Review of Economic Dynamics 14 (2011) 613–635

Contents lists available at ScienceDirect

Review of Economic Dynamics

www.elsevier.com/locate/red

The dynamics of inequality and social security in general equilibrium ☆

Zheng Song [a,b,*]

[a] *Chinese University of Hong Kong, Hong Kong, China*
[b] *Fudan University, China*

A R T I C L E   I N F O

Article history:
Received 23 June 2009
Revised 30 April 2011
Available online 14 May 2011

JEL classification:
E21
E62
H21
H55

Keywords:
Inequality
Intra-generational redistribution
Markov perfect equilibrium
Probabilistic voting
Social security

A B S T R A C T

This paper analyzes the dynamic politico-economic equilibrium of a model where repeated voting on social security and the evolution of household characteristics in general equilibrium are mutually affected over time. In particular, we incorporate within-cohort heterogeneity in a two-period Overlapping-Generation model to capture the intra-generational redistributive effect of social security transfers. Political decision-making is represented by a probabilistic voting à la Lindbeck and Weibull (1987). We analytically characterize the Markov perfect equilibrium, in which social security tax rates are shown to be increasing in wealth inequality. A dynamic interaction between inequality and social security leads to larger social security programs. In a model calibrated to the U.S. economy, the dynamic interaction is shown to be quantitatively important: It accounts for more than half of the social security growth in the dynamics. We also perform some normative analysis, showing that the politico-economic equilibrium outcomes can be fundamentally different from the Ramsey allocation.

1. Introduction

Most developed countries have large public pension programs, involving both inter-generational and intra-generational transfers. For instance, social security contributions are roughly proportional to income, while benefits have important lump-sum components. The general equilibrium effects and the welfare implications of such social security programs have been extensively studied in the literature.[1] However, the welfare state is not exogenously imposed, but endogenously determined by policy choices that reflect rich dynamic interactions between political and economic factors. For instance, the evolution of the distribution of household characteristics in general equilibrium may alter the political support for the social security system since households with different characteristics tend to have different preferences over transfers. Despite this, most of the existing literature has either assumed away politico-economic factors or, when considering them, focused on models where the size of social security is decided once and for all. As a result, the feedback of endogenous changes of household

☆ This paper is based on Chapter 3 of my dissertation at IIES, Stockholm University. I am deeply indebted to my advisor, Fabrizio Zilibotti, for his guidance and numerous discussions. I thank the editor, Matthias Doepke, and an anonymous referee for providing suggestions that improved the paper substantially. I also thank Kaiji Chen, John Hassler, Giovanni Favara, Dirk Niepelt, Jose V. Rodriguez Mora, Kjetil Storesletten and seminar participants at IIES and Oslo for helpful comments. Any remaining errors are mine.

* Correspondence to: Chinese University of Hong Kong, Hong Kong Special Administrative Region.
E-mail address: zsong@cuhk.edu.hk.

[1] See, among many others, Auerbach and Koltikoff (1987), Imrohoroglu et al. (1995) and Storesletten et al. (1999).

1094-2025/$ – see front matter
doi:10.1016/j.red.2011.05.001

characteristics on the decision of social security transfers over time has been ignored altogether (e.g., Tabellini, 2000; Cooley and Soares, 1999; Conesa and Krueger, 1999).[2]

The present paper explores the positive implications and the welfare properties of a rational-choice theory implying interactions between private intertemporal choices and repeated political decisions about social security. To this end, we construct a dynamic general equilibrium model where agents repeatedly vote on the social security system. Our analytical results show that the dynamic interaction between inequality and social security leads to larger social security programs. In a model calibrated to the U.S. economy, the dynamic interaction is shown to be quantitatively important: It accounts for more than half of the social security growth in the dynamics. We also perform some normative analysis, showing that the politico-economic equilibrium outcomes can be fundamentally different from the Ramsey allocation chosen by a benevolent planner with a commitment technology.

In our model, the incumbent government cannot commit to future transfers since they are decided by future elected governments. Instead, transfers are determined in each period by the current constituency, the wealth inequality of which is a key factor. Forward-looking households adjust their private savings when rationally anticipating the equilibrium dynamics of wealth inequality and social security. A main theoretical finding is that this interaction leads to an equilibrium where social security transfers increase over time. The underlying mechanism is twofold. On the one hand, the establishment of a social security increases future wealth inequality since within-cohort transfers discourage the private savings of low-income households more than proportionally. On the other hand, larger wealth inequality makes transfers more desirable in the future. This provides the political support for an increasing size of social security in subsequent periods.

Our workhorse is a standard two-period Overlapping-Generation model. To capture the intra-generational redistributive role of social security, we incorporate within-cohort heterogeneity by assuming young households to be born with different labor productivities. Old households are different in terms of wealth. In other words, there exists multi-dimensional heterogeneity across voters. Each group of voters has its own preferences over transfers. The political decision process is modeled by a repeated probabilistic voting framework.[3] In equilibrium, policymaker candidates respond to electoral uncertainty by proposing a policy platform that maximizes a weighted-average welfare of all groups of voters.

We focus on Markov perfect equilibria, where the size of social security is conditioned on payoff-relevant fundamental elements: the distribution of assets held by old households and the demographic structure. The Markov perfect equilibrium is obtained as one takes the limit of a finite horizon environment.[4] Moreover, under logarithm utility and Cobb–Douglas production technology, the equilibrium can be characterized analytically, making the underlying politico-economic mechanism highly transparent. In particular, we show that the equilibrium social security tax rate is increasing in wealth inequality, and this positive relationship generates growing social security over time. When calibrated to the U.S. economy from 1947 to 1969, the model predicts the initial and steady state social security tax rate of 3.5 percent and 7.2 percent, respectively. The growth of the tax rate is quantitatively close to the data, though the level of tax rate is higher: The average Old-Age and Survivors Insurance (OASI henceforth) contribution rate increases from 2 percent in the 1930s to 4.7 percent in the period from 1947 to 1969. The exercise also suggests that the dynamic interaction between inequality and social security is quantitatively important: It generates a 3.7-percentage-point increase of the tax rate in the dynamics. We then extend the model by incorporating the ageing population in the U.S. The average dependency ratio rises to 19.3 percent from 1970 to 2000, substantially higher than 15.7 percent from 1947 to 1969. Such a change will lead to a further increase in the size of social security. The extended model predicts a tax rate of 9 percent, which is roughly in line with the average OASI contribution rate of 9.8 percent in the 1970 to 2000 period. In addition, more social security benefits make wealth distribution more unequal. The Gini coefficient of wealth in the model economy increases by 3.4 percent. This might explain a significant part of the 10-percent increase in the Gini coefficient of household financial wealth in the Survey of the Financial Characteristics of Consumers (SFCC henceforth) and the Survey of Consumer Finances (SCF henceforth) from the 1960s to the 1980s.

The tractable model allows a comparison between the politico-economic equilibrium outcome and the Ramsey allocation, in which a benevolent planner with a commitment technology maximizes the discounted sum of the welfare of all current and future generations. Under logarithm utility and Cobb–Douglas production technology, the Ramsey solution can also be characterized analytically. We find that the Ramsey solution may feature a decreasing size of social security if the social discount factor is not too small. This sharply contrasts growing transfers in the political equilibrium. The basic intuition is straightforward. The initial inelastic capital stock provides the incentive for the Ramsey planner to impose high taxes for redistributive reasons.[5] However, since she can commit to future policies, low taxation will be adopted for encouraging capital accumulation in periods other than the initial one.

---

[2] A notable exception is Boldrin and Rustichini (2000), where the interaction between private intertemporal choices and political decisions may lead to a decreasing size of social security.

[3] The probabilistic voting framework is adapted from Lindbeck and Weibull (1987). See Hassler et al. (2005) and Gonzalez-Eiras and Niepelt (2008) for applications of the repeated probabilistic voting in a dynamic political economy.

[4] Previous literature has studied the sustainability and evolution of social security by assuming that voters play trigger strategies (e.g., Boldrin and Rustichini, 2000). Although trigger strategy may provide analytical convenience and have reasonable components, it is hard to provide sharp empirical predictions due to the indeterminacy of equilibria.

[5] Unlike the mechanism for high initial capital tax rates in Chamley (1986) and Judd (1985), the government here makes no attempt to confiscate the initial capital stock due to the pay-as-you-go social security system.

It is worth emphasizing that in Markov equilibria, voters not only hold rational expectations on future equilibrium outcomes, but may also strategically affect future policies via the impact of current policies on private intertemporal choices. Under logarithm utility, the current tax rate does not affect household saving rates due to a cancellation of the income and substitution effects. Thus, it cannot affect future states of the economy (wealth distribution) or future policy outcomes. In other words, strategic effects are mute in the particular case of logarithm utility. Strategic effects appear when the intertemporal elasticity of substitution is different from unity. In these cases, analytical results cannot be obtained, but we can numerically study the qualitative and quantitative impact of strategic effects. We show that if the intertemporal elasticity of substitution is smaller than unity, as many empirical studies suggest, the strategic effect is positive. A higher tax rate today widens wealth inequality tomorrow and, hence, leads to larger transfers tomorrow. Due to the positive strategic effect, current voters have the incentive to strategically raise the current social security tax rate in order to harvest larger future transfers. The calibrated economy indicates that the strategic effect in Markovian equilibria is quantitatively unimportant: The relative increase in transfers due to the strategic effect is less than 5 percent.

Although the sustainability of the social security system has been widely discussed in the literature,[6] its dynamic patterns have been much less investigated. Some pioneering studies abstracting repeated voting include Verbon (1987) and Boadway and Wildasin (1989). More recently, Forni (2005) shows that in a repeated political decision process, self-fulfilled expectations on the positive relationship between current and future social security transfers can lead to a growing pension scheme. The present paper extends the literature by linking the evolution of the system to some economic fundamentals – i.e., wealth distribution. Our model suggests that, though the inter-generational redistributive effect is key to sustain the system, the intra-generational redistributive effect plays a central role in the evolution of social security in general equilibrium. In particular, the interaction between transfers and wealth inequality can generate the growth of social security. Hassler et al. (2003) investigate the political sustainability of a welfare system with both inter- and intra-generational redistribution. In their model, the size of the rich and of the poor is determined by human capital investment. The endogenous political constituency implies strong strategic effects that may even result in multiple equilibria. With a focus on wealth accumulation, our paper shuts down the channel through which private intertemporal choices affect political constituency at the extensive margin. This results in a much weaker strategic effect.

Our work is part of a growing literature on dynamic politico-economic equilibrium, where current voting may change fundamentals in the future political environment and, hence, affect future policy outcomes. Because of the complexity of dynamic interaction between individual intertemporal choice and voting strategy, analytical results are usually unattainable, except in some small open economies (e.g., Hassler et al., 2003; Azzimonti Renzo, forthcoming). An exception is Gonzalez-Eiras and Niepelt (2008), which shows that a closed-form solution of social security transfers can be obtained in a growth model with logarithm utility and Cobb–Douglas production technology. However, the equilibrium policy rule degenerates into a constant in their model with constant population growth. The present paper generalizes Gonzalez-Eiras and Niepelt's work by incorporating within-cohort heterogeneity, keeping all results analytical.[7] The generalization gives an equilibrium policy rule which is non-trivially dependent of fundamental elements in the politico-economic environment and, hence, provides much richer implications for the dynamics of policies. This also contrasts the literature that resorts to numerical characterizations for non-trivial equilibrium policy rules in general equilibrium (e.g., Krusell et al., 1997; Krusell and Rios-Rull, 1999).

The rest of the paper is organized as follows. Section 2 presents the model. In Section 3, the dynamic politico-economic equilibrium is defined and solved under logarithm utility. Quantitative exercises are conducted in Section 4. Section 5 characterizes the Ramsey allocation. In Section 6, we solve numerically the political equilibrium under a more general CRRA utility form. Section 7 concludes.

2. The model

Consider an economy inhabited by an infinite sequence of overlapping generations. Each generation lives for two periods. Households work in the first period of their life and then retire. Labor supply is inelastic and normalized to unity. Let N_t denote the population of the cohort born at time t and $n_t \equiv N_t/N_{t-1}$ denote the gross population growth rate. We may interpret $1/n_t$ as the old dependency ratio at period t.

Young households are endowed with labor productivity γ^j with probability p^j, where $j \in \{1, 2, \dots, J\}$ and $\{\gamma^j\}_{j=1}^J$ is an increasing sequence. The mean of labor productivity is normalized to unity: $\sum_j p^j \gamma^j = 1$. Wage income is taxed at a flat rate, τ_t. The after-tax net earnings for young households of type j are $(1-\tau_t)w_t^j$. Old households receive benefits b_t from a social security system, and young households may save to finance their consumption after retirement. The corresponding intertemporal decision solves

$$\max_{k_{t+1}^j} u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right), \tag{1}$$

[6] See, for example: Boldrin and Rustichini (2000), Cooley and Soares (1999), Conesa and Krueger (1999), Mulligan and Sala-i-Martin (1999a, 1999b), Tabellini (2000), Razin et al. (2002) and Gonzalez-Eiras and Niepelt (2008).

[7] Gonzalez-Eiras and Niepelt (2008) allow random population growth. Our work is more restrictive in that dimension.

subject to

$$c_t^{y,j} = (1-\tau_t)w_t^j - k_{t+1}^j, \tag{2}$$

$$c_{t+1}^{o,j} = R_{t+1}k_{t+1}^j + b_{t+1}, \tag{3}$$

where $c_t^{i,j}$ and k_{t+1}^j denote the consumption and savings of households of type (i,j), $i \in \{y,o\}$ and $j \in \{1,2,\ldots,J\}$, respectively. The discount factor is $\beta \in (0,1)$. R_{t+1} is the gross interest rate at time $t+1$. We assume that $u(c) = \log(c)$, an assumption that will be relaxed in Section 6.

Let K_t and L_t be the aggregate capital stock and effective labor used in production at time t. The clearance of factor markets requires $K_t = N_{t-1}\sum_j p^j k_t^j$ and $L_t = N_t \sum_j p^j \gamma^j = N_t$. Assume that production follows Cobb–Douglas technology with a constant return to scale, $AK_t^\alpha L_t^{1-\alpha}$, where A denotes total factor productivity and $\alpha \in (0,1)$ is the output elasticity of capital. Factor markets are competitive. Factor prices thus correspond to marginal products

$$R_t = A\alpha(k_t/n_t)^{\alpha-1}, \tag{4}$$

$$w_t = A(1-\alpha)(k_t/n_t)^\alpha, \tag{5}$$

where $k_t \equiv \sum_j p^j k_t^j$ is the average wealth holdings of old households. The individual wage rate is $w_t^j = \gamma^j w_t$, where w_t stands for the average wage rate.

The flat-rate wage income tax rate τ_t is determined through a political process that will be specified below. τ_t is imposed on the working generation to finance social security payments. In addition to the inter-generational redistribution that defines the pay-as-you-go system, pensions entail intra-generational redistributive elements. In most systems, social security contributions are proportional to income, while benefits have lump-sum or even regressive components. According to the Old Age Insurance of the U.S. social security system, for example, a one-percent increase in lifetime earnings leads to a 0.90-, 0.32-, 0.15- and 0.00-percent increase in pension benefits from low to high income groups.[8] Following Conesa and Krueger (1999) and many others, we assume, for analytical convenience, social security benefits to be evenly distributed within old households. It is also assumed that the budget of the social security system must be balanced in each period; i.e., at any time t, social security payments, $N^{t-1}b_t$, equal social security contributions, $N_t\tau_t\sum_j p^j w_t^j$. Therefore, the balanced budget implies

$$b_t = n_t \tau_t w_t. \tag{6}$$

2.1. Households' saving choice

Under logarithm utility, households' saving choice can be analytically obtained by the Euler equation, $c_{t+1}^{o,j}/c_t^{y,j} = \beta R_{t+1}$, which solves (1). Since households are atomistic, they take factor prices, aggregate savings, the current social security tax rate and future social security benefits as given. Plugging factor prices (4), (5) and the balanced-budget rule (6) into (2) and (3), the Euler equation solves a private saving function

$$k_{t+1}^j = S^j(k_t/n_t, \tau_t, \tau_{t+1}) \equiv \theta^j(\tau_{t+1})A(1-\tau_t)(k_t/n_t)^\alpha, \tag{7}$$

where $\theta^j(\cdot)$ is defined as

$$\theta^j(\tau_{t+1}) \equiv \frac{\beta(1-\alpha)}{1+\beta}\left(\gamma^j - \frac{(1-\alpha)\tau_{t+1}}{(1+\beta)\alpha + (1-\alpha)\tau_{t+1}}\right). \tag{8}$$

It is straightforward that $S_1^j > 0$, $S_2^j < 0$ and $S_3^j < 0$, where subscript i denotes the partial derivative with respect to the ith argument of S. A high k_t increases the wage rate and, thus, private savings. The effect of a high τ_t is the opposite since it reduces the current after-tax income. $S_3^j < 0$ since a high future social security tax rate increases after-retirement income and, hence, discourages private savings.

Eq. (7) leads to the law of motion of aggregate capital

$$k_{t+1} = S(k_t/n_t, \tau_t, \tau_{t+1}) \equiv \theta(\tau_{t+1})A(1-\tau_t)(k_t/n_t)^\alpha, \tag{9}$$

where $\theta(\cdot)$ is defined as

$$\theta(\tau_{t+1}) \equiv \frac{\alpha\beta(1-\alpha)}{\alpha(1+\beta) + (1-\alpha)\tau_{t+1}}, \tag{10}$$

[8] See, for example, Storesletten et al. (2004).

with $S_1 > 0$, $S_2 < 0$ and $S_3 < 0$. Note that $S_3^j = S_3$; i.e., the effect of τ_{t+1} on private savings is homogeneous across households.

γ^j and k_{t+1}^j/k_{t+1} measure income and wealth dispersion, respectively.[9] Note that k_t or τ_t may affect the next-period interest rate and private savings through the aggregate capital accumulation. Under logarithm utility, a cancellation of the income and substitution effect implies that private savings are independent of the interest rate. Therefore, changes in k_t or τ_t will leave saving rates and future wealth dispersion unchanged. As will be seen below, this property greatly simplifies the analysis throughout the paper.

Eqs. (7) and (9) show that $k_{t+1}^j/k_{t+1} = \theta^j(\tau_{t+1})/\theta(\tau_{t+1})$. Without a social security system ($\tau_{t+1} = 0$), wealth dispersion coincides with income dispersion. With the presence of social security, $\theta^j(\tau_{t+1})/\theta(\tau_{t+1})$ is increasing in τ_{t+1} if $\gamma^j > 1$ and decreasing in τ_{t+1} if $\gamma^i < 1$. So, a high future social security tax rate τ_{t+1} would increase wealth inequality.[10] The intuition is simple. Due to the within-generation redistributive elements, high social security benefits discourage savings of the less productive more than those of the more productive. The above results are summarized by Lemma 1.

Lemma 1. *Assume that $u(c) = \log(c)$. The future wealth inequality is increasing in the future social security tax rate τ_{t+1}, but does not depend on the current social security tax rate τ_t or the aggregate capital k_t.*

3. Political equilibrium

The social security tax rate τ_t is chosen by some repeated political process at the beginning of each period. In the present paper, we assume that τ_t is determined in a probabilistic voting framework (Lindbeck and Weibull, 1987). There are two policy-maker candidates running electoral competition. The winner obtains the majority of the votes of all current voters with unobservable ideological preferences towards political candidates. Since candidates care only about winning the election, they will, in equilibrium, respond to electoral uncertainty by proposing a policy platform that maximizes a weighted-average welfare of all current voters. The weights reflect the sensitivity of different groups of voters to policy changes.[11] In the context of our model, the political decision process of τ_t can be formalized as

$$\max_{\tau_t \in [0,1]} \sum_{j=1}^{J} p^j U_t^{o,j} + n_t \sum_{j=1}^{J} p^j U_t^{y,j}, \tag{11}$$

where $U_t^{i,j}$ denotes the welfare of the households of type (i, j), with $U_t^{o,j} \equiv u(c_t^{o,j})$ and $U_t^{y,j} \equiv u(c_t^{y,j}) + \beta u(c_{t+1}^{o,j})$. For notational convenience, the weights on different groups' utility are set equal to their population size.

We focus on Markov perfect equilibria, in which the state of the economy is summarized by the distribution of assets held by old households, $\{k_t^j\}_{j=1}^J$, together with current and future population growth rates $\{n_{t+i}\}_{i=0}^{\infty}$. Here, we implicitly assume that households form perfect foresight on future population growth. This assumption is not crucial, as it will be shown below that only the current population growth rate, n_t, is payoff-relevant.[12] The Markovian policy rule of τ_t can be written as

$$\tau_t = F\big(k_t^1, k_t^2, \ldots, k_t^J, \boldsymbol{n}_t\big), \tag{12}$$

where the boldface variable is defined as $\boldsymbol{n}_t = \{n_{t+i}\}_{i=0}^{\infty}$, and F is assumed to be continuous and differentiable for technical convenience.[13] In Markov equilibria, the current political decision may affect the future asset distribution and, thus, the future social security tax rate. Forward-looking voters will adjust their intertemporal choice accordingly. To see this, we substitute (12) for τ_{t+1} in (7) and obtain the following equation:

$$k_{t+1}^j = S^j\big(k_t/n_t, \tau_t, F\big(k_{t+1}^1, k_{t+1}^2, \ldots, k_{t+1}^J, \boldsymbol{n}_{t+1}\big)\big). \tag{13}$$

This is a non-linear equation system, with J equations solving for $k_{t+1}^1, k_{t+1}^2, \ldots, k_{t+1}^J$. The solution can be written as a recursive form of private saving functions

$$k_{t+1}^j = \hat{S}^j(k_t/n_t, \tau_t, \boldsymbol{n}_{t+1}). \tag{14}$$

The expression of $\hat{S}^j$ is not available unless we know the explicit form of F. However, some properties of $\hat{S}^j$ can be obtained. Differentiating (13) with respect to τ_t gives

[9] To avoid confusion, income and wealth are hereafter referred to young households' labor earnings and old households' wealth excluding social security benefits, respectively.

[10] Wealth inequality increases if, holding k_t constant, there is an increase in any k_t^j with $k_t^j > k_t$ and a decrease in any k_t^j with $k_t^j < k_t$.

[11] See Persson and Tabellini (2000) for a more detailed discussion of probabilistic voting.

[12] Following Gonzalez-Eiras and Niepelt (2008), the analysis below can be extended to an environment with stochastic ageing.

[13] Krusell and Smith (2003) provide an example showing that discontinuous policy rules may lead to indeterminacy of Markov equilibrium.

$$\begin{bmatrix} \hat{S}_2^1 \\ \hat{S}_2^2 \\ \vdots \\ \hat{S}_2^J \end{bmatrix} = \begin{bmatrix} S_2^1 \\ S_2^2 \\ \vdots \\ S_2^J \end{bmatrix} + \begin{bmatrix} S_3^1 & S_3^1 & \cdots & S_3^1 \\ S_3^2 & S_3^2 & \cdots & S_3^2 \\ \vdots & \vdots & \vdots & \vdots \\ S_3^J & S_3^J & \cdots & S_3^J \end{bmatrix} \begin{bmatrix} F_1 \hat{S}_2^1 \\ F_2 \hat{S}_2^2 \\ \vdots \\ F_J \hat{S}_2^J \end{bmatrix},$$

which pins down the partial derivatives of saving functions $\hat{S}_2^j$:

$$\hat{S}_2^j = \frac{S_2^j (1 - S_3 \sum_{i=1}^J F_i) + S_3 \sum_{i=1}^J S_2^i F_i}{1 - S_3 \sum_{i=1}^J F_i}. \tag{15}$$

Here, we use the fact that $S_3^j = S_3$. Note that $\hat{S}_2^j$ generally differs from S_2^j. This means that the perception of the policy rule F will change the effect of τ_t on private savings. Correspondingly, the law of motion of aggregate capital becomes

$$k_{t+1} = \hat{S}(k_t/n_t, \tau_t, \boldsymbol{n}_{t+1}) \equiv \sum_j p^j \hat{S}^j (k_t/n_t, \tau_t, \boldsymbol{n}_{t+1}), \tag{16}$$

with $\hat{S}_2 = \sum_j p^j \hat{S}_2^j (k_t/n_t, \tau_t, \boldsymbol{n}_{t+1})$.

Given any F, the political decision on τ_t solves (11), subject to budget constraints (2) and (3), factor prices (4) and (5), the balanced-budget rule (6), private saving functions (14), the law of motion of aggregate capital (16), and the non-negative constraint of τ_t.[14] This yields an actual policy rule $\tau_t = \bar{F}(k_t^1, k_t^2, \ldots, k_t^J, \boldsymbol{n}_t)$, with $\bar{F}$. F is said to be a Markovian equilibrium policy rule, if and only if $\bar{F} = F$. The formal definition of a Markov perfect equilibrium is given as follows.

Definition 1. A Markov perfect political equilibrium is a set of functions $\tilde{S}^j$ and F, where private saving functions $\tilde{S}^j$, $j \in \{1, 2, \ldots, J\}$, and the policy rule F are such that:

(1) Given the policy rule F, $\tilde{S}^j(k_t^1, k_t^2, \ldots, k_t^J, \boldsymbol{n}_t) = \hat{S}^j(k_t/n_t, F(k_t^1, k_t^2, \ldots, k_t^J, \boldsymbol{n}_t), \boldsymbol{n}_{t+1})$, where $\hat{S}^j$ is the recursive private saving function (14).

(2) Given F and $\hat{S}^j$, $\bar{F}$ solves (11), subject to (2) to (6), (14), (16) and the non-negative constraint of τ_t.

(3) $\bar{F} = F$.

3.1. The equilibrium policy rule

To solve the equilibrium policy rule F, we need to know the impact of the social security tax rate τ_t on the welfare of various groups of voters. Differentiating the utility of old households with respect to τ_t yields

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = u'(c_t^{o,j}) n_t w_t > 0. \tag{17}$$

Needless to say, old households always benefit from social security transfers. Substituting for $c_t^{o,j}$ and w_t, (17) can be rewritten as

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{1-\alpha}{\alpha k_t^j / k_t + (1-\alpha)\tau_t}, \tag{18}$$

where $i \neq j$. $\partial U_t^{o,j}/\partial \tau_t$ depends on wealth distribution. This highlights the role of social security as an intra-generational redistributive policy. Specifically, the smaller is the wealth share of an old household, the more welfare gains it can get from transfers. The aggregate welfare effect of τ_t on old households, $\partial U_t^o/\partial \tau_t = \sum_{j=h,l} p^j \partial U_t^{o,j}/\partial \tau_t$, increases in wealth inequality due to the concavity of utility.

Differentiating the utility of young households with respect to τ_t yields

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = -u'(c_t^{y,j}) \gamma^j w_t + \beta u'(c_{t+1}^{o,j}) \left(k_{t+1}^j \frac{\partial R_{t+1}}{\partial k_{t+1}} + n_{t+1} \tau_{t+1} \frac{\partial w_{t+1}}{\partial k_{t+1}} \right) \hat{S}_2 + \beta u'(c_{t+1}^{o,j}) n_{t+1} w_{t+1} \frac{\partial \tau_{t+1}}{\partial \tau_t}. \tag{19}$$

Note that the effect of τ_t via k_{t+1}^j cancels out due to the Euler equation. The first term in (19) reflects the direct cost of social security contributions. The second term captures the general equilibrium effect of τ_t via its impact on capital accumulation $\hat{S}_2$. The general equilibrium effect is twofold. On the one hand, a high τ_t reduces private savings at time t and, thus, reduces the tax base of social security at time $t+1$. On the other hand, young households at time t benefit from

[14] The constraint that $\tau_t \leqslant 1$ is never binding since, otherwise, it delivers zero consumption to young households.

a higher interest rate R_{t+1}. As long as τ_{t+1} or wealth inequality is not very large, the interest rate effect dominates the first effect.[15] Hence, the general equilibrium effect can benefit young households.[16] The third term is the "strategic effect", which captures the fact that voters can affect the future tax rate τ_{t+1} by their current choice of τ_t. The sign and size of the strategic effect are determined by $\partial\tau_{t+1}/\partial\tau_t$, which follows

$$\frac{\partial\tau_{t+1}}{\partial\tau_t}=\sum_{j=1}^{J}\digamma_j\big(k^1_{t+1},\dots,k^J_{t+1},\boldsymbol{n}_{t+1}\big)\hat{S}^j_2(k_t/n_t,\tau_t,\boldsymbol{n}_{t+1}). \tag{20}$$

If $\partial\tau_{t+1}/\partial\tau_t>0$, young households know that a higher current social security tax rate leads to more social security benefits in the future. Thus, they may strategically increase τ_t as compared to the case where the current political choice does not affect future policy outcomes.

Then, the first-order condition of (11) can be written as

$$\sum_{j=1}^{J}p^j\frac{\partial U^{o,j}_t}{\partial\tau_t}+n_t\sum_{j=1}^{J}p^j\frac{\partial U^{y,j}_t}{\partial\tau_t}+\lambda_t=0, \tag{21}$$

where λ_t denotes the multiplier on the non-negative constraint of τ_t, $\lambda_t=0$ for $\tau_t>0$ and $\lambda_t>0$ for $\tau_t=0$. Eq. (21) implies a functional equation for $\digamma$.

Under logarithm utility, the fixed-point can be characterized as the limit of finite-horizon solutions. The corresponding political equilibrium is, thus, unique within the class of equilibria that are limits of equilibria in a finite-horizon economy (see Appendix A.1). Proposition 1 summarizes two important results.

Proposition 1. *Assume that $u(c)=\log(c)$. In the Markov perfect equilibrium,*

(i) *the welfare effect of social security tax on the young is equal to*

$$\frac{\partial U^{y,j}_t}{\partial\tau_t}=-\frac{1+\beta\alpha}{1-\tau_t}; \tag{22}$$

(ii) $\tau_t=\digamma(k^1_t,k^2_t,\dots,k^J_t,n_t)$, *with $\digamma: R^{J+1}\to[0,1]$; i.e., given $\{k^j_t\}^J_{j=1}$, τ_t is independent of $\boldsymbol{n}_{t+1}$*;
(iii) $\partial\tau_{t+1}/\partial\tau_t=0$*; i.e., the strategic effect is mute*;
(iv) *the equilibrium social security tax, though independent of aggregate wealth, is increasing in wealth dispersion and decreasing in n_t.*

A formal proof of the proposition, based on backward induction, is provided in Appendix A.1. The underlying mechanism can be illustrated in a less rigorous but more intuitive way. First, it is immediate that the absence of the strategic effect yields (22). Although $\partial U^{o,j}_t/\partial\tau_t$ is different among old households, τ_t delivers the same welfare effect on young households with different labor productivity. In addition, (22) implies that $\partial U^{y,j}_t/\partial\tau_t$ is independent of τ_{t+1} and, thus, $\boldsymbol{n}_{t+1}$. The independence, together with the fact that future population growth has no impact on the welfare of current old households, establishes the second part of the proposition. The fact that $\boldsymbol{n}_{t+1}$ is redundant for τ_t will greatly relieve the computation burden in the following quantitative exercises.

A more important question is why the strategic effect is mute. Substituting (18) and (22) back into (21) leads to

$$\sum_{j=1}^{J}p^j\frac{1-\alpha}{\alpha k^j_t/k_t+(1-\alpha)\tau_t}-n_t\frac{1+\beta\alpha}{1-\tau_t}=0. \tag{23}$$

Here, we assume interior solution. The following argument trivially extends to corner solutions. On the one hand, Eq. (23) implies that the social security tax rate depends only on the current wealth dispersion, k^j_t/k_t, and the current ratio of young to old households, n_t. Leading the argument by one period suggests that proportional changes to the level of k^j_{t+1} would not affect τ_{t+1}. On the other hand, Lemma 1 has shown that τ_t does not affect the future wealth dispersion. A joint of the two properties breaks down the dynamic linkage between τ_t and τ_{t+1}, making our analysis much simpler. Both properties apply to logarithm preferences only. The strategic effect will appear under more general utility forms. Nevertheless, numerical experiments in Section 6 suggest that our main findings are robust.

The last part of Proposition 1 is intuitive. Eq. (23) has illustrated that aggregate wealth has no impact on τ_t. The relative wealth plays an important role, though. Due to the even distribution of social security transfers among the old, an increase of wealth dispersion weakens the welfare effect of transfers on the rich (with $k^j_t>k_t$) but strengthens the effect on the

[15] This can be seen by the fact that $\mathrm{sgn}(k^j_{t+1}\partial R_{t+1}/\partial k_{t+1}+n\tau_{t+1}\partial w_{t+1}/\partial k_{t+1})=\mathrm{sgn}(-k^j_{t+1}/k_{t+1}+\tau_{t+1})$.
[16] Gonzalez-Eiras and Niepelt (2008) show that the interest rate effect plays an important role in sustaining the social security system in an economy without within-cohort heterogeneity.

poor (with $k_t^j < k_t$). The concavity of utility, therefore, yields a higher tax rate. n_t affects the political decision through a different channel. A lower n_t implies a larger population share of the old who benefit from the social security system. As the old gain more political weight in the decision process, the tax rate will increase accordingly.

Finally, Eq. (23) implies that τ_t decreases in α. A lower α implies that the interest rate, R_{t+1}, is more elastic to aggregate capital k_{t+1}. This tends to amplify the general equilibrium effect and, hence, mitigates the negative welfare effect of τ_t on the young, as shown by Eq. (22).

3.2. Dynamics

Eq. (23) also reveals how the equilibrium tax rate evolves over time. Leading (23) by one period and substituting (7) and (9) for k_{t+1}^j/k_{t+1} pin down τ_{t+1}:

$$\tau_{t+1} = \max\{0, \Gamma(n_{t+1})\}, \tag{24}$$

where $\Gamma(n_{t+1})$ solves the following equation:

$$\sum_{j=1}^{J} p^j \frac{1-\alpha}{\alpha\theta^j(\Gamma(n_{t+1}))/\theta(\Gamma(n_{t+1})) + (1-\alpha)\Gamma(n_{t+1})} - n_{t+1}\frac{1+\beta\alpha}{1-\Gamma(n_{t+1})} = 0. \tag{25}$$

Three remarks are in order. First, (24) and (25) suggest that for households at period t, τ_{t+1} can be considered a rational expectation that depends only on n_{t+1}. The expectation is formed as follows. Given an expected future tax rate τ_{t+1}, young households at period t make intertemporal choices so that k_{t+1}^j/k_{t+1} is equal to $\theta^j(\tau_{t+1})/\theta(\tau_{t+1})$. For the expectation to be self-fulfilled, τ_{t+1} must equal the one implied by the policy rule solving (25). Second, (25) holds true for any tax rate other than the one in the initial period. Although the future tax rate still follows the equilibrium policy rule F, it can be rewritten as a function of the future demographic structure only. Finally, Eq. (25) does not allow us to analytically characterize $\Gamma(n_{t+1})$ – i.e., the self-fulfilled expectation of social security tax rate. Multiple equilibria are possible, though extensive numerical experiments suggest that $\Gamma(n_{t+1})$ is indeed a well-defined function.

We then provide a simple characterization of the equilibrium dynamics. Suppose that voting for social security is unanticipatedly launched at period 1. Eq. (7) implies a wealth dispersion of $k_1^j/k_1 = \theta^j(0)/\theta(0)$ and the tax rate of $\tau_1 = F(k_1^1, k_1^2, \ldots, k_1^J, n_1)$. For $t > 1$, the wealth dispersion turns to $\theta^j(\tau_t)/\theta(\tau_t)$ and τ_t follows (24). To conclude, we have

Proposition 2. *Assume that $u(c) = \log(c)$. In the Markov perfect equilibrium, the initial social security tax rate, τ_1, is determined by $F(k_1^1, k_1^2, \ldots, k_1^J, n_1)$, while τ_t follows* (24) *and is entirely determined by n_t for any $t > 1$.*

The dynamics of social security and wealth inequality become more explicit when considering an economy with stationary population growth – i.e., $n_t = n$ for all t. The above analysis implies a constant tax rate and, thus, a stationary wealth distribution after the initial period. Let $\hat{\tau}$ be the constant tax rate that solves (25) with $n_t = n$. Since $\theta^j(\hat{\tau})/\theta(\hat{\tau}) > \theta^j(0)/\theta(0)$ for $\gamma^j > 1$ and $\theta^j(\hat{\tau})/\theta(\hat{\tau}) < \theta^j(0)/\theta(0)$ for $\gamma^j < 1$, transfers increase wealth dispersion after the initial period, leading to a growing size of social security. To conclude, we have

Corollary 2. *Assume that $u(c) = \log(c)$. In the Markov perfect equilibrium with stationary population growth n,*

(i) *Wealth inequality and the social security tax rate converge to the steady state in two periods.*
(ii) *Wealth inequality and the social security tax rate in any subsequent period are higher than those in the initial one.*

Note that the dynamics of social security are not decided by the government with a commitment technology. Instead, the system is repeatedly determined by its current constituency, of which wealth dispersion is a key factor. Forward-looking households, rationally perceiving the link between wealth distribution and social security, will adjust their private savings accordingly. This alters the constituency for social security in the future. In particular, Corollary 2 shows that this interaction leads to a growing size of social security in the dynamic politico-economic equilibrium with stationary demography. The underlying mechanism is twofold. On the one hand, the establishment of a social security system increases future wealth dispersion since within-cohort transfers discourage private savings of the poor more than those of the rich. On the other hand, the larger wealth dispersion makes transfers more desirable in the future. This provides the political support for a larger size of social security in the following periods.

3.3. A closed-form solution for the two-group case

In this subsection, we provide a closed-form solution of the Markov perfect equilibrium. A complete characterization of the equilibrium explicitly reveals the dynamic interaction between policy decision-making and individuals' intertemporal choice. To this end, we let $J = 2$ and $p^j = 1/2$. Households with type $j = 1(2)$ are referred to as the poor (rich).

Proposition 3 states the main results in the two-group case.

Proposition 3. *Assume $u(c) = \log(c)$, $J = 2$ and $p^j = 1/2$. In the Markov perfect equilibrium,*

(i) *The policy rule $F(k_t^1, k_t^2, n_t)$ follows*

$$F\left(k_t^1, k_t^2\right) = \begin{cases} H(k_t^2/k_t^1, n_t) > 0, & \text{if } \alpha\upsilon_t < 1, \text{ or if } \alpha\upsilon_t \geqslant 1 \text{ and } k_t^2/k_t^1 > \Theta(\upsilon_t), \\ 0, & \text{otherwise}, \end{cases} \tag{26}$$

where

$$H\left(k_t^2/k_t^1, n_t\right) \equiv \frac{-\Phi(\upsilon_t) + \sqrt{\Phi(\upsilon_t)^2 + 4\Delta(\upsilon_t)\left(\alpha - \frac{4\alpha^2\upsilon_t k_t^2/k_t^1}{(1+k_t^2/k_t^1)^2}\right)}}{2\Delta(\upsilon_t)} \tag{27}$$

with $\upsilon_t \equiv ((1+\alpha\beta)/(1-\alpha))n_t$, $\Delta(\upsilon_t) \equiv (1-\alpha) + (1-\alpha)^2\upsilon_t$, $\Phi(\upsilon_t) \equiv -1 + 2\alpha + 2\alpha(1-\alpha)\upsilon_t$ and $\Theta(\upsilon\upsilon_t) \equiv 2\alpha\upsilon_t - 1 + 2\sqrt{\alpha\upsilon_t(\alpha\upsilon_t - 1)}$.

The proof is given in Appendix A.2. When $J = 2$, the political decision on the social security tax rate depends on the wealth ratio, k_t^2/k_t^1. Eqs. (26) and (27) show explicitly that τ_t is increasing in wealth inequality. More importantly, the conditions in Proposition 3 characterize the politico-economic environment in which the social security system can be sustained in the Markov equilibrium. For $\alpha\upsilon_t < 1$ to hold, a low n_t or α is needed to reinforce the political constituency for the system, either by increasing the political weight of the old or by mitigating the welfare cost of the young (see the intuition discussed above). When $\alpha\upsilon_t \geqslant 1$, the intra-generational redistribution becomes the key. Social security survives only in societies with sufficiently unequal wealth distribution. There would be no social security if no within-cohort heterogeneity existed. Therefore, when $\alpha\upsilon_t \geqslant 1$, the political support for social security essentially comes from intra-generational redistribution.

4. Quantitative exercises

Although the two-period OG model is very simple, we would like to quantitatively see the size of social security and the importance of the dynamic interaction between inequality and social security. To this end, two quantitative exercises are conducted in this section. First, we calibrate an economy with a stationary demographic structure to the 1947–1969 U.S. economy as the benchmark case. The exercise suggests a quantitatively important contribution of the dynamic interaction between inequality and social security to the growth of social security after World War II. We then extend the model by incorporating ageing population in the U.S. We will see that the model not only can match the rise in the OASI contribution rate in the 1970s and 1980s, but also explain a significant part of the increase in wealth inequality in the data.

4.1. Stationary demographic structure

We first consider an economy with a stationary demographic structure. The parameter values are set as follows. The stationary population growth is calibrated to match the average dependency ratio of 15.7 percent in the U.S. between 1947 and 1969.[17] A is normalized to unity. $\alpha = 0.3$ so that the labor income share of 0.7 matches that in the U.S. in the two decades after World War II (see Gomme and Rupert, 2004). We let $J = 5$ and $p^j = 0.20$, and set $\{\gamma^j\}_{j=1}^5$ equal to average labor earnings by quintiles in the 1962 SFCC.[18] Finally, each period in the OG model is assumed to contain 30 years, and $\beta = 0.994^{30}$ is calibrated so that the steady state annual interest rate equals the average real return of 2.24 percent of ten-year U.S. treasury bills from 1962 to 1969.[19]

Consider the scenario in Section 3.2, where no social security exists before period 1, and voting for social security is unanticipatedly launched at period 1. The period-1 wealth distribution is identical to the earnings distribution, which determines the period-1 social security tax rate of 3.49 percent by solving (21). Note that the initial capital stock is irrelevant (Proposition 1). Since the original Social Security Act was signed into law in 1935, it would be natural to refer to period 1 as the 1930s. Then, the period-1 tax rate of 3.49 percent is not far from the initial OASI contribution rate of 2 percent (Martin and Weaver, 2005).

Following the first part of Corollary 2, the economy converges to the steady state in period 2. Recall that our model is calibrated to the U.S. economy in the period from 1947 to 1969. So period 2 can be referred to as 1947–1969. The second part of Corollary 2 implies a growing size of social security. In our model, the tax rate increases to 7.19 percent, 3.7 percentage points higher than the initial rate. The increase of τ_t in period 2, driven purely by the dynamic interaction

[17] Data source: U.S. Census Bureau.
[18] The earnings quintiles are provided in the technical appendix, which is available upon request.
[19] Data source: Bureau of Labor Statistics. The return data start from 1962. The main findings below are robust if we choose to target the average real return of 2.86% of 10-year U.S. treasure securities from 1962 to 2000.

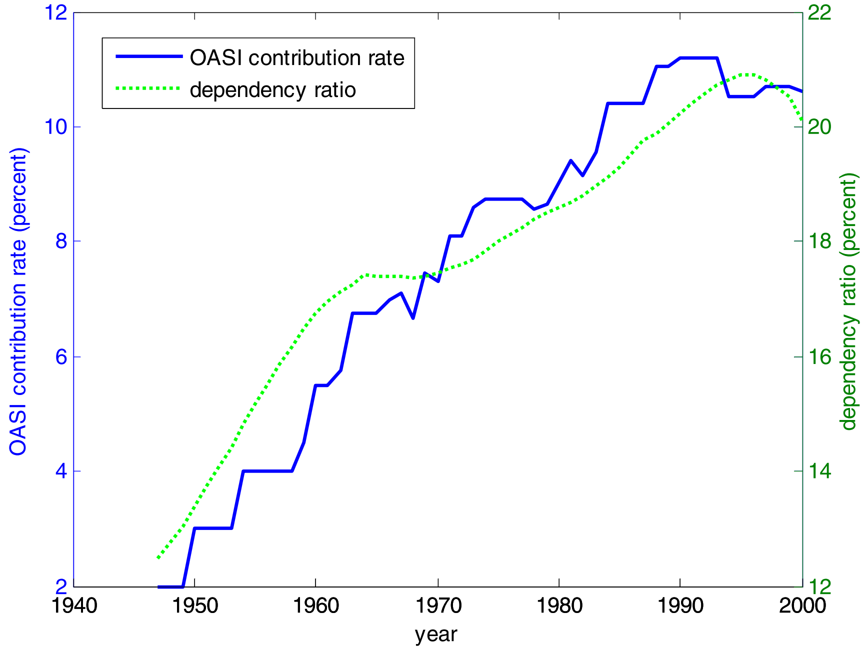

Fig. 1. Social security contribution rates and dependency ratios in the U.S. The solid line stands for Old-Age and Survivors Insurance contribution rates (combined with employee and employer). Data source: Annual Statistical Supplement to the Social Security Bulletin, Table 2.A3. The dotted line is the ratio of population above 65 over population from 15 to 64. Data source: U.S. Census Bureau.

between social security and wealth inequality, turns out to be quantitatively sizable. The increase is one percentage point larger than the 2.68 percentage points increase in the data.[20]

4.2. Ageing population

Despite being highly stylized, the calibrated model economy performs reasonably well in matching the dynamics of social security in the U.S. in the two decades after World War II. A salient feature of the U.S. social security system is that it has continued to grow after 1970. The solid line in Fig. 1 plots the OASI contribution rate, which increased to 10 percent in the 1980s. This is obviously inconsistent with the model with a stationary demographic structure, in which social security stops growing in two periods. Nevertheless, we may easily extend the model by incorporating some non-stationary feature of the U.S. economy and then investigate qualitative and quantitative implications.

The equilibrium policy rule has already suggested the importance of the old population in determining social security (see, also, Gonzalez-Eiras and Niepelt, 2008). Interestingly, the dependency ratio in the U.S., depicted by the dotted line in Fig. 1, exhibits an increasing trend similar to that of the OASI contribution rate. The average dependency ratio reached 17.9 percent in the 1980s, more than two percentage points higher than that between 1947 and 1969. The ageing population may also affect wealth inequality through social security, as illustrated by our model. Table 1 reports the average household financial wealth by quintiles in the 1962, 1983, 1986 and 1989 SFCC/SCF.[21] In line with earlier findings (e.g., Wolff, 1987), household financial wealth became more unequal in the 1980s. Households in the top quintile increased their wealth from USD40,088 in 1962 to an average of USD53,784 in the 1980s, while households in the bottom quintile nearly tripled their debt, which increased from USD773 in 1962 to an average of USD2,374 in the 1980s. We also compute the Gini coefficient of wealth quintiles. The average Gini coefficient in the 1980s was 0.638, substantially higher than 0.580 in 1962.

The observed co-movement of social security and wealth inequality can naturally occur in our model economy with an ageing population. An increase of the dependency ratio tends to raise the tax rate by assigning more weight to the old in the political decision process. Moreover, a higher tax rate will make wealth more unequal, as illustrated by Lemma 1. In addition to the qualitative features, which are consistent with the observations, the model can generate a quantitatively

[20] The average OASI contribution rate increased from 2 percent to 4.68 percent between 1947 and 1969. Data source: Annual Statistical Supplement to the Social Security Bulletin, Table 2.A3.

[21] Financial wealth is referred to as total household wealth net of consumer durables, household inventories, net equity in owner-occupied housing and pension wealth. The concept of financial wealth is perhaps closest to that of k in our model, which abstracts away many realistic elements such as durables and housing. We follow Wolff (1987) to construct financial wealth, except that pension wealth is excluded from our calculation. The details are provided in the technical appendix, which is available upon request. Nominal wealth is deflated by CPI from the Bureau of Labor Statistics.

Table 1
Average household financial wealth by quintiles (in 1962 dollars).

	1962	1983	1986	1989
Top quintile	40 088	49 429	62 607	49 315
2nd quintile	3398	2400	4237	3507
3rd quintile	995	218	1034	534
4th quintile	106	−142	64	13
Bottom quintile	−773	−2592	−2172	−2357
Gini coefficient	0.580	0.665	0.613	0.636

Data source: SFCC/SCF 1962, 1983, 1986 and 1989. Gini coefficient is based on the author's calculation.

Table 2
Model and data.

	Period 0	Period 1	Period 2
	Exogenous variable		
$1/n_t$	0.157	0.157	0.193
	Endogenous variable		
τ_t	0.0349	0.0719	0.0897
Gini coefficient of wealth	0.2762	0.3198	0.3306
	Data		
OASI contribution rate	0.0200	0.0468	0.0981
Gini coefficient of wealth	–	0.5797	0.6380

sizable increase in the social security tax rate and wealth dispersion in response to the observed increase in the dependency ratio.

Consider the following experiment as a simple extension to the benchmark case. Suppose that the economy in periods 1 and 2 is exactly the same as before. The only difference is that in the present experiment, the dependency ratio increases in period 3, which corresponds to 1970–2000 in the U.S. Anticipating the higher dependency ratio and the associated higher social security tax rate, young households at period 2 make saving decisions according to (7), such that wealth distribution in period 3 becomes more dispersed. The period-2 tax rate increases according to Proposition 1. When calibrating the period-2 dependency ratio to 19.3 percent, the average ratio in the U.S. from 1970 to 2000, and maintaining all the other parameter values in the previous case, we find that the social security tax rate will increase to 8.97 percent, about 0.8 percentage points lower than the average OASI contribution rate of 9.81 percent between 1970 and 2000.[22] The increase may account for a significant proportion of the rise in the OASI contribution rate. The ageing population also widens wealth inequality through its interaction with social security. The Gini coefficient of wealth increases by 3.4 percent, rising from 0.320 to 0.331. Although the model fails to match the observed level of wealth inequality, the above result suggests that the ageing population through social security might account a significant part of the ten-percent increase in the Gini coefficient of household financial wealth between the 1962 and 1983/1986/1989 SFCC/SCF. Table 2 summarizes the results.[23]

It is also interesting to know how the increase in social security and wealth inequality reinforces each other in the political equilibrium. To this end, we shut down the interaction by letting the period-3 wealth distribution be identical to that in period 2. The corresponding period-3 tax rate is 8.07 percent, substantially lower than 8.97 percent in the equilibrium. In other words, with the absence of endogenous response of wealth dispersion, the increase in the dependency ratio alone accounts for only 49 percent of the increase of social security tax rate. The dynamic interaction between social security and wealth inequality doubles the effect of the ageing population on the size of social security. This confirms the quantitative importance of this dynamic interaction for social security growth, as shown in the benchmark case. Finally, some caveats are worth mentioning. The rise in wealth inequality in the 1980s might be partly related to the rise in income inequality in the 1980s. This channel is missing here since changes to income inequality can affect only the next-period wealth inequality in the model. In addition, the calibrated model fails to match the level of wealth inequality in the data. So, the above quantitative exercise can provide only some rough clues regarding the relative change of wealth inequality in response to an increase of the dependency ratio through social security.

Given the simplicity of the model, our calibrated political economy has the ability to match the size of social security in the U.S. Moreover, the quantitative exercise suggests an important role of the endogenous change in wealth inequality;

[22] The rising life expectancy is associated with a relative fall in the population of the old rich. If we assume that the increase in the dependency ratio of the U.S. from the 1947–1969 period to the 1970–2000 period is driven entirely by a rise in life expectancy of households in the bottom two income quintiles, the steady state social security tax will increase to 9.77 percent.

[23] The consumption Gini coefficient of the old increases from 0.314 to 0.323. The 2.9-percent increase in consumption inequality is smaller than the 3.4-percent increase in wealth inequality. This is because the old receive lump-sum social security benefits, which partially offset the effect of a more unequal wealth distribution on consumption inequality.

more than half of the increase in the social security tax rate is driven by the dynamic interaction between wealth inequality and social security. This may shed some light on the growth of the welfare state in other OECD countries.[24]

5. Ramsey solution

We have characterized the Markov political equilibrium. It is instructive to compare the outcomes with the Ramsey solution. To this end, we characterize the efficient allocation, where a benevolent planner with a commitment technology sets the sequence of tax rates $\{\tau_t\}_{t=1}^{\infty}$ so as to maximize the sum of the discounted utilities of all generations. The planner's constraint is that the chosen policy should be implementable as a competitive equilibrium. The corresponding Ramsey problem is

$$\max_{\{\tau_t\}_{t=1}^{\infty}} \beta \sum_{j=1}^{J} p^j U_1^{o,j} + \sum_{t=1}^{\infty} \rho^t \left(\sum_{j=h,l} p^j U_t^{y,j} \right), \tag{28}$$

subject to individuals' budget constraints (2) and (3), factor prices (4) and (5), the balanced-budget rule (6), private saving function (7), the law of motion of aggregate capital (9) and the non-negative constraint of τ_t. $\rho \in (0,1)$ is the intergenerational discount factor. $\rho = \beta$ in the case of geometric discounting, in which the planner weighs cohorts' welfare by households' discount factor. When population growth is constant, $\rho = \beta n$ in the case of dynastic discounting in which the planner weighs cohorts by their size and discounts their welfare by β.[25] Compared with the political decision problem (11), the Ramsey allocation problem (28) has two distinctive features. First, the Ramsey planner cares about the welfare of all future generations and, second, she has the ability to commit to future policies.[26]

For notational convenience, $I_{t,t+i} \equiv \partial k_{t+i}/\partial \tau_t$ is denoted as the impact of τ_t on the future capital stock k_{t+i} for $i \geqslant 1$, as implied by the law of motion of capital (9):

$$I_{t,t+i} = \begin{cases} \frac{\partial k_{t+i}}{\partial k_{t+i-1}} \frac{\partial k_{t+i-1}}{\partial k_{t+i-2}} \cdots \left(\frac{\partial k_{t+1}}{\partial \tau_t} + \frac{\partial k_{t+1}}{\partial k_t}\frac{\partial k_t}{\partial \tau_t}\right) < 0, & \text{for } t > 1, \\ \frac{\partial k_i}{\partial k_{i-1}} \frac{\partial k_{i-1}}{\partial k_{i-2}} \cdots \frac{\partial k_2}{\partial \tau_1} < 0, & \text{for } t = 1. \end{cases} \tag{29}$$

The second line of (29) is due to the fact that k_1 is predetermined. τ_t also affects the capital stock at time t, since τ_t may influence private savings in the preceding period. Its impact, denoted by $I_{t,t}$, is equal to

$$I_{t,t} = \begin{cases} \frac{\partial k_t}{\partial \tau_t} < 0, & \text{for } t > 1, \\ 0, & \text{for } t = 1. \end{cases} \tag{30}$$

$I_{1,1} = 0$ since k_1 is predetermined. Note that τ_t directly influences the welfare of households born at time t and $t-1$ by affecting their after-tax net earnings and social security benefits, respectively. In addition, τ_t indirectly influences the welfare of households born at time t and afterwards via its impact on capital accumulation captured by $I_{t,t+i}$. τ_t has no effect on households born before time $t-1$.

Following the same procedure as in the preceding section, let us look at the impact of the social security tax rate τ_t on the welfare of various groups of households. Due to the envelope argument based on the Euler equation, the welfare effect of τ_t on agents born at time $t-1$, $\partial U_{t-1}^{y,j}/\partial \tau_t$, parallels its effect on old households at time t, $\partial U_t^{o,j}/\partial \tau_t$. Specifically,

$$\frac{\partial U_{t-1}^{y,j}}{\partial \tau_t} = \beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \beta \left(u'\left(c_t^{o,j}\right) n_t w_t + u'\left(c_t^{o,j}\right) \left(k_t^j \frac{\partial R_t}{\partial k_t} + n_t \tau_t \frac{\partial w_t}{\partial k_t} \right) I_{t,t} \right), \tag{31}$$

where $I_{t,t}$ follows (30). The first term on the RHS of (31) reflects the direct effect of τ_t, which increases social security transfers and, thus, benefits old households at time t. The second term captures the general equilibrium effect of τ_t through $I_{t,t}$. Comparing (31) with (17), we see that the general equilibrium effect is absent in the political decision process, where voters take k_t as given. In the Ramsey problem, the planner has the ability to commit to future policies. Thus, she must take into account the impact of τ_t on k_t, for $t > 1$. As shown in Section 3, the general equilibrium effect is twofold. The negative $I_{t,t}$ reduces k_t and, thus, the social security tax base. But a low k_t increases the interest rate. The interest rate effect dominates if τ_t or wealth inequality is not too large. In this case, the overall general equilibrium effect would be positive, implying that the marginal benefit of τ_t to the current old households in the Ramsey problem tends to be larger than its counterpart in the political decision process. A special case is that for $t = 1$, the welfare effect of τ_1 on old households equals that in (17), since the capital in the initial period is predetermined ($I_{1,1} = 0$). More specifically, we have

[24] See, for example, Breyer and Craig (1997) for the description of growing social security benefits in OECD countries.

[25] The analysis below can be trivially extended to the case of dynastic discounting with non-stationary demographic structures.

[26] Gonzalez-Eiras and Niepelt (2008) show that if there is no within-cohort heterogeneity, the Ramsey solution coincides with the first-best allocation, which makes the calculation much simpler. However, the equivalence does not carry over into the present model. It is straightforward that a social planner would like to eliminate within-cohort consumption inequality. This outcome cannot be implemented as a competitive equilibrium, since it implies 100 percent tax rate and zero capital stock. Therefore, the social-planner approach cannot be adopted here.

Lemma 2. *Assume that $u(c) = \log(c)$. In the Ramsey problem, the welfare effect of τ_t on old households at time t equals*

$$\beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \begin{cases} \frac{(1+\beta)(\theta(\tau_t)+\tau_t\theta'(\tau_t))}{\alpha\gamma^j+\tau_t\theta(\tau_t)} - \frac{\beta(1-\alpha)\theta'(\tau_t)}{\theta(\tau_t)} > 0, & \text{for } t > 1, \\ \beta \frac{1-\alpha}{\alpha k_1^j/k_1+(1-\alpha)\tau_1} > 0, & \text{for } t = 1, \end{cases} \tag{32}$$

where $\theta(\cdot)$ is defined by (10)*.*

The first line of (32) is derived in Appendix A.3 and the second line simply follows (18). Note that for $t > 1$, the marginal welfare gain is decreasing in γ^j. The intra-generational redistributive components of social security imply that the higher a household's labor income, the less it can benefit from the pension system.

The social security tax rate τ_t also affects the welfare of all generations born at time t and afterwards. The welfare effect of τ_t on young households at time $t+i$, for $i \geqslant 0$, equals

$$\begin{aligned} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} &= -u'\left(c_{t+i}^{y,j}\right)\gamma^j w_{t+i} + u'\left(c_{t+i}^{y,j}\right)\gamma^j \frac{\partial w_{t+i}}{\partial k_{t+i}}(1-\tau_{t+i})I_{t,t+i} \\ &\quad + \beta u'\left(c_{t+i+1}^{o,j}\right)\left(k_{t+i+1}^j \frac{\partial R_{t+i+1}}{\partial k_{t+i+1}} + n_{t+i+1}\tau_{t+i+1}\frac{\partial w_{t+i+1}}{\partial k_{t+i+1}}\right)I_{t,t+i+1}. \end{aligned} \tag{33}$$

As in (19), the first term in (33) reflects the direct cost of social security taxes for young households. The second and third terms are the general equilibrium effects via I_{t+i} and I_{t+i+1}. Note that for $i \geqslant 1$, the welfare effect $\partial U_{t+i}^{y,j}/\partial \tau_t$ does not enter the political decision on τ_t, since the welfare of future generations is ignored in electoral competition. For $i = 0$, a comparison between (33) and (19) reveals that $\partial U_t^{y,j}/\partial \tau_t$ in the Ramsey problem differs from its counterpart in the political equilibrium in two respects.[27] First, the planner takes into account the negative impact of τ_t on k_t, which reduces the social security tax base at time t. This effect is captured by the second term on the RHS of (33). In the political equilibrium, voters at time t take k_t as given and, hence, ignore this negative impact. Second, there is no strategic effect in the Ramsey problem since the planner can commit to future policies. However, we have shown that the strategic effect is mute under logarithm utility. Therefore, the welfare loss of τ_t to the current young households in the Ramsey problem is greater than that in the political equilibrium, due to the negative $I_{t,t+i}$. An exception is that for $t = 1$, since $I_{1,1} = 0$, the welfare effect of τ_1 on the young is exactly the same as that in the political decision.

Lemma 3. *Assume that $u(c) = \log(c)$. In the Ramsey problem, the welfare effect of τ_t on young households at time $t+i$, for $i \geqslant 0$, is equal to*

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} + \frac{(1+\beta\alpha)\alpha^{i+1}\theta'(\tau_t)}{\theta(\tau_t)} < 0, & \text{if } t > 1, \\ -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} < 0, & \text{if } t = 1, \end{cases} \tag{34}$$

where $\theta(\cdot)$ is defined by (10)*.*

The proof is given in Appendix A.4. Three remarks are in order. First, $\partial U_{t+i}^{y,j}/\partial \tau_t < 0$ shows that τ_t incurs a net welfare loss to all generations born at time t and afterwards. Second, the magnitude of the loss depends only on the current tax rate τ_t, due to the additive separability implied by logarithm utility. The irrelevance of future capital stocks and future tax rates remarkably simplifies the characterization of the Ramsey allocation. Third, τ_t has the same effect on the welfare of the poor and the rich, due to the symmetric effect of τ_t on private savings k_{t+1}^j, as discussed in Section 2.

Now, the first-order conditions of (28) with respect to τ_t can be written as:

$$\beta \sum_{j=h,l} p^j \frac{\partial U_t^{o,j}}{\partial \tau_t} + \sum_{i=0}^{\infty}\left(\rho^{i+1}\sum_{j=h,l} p^j \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t}\right) + \lambda_t = 0, \tag{35}$$

where λ_t is the multiplier on the non-negative constraint of τ_t. Plugging (32) and (34) into (35) solves τ_t. An immediate observation is that the social security tax rate converges to a steady state in two periods, exactly the same as the dynamics of the political equilibrium illustrated by Corollary 2. This is, again, because of the property in Lemma 1 that future wealth inequality depends solely on the future tax rate and is independent of the current state. Appendix A.6 explores some quantitative implications of the Ramsey allocation in the same calibrated economy as in Section 4.1. The main finding is that when $\rho = \beta$, the Ramsey tax rate is much higher than that in the political equilibrium. Comparing (35) with (21), one can find that the weight on the current young relative to the current old in (35) equals ρ/β, while it equals n in the political equilibrium. Therefore, as long as $\rho/\beta < n$, the Ramsey planner would like to place less value on the young. This leads to higher Ramsey tax rates.

[27] Note that for $i = 0$, the direct effect (the first term on the RHS of Eq. (33)) is the same in both the political equilibrium and the Ramsey problem.

5.1. The two-group case

We now turn to the two-group case, which allows an analytical characterization and facilitates comparison with the Markov perfect political equilibrium. The main results are summarized by Proposition 4.

Proposition 4. *Assume that $u(c) = \log(c)$, $J = 2$ and $p^j = 1/2$. In the Ramsey solution,*

(i) *The initial social security tax rate*

$$\tau_1 = \begin{cases} H(k_1^h/k_1^l) > 0, & \text{if } \upsilon\alpha < 1, \text{ or if } \upsilon\alpha \geqslant 1 \text{ and } k_1^h/k_1^l > \Theta(\upsilon), \\ 0, & \text{otherwise}, \end{cases} \tag{36}$$

where $H(\cdot)$ follows (27) with $\upsilon \equiv \rho(1+\alpha\beta)/(\beta(1-\rho\alpha)(1-\alpha))$.

(ii) *$\tau_1^R \gtreqless \tau_1^M$ if and only if $\rho \lesseqgtr \beta n_t/(1+\alpha\beta n_t)$, where τ_1^R and τ_1^M denote the initial social security tax rate in the Ramsey solution and the Markov political equilibrium, respectively.*

(iii) *The social security tax rate converges to a unique steady state in two periods. The steady state tax rate $\bar\tau > 0$ if $\Omega > 0$ and $\bar\tau = 0$ if $\Omega \leqslant 0$, where*

$$\Omega \equiv \frac{(1-\alpha)\beta(1+\gamma^2/\gamma^1)^2}{2\alpha\gamma^2/\gamma^1} - \frac{2\rho(1+\beta\alpha)(2+\beta-\alpha)}{(1-\alpha\rho)(1+\beta)} + \frac{2\beta(1-\alpha)^2}{\alpha(1+\beta)}. \tag{37}$$

(iv) *If $\rho > \frac{\beta(1-\alpha)}{1+\alpha\beta}$, the steady state wealth inequality and social security tax rate are lower than those in the initial levels.*

The proof is given in Appendix A.5. The first part of Proposition 4 states that the initial tax rate τ_1 is determined by the initial wealth inequality k_1^h/k_1^l, which parallels Proposition 3 in the political equilibrium. A high k_1^h/k_1^l leads to a high τ_1, due to the within-cohort redistributive effects of τ_1. The second part of the proposition compares the initial tax rate in the Ramsey solution with that in the political equilibrium. Two effects drive the political outcome τ_1^M to deviate from the efficient allocation τ_1^R. To see this, we rewrite the first-order condition of τ_1 (35) as

$$\beta \sum_{j=h,l} p^j \frac{\partial U_1^{o,j}}{\partial \tau_1} + \beta n_t \sum_{j=h,l} p^j \frac{\partial U_1^{y,j}}{\partial \tau_1} + \sum_{i=1}^{\infty}\left(\rho^{i+1} \sum_{j=h,l} p^j \frac{\partial U_{i+1}^{y,j}}{\partial \tau_1}\right) - (\beta n_t - \rho) \sum_{j=h,l} p^j \frac{\partial U_1^{y,j}}{\partial \tau_1} + \lambda_1 = 0. \tag{38}$$

The first two terms on the LHS of (38) capture the same trade-off in the political decision process (see Eq. (21)).[28] The third term reflects the negative impact of τ_1 on the welfare of households born after the initial period via capital accumulation (see Lemma 3). This negative impact, which makes τ_1^M higher than τ_1^R, is ignored in the political decision process since non-altruistic voters do not care about future generations. The fourth term on the LHS of (38) illustrates the discrepancy between the weight on the current young in the political decision process and that in the Ramsey problem. If $\beta n_t > \rho$, the second effect is opposite to the first effect; the Ramsey planner would like to impose a higher τ_1 since the weight on the young is lower than that in the political decision process.[29] The proposition shows that the second effect dominates the first effect for sufficiently small social discount factors – i.e., $\rho < \beta n_t/(1+\alpha\beta n_t)$.

The third part of the proposition characterizes the steady state. The uniqueness of the steady state in the Ramsey solution can be analytically established. We further provide the sustainability condition of the social security system in the Ramsey allocation. It is immediate that Ω increases in γ^2/γ^1 but decreases in ρ. Intuitively, a high income inequality γ^2/γ^1 increases the within-cohort redistributive benefit of transfers. A high ρ, on the contrary, increases the relative weight on the welfare of future generations and, thus, makes social security, as an inter-generational redistributive policy, less desirable.

The last part of Proposition 4 gives a sufficient condition for decreasing sizes of social security over time. This contrasts the political equilibrium that features growing social security in Corollary 2. The somewhat surprising result primarily comes from the fact that lowering future tax rates are in favor of future generations by encouraging capital accumulation. In addition, as discussed above, forward-looking households will adjust their intertemporal choices according to future transfers. The expectation of lower social security benefits in the future will lead to a lower level of wealth inequality, which considerably offsets the within-cohort redistributive effects of social security.

6. The strategic effect under CRRA utility

So far, we have focused on logarithm utility, which substantially simplifies the analysis and makes explicit solutions available in some special cases. However, many empirical studies suggest the elasticity of intertemporal substitution to be less than unity. It remains unclear to what extent our results would be affected by the deviation from logarithm utility. In

---

[28] Note that for $t = 1$, $\partial U_1^{o,j}/\partial\tau_1$ is the same in both of the Markov political equilibrium and the Ramsey problem, so as $\partial U_1^{y,j}/\partial\tau_1$ (see Lemmas 2 and 3).

[29] Social security always causes welfare loss to the young in both the political equilibrium and the Ramsey allocation.

particular, the strategic effect $\partial\tau_{t+1}/\partial\tau_t$ may arise under a less restrictive utility form. This section adopts a more general CRRA utility function, to see whether the analytical results in the preceding sections are robust with the presence of strategic effect. Assume that

$$u(c)=\frac{c^{1-\sigma}-1}{1-\sigma}, \tag{39}$$

where $\sigma>0$ is the inverse of the intertemporal elasticity of substitution. Households' intertemporal choices and the political decision on social security tax rates are characterized in Appendix A.7. Analytical solutions cannot be obtained for $\sigma\neq 1$, so we resort to numerical methods.

The computational strategy for the Markov perfect equilibrium adopts a standard projection method with Chebyshev collocation (Judd, 1992). The basic idea of the projection method is to approximate unknown functions by finite, weighted sum of a simple basis function such as a polynomial. This method is applied for time-consistent problems in some recent research (Judd, 2003; Ortigueira, 2006).

For simplicity, we focus on an economy with constant population growth and $J=2$, in which all parameters are re-calibrated under the log preference.[30] Let $p^1=p^2=0.5$, $\gamma^1=0.4$ and $\gamma^2=1.6$ such that the labor earnings ratio is equal to the ratio of the mean of the bottom three earnings quintiles to the mean of the top three earnings quintiles in the 1962 SFCC. As in the benchmark calibrated economy, $\alpha=0.30$ and each period contains 30 years. $\beta=0.979^{30}$ is recalibrated to match the annual interest rate of 2.24 percent. Finally, we set $n=2.38$ to target the steady state social security tax rate of 7.19 percent in the benchmark calibrated economy. Lowering n can be interpreted as adjusting the weight of the old in the political decision process downward. We then solve the equilibrium policy rule in the calibrated economy with $\sigma=2$. The result is plotted by Panel A of Fig. 2. τ_t is increasing in k_t^2/k_t^1, implying a positive relationship between social security and wealth inequality, as in the logarithm case.[31]

The solid line in Panel B of Fig. 2 plots the dynamics of social security. The initial aggregate capital stock is set equal to the steady state stock with zero tax rate. The initial wealth ratio is set equal to the earnings ratio. Note that, different from the logarithm case, social security tax rate now depends on the aggregate capital stock. When $\sigma>1$, τ_t turns out to be decreasing in k_t (see Panel A of Fig. 2), consistent with the finding in Gonzalez-Eiras and Niepelt (2008). The impact of k_t on τ_t is quantitatively small, though. A one-percent increase in k_t leads to a less-than-0.1-percent decrease in τ_t. Moreover, although the convergence becomes asymptotic with $\sigma\neq 1$, the quantitative implication for the speed of convergence is actually similar to that in the benchmark case; over 99 percent of the gap between the initial and steady state tax rates will be closed in two periods.

The strategic effect (20) arises with $\sigma\neq 1$. When $\sigma=1$, τ_t does not affect private saving rates and, therefore, the future wealth inequality, due to a cancellation of the income and substitution effects. A lower intertemporal elasticity of substitution ($\sigma>1$) weakens the substitution effect, implying lower saving rates in response to an increase of τ_t which reduces k_{t+1} and increases R_{t+1}. Moreover, the substitution effect becomes even weaker for households with lower earnings. For this reason, the poor reduce their saving rate more than the rich do. The asymmetric effect of τ_t enlarges future wealth inequality k_{t+1}^h/k_{t+1}^l and, thus, raises the future social security tax rate τ_{t+1} via the equilibrium policy rule F. This gives rise to a positive strategic effect of τ_t on τ_{t+1}. Hence, the current young households would like to strategically vote for a higher τ_t since it incurs higher future social security benefits. By constructing a myopic voting equilibrium – where voters can rationally expect future policy outcomes but assume, incorrectly, no strategic interaction between the current and future policies – we can isolate the strategic effect.[32] The strategic effect is quantitatively unimportant: The relative increase in the social security tax rate due to the strategic effect is less than 5 percent.[33]

7. Conclusion

Redistributive transfers in the pay-as-you-go social security system create conflicts of interest among various groups of households. The evolution of household characteristics may change the political support for the system over time. Despite extensive studies of the aggregate and distributive effects of social security, most of the existing literature is silent on how the public decision on social security responds to time-varying political supports in dynamic general equilibrium. In this paper, we analytically characterize the Markov perfect political equilibrium, in which private intertemporal choices and the repeated political decision on social security are mutually affected over time. The main finding is that the dynamic

[30] Although the two-group case can, in principle, be generalized, the projection methods are typically hard to apply for models with more than two state variables.

[31] Different from the logarithm case, the social security tax rate now depends on the aggregate capital stock. When $\sigma>1$, τ_t turns out to be decreasing in k_t, in line with the finding in Gonzalez-Eiras and Niepelt (2008).

[32] A similar notion of pseudo-equilibrium is used by Alesina and Rodrik (1994). The details can be found in an earlier version of this paper (Chen and Song, 2005).

[33] Chen and Song (2005) also provide a detailed description of the Ramsey solution with CRRA preference. We transform the infinite-horizon problem into a finite-horizon problem by the truncated method (e.g., Jones et al., 1993). The benchmark calibrated economy has a steady state Ramsey tax rate of 12.93 percent, much higher than that in the political equilibrium. This is, again, in line with the finding in the logarithm case. The algorithm is provided in a technical appendix, which is available upon request.

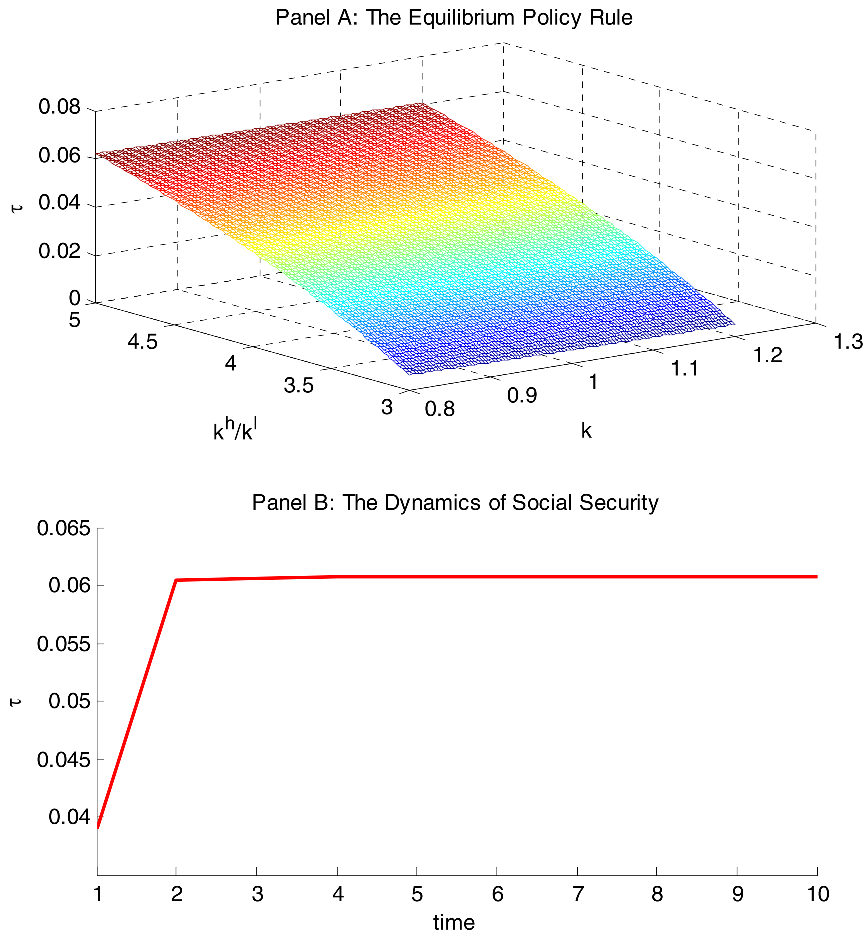

Fig. 2. Equilibrium results under CRRA preference ($\sigma = 2$). Panel A and B plot the Markovian equilibrium policy rule and the dynamics of the social security tax rate, respectively. $\sigma = 2$ and the other parameter values are reported in the text.

interaction between social security and wealth inequality leads to the growing size of social security. This result may shed light on the increasing generosity of social security in OECD countries during the post-war period (Breyer and Craig, 1997). In addition, the dynamic interaction is shown to be quantitatively important: It accounts for more than half of the social security growth in the dynamics.

Our analysis is subject to a number of caveats. For instance, the theory is completely silent on the structure of social security. An interesting extension is to analyze the determination of the size of social security and the degree of its redistributiveness simultaneously.[34] For analytical convenience, we impose a balanced budget on social security transfers. A natural extension of the model would be to relax this assumption. In a related work, Song et al. (2007) analyze the determination of public debt in a small open economy without social security. It will be interesting for future research to incorporate government borrowing into the current setup, to see how public debt interacts with social security.

Appendix A

A.1. Proof of Proposition 1

This subsection proves the first and second parts of Proposition 1. To characterize the equilibrium policy rule, we first investigate a finite-period version of the model. It will be shown that the limit of finite-horizon equilibria turns out to be equivalent to the infinite-horizon equilibrium. Suppose that the economy terminates at time T and that young households born at time T live only one period.

Consider the terminal period T. Since young households do not have any intertemporal trade-off and $c_T^{y,j}$ simply equals the net earning $(1-\tau_T)\gamma^j w_T$, the welfare effect of τ_T on young households at time T is equal to

[34] Cremer et al. (2006) recently deliver such an analysis in a static environment.

$$\frac{\partial U_T^{y,j}}{\partial \tau_T} = -\frac{\gamma^j w_T}{c_T^{y,j}} = -\frac{1}{1-\tau_T}. \tag{40}$$

Plugging (40) and (18) into the first-order condition yields

$$\sum_{j=1}^{J} p^j \frac{1-\alpha}{\alpha k_T^j / k_T + (1-\alpha)\tau_T} - n_T \frac{1}{1-\tau_T} = 0. \tag{41}$$

The second-order condition always holds. Here, we assume interior solution. The following analysis would be trivial if τ_T is bound by 0. Eq. (41) solves a Markovian policy rule at time T:

$$\tau_T = F^T\big(k_T^1, k_T^2, \ldots, k_T^J, \boldsymbol{n}_T\big).$$

Eq. (41) implies that F^T is homogeneous of degree zero for the first J arguments; i.e., $F^T(k_T^1, k_T^2, \ldots, k_T^J, \boldsymbol{n}_T) = F^T(\psi k_T^1, \psi k_T^2, \ldots, \psi k_T^J, \boldsymbol{n}_T)$. Therefore, we have

$$\sum_{j=1}^{J} F_j^T k_T^j = 0. \tag{42}$$

Substituting (7) back into (42) yields an important property that will play a central role in the following proof:

$$\sum_{j=1}^{J} F_j^T \theta^j(\tau_T) = \sum_{j=1}^{J} F_j^T S_2^j(k_{T-1}, \tau_{T-1}, \tau_T) = 0. \tag{43}$$

Now, use (20):

$$\frac{\partial \tau_T}{\partial \tau_{T-1}} = \sum_{j=1}^{J} F_j^T \frac{S_2^j\big(1 - S_3 \sum_{j=1}^{J} F_j^T\big) + S_3 \sum_{i=1}^{J} S_2^j F_j^T}{1 - S_3 \sum_{j=1}^{J} F_j^T} = 0. \tag{44}$$

The second equality comes from (43).

We move to period $T-1$. The absence of the strategic effect makes the rest of the derivation fairly straightforward. The welfare effect of τ_{T-1} on young households in period $T-1$ follows (19). Using (44) and the indirect utility approach, which will be specified below (see Eq. (55)), we find that

$$\frac{\partial U_{T-1}^{y,j}}{\partial \tau_{T-1}} = -\frac{1+\beta\alpha}{1-\tau_{T-1}}. \tag{45}$$

The welfare effect of τ_{T-1} on old households still follows (18). Plugging (40) and (18) into the first-order condition and assuming interior solution, we have

$$\sum_{j=1}^{J} p^j \frac{1-\alpha}{\alpha k_{T-1}^j / k_{T-1} + (1-\alpha)\tau_{T-1}} - n_{T-1} \frac{1+\beta\alpha}{1-\tau_{T-1}} = 0. \tag{46}$$

Again, the second-order condition holds true. Eq. (46) solves a Markovian policy rule at time $T-1$:

$$\tau_{T-1} = F^{T-1}\big(k_{T-1}^1, k_{T-1}^2, \ldots, k_{T-1}^J, \boldsymbol{n}_{T-1}\big).$$

Like F^T, Eq. (46) implies that F^{T-1} is also homogeneous of degree zero for the first J arguments. Applying the same procedure as above, we can prove that

$$\frac{\partial \tau_{T-1}}{\partial \tau_{T-2}} = 0.$$

The only difference between (41) and (46) is the welfare effect of social security tax on the young: $\partial U_{T-1}^{y,j}/\partial \tau_{T-1}$ in (45) differs from $\partial U_T^{y,j}/\partial \tau_T$ in (40). This is because young households born at time $T-1$ live for two periods. Moreover, it can easily be seen that the political decision on τ_t for $t < T-1$ is exactly the same as in time $T-1$. The equivalence boils down to two features: (i) the independence of $\partial U_t^{y,j}/\partial \tau_t$ on the future tax rate; and (ii) the mute strategic effect, as shown by (44) and (45), respectively. □

A.2. Proof of Proposition 3

When $J = 2$, the first-order condition (21) can be rewritten as

$$\sum_{j=1,2} \frac{1-\alpha}{2\alpha \frac{k_T^j/k_T^i}{1+k_T^j/k_T^i} + (1-\alpha)\tau_T} - 2n_T \frac{1}{1-\tau_T} = 0, \tag{47}$$

where $i \neq j$. We assume interior solution. The non-negative constraint of τ_T will be discussed below. Eq. (47) gives a quadratic equation of τ_T

$$\Delta(\upsilon_T)\tau_T^2 + \Phi(\upsilon_T)\tau_T + \frac{4\upsilon_T \alpha^2 k_T^h/k_T^l}{(1+k_T^h/k_T^l)^2} - \alpha = 0, \tag{48}$$

where $\upsilon_T \equiv n/(1-\alpha)$.

Now consider the non-negative constraint. The first case is that $\Phi(\upsilon_T) \geqslant 0$. Since $\Delta(\upsilon_T) > 0$, there is a unique positive τ_T if and only if

$$\left(k_T^h/k_T^l\right)^2 + (2 - 4\upsilon_T\alpha)\left(k_T^h/k_T^l\right) + 1 > 0. \tag{49}$$

For $\upsilon_T\alpha < 1$, the condition always holds. Otherwise, we need

$$k_T^h/k_T^l > 2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha - 1)} \quad \text{or} \quad k_t^h/k_t^l < 2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha - 1)}. \tag{50}$$

The first inequality in (50) is binding since $2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha - 1)} > 1$ for $\upsilon_T\alpha \geqslant 1$. The other inequality in (50) cannot be satisfied since $2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha - 1)} < 1$ for $\upsilon_T\alpha \geqslant 1$. The second case is that $\Phi(\upsilon_T) < 0$. For $\upsilon_T\alpha < 1$, Eq. (49) ensures a unique positive τ_T. For $\upsilon_T\alpha \geqslant 1$, there can be two positive roots if the LHS of (49) is non-positive. This implies $n_T < 1/(2\alpha) - 1$ and contradicts the condition that $n_T \geqslant 1/\alpha - 1$, as given by $\upsilon_T\alpha \geqslant 1$.

To conclude, for $\upsilon_T\alpha < 1$, the Markovian policy rule at time T follows

$$\tau_T = F^T\left(k_T^h, k_T^l\right) = \frac{-\Phi(\upsilon_T) + \sqrt{\Phi(\upsilon_T)^2 + 4\Delta(\upsilon_T)\left(\alpha - \frac{4\upsilon_T\alpha^2 k_T^h/k_T^l}{(1+k_T^h/k_T^l)^2}\right)}}{2\Delta(\upsilon_T)} > 0. \tag{51}$$

For $\upsilon_T\alpha \geqslant 1$, $F_T(k_T^h, k_T^l)$ follows (51) if k_T^h/k_T^l satisfies the first inequality in (50) and is equal to zero otherwise.

We turn to period $T-1$. The proof in Appendix A.1 has shown that $\partial\tau_T/\partial\tau_{T-1} = 0$ and $\partial U_{T-1}^{y,j}/\partial\tau_{T-1} = -(1+\beta\alpha)/(1-\tau_{T-1})$. Then, the first-order condition yields, again, a quadratic equation of τ_{T-1}:

$$\Delta(\upsilon_{T-1})\tau_{T-1}^2 + \Phi(\upsilon_{T-1})\tau_{T-1} + \frac{4\upsilon_{T-1}\alpha^2 k_{T-1}^h/k_{T-1}^l}{(1+k_{T-1}^h/k_{T-1}^l)^2} - \alpha = 0, \tag{52}$$

where $\upsilon_{T-1} \equiv n_{T-1}(1+\alpha\beta)/(1-\alpha)$. The conditions for corner solutions can easily be derived following the above procedures.

To conclude, for $\upsilon_{T-1}\alpha < 1$, the Markovian policy rule at time $T-1$ follows

$$\tau_{T-1} = F^{T-1}\left(k_{T-1}^h, k_{T-1}^l\right) = \frac{-\Phi(\upsilon_{T-1}) + \sqrt{\Phi(\upsilon_{T-1})^2 + 4\Delta(\upsilon_{T-1})\left(\alpha - \frac{4\upsilon_{T-1}\alpha^2 k_{T-1}^h/k_{T-1}^l}{(1+k_{T-1}^h/k_{T-1}^l)^2}\right)}}{2\Delta(\upsilon_{T-1})} > 0. \tag{53}$$

For $\upsilon_{T-1}\alpha \geqslant 1$, τ_{T-1} follows (53) if k_{T-1}^h/k_{T-1}^l satisfies

$$k_{T-1}^h/k_{T-1}^l > 2\upsilon_{T-1}\alpha - 1 + 2\sqrt{\upsilon_{T-1}\alpha(\upsilon_{T-1}\alpha - 1)} \tag{54}$$

and τ_{T-1} is equal to zero otherwise. As discussed in Appendix A.1, the political decision on τ_t for $t < T-1$ is exactly the same as in time $T-1$. Consequently, the key parameter υ_t is exactly the same as υ_{T-1} for $t < T-1$. □

A.3. Proof of Lemma 2

We use the indirect utility approach to simplify the derivation of the welfare effect of social security tax rates. Using individuals' budget constraints (2) and (3), factor prices (4) and (5), the balanced budget (6), private saving function (7) and the law of motion of aggregate capital (9), after some algebra, we can obtain the indirect utility of all generations born at time t in terms of k_t, τ_t and τ_{t+1}:

$$V_t^j(k_t,\tau_t,\tau_{t+1}) = (1+\beta\alpha)\alpha\log k_t + (1+\beta\alpha)\log(1-\tau_t) \\ + (1+\beta)\log\big(\alpha\gamma^j + \tau_{t+1}\theta(\tau_{t+1})\big) - \beta(1-\alpha)\log\theta(\tau_{t+1}), \tag{55}$$

where $\theta(\cdot)$ is defined by (10). The indirect utility of old households at time 1 is

$$U_1^{o,j} = \log\left(\alpha\frac{k_1^j}{k_1} + (1-\alpha)\tau_1\right) + \alpha\log k_1. \tag{56}$$

Differentiating (55), the welfare effect of τ_t on old households at time t equals

$$\beta\frac{\partial U_t^{o,j}}{\partial\tau_t} = \frac{\partial V_{t-1}^j}{\partial\tau_t} = (1+\beta)\frac{\theta(\tau_t)+\tau_t\theta'(\tau_t)}{\alpha\gamma^j+\tau_t\theta(\tau_t)} - \beta(1-\alpha)\frac{\theta'(\tau_t)}{\theta(\tau_t)}, \tag{57}$$

for $t>1$. Differentiating (56) with respect to τ_1 yields the second line of (32). This proves the lemma. □

A.4. Proof of Lemma 3

By (9), we know $\partial k_{t+i}/\partial k_{t+i-1} = \alpha k_{t+i}/k_{t+i-1}$, $\partial k_{t+1}/\partial\tau_t = -k_{t+1}/(1-\tau_t)$ and $\partial k_t/\partial\tau_t = \theta'(\tau_t)k_t/\theta(\tau_t)$. Thus, $I_{t,t+i}$ can be written as

$$I_{t,t+i} = \begin{cases} \alpha^{i-1}k_{t+i}\left(-\frac{1}{1-\tau_t} + \alpha\frac{\theta'(\tau_t)}{\theta(\tau_t)}\right), & \text{for } t>1,\\ -\alpha^{i-1}\frac{k_i}{1-\tau_1}, & \text{for } t=1. \end{cases} \tag{58}$$

According to the indirect utility function (55), the welfare effect of τ_t on young households at time t, for $t>1$, equals

$$\frac{\partial U_t^{y,j}}{\partial\tau_t} = \frac{\partial V_t^j}{\partial\tau_t} + \frac{\partial V_t^j}{\partial k_t}\frac{\partial k_t}{\partial\tau_t} = (1+\beta\alpha)\left(-\frac{1}{1-\tau_t} + \alpha\frac{\theta'(\tau_t)}{\theta(\tau_t)}\right). \tag{59}$$

The welfare effect of τ_t on households born after time t is

$$\frac{\partial U_{t+i}^{y,j}}{\partial\tau_t} = \frac{\partial V_{t+i}^j}{\partial k_{t+i}}I_{t,t+i} = (1+\beta\alpha)\alpha^i\left(-\frac{1}{1-\tau_t} + \alpha\frac{\theta'(\tau_t)}{\theta(\tau_t)}\right), \tag{60}$$

for $i=1,2,\ldots$. The second equality in (60) comes from the first line in (58). Eqs. (59) and (60) give the first line of (34). Finally, for $t=1$, we have

$$\frac{\partial U_1^{y,j}}{\partial\tau_1} = \frac{\partial V_1^j}{\partial\tau_1} = -\frac{1+\beta\alpha}{1-\tau_1} \tag{61}$$

and

$$\frac{\partial U_{i+1}^{y,j}}{\partial\tau_1} = \frac{\partial V_{i+1}^j}{\partial k_{i+1}}I_{1,i+1} = -\alpha^i\frac{1+\beta\alpha}{1-\tau_1}, \tag{62}$$

for $i=1,2,\ldots$. The second equality in (62) comes from the second line in (58). Eqs. (61) and (62) give the second line of (34). □

A.5. Proof of Proposition 4

Assuming interior solution, the first-order condition of (28) with respect to τ_1 is

$$\beta\sum_{j=h,l}\frac{\partial U_1^{o,j}}{\partial\tau_1} + \sum_{i=0}^{\infty}\rho^{i+1}\left(\sum_{j=h,l}\frac{\partial U_{i+1}^{y,j}}{\partial\tau_1}\right) = 0, \tag{63}$$

where $\partial U_i^{y,j}/\partial\tau_1$ follows from (61) and (62). This leads to

$$\beta\sum_{j=h,l}\frac{1-\alpha}{2\alpha\frac{k_1^j/k_1^i}{1+k_1^j/k_1^i} + (1-\alpha)\tau_1} - 2\rho\frac{1+\beta\alpha}{1-\rho\alpha}\frac{1}{1-\tau_1} = 0, \tag{64}$$

which gives a quadratic equation of τ_1. Since $\partial^2 U_1^{o,j}/\partial\tau_1^2$ and $\partial^2 U_{i+1}^{y,j}/\partial\tau_1^2$ are negative by (56), (61) and (62), the second-order condition is always satisfied.

Comparing (64) with (46) and incorporating corner solutions, it is immediate that the closed-form solution of τ_1 follows (36) with $\upsilon \equiv \rho(1+\alpha\beta)/(\beta(1-\rho\alpha)(1-\alpha))$. A comparison of the two first-order conditions shows that $\tau_1^R \gtreqless \tau_1^M$ if and only if $\rho \lesseqgtr \beta n/(1+\alpha\beta n)$.

The first-order conditions of (28) with respect to τ_t for $t>1$ are

$$\beta \sum_{j=h,l} \frac{\partial U_t^{o,j}}{\partial \tau_t} + \sum_{i=0}^{\infty} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} \right) = 0. \tag{65}$$

Substituting (57), (59) and (60) for $\partial U_t^{o,j}/\partial \tau_t$ and $\partial U_{t+i}^{y,j}/\partial \tau_t$, respectively, (65) leads to

$$\sum_{j=h,l} \left(\frac{(1+\beta)(\theta(\tau_t)+\tau_t\theta'(\tau_t))}{\alpha\gamma^j + \tau_t\theta(\tau_t)} \right) - 2\rho \frac{1+\beta\alpha}{1-\rho\alpha} \frac{1}{1-\tau_t} + 2\left(\rho \frac{(1+\beta\alpha)\alpha}{1-\rho\alpha} - \beta(1-\alpha) \right) \frac{\theta'(\tau_t)}{\theta(\tau_t)} = 0. \tag{66}$$

Eq. (66) solves a constant τ_t for $t>1$; i.e., the Ramsey tax rate converges to a steady state in two periods.

Note that the second-order conditions are always satisfied. To see this, Eq. (57) shows that $\partial^2 U_t^{o,j}/\partial \tau_t^2 < 0$. Differentiating (59) and (60) with respect to τ_t establishes

$$\operatorname{sgn}\left(\frac{\partial^2 U_{t+i}^{y,j}}{\partial \tau_t^2} \right) = \operatorname{sgn}\left(\frac{\alpha(1-\alpha)^2}{(\alpha(1+\beta)+(1-\alpha)\tau_t)^2} - \frac{1}{(1-\tau_t)^2} \right).$$

Since $\tau_t \in [0,1]$, it can easily be found that $\partial^2 U_{t+i}^{y,j}/\partial \tau_t^2 < 0$ always holds. The second-order condition implies that the solution of (66) is unique.

Denote $L(\tau_t)$ the LHS of (66). After some algebra, $L(\tau_t)$ can be written as

$$\begin{aligned} L(\tau_t) = {} & \sum_{j=h,l} \left(\frac{(1-\alpha)\beta\alpha^2(1+\beta)^2}{(\alpha(1+\beta)+(1-\alpha)\tau_t)(\gamma^j\alpha^2(1+\beta)+\alpha(1-\alpha)(\gamma^j+\beta)\tau_t)} \right) \\ & - 2\rho \frac{1+\beta\alpha}{1-\rho\alpha} \frac{1}{1-\tau_t} - 2\left(\rho \frac{(1+\beta\alpha)\alpha}{1-\rho\alpha} - \beta(1-\alpha) \right) \frac{1-\alpha}{\alpha(1+\beta)+(1-\alpha)\tau_t}. \end{aligned}$$

It immediately follows that $\lim_{\tau_t \to 1} L(\tau_t) = -\infty$. Since $L'(\tau_t) < 0$ by the second-order condition, there is a strictly positive τ_t if and only if $L(0) > 0$. This establishes (37).

Finally, turn to the last part of the proposition. Comparing (66) and (64), we only need to show that for any τ,

$$\sum_{j=h,l} \left(\frac{(1+\beta)(\theta(\tau)+\tau\theta'(\tau))}{\alpha\gamma^j + \tau\theta(\tau)} \right) + 2\chi \frac{\theta'(\tau)}{\theta(\tau)} < \beta \sum_{j=h,l} \frac{1-\alpha}{\alpha\gamma^j + (1-\alpha)\tau},$$

where $\chi \equiv \rho \frac{(1+\beta\alpha)\alpha}{1-\rho\alpha} - \beta(1-\alpha)$.[35] The condition that $\rho > \frac{\beta(1-\alpha)}{(1+\alpha\beta)}$ implies that $\chi > 0$. Therefore, it is sufficient to have

$$\sum_{j=h,l} \left(\frac{(1+\beta)(\theta(\tau)+\tau\theta'(\tau))}{\alpha\gamma^j + \tau\theta(\tau)} \right) \leqslant \beta \sum_{j=h,l} \frac{1-\alpha}{\alpha\gamma^j + (1-\alpha)\tau}.$$

Since $\tau\theta(\tau) > (1-\alpha)\tau$, we are left to prove that

$$(1+\beta)\big(\theta(\tau)+\tau\theta'(\tau)\big) \leqslant \beta(1-\alpha).$$

Some algebra establishes that the inequality always holds. □

A.6. A quantitative exercise for the Ramsey solution

This subsection quantitatively compares the Ramsey allocation with the political equilibrium in the calibrated economy in Section 4.1. We assume geometric discounting with $\rho = \beta$.[36] The Ramsey allocation will favor the old with a smaller population size since geometric discounting does not count population shares. As illustrated by the second part of Proposition 4, if n is sufficiently large, this effect will dominate the negative welfare effect on future generations, resulting in a higher tax rate. In our benchmark calibrated economy, the initial and steady state Ramsey tax rates are equal to 22.89 percent and 28.12 percent, respectively, much higher than their counterparts in the political equilibrium.[37]

[35] We use the fact that $k_1^j/k_1^i = \gamma^j/\gamma^i$ and $\gamma^j + \gamma^i = 2$.

[36] Assuming geometric discounting; i.e., $\rho = \beta n$, is not feasible here as $\beta n > 1$ in the calibrated economy.

[37] Although the Ramsey allocation also features a growing social security in our calibrated economy, it is worth mentioning that the result highly depends on parameter values as illustrated by the last part of Proposition 4. An earlier version of this paper (Chen and Song, 2005) finds a decreasing sequence of tax rate in the Ramsey solution with reasonable parameter values.

To evaluate the welfare implications of the Ramsey solution, we use a consumption equivalent variation measure. Specifically, the aggregate welfare gain of the Ramsey solution to the political equilibrium for cohort i is measure by

$$\Delta(i) = \exp\left(\frac{\sum_j p^j U^R(i,j)}{\sum_j p^j U^M(i,j)}\right) - 1,$$

where $U^R(i,j)$ and $U^M(i,j)$ stand for utility of households of group j and cohort i in the Ramsey solution and the political equilibrium, respectively. Moving from the political equilibrium to the Ramsey allocation can increase the welfare of the first two cohorts (i.e., the old and young in the initial period) by 101.99 percent and 4.31 percent. Cohorts born after the initial period will experience welfare loss; their welfare drops by 15.75 percent. The first cohort benefits the most since they have retired and, thus, do not suffer from the higher tax rate. The aggregate welfare of the second cohort also improves because of the even higher social security tax rate when they retire. Later cohorts suffer from the Ramsey allocation since higher tax rates discourage capital accumulation and, therefore, lower long-run consumption levels.

A.7. CRRA utility

Given (39), households' problem (1) becomes

$$\max_{k^j_{t+1}} \frac{(c_t^{y,j})^{1-\sigma} - 1}{1-\sigma} + \beta \frac{(c_{t+1}^{o,j})^{1-\sigma} - 1}{1-\sigma}, \tag{67}$$

subject to (2) and (3). Households' saving choice follows the Euler equation $c_{t+1}^{o,j}/c_t^{y,j} = (\beta R_{t+1})^{1/\sigma}$. Using budget constraints (2) and (3), factor prices (4) and (5) and the balanced-budget rule (6), k^j_{t+1} follows

$$\begin{aligned} k^j_{t+1} &= G^j(k_t, \tau_t, \tau_{t+1}, k_{t+1}) \\ &\equiv \frac{\gamma^j (A\alpha(k_{t+1}/n)^{\alpha-1}\beta)^{1/\sigma} A(1-\alpha)(1-\tau_t)(k_t/n)^\alpha - A(1-\alpha)\tau_{t+1}k_{t+1}^\alpha n^{1-\alpha}}{(A\alpha(k_{t+1}/n)^{\alpha-1}\beta)^{1/\sigma} + A\alpha(k_{t+1}/n)^{\alpha-1}}. \end{aligned} \tag{68}$$

By $k_{t+1} = \sum_{j=h,l} k^j_{t+1}/2$, Eq. (68) solves private saving functions

$$k^j_{t+1} = S^j(k_t, \tau_t, \tau_{t+1}), \tag{69}$$

with

$$S_i^j = \frac{G_i^j + (G_i^l G_4^h - G_i^h G_4^l)/2}{1 - \sum_{j=h,l} G_4^j/2}, \tag{70}$$

for $i = 1, 2, 3$. Correspondingly, the aggregate saving function can be written as

$$k_{t+1} = S(k_t, \tau_t, \tau_{t+1}), \tag{71}$$

with

$$S_i = \frac{\sum_{j=h,l} S_i^j}{2}. \tag{72}$$

Given the Markovian policy rule (12), a recursive form of private and aggregate saving functions can be solved.

$$k^j_{t+1} = \hat{S}^j(k_t, \tau_t), \tag{73}$$

$$k_{t+1} = \hat{S}(k_t, \tau_t), \tag{74}$$

with $\hat{S}_i^h$, $\hat{S}_i^l$ and $\hat{S}_i$ pinned down by the same method in Section 3, for $i = 1, 2$. These derivatives will be used in the numerical solution, as will be seen in the next subsection. The welfare effect, $\partial U_t^{o,j}/\partial \tau_t$ and $\partial U_t^{y,j}/\partial \tau_t$, as well as the first-order conditions of (11) still follow (17), (19) and (21), respectively.

Now, we turn to the Ramsey problem. The indirect utility of young households at time t can be expressed as follows:

$$W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1}) \equiv \left(\gamma^j A(1-\tau_t)(k_t/n)^\alpha + \tau_{t+1}k_{t+1}/\alpha\right)^{1-\sigma} \left(1 + \beta^{1/\sigma}\left(A\alpha(k_{t+1}/n)^{\alpha-1}\right)^{1/\sigma-1}\right)^\sigma. \tag{75}$$

Eqs. (69) and (75) give the indirect utility function $V_t^j(k_t, \tau_t, \tau_{t+1})$, with

$$\frac{\partial V_t^j}{\partial k_t} = \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial k_t} + \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t},$$
$$\frac{\partial V_t^j}{\partial \tau_t} = \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial \tau_t} + \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial \tau_t},$$
$$\frac{\partial V_t^j}{\partial \tau_{t+1}} = \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial \tau_{t+1}} + \frac{\partial W^j(k_t, \tau_t, \tau_{t+1}, k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial \tau_{t+1}}.$$

The welfare effects can be written as follows:

$$\beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{\partial V_{t-1}^j}{\partial \tau_t}, \tag{76}$$

and

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i}, & \text{if } i \geqslant 1, \\ \frac{\partial V_t^j}{\partial \tau_t} + \frac{\partial V_t^j}{\partial k_t} I_{t,t}, & \text{if } i = 0. \end{cases} \tag{77}$$

The first-order conditions of the Ramsey problem still follow (35).

A.8. Numerical method for the Markovian political equilibrium

A direct application of the projection method for the present problem with heterogeneous agents is to approximate $\mathcal{F}$, $\hat{S}^h$ and $\hat{S}^l$ by three two-dimensional n-order Chebyshev polynomials with tensor products. Consequently, we need to pin down $3 \times n^2$ coefficients of the polynomials that satisfy the Euler equation and the first-order condition (21). That is to say, the computation will be involved in solving $3 \times n^2$ non-linear equations.

However, the analysis in the preceding subsection suggests that computing functions $\hat{S}^j$ is not necessary. In fact, only the derivatives $\hat{S}_i^j$, rather than the function $\hat{S}^j$, are of importance for the equilibrium policy rule $\mathcal{F}$. The following strategy substantially reduces the computational cost: the number of non-linear equations drops from $3 \times n^2$ to n^2. First, we approximate $\mathcal{F}$ by

$$\mathcal{F}(k^h, k^l) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} \theta_{ij}(k^h, k^l), \tag{78}$$

where $\theta_{ij}(k^h, k^l)$ are the tensor products of one-dimensional Chebyshev polynomials. The second step is to pin down the partial derivatives appearing in the first-order condition (21). S_i^j is easy to compute. Plugging $\mathcal{F}_1$, $\mathcal{F}_2$ and S_i^j into the derivatives implied by (73), $\hat{S}_i^j$ can be solved. Finally, choose n points in the state space $[k^{h,\min}, k^{h,\max}]$ and $[k^{l,\min}, k^{l,\max}]$, respectively, by Chebyshev collocation. The first-order condition (21) has to be satisfied for each point. Thus, the functional equation is transformed into n^2 non-linear equations, which solve n^2 unknown coefficients a_{ij} in (78).

Following Judd (1992), the accuracy of the approximation can be indirectly assessed by the Euler equation error. Let $\tilde{\mathcal{F}}$ be the approximated $\mathcal{F}$. The Euler equation error on any given pair (k^h, k^l) is measured by the percentage deviation from τ_t implied by the approximated equilibrium policy rule $\mathcal{F}(k^h, k^l)$ to the "true" optimal τ_t that solves (21) as if $\mathcal{F} = \tilde{\mathcal{F}}$. The accuracy increases with the order of Chebyshev polynomial. However, the improvement tends to be less significant with higher degrees, which increase the computation cost exponentially. In our case, the polynomial of 8-order turns out to be sufficient. We compute the Euler equation errors over 900 points that are uniformly collected in the state space. The maximum errors in all numerical experiments are below 10^{-3}.

A common problem associated with the projection method is that the convergence of the solution for unknown coefficients highly depends on the initial guess. In a standard growth model, a good initial guess can be obtained by linearizing the policy function around the steady state. This problem turns out to be much more serious in the present environment since we essentially have no idea about the steady state. Fortunately, we know the closed-form solution $\mathcal{F}$ under logarithm utility. So, we adopt a simple continuation method – i.e., using the analytical solution $\mathcal{F}$ as an initial guess for $\sigma = 1 + \varepsilon$. Some perturbations on the initial guess are used to check the local convergence of the solution. The equilibrium policy rule $\mathcal{F}$ turns out to be unique in the numerical experiments so far.

References

Alesina, Alberto, Rodrik, Dani, 1994. Distributive politics and economic growth. Quarterly Journal of Economics 109, 465–490.
Auerbach, Alan J., Koltikoff, Laurence J., 1987. Dynamic Fiscal Policy. Cambridge University Press.

Azzimonti Renzo, Marina, forthcoming. Barriers to investment in polarized societies, American Economic Review, 2009.
Boadway, Robin W., Wildasin, David E., 1989. A median voter model of social security. International Economic Review 30, 75–96.
Boldrin, Michele, Rustichini, Aldo, 2000. Political equilibria with social security. Review of Economic Dynamics 3, 41–78.
Breyer, Friedrich, Craig, Ben, 1997. Voting on social security: Evidence from OECD countries. European Journal of Political Economy 13, 705–724.
Chamley, Christophe, 1986. Optimal taxation of capital income in general equilibrium with infinite lives. Econometrica 54, 607–622.
Chen, Kaiji, Song, Zheng, 2005. A Markovian social contract of social security. Mimeo.
Conesa, Juan C., Krueger, Dirk, 1999. Social security reform with heterogeneous agents. Review of Economic Dynamics 2, 757–795.
Cooley, Thomas F., Soares, Jorge, 1999. A positive theory of social security based on reputation. Journal of Political Economy 107, 135–160.
Cremer, Helmuth, De Donder, Philippe, Maldonado, Dario, Pestieau, Pierre, 2006. Voting over type and size of a pension system when some individuals are myopic. Mimeo.
Forni, Lorenzo, 2005. Social security as Markov equilibrium in OLG models. Review of Economic Dynamics 8, 178–194.
Gomme, Paul, Rupert, Peter, 2004. Measuring labor's share of income. Policy discussion papers. Federal Reserve Bank of Cleveland.
Gonzalez-Eiras, Martin, Niepelt, Dirk, 2008. The future of social security. Journal of Monetary Economics 55, 197–218.
Hassler, John, Rodriguez Mora, Jose V., Storesletten, Kjetil, Zilibotti, Fabrizio, 2003. The survival of the welfare state. American Economic Review 93, 87–112.
Hassler, John, Krusell, Per, Storesletten, Kjetil, Zilibotti, Fabrizio, 2005. The dynamics of government. Journal of Monetary Economics 52, 1331–1358.
Imrohoroglu, Ayse, Imrohoroglu, Selahattin, Joines, Douglas H., 1995. A life cycle analysis of social security. Economic Theory 6, 83–114.
Jones, Larry E., Manuelli, Rodolfo E., Rossi, Peter E., 1993. Optimal taxation in models of endogenous growth. Journal of Political Economy 101, 485–517.
Judd, Kenneth L., 1985. Redistributive taxation in a perfect foresight model. Journal of Public Economics 28, 59–84.
Judd, Kenneth L., 1992. Projection method for solving aggregate growth models. Journal of Economic Theory 58, 410–452.
Judd, Kenneth L., 2003. Existence, uniqueness, and computational theory for time consistent equilibria: A hyperbolic discounting example. Mimeo, Hoover Institution.
Krusell, Per, Quadrini, Vincenzo, Rios-Rull, Jose Victor, 1997. Politico-economic equilibrium and economic growth. Journal of Economic Dynamic and Controls 21, 243–272.
Krusell, Per, Rios-Rull, Jose-Victor, 1999. On the size of US government: Political economy in the neoclassical growth model. American Economic Review 89, 1156–1181.
Krusell, Per, Smith, Anthony A., 2003. Consumption-savings decisions with quasi-geometric discounting. Econometrica 71, 365–375.
Lindbeck, Assar, Weibull, Jorgen W., 1987. Balanced-budget redistribution as the outcome of political competition. Public Choice 52, 273–297.
Martin, Patricia P., Weaver, David B., 2005. Social Security: A program and policy history. Social Security Bulletin 66 (1), 1–15.
Mulligan, Casey B., Sala-i-Martin, Xavier, 1999a. Gerontocracy, retirement, and social security. NBER working paper, No. 7117.
Mulligan, Casey B., Sala-i-Martin, Xavier, 1999b. Social security in theory and practice (I): Facts and political theories. NBER working paper, No. 7118.
Persson, Tortsen, Tabellini, Guido, 2000. Political Economics. MIT Press, Cambridge, MA.
Ortigueira, Salvador, 2006. Markov-perfect optimal taxation. Review of Economic Dynamics 9, 153–178.
Razin, Assaf, Sadka, Efraim, Swagel, Phillip, 2002. The aging population and the size of the welfare state. Journal of Political Economy 110, 900–918.
Song, Zheng, Storesletten, Kjetil, Zilibotti, Fabrizio, 2007. Rotten parents and disciplined children: A politico-economic theory of debt and public expenditure. Mimeo.
Storesletten, Kjetil, Telmer, Christopher I., Yaron, Amir, 1999. The risk sharing implications of alternative social security arrangements. Carnegie Rochester Conference Series on Public Policy 50, 213–259.
Storesletten, Kjetil, Telmer, Christopher I., Yaron, Amir, 2004. Consumption and risk sharing over the life cycle. Journal of Monetary Economics 51, 609–633.
Tabellini, Guido, 2000. A positive theory of social security. Scandinavian Journal of Economics 102, 523–545.
Verbon, Harrie A.A., 1987. The rise and evolution of public pension systems. Public Choice 52, 75–100.
Wolff, Edward N., 1987. Estimates of household wealth inequality in the U.S., 1962–1983. Review of Income and Wealth 33, 231–256.