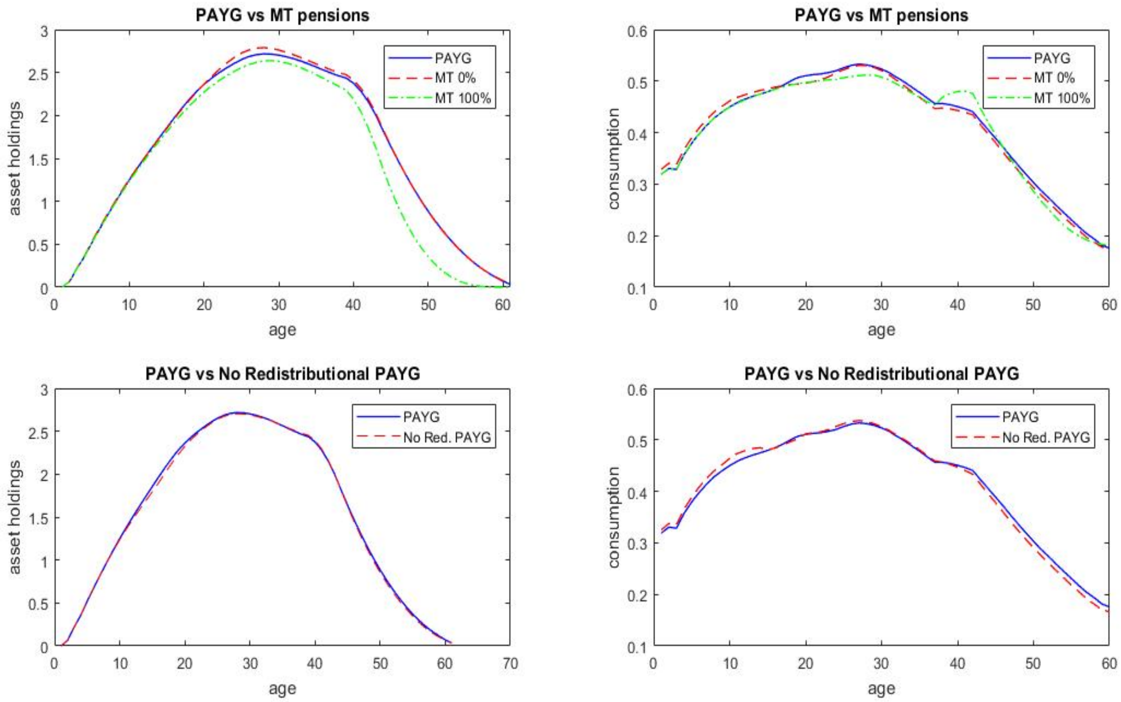

Figure 3: Differential mortality - Life-cycle profiles

gain.

	τ_p	L	K	Y	CEV (%)
PAYG	0.197	100.000	100.000	100.000	
No Red 40%	0.196	100.283	99.028	98.123	0.87
No Red 45%	0.221	100.657	94.023	95.903	-2.09

Table 8: Differential mortality - PAYG vs No redistributional PAYG programs

In sum, aggregate capital and labor supply in general decrease when we switch from current PAYG to a means tested pension program with a fixed pension tax both in differential and no differential mortality cases.

The amount of decreases in labor supply is pretty close to each other (compare tables 3 and tables 6). In both models, an increase in the benefit reduction rate generates disincentives on labor supply. Since pension benefits depend on private retirement income, an increase in benefit reduction rate with the accompanied increase in pension benefit make individuals to decrease their labor supply to be eligible for generous benefits. Yet, the existence of the differential mortality leads to more pronounced drop in the capital stock. The intuition is as follows. Due to longer life span, higher income groups have a larger retirement population share. Hence, they adjust their savings to be eligible to generous pensions. In a similar fashion, during the retirement periods, high income groups prefer to decumulate their savings as quickly as possible. Since they have a relatively larger share in the retirement population and higher savings, the drop in savings will be much larger compared to the no-differential mortality case. Similarly, welfare gains are larger in differential mortality case. When we kept the maximum means-tested pension benefit constant and vary pension tax rate, we see substantial welfare improvements in both no differential and differential mortality cases. Again, welfare gains are substantially higher in the differential mortality case. More interestingly, we observe that replacing the current PAYG with a non redistributive pension program generates completely opposite results in both cases. A non redistributive PAYG program can be welfare improving when we take differential mortality into our account.

6 Conclusion

Most developed countries have nominally progressive PAYG social security programs as their benefits are. The US PAYG Social Security has a highly progressive benefit formula to determine monthly payments. Hence, individuals with low lifetime earnings get much higher benefits than those with high lifetime earnings. For instance, Social Security might replace 70 percent of earnings for someone with a full- length career in the bottom quantile of the earnings distribution (see Goda et al. (2011) for a detailed discussion). Since benefits are paid as annuity, the total amount of benefits an individual receives depends on the that individual's longevity. If individuals from high income groups can relatively live long enough, the progressive structure of the PAYG system would disappear.

Starting with Kitagawa and Hauser (1973), the extent, causes, and trends of differential mortality in the US has been well analyzed empirically. Cristia (2009) finds large differentials in age-adjusted mortality rates across individuals in different quintiles of the individual lifetime earnings distribution. The existence of strong empirical evidence regarding mortality differentials across different earning quintiles requires evaluating social security programs once again. In this paper we analyze the implications of social security programs taking differential mortality rates across different earnings quintiles into our account.

We first generate a simple two period partial-equilibrium OLG model with differential mortality to lay out the conditions in which a PAYG program can be regressive despite its progressive benefits design. Then, we generate a large scale general equilibrium incomplete market OLG model that is calibrated to the US economy. The model mimics the features of the US income tax system and PAYG Social Security program. We then generate models in which a means-tested pension program and a non-progressive PAYG program replaces the current US PAYG program. We show that once we take into account differential mortality risks welfare rankings of the PAYG and means-tested programs do not change. Yet, we show that welfare rankings dramatically change when we replace the current PAYG with a non-redistributional PAYG across no-differential and differential mortality cases. When differential mortality is taken into account, non progressive no redistributional PAYG program dominates the current PAYG.

In sum, both analytical and computational models imply that the existence of mortality differences have important aggregate and behavioral implications and should have been taken into account seriously. This is because low income individuals receive pension benefits relatively shorter period of times. As a result, the progressive benefits would be outweighed by differential mortality risks, and hence the social security becomes regressive in terms of welfare.

References

Auerbach, A. J. and Kotlikoff, L. J. (1987). *Dynamic Fiscal Policy.* Cambridge University Press, New York, NY, USA.

Bagchi, S. (2017). Differential mortality and the progressivity of social security. *Working Paper.*

Bell, F. and Miller, M. (2002). Life tables for the united states social security area 1900-2100. *Actuarial Study 120.*

Bommier, A., Leroux, M.-L., and Lozachmeur, J.-M. (2011). Differential mortality and social security. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(1):273–289.

Conesa, J. C., Kitao, S., and Krueger, D. (2009). Taxing capital? not a bad idea after all! *American Economic Review*, 99:25–48.

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.

Diamond, P. A. (1965). National debt in a neoclassical growth model. *American Economic Review*, 55:1126–1150.

Fehr, H., Kallweit, M., and Kindermann, F. (2013). Should pensions be progressive? *European Economic Review*, 63(Supplement C):94 – 116.

Fehr, H. and Uhde, J. (2013). Means-testing and economic efficiency in pension design. *Working Paper.*

Goda, G. S., Shoven, J. B., and Slavov, S. N. (2011). Implicit taxes on work from social security and medicare. *Tax Policy and the Economy*, 25:69–88.

Hadden, W. C. and Rockswold, P. D. (2008). Increasing differential mortality by educational attainment in adults in the united states international journal of health services. *International Journal of Health Services*.

Huggett, M. and Parra, J. C. (2010). How well does the u.s. social insurance system provide social insurance? *Journal of Poliitcal Economy*, 118:76–112.

Huggett, M. and Ventura, G. (1999). On the distributional effects of social security reform. *Review of Economic Dynamics*, 2:498–531.

Imrohoroglu, A., Imrohoroglu, S., and Joines, D. H. (1995). A life cycle analysis of social security. *Economic Theory*, 6:83–114.

Kaplan, G. (2012). Inequality and the life cycle. *Quantitative Economics*, 3:471–525.

Kitagawa, E. and Hauser, P. (1973). *Differential Mortality in the United States: A Study in Socioeconomic Epidemiology*. Harvard University Press.

Kitao, S. (2014). Sustainable social security: Four options. *Review of Economic Dynamics*.

Le Garrec, G. and Lhuissier, S. (2017). Differential mortality, aging and social security: delaying the retirement age when educational spillovers matter. *Journal of Pension Economics and Finance*, 16(03):395–418.

Meara, E. R., Richards, S., and Cutler, D. (2008). The gap gets bigger: Changes in mortality and life expectancy, by education, 1981-2000. *Health Affairs*.

Nishiyama, S. and Smetters, K. (2007). Does social security privatization produce efficiency gains?*. *The Quarterly Journal of Economics*, 122(4):1677–1719.

Peterman, W. B. (2016). The effect of endogenous human capital accumulation on optimal taxation. *Review of Economic Dynamics*, 21(Supplement C):46 – 71.

Tan, O. (2015). Differential mortality and its implications on pension system analysis. *Unpublished Honours Thesis.*

Waldron, H. (2007). Trends in mortality differentials and life expectancy for male social security-covered workers, by average relative earnings. *Working Paper 108. Washington, DC, United States: Social Security Administration, Office of Research, Evaluation, and Statistics, Office of Policy.*

Appendix

Remaining Proofs

Lemma: $\Pi > 0$

$$
\begin{aligned}
\Pi &\equiv 1 - \frac{\alpha(1-s^h)(s^h)^{\frac{1}{\sigma}}}{1+(s^h)^{\frac{1}{\sigma}}} - \frac{(1-\alpha)(1-s^l)(s^l)^{\frac{1}{\sigma}}}{1+(s^l)^{\frac{1}{\sigma}}} \\
&= 1 - \frac{\begin{array}{c}\alpha\left((s^h)^{\frac{1}{\sigma}} + (s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}} - s^h(s^h)^{\frac{1}{\sigma}} - s^h(s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}\right) \\ +(1-\alpha)\left((s^l)^{\frac{1}{\sigma}} + (s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}} - s^l(s^l)^{\frac{1}{\sigma}} - s^l(s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}}\right)\end{array}}{\left(1+(s^h)^{\frac{1}{\sigma}}\right)\left(1+(s)^{\frac{1}{\sigma}}\right)} \\
&= \frac{\begin{array}{c}1+(s^h)^{\frac{1}{\sigma}} + (s^l)^{\frac{1}{\sigma}} + (s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}} \\ -\alpha\left((s^h)^{\frac{1}{\sigma}} + (s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}} - s^h(s^h)^{\frac{1}{\sigma}} - s^h(s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}\right) - \\ (1-\alpha)\left((s^l)^{\frac{1}{\sigma}} + (s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}} - s^l(s^l)^{\frac{1}{\sigma}} - s^l(s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}}\right)\end{array}}{\left(1+(s^h)^{\frac{1}{\sigma}}\right)\left(1+(s^l)^{\frac{1}{\sigma}}\right)} \\
&= \frac{\begin{array}{c}1+(s^h)^{\frac{1}{\sigma}}(1-\alpha) + (s^l)^{\frac{1}{\sigma}}(\alpha) + (s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}(1-\alpha) \\ -\alpha\left(-s^h(s^h)^{\frac{1}{\sigma}} - s^h(s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}\right) - (1-\alpha)\left(-s^l(s^l)^{\frac{1}{\sigma}} - s^l(s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}}\right)\end{array}}{\left(1+(s^h)^{\frac{1}{\sigma}}\right)\left(1+(s^l)^{\frac{1}{\sigma}}\right)} \\
&= \frac{\begin{array}{c}1+(s^h)^{\frac{1}{\sigma}}(1-\alpha) + (s^l)^{\frac{1}{\sigma}}(\alpha) + (s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}(1-\alpha) + \\ \alpha\left(s^h(s^h)^{\frac{1}{\sigma}} + s^h(s^h)^{\frac{1}{\sigma}}(s^l)^{\frac{1}{\sigma}}\right) + (1-\alpha)\left(s^l(s^l)^{\frac{1}{\sigma}} + s^l(s^l)^{\frac{1}{\sigma}}(s^h)^{\frac{1}{\sigma}}\right)\end{array}}{\left(1+(s^h)^{\frac{1}{\sigma}}\right)\left(1+(s^l)^{\frac{1}{\sigma}}\right)}
\end{aligned}
$$

Since all the terms both in the numerator and in the denominator are positive, $\Pi > 0$.

Solution Algorithm

1. Fix the pension program parameters.

2. Guess prices r, w, amount of lump-sum transfer η, and the pension tax rate in economies.

3. Solve the individual's maximization problem by the backward induction and calculate the optimal decision rules for consumption, asset holdings, and labor supply.

4. After obtaining the optimal decision rules, calculate the distribution of individuals through forward recursion.

5. By using the results in step 4, compute the aggregate variables.

6. Use the aggregate variables calculated in step 5 to construct new guesses for the variables in step 2.

7. Compare the old and new guess values. If the distance between the old and new guess values is smaller than the pre-determined tolerance value, an equilibrium is found. Otherwise, update the guess values and go to step 3.