The Dynamics of Inequality and Social Security in General Equilibrium

Song, Zheng

8 April 2008

The Dynamics of Inequality and Social Security in General Equilibrium*

Zheng Song
Fudan University

April 13, 2008

Abstract

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. The dynamic interaction between inequality and social security leads to growing social security programs. We also perform some normative analysis, showing that the politico-economic equilibrium outcomes are fundamentally different from the Ramsey allocation.

JEL Classification: E21 E62 H21 H55

Keywords: inequality, intra-generational redistribution, Markov perfect equilibrium, probabilistic voting, social security

*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 Kaiji Chen, John Hassler, Giovanni Favara, Dirk Niepelt, Jose V. Rodriguez Mora, Kjetil Storesletten and seminar participants at IIES and Oslo for helpful comments. I also thank Christina Lonnblad for editorial assistance. Any remaining errors are mine.

1 Introduction

Most developed countries have large public pension programs, involving not only inter-generational but also 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 a social security 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, it has 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 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 on social security. To this end, we construct a dynamic general equilibrium model where agents repeatedly vote over the social security system. We also analyze normative implications by comparing the political equilibrium with 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, of which the extent of wealth inequality 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 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 private savings of the poor more than those

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

[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.

of the rich. On the other hand, the larger wealth inequality makes transfers more desirable in the future. This provides the political support for an increasing size of social security in the following 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. 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. The model calibrated to the U.S. economy predicts a long-run size of social security of 9.32%, which is roughly in line with the data.[5]

The tractable model allows a comparison between the politico-economic equilibrium outcomes 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 features a decreasing size of social security if the social discount factor is not too small. This sharply contrasts increasing 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

[3] The probabilistic voting framework is adapted from Lindbeck and Weibull (1987). See Hassler *et al.* (2005) and Gonzalez-Eiras and Niepelt (2005) for applications of the repeated probabilistic voting in 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] The size is measured by social security transfers as a percentage of GDP. The average size of social security from 1960 to 1997 is 9.53% in the U.S. Data source: Brady *et al.* (2004).

redistributive reasons.[6] However, since she can commit to future policies, low taxation will be adopted for encouraging capital accumulation in periods other than the initial one. In our calibrated economy, the Ramsey allocation gives a long-run size of social security of 3.16%, much lower than the political equilibrium outcome.

It is worth emphasizing that in Markov equilibria, voters do not only hold rational expectations on future equilibrium outcomes, but may strategically affect future policies via the impact of current policies on private intertemporal choices. Under logarithm utility, the current tax rate has symmetric effects on private savings of the rich and poor. Thus, it cannot affect future states of the economy (wealth distribution), nor 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. To this end, it is useful to compare Markov perfect equilibria with an environment (referred to as "myopic voting equilibria"), where voters can rationally expect future policy outcomes but assume, incorrectly, that there are no strategic interaction between the current and future policies.[7] We show that if the intertemporal elasticity of substitution is smaller than unity, as suggested by many empirical studies, the strategic effect is positive. A higher tax rate today leads to a higher wealth inequality and hence larger transfers tomorrow. Due to the positive strategic effect, current voters have the incentive to strategically raise current social security taxes, in order to obtain larger transfers in the future. The calibrated economy indicates that the strategic effect in Markovian equilibria is quantitatively not important: the relative increase in transfers due to the strategic effect is less than 4%.

The sustainability of the social security system has been widely discussed in the literature.[8] However, the dynamic patterns of social security are much less investigated. Some pioneer 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,

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

[7] A similar notion of pseudo-equilibrium is used by Alesina and Rodrik (1994).

[8] 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), Gonzalez-Eiras and Niepelt (2005).

the intra-generational redistributive effect plays a central role in the evolution of social security in general equilibrium. In particular, the growing sizes of social security can be generated by the interaction between transfers and wealth inequality.

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 implausible except in some small open economies (e.g. Hassler *et al.*, 2003, Azzimonti Renzo, 2005). An exception is Gonzalez-Eiras and Niepelt (2005). They show 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 in their model is a constant, and therefore silent on the dynamics of social security. The present paper generalizes Gonzalez-Eiras and Niepelt's work by incorporating within-cohort heterogeneity, with all results keeping analytical. The generalization gives an equilibrium policy rule which is nontrivially dependent of fundamental elements in the politico-economic environment and hence, provides much richer implications on the dynamics of policies. This also contrasts the literature that resorts to numerical characterizations for nontrivial 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. Section 4 characterizes the Ramsey allocation. In Section 5, we solve numerically the political equilibrium and the Ramsey allocation under a more general CRRA utility form. Section 6 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. Assume the gross population growth rate N^t/N^{t-1} to be a constant n, where N^t denotes the population of the cohort born at time t.

Young households are endowed with high labor productivity γ^h with probability P and with low productivity γ^l ($\gamma^h > \gamma^l$) with probability $1 - P$. For simplicity, let $P = 1/2$.[9] Households with type $j = l$ (h) are referred to as poor (rich). Wage income is taxed at the flat rate, τ_t. The after-tax net earning for young households of type j is $(1 - \tau_t) w_t^j$. Old

[9] P has no effect on the main results below.

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

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

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 $j \in \{l,h\}$, 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 which will be relaxed in Section 5.

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 $K_t = N^{t-1}\left(k_t^l + k_t^h\right)/2$ and $L_t = N^t\left(\gamma^l + \gamma^h\right)/2$. Without loss of generality, the average productivity $\left(\gamma^l + \gamma^h\right)/2$ is normalized to unity so that $\gamma^h = 2 - \gamma^l$ and $L_t = N^t$. 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

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

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

where $k_t \equiv \left(k_t^h + k_t^l\right)/2$ is the average wealth holdings of old households. The individual wage rate is $w_t^j = \gamma^j w_t$. The average wage rate equals w_t.

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 which 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 1% increase in lifetime earnings leads to a 0.90%, 0.32%, 0.15% and 0.00% increase in pension benefits from low to high income groups.[10] Following Conesa and Krueger (1999) and many others, we assume, for analytical convenience, social security benefits to be evenly

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

distributed within old households. It is also assumed that the budget of the social security system must be balanced in each period. This implies that at any time t, social security payments $b_t N^{t-1}$ equal social security contributions $\tau_t \left(w_t^l + w_t^h\right) N^t/2$:

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

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 doublet of private saving functions

$$k_{t+1}^h = S^h\left(k_t, \tau_t, \tau_{t+1}\right) \equiv \omega\left(\tau_{t+1}\right)\psi\left(\tau_{t+1}\right)A\left(1-\tau_t\right)\left(k_t/n\right)^\alpha, \tag{7}$$

$$k_{t+1}^l = S^l\left(k_t, \tau_t, \tau_{t+1}\right) \equiv \psi\left(\tau_{t+1}\right)A\left(1-\tau_t\right)\left(k_t/n\right)^\alpha, \tag{8}$$

where $\psi(\cdot)$ and $\omega(\cdot)$ are defined as:

$$\omega\left(\tau_{t+1}\right) \equiv \frac{\theta\alpha\left(1+\beta\right)+\left(\theta-1\right)\left(1-\alpha\right)\tau_{t+1}/2}{\alpha\left(1+\beta\right)-\left(\theta-1\right)\left(1-\alpha\right)\tau_{t+1}/2}, \tag{9}$$

$$\psi\left(\tau_{t+1}\right) \equiv \frac{\gamma^l\left(1-\alpha\right)\beta\left(\alpha\left(1+\beta\right)-\left(\theta-1\right)\left(1-\alpha\right)\tau_{t+1}/2\right)}{\left(1+\beta\right)\left(\alpha\left(1+\beta\right)+\left(1-\alpha\right)\tau_{t+1}\right)}, \tag{10}$$

where $\theta \equiv \gamma^h/\gamma^l$ denotes the ratio of labor productivity of the rich to that of the poor. It is easy to show 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. Social security benefits increase the income after retirement and hence, discourage private savings.

Note that $\theta = w_t^h/w_t^l$ and $\omega\left(\tau_t\right) = k_t^h/k_t^l$ measure young households' income inequality and old households' wealth inequality (excluding social security benefits), respectively.[11] Without social security system ($\tau_t = 0$), wealth inequality $\omega(0)$ coincides with income equality θ. The establishment of a social security system can affect future wealth inequality k_{t+1}^h/k_{t+1}^l via τ_t and τ_{t+1}. First, under logarithm utility, (7) and (8) imply that τ_t has a symmetric impact on k_{t+1}^h and k_{t+1}^l and thus, does not affect k_{t+1}^h/k_{t+1}^l. Second, since $\omega\left(\tau_{t+1}\right)$ increases in τ_{t+1}, a high future social security tax rate τ_{t+1} enlarges future wealth inequality. The intuition is the following. The poor receive the same amount of social security benefits as the rich after

[11] To avoid confusion, wealth inequality is hereinafter referred to as inequality in terms of old households' wealth, excluding social security benefits.

retirement, while their earnings are smaller than those of the rich. Therefore, high social security benefits discourage savings of the poor more than those of the rich. The results are written in Lemma 1.

Lemma 1 *Assume that* $u(c) = \log(c)$*.*

(i) Future wealth inequality k_{t+1}^h / k_{t+1}^l *is increasing in the future social security tax rate* τ_{t+1}*.*

(ii) Given τ_{t+1}*,* k_{t+1}^h / k_{t+1}^l *does not depend on the current social security tax rate* τ_t *and the aggregate capital* k_t*.*

The second part of Lemma 1 states an important property that will be frequently used in the following analysis: the choice of the current tax rate has no effect on future wealth inequality. This property is due to the assumption of logarithm utility, which implies the cancellation of a substitution and an income effect and thus makes private savings proportional to labor income. As will be seen below, Lemma 1 substantially simplifies the analysis throughout the paper.

(7) and (8) lead to the law of motion of aggregate capital

$$k_{t+1} = S(k_t, \tau_t, \tau_{t+1}) \equiv \phi(\tau_{t+1}) A (1 - \tau_t) (k_t / n)^{\alpha}, \tag{11}$$

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

$$\phi(\tau_{t+1}) \equiv \frac{\alpha\beta(1-\alpha)}{\alpha(1+\beta) + (1-\alpha)\tau_{t+1}}. \tag{12}$$

It immediately follows that $S_1 > 0$, $S_2 < 0$ and $S_3 < 0$. These aggregate results come from $S_1^j > 0$, $S_2^j < 0$ and $S_3^j < 0$ implied by private saving functions (7) and (8).

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 only care 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.[12] In the

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

context of our model, the political decision process of τ_t can be formalized as

$$\max_{\tau_t} \sum_{j=h,l} U_t^{o,j} + n \sum_{j=h,l} U_t^{y,j}, \tag{13}$$

where $U_t^{i,j}$ denotes the welfare of the households of type (i, j), $i \in \{y, o\}$ and $j \in \{l, h\}$, with $U_t^{o,j} \equiv u\left(c_t^{o,j}\right)$ and $U_t^{y,j} \equiv u\left(c_t^{y,j}\right) + \beta u\left(c_{t+1}^{o,j}\right)$. For notational convenience, the weights on different groups' utility are set equal.[13]

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^h and k_t^l. Hence, the Markovian policy rule of τ_t can be written as

$$\tau_t = \digamma\left(k_t^h, k_t^l\right), \tag{14}$$

where $\digamma : R \times R \rightarrow [0, 1]$ is assumed to be continuous and differentiable for technical convenience.[14] 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 (14) for τ_{t+1} in (7) and (8) and solve a recursive form of private saving functions, which can be written as

$$k_{t+1}^j = \hat{S}^j\left(k_t, \tau_t\right). \tag{15}$$

The expression of $\hat{S}^j$ is not available, unless we know the explicit form of $\digamma$. However, some properties of $\hat{S}^j$ can be obtained. Differentiating (7) and (8) with respect to τ_t gives

$$\begin{bmatrix} \hat{S}_2^h \\ \hat{S}_2^l \end{bmatrix} = \begin{bmatrix} S_2^h \\ S_2^l \end{bmatrix} + \begin{bmatrix} S_3^h \digamma_1 & S_3^h \digamma_2 \\ S_3^l \digamma_1 & S_3^l \digamma_2 \end{bmatrix} \begin{bmatrix} \hat{S}_2^h \\ \hat{S}_2^l \end{bmatrix},$$

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

$$\hat{S}_2^h = \frac{S_2^h + \digamma_2\left(S_2^l S_3^h - S_2^h S_3^l\right)}{1 - \digamma_1 S_3^h - \digamma_2 S_3^l}, \tag{16}$$

$$\hat{S}_2^l = \frac{S_2^l + \digamma_1\left(S_2^h S_3^l - S_2^l S_3^h\right)}{1 - \digamma_1 S_3^h - \digamma_2 S_3^l}. \tag{17}$$

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

$$k_{t+1} = \hat{S}\left(k_t, \tau_t\right) \equiv \frac{\hat{S}^h\left(k_t, \tau_t\right) + \hat{S}^l\left(k_t, \tau_t\right)}{2}, \tag{18}$$

[13] Deviation from equal weights does not affect the main results below.

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

with $\hat{S}_2 = \left(\hat{S}_2^h (k_t, \tau_t) + \hat{S}_2^l (k_t, \tau_t) \right) /2$.

Given any $\digamma$, the political decision on τ_t solves (13), subject to budget constraints (2) and (3), factor prices (4) and (5), the balanced-budget rule (6), private saving functions (15), the law of motion of aggregate capital (18), and the non-negative constraint of τ_t.[15] This yields an actual policy rule $\tau_t = \bar{\digamma} \left(k_t^h, k_t^l \right)$, with $\bar{\digamma} : R \times R \to [0,1]$. $\digamma$ is said to be a Markovian equilibrium policy rule, if and only if $\bar{\digamma} = \digamma$. The formal definition of a Markov perfect equilibrium is given as follows.

Definition 1 *A Markov perfect political equilibrium is a triplet of functions $\tilde{S}^h$, $\tilde{S}^l$ and $\digamma$, where private saving functions $\tilde{S}^j : R \times R \to R$, $j \in \{h, l\}$, and the policy rule $\digamma : R \times R \to [0,1]$ are such that:*

(1) Given the policy rule $\digamma$, $\tilde{S}^j \left(k_t^h, k_t^l \right) = \hat{S}^j \left(k_t^j, \digamma \left(k_t^h, k_t^l \right) \right)$, where $\hat{S}^j$ is the recursive private saving function (15).

(2) Given $\digamma$ and $\hat{S}^j$, $\bar{\digamma}$ solves (13), subject to (2) to (6), (15), (18) and the non-negative constraint of τ_t.

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

To solve the equilibrium policy rule $\digamma$, 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

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = u' \left(c_t^{o,j} \right) n w_t > 0. \tag{19}$$

Needless to say, old households always benefit from social security transfers. Substituting for $c_t^{o,j}$ and w_t, (19) can be rewritten as

$$\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{1-\alpha}{2\alpha \frac{k_t^j / k_t^i}{1 + k_t^j / k_t^i} + (1-\alpha) \tau_t}, \tag{20}$$

where $i, j \in \{h, l\}$, $i \neq j$. $\partial U_t^{o,j} / \partial \tau_t$ depends on wealth distribution. This highlights the role of social security as an intra-generational redistributive policy. Specifically, the smaller is the old household share of total wealth, the more welfare gains can they get from transfers. Although the rich gain less, the aggregate welfare effect of τ_t on old households, $\partial U_t^o / \partial \tau_t = \sum_{j=h,l} \left(\partial U_t^{o,j} / \partial \tau_t \right) /2$, increases in wealth inequality due to the concavity of utility.[16]

[15] The constraint that $\tau_t \leq 1$ is never binding since otherwise it delivers zero consumption to young households.
[16] This can be formally derived by showing that $\left(\partial U_t^o / \partial \tau_t \right) / \partial \left(k_t^h / k_t^l \right) > 0$ for $k_t^h / k_t^l > 1$.

Differentiating the utility of young households with respect to τ_t yields

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = -u'\left(c_t^{y,j}\right)\gamma^j w_t + \beta u'\left(c_{t+1}^{o,j}\right)\left(k_{t+1}^j \frac{\partial R_{t+1}}{\partial k_{t+1}} + n\tau_{t+1}\frac{\partial w_{t+1}}{\partial k_{t+1}}\right)\hat{S}_2 + \beta u'\left(c_{t+1}^{o,j}\right) n w_{t+1}\frac{\partial \tau_{t+1}}{\partial \tau_t}. \tag{21}$$

Note that the effect of τ_t via k_{t+1}^j cancels out due to the Euler equation. The first term in (21) 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 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.[17] Hence, the general equilibrium effect can benefit young households.[18] 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

$$\frac{\partial \tau_{t+1}}{\partial \tau_t} = \digamma_1\left(k_{t+1}^h, k_{t+1}^l\right)\hat{S}_2^h\left(k_t, \tau_t\right) + \digamma_2\left(k_{t+1}^h, k_{t+1}^l\right)\hat{S}_2^l\left(k_t, \tau_t\right). \tag{22}$$

If $\partial\tau_{t+1}/\partial\tau_t > 0$ (< 0), young households know that a higher current social security tax rate leads to more (less) social security benefits in the future. Thus, they may strategically increase (reduce) τ_t as compared to the case where the current political choice does not affect future policy outcomes.[19]

Then, the first-order condition of (13) can be written as

$$\sum_{j=h,l}\frac{\partial U_t^{o,j}}{\partial \tau_t} + n\sum_{j=h,l}\frac{\partial U_t^{y,j}}{\partial \tau_t} + \lambda_t = 0, \tag{23}$$

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 $\tau_t = 0$. (23) implies a functional equation for $\digamma$. Under logarithm utility, the fixed-point can be analytically obtained 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.

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

[17] This can be seen by the fact that $\mathrm{sgn}\left(k_{t+1}^j \partial R_{t+1}/\partial k_{t+1} + n\tau_{t+1}\partial w_{t+1}/\partial k_{t+1}\right) = \mathrm{sgn}\left(-k_{t+1}^j/k_{t+1} + \tau_{t+1}\right)$.

[18] Gonzalez-Eiras and Niepelt (2004) show that the interest rate effect plays an important role in sustaining the social security system in an economy without within-cohort heterogeneity.

[19] In Section 5, we will study "the myopic voting equilibrium", where voters can rationally expect future policy outcomes but assume there to be no strategic interaction between the current and future policies.

(i) The policy rule $F\left(k_t^h, k_t^l\right)$ *follows*

$$F\left(k_t^h, k_t^l\right) = \begin{cases} H\left(k_t^h/k_t^l\right) > 0 & \text{if } \upsilon\alpha < 1, \text{ or if } \upsilon\alpha \geq 1 \text{ and } k_t^h/k_t^l > \Theta(\upsilon) \\ 0 & \text{otherwise} \end{cases}, \tag{24}$$

where

$$H\left(k_t^h/k_t^l\right) \equiv \frac{-\Phi(\upsilon) + \sqrt{\Phi(\upsilon)^2 + 4\Delta(\upsilon)\left(\alpha - \frac{4\upsilon\alpha^2 k_t^h/k_t^l}{\left(1+k_t^h/k_t^l\right)^2}\right)}}{2\Delta(\upsilon)} \tag{25}$$

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

(ii) Private saving functions follow

$$\tilde{S}^h\left(k_t^h, k_t^l\right) = \omega(\hat{\tau})\,\psi(\hat{\tau})\,A\left(1 - F\left(k_t^h, k_t^l\right)\right)\left(\left(k_t^h + k_t^l\right)/(2n)\right)^{\alpha}, \tag{26}$$

$$\tilde{S}^l\left(k_t^h, k_t^l\right) = \psi(\hat{\tau})\,A\left(1 - F\left(k_t^h, k_t^l\right)\right)\left(\left(k_t^h + k_t^l\right)/(2n)\right)^{\alpha}, \tag{27}$$

where $\hat{\tau}$ *is a constant solving*

$$\hat{\tau} = \max\{0, H(\omega(\hat{\tau}))\}. \tag{28}$$

Proof: See the appendix.

Four remarks about this proposition are in order. First, the political decision on the social security tax rate solely depends on wealth inequality. Moreover, it is easily seen that τ_t increases in k_t^h/k_t^l. That is to say, the larger the wealth inequality, the more political support the social security program receives. Social security as an inter-generational redistribution policy has been widely studied in the literature. The within-cohort redistributive components of such a system are often neglected, however. In the context of the present model, the aggregate welfare effect of τ_t on old households is increasing in wealth inequality. This reveals the underlying mechanism for the positive relationship between τ_t and k_t^h/k_t^l. Although $\partial U_t^{o,j}/\partial\tau_t$ is different among old households, τ_t delivers the same welfare effect on young households with different labor productivity. In the appendix, we show that

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = -\frac{1+\beta\alpha}{1-\tau_t}, \tag{29}$$

i.e., social security has no intra-generational redistributive effect on young households. This is primarily due to the symmetric effect of τ_t on private savings k_{t+1}^j, as discussed in the preceding section. In addition, due to the logarithmic specification, aggregate capital k_t is additively separable in the utility. Hence, τ_t is independent of k_t. Although these properties does not carry over into the case with a more general utility form, numerical experiments in

Section 5 shows that our main results are robust: wealth inequality plays an essential role in the decision of τ_t and τ_t is increasing in k_t^h/k_t^l.

Second, the conditions in Proposition 1 characterize the politico-economic environment where the social security system can be sustained in the Markov equilibrium. For $\upsilon\alpha < 1$ to hold, a small n or α is needed. A small n implies a large share of old in the population and hence, a large number of agents benefiting from the pension system. A low α implies that the interest rate R_{t+1} is rather elastic to aggregate capital k_{t+1}. This amplifies the general equilibrium effect and hence mitigates the negative welfare effect of τ_t on young households, which can directly be seen from (29). So a small n and α reinforce the political constituency of the system. When $\upsilon\alpha \geq 1$, the intra-generational redistribution becomes the key. There would be no social security system in an economy without within-cohort heterogeneity. However, social security can be sustained as long as there exists a sufficiently high level of inequality within cohorts. Therefore, when $\upsilon\alpha \geq 1$, the political support largely comes from intra-generational redistribution.

Third, (26) and (27) imply that τ_t does not affect future wealth inequality in the Markov equilibrium, exactly the same as in the competitive equilibrium. Since the social security tax rate is determined by wealth inequality as shown in (24), this property implies that the strategic effect under logarithm utility is mute, i.e., $\partial\tau_{t+1}/\partial\tau_t = 0$.[20] That is to say, although current voters can in principle influence future political outcomes by affecting future wealth inequality, they are actually unable to do this. The lack of any strategic effect is due to the fact that future wealth inequality is independent of the current social security tax rate, as stated in Lemma 1. This independence breaks down the dynamic link between τ_t and τ_{t+1} in the Markov equilibrium. The strategic effect arises under a more general utility case, where the choice of τ_t may affect future wealth inequality and thus, future policy outcomes. However, as will be seen in Section 5, the strategic effect turns out to be quantitatively unimportant.

Finally, $\hat{\tau}$ which satisfies (28) is the rational expectation of future tax rates. Given any expectation of $\hat{\tau}$, agents make intertemporal choices so that the future wealth inequality will be equal to $\omega(\hat{\tau})$. For the expectation to be self-fulfilled, $\hat{\tau}$ must equal that implied by the policy rule, i.e., $\hat{\tau} = H(\omega(\hat{\tau}))$ for positive $\hat{\tau}$. Due to the rather complicated expression of $H(\omega(\hat{\tau}))$, we are unable to characterize analytically the solution of (28). Extensive numerical experiments show that the self-fulfilled expectation $\hat{\tau}$ is unique. Note that the formation of the rational expectation on future tax rates holds for any time other than the initial period.

[20] (24) implies that $F_1\left(k_t^h, k_t^l\right)/F_2\left(k_t^h, k_t^l\right) = -k_{t+1}^l/k_{t+1}^h$. (26) and (27) give $\hat{S}_2^h(k_t, \tau_t)/\hat{S}_2^l(k_t, \tau_t) = k_{t+1}^h/k_{t+1}^l$. Plugging these two results into (22) establishes that $\partial\tau_{t+1}/\partial\tau_t = 0$.

Hence, all future tax rates are a constant and independent of the initial wealth inequality and transfers. It is worthy emphasizing that future tax rates do follow the equilibrium policy rule $\digamma$. The constant tax rates are due to the fact that wealth inequality becomes a constant $\omega(\hat{\tau})$ after the initial period.

Now we can characterize the dynamics of wealth inequality and social security. Suppose that voting for social security is unanticipatedly launched at time 1. So (9) implies that the initial wealth inequality k_1^h/k_1^l equals income inequality θ, which gives τ_1 by the policy rule $\digamma$. In periods after the initial one, wealth inequality and tax rates are equal to $\omega(\hat{\tau})$ and $\hat{\tau}$, respectively, as shown above. Therefore, k_t^h/k_t^l and τ_t converge to the steady state in two periods. Moreover, since $\omega(\hat{\tau}) \geq \omega(0) = \theta$, the expected transfers increase future wealth inequality. This leads to a growing size of social security.[21] To conclude, we have

Proposition 2 *Assume that $u(c) = \log(c)$. In the Markov perfect equilibrium,*

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

3.1 A Quantitative Exercise

Although the two-period OG model is very stylized, we would like to see quantitatively the size of social security in a calibrated economy, and then assess the importance of the dynamic

[21] Formally, the tax rate τ_t at any time $t > 1$ is equal to $H(\omega(\hat{\tau}))$, which is greater than the initial tax rate $\tau_1 = H(\omega(0))$.

interaction between k_t^h/k_t^l and τ_t. The parameter values are set as follows. $\alpha = 0.36$, as widely adopted in the literature of macroeconomics (e.g. Prescott, 1986). The ratio of income of the rich to income of the poor, θ, is set to 3 by the U.S. data.[22] Each period in the OG model is assumed to contain 30 years. Then, the gross growth rate of the U.S. population between 1970 and 2000 gives $n = 1.384$ (Gonzalez-Eiras and Niepelt, 2005). The discount factor is calibrated such that the capital output ratio equals 2.5, which gives $\beta = 0.965^{30}$. A is chosen such that the steady state k without social security is equal to unity. Finally, we measure the size of social security by social security transfers as a percentage of GDP, which is equal to $(1-\alpha)\,\tau_t$.

Suppose that there is no social security at time 0 and voting for the welfare state is unanticipatedly launched at time 1. Let $k_0 = 1$, i.e., the economy is in the steady state before the establishment of social security.[23] Table 1 provides details on the dynamics of wealth inequality, social security tax rates, and the consumption of different groups of households. Wealth inequality k_t^h/k_t^l increases by 33% (from 3 to 3.96) in the first period and remains at that level afterwards.[24] Policy rule F gives that $\tau_1 = 11.96\%$ and $\tau_t = 14.57\%$ for $t > 1$. The implied size of social security is thus equal to 7.65% in the first period and 9.32% afterwards, respectively. The numbers are roughly in line with the average size of social security of 9.53% in the U.S. between 1960 and 1997.[25]

[Insert Table 1]

Given the simplicity of the model, our calibrated political economy performs reasonably well in terms of matching the size of social security in the U.S. Moreover, the quantitative exercise suggests that the impact of the endogenous change in k_t^h/k_t^l on τ_t is sizable; the social security tax rate increases by nearly 20% between time 1 and 2. The nontrivial dynamic interaction between wealth inequality and social security can thus help explain growing sizes of the welfare state in OECD countries.[26]

It is worthy emphasizing that though social security enlarges wealth inequality (excluding social security transfers), it does narrow the within-cohort consumption inequality. Without

[22] We use quintile income shares in Deininger and Squire (1996). Specifically, we calculate the income share of the rich (poor) by summing incomes shares of the top (bottom) two quintiles and half of income share of the middle quintile. The income ratio is rather stable, ranged between 2.8 and 3.2 from 1948 to 1991.

[23] By Proposition 1, k_0 has no effect on k_t^h/k_t^l and τ_t.

[24] This is close to the result in a recent quantitative study by Fuster *et al.* (2003), which shows that introducing social security increases the Gini coefficient of asset distribution by 27% (from 0.51 to 0.65).

[25] Note that social security transfers data also include benefits for sickness and family allowance (see Brady *et al.*, 2004). So the average size of pension benefits in the U.S. should be lower than 9.53%.

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

social security system, consumption inequality is equal to income inequality, i.e., $c_t^{o,h}/c_t^{o,l} = c_t^{y,h}/c_t^{y,l} = 3\ \forall t \leq 0$. After the establishment of social security, there is a significant decline in the inequality of consumption, with $c_t^{y,r}/c_t^{y,p} = c_t^{o,r}/c_t^{o,p} = 2.80$ for $t > 1$. This reflects the role of social security as an intra-generational redistributive policy.

4 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

$$\max_{\{\tau_t\}_{t=1}^{\infty}} \beta \sum_{j=h,l} U_1^{o,j} + \sum_{t=1}^{\infty} \rho^t \left(\sum_{j=h,l} U_t^{y,j} \right), \tag{30}$$

subject to individuals' budget constraints (2) and (3), factor prices (4) and (5), the balanced-budget rule (6), private saving functions (7) and (8), the law of motion of aggregate capital (11) and the non-negative constraint of τ_t. $\rho \in (0, 1)$ is the intergenerational discount factor. Compared with the political decision problem (13), the Ramsey allocation problem (30) 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.[27]

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 $i \geq 1$, as implied by the law of motion of capital (11):

$$I_{t,t+i} = \begin{cases} \frac{\partial k_{t+i}}{\partial k_{t+i-1}} \frac{\partial k_{t+i-1}}{\partial k_{t+i-2}} \cdots \left(\frac{\partial k_{t+1}}{\partial \tau_t} + \frac{\partial k_{t+1}}{\partial k_t} \frac{\partial k_t}{\partial \tau_t} \right) < 0 & \text{for } t > 1 \\ \frac{\partial k_i}{\partial k_{i-1}} \frac{\partial k_{i-1}}{\partial k_{i-2}} \cdots \frac{\partial k_2}{\partial \tau_1} < 0 & \text{for } t = 1 \end{cases}. \tag{31}$$

The second line of (31) 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

$$I_{t,t} = \begin{cases} \frac{\partial k_t}{\partial \tau_t} < 0 & \text{for } t > 1 \\ 0 & \text{for } t = 1 \end{cases}. \tag{32}$$

[27] Gonzalez-Eiras and Niepelt (2004) shows 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% tax rate and zero capital stock. Therefore, the social planner approach cannot be adopted here.

$I_{1,1} = 0$ since k_1 is predetermined. Note that τ_t directly influences the welfare of agents 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 agents born at time t and afterwards via its impact on capital accumulation $I_{t,t+i}$. τ_t has no effect on agents 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,

$$\frac{\partial U_{t-1}^{y,j}}{\partial \tau_t} = \beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \beta \left(u'\left(c_t^{o,j}\right) n w_t + u'\left(c_t^{o,j}\right) \left(k_t^j \frac{\partial R_t}{\partial k_t} + n \tau_t \frac{\partial w_t}{\partial k_t} \right) I_{t,t} \right), \tag{33}$$

where $I_{t,t}$ follows (32). The first term on the RHS of (33) 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 $I_{t,t}$. Compare (33) with (19), 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 the preceding section, 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. The positive overall general equilibrium effect implies 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. An exception is that for $t = 1$, the welfare effect of τ_1 on old households equals that in (19), since the capital in the initial period is predetermined ($I_{1,1} = 0$). More specifically, we have

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*

$$\beta \frac{\partial U_t^{o,j}}{\partial \tau_t} = \begin{cases} \frac{(1+\beta)(\phi(\tau_t)+\tau_t\phi'(\tau_t))}{\alpha\gamma^j+\tau_t\phi(\tau_t)} - \frac{\beta(1-\alpha)\phi'(\tau_t)}{\phi(\tau_t)} > 0 & \text{for } t > 1 \\ \beta \frac{1-\alpha}{2\alpha\frac{k_1^j/k_1^i}{1+k_1^j/k_1^i}+(1-\alpha)\tau_1} > 0 & \text{for } t = 1 \end{cases}, \tag{34}$$

where $\phi(\cdot)$ is defined by (12).

The first line of (34) is derived in the appendix and the second line simply follows (20). 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 labor income a household has, the less can she 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 $i \geq 0$, equals

$$\begin{aligned} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} &= -u'\left(c_{t+i}^{y,j}\right)\gamma^j w_{t+i} + u'\left(c_{t+i}^{y,j}\right)\gamma^j \frac{\partial w_{t+i}}{\partial k_{t+i}}\left(1-\tau_{t+i}\right) I_{t,t+i} \\ &\quad + \beta u'\left(c_{t+i+1}^{o,j}\right)\left(k_{t+i+1}^j \frac{\partial R_{t+i+1}}{\partial k_{t+i+1}} + n\tau_{t+i+1}\frac{\partial w_{t+i+1}}{\partial k_{t+i+1}}\right) I_{t,t+i+1}. \end{aligned} \tag{35}$$

As in (21), the first term in (35) 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 $i \geq 1$, 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 (35) and (21) 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. First, the planner takes into account the negative impact of τ_t on k_{t+i}, which reduces the social security tax base at time $t+i$. This effect is captured by the second term on the RHS of (35). In the political equilibrium, voters at time t takes 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 $i \geq 0$, is equal to*

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} + \frac{(1+\beta\alpha)\alpha^{i+1}\phi'(\tau_t)}{\phi(\tau_t)} < 0 & \text{if } t > 1 \\ -\frac{(1+\beta\alpha)\alpha^i}{1-\tau_t} < 0 & \text{if } t = 1 \end{cases}, \tag{36}$$

where $\phi(\cdot)$ is defined by (12).

The proof is given in the appendix. 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 only depends 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 (30) with respect to τ_t can be written as:

$$\beta \sum_{j=h,l} \frac{\partial U_t^{o,j}}{\partial \tau_t} + \sum_{i=0}^{\infty} \left(\rho^{i+1} \sum_{j=h,l} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} \right) + \lambda_t = 0, \tag{37}$$

where λ_t is the multiplier on the non-negative constraint of τ_t. Let us first solve τ_1. Plugging (34) and (36) into (37), we have

Proposition 3 *Assume that $u(c) = \log(c)$. In the Ramsey solution,*

(i) The initial social security tax rate

$$\tau_1 = \begin{cases} H\left(k_1^h/k_1^l\right) > 0 & \text{if } \upsilon\alpha < 1, \text{ or if } \upsilon\alpha \geq 1 \text{ and } k_1^h/k_1^l > \Theta(\upsilon) \\ 0 & \text{otherwise} \end{cases}, \tag{38}$$

where $H(\cdot)$ follows (25) 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 $\rho \lesseqgtr \beta n/(1+\alpha\beta n)$, 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.

Proof is given in the appendix. The first part of Proposition 3 states that the initial tax rate τ_1 is determined by the initial wealth inequality k_1^h/k_1^l, which parallels Proposition 1 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. There are two effects which 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 (37) as

$$\beta \sum_{j=h,l} \frac{\partial U_1^{o,j}}{\partial \tau_1} + \beta n \sum_{j=h,l} \frac{\partial U_1^{y,j}}{\partial \tau_1} + \sum_{i=1}^{\infty} \left(\rho^{i+1} \sum_{j=h,l} \frac{\partial U_{i+1}^{y,j}}{\partial \tau_1} \right) - (\beta n - \rho) \sum_{j=h,l} \frac{\partial U_1^{y,j}}{\partial \tau_1} + \lambda_1 = 0. \tag{39}$$

The first two terms on the LHS of (39) capture the same trade-off in the political decision process (see 23).[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 (39) illustrates the discrepancy between the weight on the current young in the political decision process and

[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 Lemma 2 and 3).

that in the Ramsey problem. If $\beta n > \rho$, this 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 second part of Proposition 3 shows that the second effect dominates the first effect for sufficiently small social discount factors, i.e., $\rho < \beta n / (1 + \alpha\beta n)$.

Now we proceed to τ_t for $t > 1$. Plugging (34) and (36) into (37), one can find that τ_t is a constant over time. In the appendix, we prove the following proposition.

Proposition 4 *Assume that $u(c) = \log(c)$. In the Ramsey solution,*

(i) The social security tax rate converges to a unique steady state in two periods.

(ii) If $\rho > \frac{\beta(1-\alpha)}{1+\alpha\beta}$, wealth inequality and the social security tax rate in any subsequent period are lower than those in the initial one.

(iii) Define

$$\Omega \equiv \frac{(1-\alpha)\beta(1+\theta)^2}{2\alpha\theta} - \frac{2\rho(1+\beta\alpha)(2+\beta-\alpha)}{(1-\alpha\rho)(1+\beta)} + \frac{2\beta(1-\alpha)^2}{\alpha(1+\beta)}. \tag{40}$$

If $\Omega > 0$, the steady state tax rate $\bar{\tau} > 0$. If $\Omega \leq 0$, $\bar{\tau} = 0$.

The first part of the proposition parallels Proposition 2 in the political equilibrium, except that the uniqueness of the steady state can be established analytically in the Ramsey solution. The common feature is that the effect of the initial wealth inequality on social security tax rates only lasts one period. This should not be surprising. As shown in Lemma 1, future wealth inequality solely depends on the future tax rate.

The second part of Proposition 4 gives a sufficient condition for decreasing sizes of social security over time. The condition holds for a wide range of parameter values.[30] This sharply contrasts the prediction of the political equilibrium. The somewhat surprising result primarily comes from the fact that low tax rates substantially encourage capital accumulation. Moreover, 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.

The third part of the proposition gives the condition that the social security system can be sustained or not in the Ramsey allocation. It is immediate that Ω increases in θ but decreases in ρ. Intuitively, a high income inequality θ increases the within-cohort redistributive benefit of transfers. A high ρ, on the contrary, increases the relative weight on the welfare of future

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

[30] In our calibrated economy with $\alpha = 0.36$ and $\beta = 0.965^{30}$, the condition implies that $\rho > 0.947^{30}$.

generations, and thus makes social security as an inter-generational redistributive policy less desirable. Figure 1 plots the threshold condition of θ implied by (40) under $\alpha = 0.36$ and $\beta = 0.965$[30]. $\Omega > 0$ is satisfied for the region above the line in the figure. It can directly be seen that a high ρ requires a high θ to sustain social security in the Ramsey allocation.

[Insert Figure 1]

4.1 A Quantitative Exercise

Now we compare quantitatively the Ramsey allocation with the political equilibrium in the same calibrated economy as in Section 3.1. We assume that $\rho = \beta n$, i.e., the planner weighs generations by their sizes and discounts their welfare by households' discount factor. Results are presented in Table 2. The parameter values give $\rho = 0.48 > \frac{\beta n}{1+\alpha\beta n} = 0.41$. By the second part of Proposition 3, we know that the initial Ramsey social security tax rate is lower than that in the political equilibrium. Specifically, $\tau_1^R = 6.32\% < \tau_1^M = 11.96\%$. The political decision process ignores the welfare of future generations and thus leaves too larger transfers to the old in the initial period.[31].

[Insert Table 2]

In terms of the long-run tax rate, it is straightforward that $\rho > \frac{\beta(1-\alpha)}{1+\alpha\beta} = 0.20$. Thus, the second part of Proposition 4 implies a decreasing size of social security in the Ramsey solution. The long-run Ramsey tax rate is 4.94%, much lower than the political equilibrium tax rate of 14.57%. The move from the political equilibrium to the Ramsey solution increases the steady state consumption of the young rich and poor by 23% and 17%, respectively. The decline in the steady state consumption of the old rich and poor is relatively modest, amounting to 1% and 5%, respectively.

To conclude, the political equilibrium on social security transfers is fundamentally different from the Ramsey allocation. The political decision process delivers too large transfers in the long run. Moreover, social security is converging to the steady state along an increasing path in the political equilibrium. In contrast, the Ramsey allocation implies a decreasing size of social security. These results hold true under a wide range of parameter values.

5 The Strategic Effect under CRRA Utility

So far, we have focused on logarithm utility, which substantially simplifies the analysis and makes explicit solutions available. However, many empirical studies suggest the elasticity of

[31]Since $\rho = \beta n$, the fourth term in (39) goes away.

intertemporal substitution to be less than unity. It is an open question to what extent our results would be affected by the deviation from logarithm utility. Particularly, the strategic effect $\partial\tau_{t+1}/\partial\tau_t$ in the Markovian political equilibrium may arise under a less restrictive utility form. This section adopts a more general CRRA utility function, to see whether analytical results in the preceding sections are robust with the presence of strategic effect. Assume that

$$u(c) = \frac{c^{1-\sigma} - 1}{1 - \sigma} \tag{41}$$

where $\sigma > 0$ is the inverse of the intertemporal elasticity of substitution. Households' intertemporal choices, the political decision as well as the Ramsey problem on social security tax rates are characterized in the appendix. 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 simple basis function such as polynomial. This method is applied for time-consistent problems in some recent research (Judd, 2003 and Ortigueira, 2004). As for the Ramsey solution, we transform the infinite-horizon problem into a finite-horizon problem by the truncated method (e.g. Jones, Manuelli and Rossi, 1993). The corresponding algorithms are provided in the appendix.

Let us first look at the Markovian political equilibrium. The equilibrium policy outcomes under $\sigma = 2$ are plotted in Figure 2.[32] It can directly be seen that τ_t is increasing in wealth inequality k_t^h/k_t^l. Different from the logarithm case, now the aggregate capital can affect social security tax rates; τ_t turns out to be decreasing in k_t. The solid line in Figure 3 plots the evolution of the social security tax rate in the Markovian political equilibrium.[33] As in the log case, sizes of social security converge to the steady state along an increasing path. The main difference is that convergence is asymptotic, instead of in two periods under logarithm utility. However, the quantitative implication on the speed of convergence is robust; over 92% of the gap between τ_1 and τ_∞ will be closed in one period.

[Insert Figure 2]

[Insert Figure 3]

The strategic effect (22) arises for $\sigma \neq 1$. To identify the strategic effect, it is useful to study the myopic voting equilibrium where voters at time t have rational expectation on future

[32] The other parameters are set to the values in Section 3.1.

[33] As in Section 3, we assume that $k_0 = 1$ and voting for social security is unanticipatedly launched at time 1.

tax rates but (incorrectly) disregard the strategic effect of the current political decision τ_t on future taxes. The formal definition is given as follows.

Definition 2 *A myopic voting political equilibrium is a doublet of private saving functions $S^j : R \times [0,1] \times [0,1] \to R$, $j \in \{h, l\}$, and a sequence of social security tax rates $\{\tau_t\}_{t=1}^{\infty}$ such that*

(1) S^j is solved by the Euler equation.

(2) $\{\tau_t\}_{t=1}^{\infty}$ is solved by (13), subject to budget constraints (2), (3), private saving functions S^j and the non-negative constraint of τ_t.

Recall that in the log case, the strategic effect is mute and thus the myopic equilibrium and the Markov equilibrium give the same outcomes. The dotted line in Figure 3 plots social security tax rates in the myopic voting equilibrium, which turn out to be slightly lower than those in the Markov perfect equilibrium. The discrepancy has its roots in the strategic effect. Under logarithm utility, private savings of the poor and those of the rich decrease in τ_t by the same proportions. Thus, the current social security tax rate does not affect future wealth inequality. For $\sigma > 1$, τ_t affects asymmetrically the incentive for income-smoothing between the rich and poor. The asymmetric effects enlarge future wealth inequality k_{t+1}^h / k_{t+1}^l and thus raise 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. The strategic effect is quantitatively not important, however. The relative increase in the social security tax rate due to the strategic effect is less than 4%.

Finally, we turn to the Ramsey solution. The dashed line in Figure 3 plots the Ramsey tax rates over time. As in the logarithm case, the Ramsey tax rates converge to the steady state along a decreasing path and the long-run Ramsey tax rate is substantially lower than that in the Markov political equilibrium. The underlying mechanism is essentially the same: taxing more the initial inelastic tax base and lowering taxes afterwards to encourage capital accumulation.

6 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 interaction between social security and wealth inequality leads to growing sizes of social security, which may shed light on the increasing generosity of social security in OECD countries during the post-war period (Breyer and Craig, 1997).

We compare the political equilibrium with the Ramsey allocation. It turns out that they are fundamentally different. In particular, social security is converging to the steady state along an increasing path in the politico-economic equilibrium, while the Ramsey allocation implies a decreasing size of social security in the calibrated economy. Moreover, the political decision process induces too large social security transfers in the long run, since non-altruistic voters ignore the negative effect of taxation on the welfare of future generations via capital accumulation.

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, Storesletten and Zilibotti (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 is interacted with social security.

References

[1] Alesina, Alberto and Dani Rodrik, 1994, Distributive Politics and Economic Growth, Quarterly Journal of Economics, 109, 465-490.

[2] Auerbach, Alan J. and Laurence J. Koltikoff, 1987, Dynamic Fiscal Policy, Cambridge University Press.

[3] Azzimonti Renzo, Marina, 2005, On the Dynamic Inefficiency of Governments, mimeo.

[4] Boadway, Robin W. and David E. Wildasin, 1989, A Median Voter Model of Social Security, International Economic Review, 30, 75-96.

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

[5] Boldrin, Michele and Aldo Rustichini, 2000, Political Equilibria with Social Security, Review of Economic Dynamics, 3, 41-78.

[6] Brady, David, Jason Beckfield and John Stephens, 2004, Comparative Welfare States Data Set, mimeo.

[7] Breyer, Friedrich and Ben Craig, 1997, Voting on Social Security: Evidence from OECD countries, European Journal of Political Economy, 13, 705-724.

[8] Chamley, Christophe, 1986, Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives, Econometrica, 54, 607-22.

[9] Chen, Kaiji and Zheng Song, 2005, A Markovian Social Contract of Social Security, mimeo.

[10] Conesa, Juan C. and Dirk Krueger, 1999, Social Security Reform with Heterogeneous Agents, Review of Economic Dynamics, 2, 757-795.

[11] Cooley, Thomas F. and Jorge Soares, 1999, A Positive Theory of Social Security Based on Reputation, Journal of Political Economy, 107, 135-160.

[12] Cremer, Helmuth, Philippe De Donder, Dario Maldonado and Pierre Pestieau, 2006, Voting over Type and Size of a Pension System when Som Individuals are Myopic, mimeo.

[13] De Nardi, Mariacristina, 2004, Wealth Inequality and Intergenerational Links, Review of Economic Studies, 71, 743-768.

[14] Deininger, Klaus and Lyn Squire, 1996, A New Data Set Measuring Income Inequality, World Bank Economic Review, 10, 565-591.

[15] Domeij, David and Paul Klein, 2002, Public Pension: To What Extent Do They Account for Swedish Wealth Distribution, Review of Economic Dynamics, 5, 503-534.

[16] Feldstein, Martin and Liebman, Jeffrey B., 2001, Social Security, NBER Working Paper, No. 8451.

[17] Forni, Lorenzo, 2005, Social Security as Markov Equilibrium in OLG Models, Review of Economic Dynamics, 8, 178-194.

[18] Fuster, Luisa, Ayse Imrohoroglu and Selahattin Imrohoroglu, 2003, A Welfare Analysis of Social Security in a Dynastic Framework, International Economic Review, 44, 1247-1274.

[19] Gonzalez-Eiras, Martin and Dirk Niepelt, 2005, Sustaining Social Security, mimeo.

[20] Hassler, John, Jose V. Rodriguez Mora, Kjetil Storesletten and Fabrizio Zilibotti, 2003, The Survival of the Welfare State, American Economic Review, 93, 87-112.

[21] Hassler, John, Kjetil Storesletten and Fabrizio Zilibotti, 2007, Democratic Public Good Provision, Journal of Economic Theory, 133, 127-151.

[22] Hassler, John, Per Krusell, Kjetil Storesletten and Fabrizio Zilibotti, 2005, The Dynamics of Government, Journal of Monetary Economics, 52, 1331-1358.

[23] Imrohoroglu, Ayse, Imrohoroglu, Selahattin and Joines, Douglas H, 1995, A Life Cycle Analysis of Social Security, Economic Theory, 6, 83-114.

[24] Jones, Larry E., Rodolfo E. Manuelli and Peter E. Rossi, 1993, Optimal Taxation in Models of Endogenous Growth, Journal of Political Economy, 101, 485-517.

[25] Judd, Kenneth L., 1985, Redistributive Taxation in a Perfect Foresight Model, Journal of Public Economics, 28, 59-84.

[26] Judd, Kenneth L., 1992, Projection Method for Solving Aggregate Growth Models, Journal of Economic Theory, 58, 410-452.

[27] Judd, Kenneth L., 2003, Existence, Uniqueness, and Computational Theory for Time Consistent Equilibria: A Hyperbolic Discounting Example, mimeo, Hoover Institution.

[28] Krusell, Per, Vincenzo Quadrini and Jose Victor Rios-Rull, 1997, Politico-Economic Equilibrium and Economic Growth, Journal of Economic Dynamic and Controls, 21, 243-272.

[29] Krusell, Per and Jose-Victor Rios-Rull, 1999, On the Size of US Government: Political Economy in the Neoclassical Growth Model, American Economic Review, 89, 1156-1181.

[30] Krusell, Per, and Anthony A. Smith, 2003, Consumption-Savings Decisions with Quasi-Geometric Discounting, Econometrica, 71, 365–375.

[31] Lindbeck, Assar and Jorgen W. Weibull, 1987, Balanced-Budget Redistribution as the Outcome of Political Competition, Public Choice, 52, 273-297.

[32] Mulligan, Casey B. and Xavier Sala-i-Martin, 1999a, Gerontocracy, Retirement, and Social Security, NBER Working Paper, No. 7117.

[33] Mulligan, Casey B. and Xavier Sala-i-Martin, 1999b, Social Security in Theory and practice(I): Facts and Political Theories, NBER Working Paper, No. 7118.

[34] Persson, Tortsen and Guido Tabellini, 2000, Political Economics, MIT Press, Cambridge, MA.

[35] Prescott, Edward C., 1986, Theory ahead of Business Cycle Measurement, Federal Reserve Bank of Mineapolis Quarter Review, 10, 9-22.

[36] Ortigueira, Salvador, 2004, Markovian Optimal Taxation: The Macroeconomic Implications of Two Notions of Time Consistency, mimeo, European University Institute.

[37] Razin, Assaf, Efraim Sadka and Phillip Swagel, 2002, The Aging Population and the Size of the Welfare State, Journal of Political Economy, 110, 900-918.

[38] Song, Zheng, Kjetil Storesletten and Fabrizio Zilibotti, 2007, Rotten Parents and Disciplined Children: A Politico-Economic Theory of Debt and Public Expenditure, mimeo.

[39] Storesletten, Kjetil, Christopher I. Telmer and Amir Yaron, 1999, The Risk Sharing Implications of Alternative Social Security Arrangements, Carnegie Rochester Conference Series on Public Policy, 50, 213-259.

[40] Storesletten, Kjetil, Christopher I. Telmer and Amir Yaron, 2004, Consumption and Risk Sharing over the Life Cycle, Journal of Monetary Economics, 51, 609-633.

[41] Tabellini, Guido, 2000, A Positive Theory of Social Security, Scandinavian Journal of Economics, 102, 523-545.

[42] Verbon, Harrie A. A., 1987, The Rise and Evolution of Public Pension Systems, Public Choice, 52, 75-100.

7 Appendix

7.1 Proof of Proposition 1

To solve 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 only live one period.

First 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

$$\frac{\partial U_T^{y,j}}{\partial \tau_T} = -\frac{\gamma^j w_T}{c_T^{y,j}} = -\frac{1}{1-\tau_T}. \tag{42}$$

The welfare effect of τ_T on old households follows (20). Plugging (42) and (20) into the first-order condition and assuming interior solution,

$$\sum_{j=h,l} \frac{1-\alpha}{2\alpha \frac{k_T^j/k_T^i}{1+k_T^j/k_T^i} + (1-\alpha)\tau_T} - 2n\frac{1}{1-\tau_T} = 0. \tag{43}$$

Note that the second order condition always holds. (43) gives a quadratic equation of τ_T

$$\Delta(\upsilon_T)\tau_T^2 + \Phi(\upsilon_T)\tau_T + \frac{4\upsilon_T\alpha^2 k_T^h/k_T^l}{\left(1+k_T^h/k_T^l\right)^2} - \alpha = 0, \tag{44}$$

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

Now we consider the non-negative constraint. First consider the case where $\Phi(\upsilon_T) \geq 0$. Since $\Delta(\upsilon_T) > 0$, there is a unique positive τ_T if and only if

$$\left(k_T^h/k_T^l\right)^2 + (2-4\upsilon_T\alpha)\left(k_T^h/k_T^l\right) + 1 > 0. \tag{45}$$

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

$$k_T^h/k_T^l > 2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} \text{ or } k_t^h/k_t^l < 2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)}. \tag{46}$$

The first inequality in (46) is binding since $2\upsilon_T\alpha - 1 + 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} > 1$ for $\upsilon_T\alpha \geq 1$. The other inequality in (46) cannot be satisfied since $2\upsilon_T\alpha - 1 - 2\sqrt{\upsilon_T\alpha(\upsilon_T\alpha-1)} < 1$ for $\upsilon_T\alpha \geq 1$.

Second consider the case where $\Phi(\upsilon_T) < 0$. For $\upsilon_T\alpha < 1$, (45) ensures a unique positive τ_T. For $\upsilon_T\alpha \geq 1$, there can be two positive roots if the LHS of (45) is non-positive. This implies $n < 1/(2\alpha) - 1$ and contradicts the condition that $n \geq 1/\alpha - 1$, as given by $\upsilon_T\alpha \geq 1$.

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

$$\tau_T = F^T\left(k_T^h, k_T^l\right) = \frac{-\Phi(\upsilon_T) + \sqrt{\Phi(\upsilon_T)^2 + 4\Delta(\upsilon_T)\left(\alpha - \frac{4\upsilon_T\alpha^2 k_T^h/k_T^l}{\left(1+k_T^h/k_T^l\right)^2}\right)}}{2\Delta(\upsilon_T)} > 0. \tag{47}$$

For $\upsilon_T\alpha \geq 1$, $F_T\left(k_T^h, k_T^l\right)$ follows (47) if k_T^h/k_T^l satisfies the first inequality in (46) and is equal to zero otherwise.

Next we consider period $T-1$. The policy rule $F^T\left(k_T^h, k_T^l\right)$ at time T implies $F_1^T/F_2^T = -k_T^l/k_T^h$. Some algebra manipulations establish

$$\frac{\partial k_T^h}{\partial \tau_{T-1}} = \hat{S}_2^h(k_{T-1}, \tau_{T-1}) = -\omega(\tau_T)\psi(\tau_T)A(k_{T-1}/n)^\alpha, \tag{48}$$

$$\frac{\partial k_T^l}{\partial \tau_{T-1}} = \hat{S}_2^l(k_{T-1}, \tau_{T-1}) = -\psi(\tau_T)A(k_{T-1}/n)^\alpha, \tag{49}$$

where $\omega(\cdot)$ and $\psi(\cdot)$ are defined by (9) and (10), respectively. This gives $\hat{S}_2^j(k_{T-1}, \tau_{T-1}) = S_2^j(k_{T-1}, \tau_{T-1}, \tau_T)$. Moreover, substituting (48) and (49) for $\hat{S}_2^j$ in (22) leads to

$$\frac{\partial \tau_T}{\partial \tau_{T-1}} = 0. \tag{50}$$

The intuition is the following. According to Lemma 1, for any given τ_T, τ_{T-1} does not affect the future wealth inequality k_T^h/k_T^l. Therefore, given the policy rule F^T as a function of k_T^h/k_T^l, τ_{T-1} has no impact on future policy outcome τ_T. The dynamic link between τ_{T-1} and τ_T breaks down and the strategic effect does not exist.

The absence of the strategic effect makes the rest of the derivation fairly straightforward. The welfare effect of τ_{T-1} on young households follows (21). Using (50) and the indirect utility approach, which will be specified in the next subsection, we find that

$$\frac{\partial U_{T-1}^{y,j}}{\partial \tau_{T-1}} = -\frac{1+\beta\alpha}{1-\tau_{T-1}}. \tag{51}$$

The welfare effect of τ_{T-1} on old households still follows (20). Plugging (42) and (20) into the first-order condition and assuming interior solution, we have

$$\sum_{j=h,l} \frac{1-\alpha}{2\alpha\frac{k_{T-1}^j/k_{T-1}^i}{1+k_{T-1}^j/k_{T-1}^i} + (1-\alpha)\tau_{T-1}} - 2n\frac{1+\beta\alpha}{1-\tau_{T-1}} = 0, \tag{52}$$

which gives a quadratic equation of τ_{T-1}

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

where $\upsilon_{T-1} \equiv n(1+\alpha\beta)/(1-\alpha)$. 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

$$\tau_{T-1} = F^{T-1}\left(k_{T-1}^h, k_{T-1}^l\right) = \frac{-\Phi(\upsilon_{T-1}) + \sqrt{\Phi(\upsilon_{T-1})^2 + 4\Delta(\upsilon_{T-1})\left(\alpha - \frac{4\upsilon_{T-1}\alpha^2 k_{T-1}^h/k_{T-1}^l}{\left(1+k_{T-1}^h/k_{T-1}^l\right)^2}\right)}}{2\Delta(\upsilon_{T-1})} > 0. \tag{54}$$

For $\upsilon_{T-1}\alpha \geq 1$, τ_{T-1} follows (54) if k_{T-1}^h/k_{T-1}^l satisfies

$$k_{T-1}^h/k_{T-1}^l > 2\upsilon_{T-1}\alpha - 1 + 2\sqrt{\upsilon_{T-1}\alpha(\upsilon_{T-1}\alpha - 1)} \tag{55}$$

and τ_{T-1} is equal to zero otherwise.

It immediately follows that the only difference in F^{T-1} and F^T lies in $\upsilon_{T-1} = n(1+\alpha\beta)/(1-\alpha)$ and $\upsilon_T = n/(1-\alpha)$. Young households born at time $T-1$ live for two periods and thus

$\partial U_{T-1}^{y,j}/\partial \tau_{T-1}$ in (51) differs from $\partial U_T^{y,j}/\partial \tau_T$ in (42). 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 the independence of $\partial U_t^{y,j}/\partial \tau_t$ on the future tax rate and the mute strategic effect, as shown in (51) and (50), respectively. These two features transform the dynamic problem into a static one. Consequently, the key parameter υ_t is exactly the same as υ_{T-1} for $t < T-1$. The finite-horizon equilibria thus converge to the infinite-horizon equilibrium in two periods.

Finally, we need to solve the private saving function $\tilde{S}^j$. (24) implies that $\digamma_1\left(k_t^h, k_t^l\right)/\digamma_2\left(k_t^h, k_t^l\right) = -k_{t+1}^l/k_{t+1}^h$. (16) and (17) can thus be rewritten as

$$\hat{S}_2^h = S_2^h \frac{1 + \digamma_2\left(\left(S_2^l/S_2^h\right) S_3^h - S_3^l\right)}{1 + \digamma_2\left(\left(S^l/S^h\right) S_3^h - S_3^l\right)}, \tag{56}$$

$$\hat{S}_2^l = S_2^l \frac{1 + \digamma_1\left(\left(S_2^h/S_2^l\right) S_3^l - S_2^l S_3^h\right)}{1 + \digamma_1\left(\left(S^h/S^l\right) S_3^h - S_3^l\right)}. \tag{57}$$

Since $S_2^l/S_2^h = S^l/S^h$, (56) and (57) give $\hat{S}_2^j = S_2^j$. The same argument establishes that $\hat{S}_1^j = S_1^j$, which implies that $\hat{S}^h$ and $\hat{S}^l$ follow (26) and (27), respectively, with a constant $\hat{\tau}$ to be determined. Since future wealth inequality equals $\omega(\hat{\tau})$, the equilibrium policy rule (24) implies that $\hat{\tau}$ solves (28). Substituting $\digamma$ for τ_t in $\hat{S}^j$ establishes $\tilde{S}^j$. $\square$

7.2 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 functions (7) and (8) and the law of motion of aggregate capital (11), 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}:

$$\begin{aligned} V_t^j\left(k_t, \tau_t, \tau_{t+1}\right) &= (1+\beta\alpha)\,\alpha \log k_t + (1+\beta\alpha)\log\left(1-\tau_t\right) \\ &\quad + (1+\beta)\log\left(\alpha\gamma^j + \tau_{t+1}\phi\left(\tau_{t+1}\right)\right) - \beta(1-\alpha)\log\phi\left(\tau_{t+1}\right), \end{aligned} \tag{58}$$

where $\phi(\cdot)$ is defined by (12). The indirect utility of old households at time 1 is

$$U_1^{o,j} = \log\left(2\alpha\frac{k_1^j/k_1^i}{1+k_1^j/k_1^i} + (1-\alpha)\,\tau_1\right) + \alpha\log k_1. \tag{59}$$

Differentiating (58), the welfare effect of τ_t on old households at time t equals

$$\beta\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{\partial V_{t-1}^j}{\partial \tau_t} = (1+\beta)\frac{\phi\left(\tau_t\right) + \tau_t\phi'\left(\tau_t\right)}{\alpha\gamma^j + \tau_t\phi\left(\tau_t\right)} - \beta(1-\alpha)\frac{\phi'\left(\tau_t\right)}{\phi\left(\tau_t\right)}, \tag{60}$$

for $t > 1$. Differentiating (59) with respect to τ_1 yields the second line of (34). This proves the lemma. $\square$

7.3 Proof of Lemma 3

By (11), we know $\partial k_{t+i}/\partial k_{t+i-1} = \alpha k_{t+i}/k_{t+i-1}$, $\partial k_{t+1}/\partial \tau_t = -k_{t+1}/\left(1-\tau_t\right)$ and $\partial k_t/\partial \tau_t = \phi'\left(\tau_t\right) k_t/\phi\left(\tau_t\right)$. Thus, $I_{t,t+i}$ can be written as

$$I_{t,t+i} = \begin{cases} \alpha^{i-1} k_{t+i}\left(-\frac{1}{1-\tau_t} + \alpha\frac{\phi'(\tau_t)}{\phi(\tau_t)}\right) & \text{for } t > 1 \\ -\alpha^{i-1}\frac{k_i}{1-\tau_1} & \text{for } t = 1 \end{cases}. \tag{61}$$

According to the indirect utility function (58), the welfare effect of τ_t on young households at time t, for $t > 1$, equals

$$\frac{\partial U_t^{y,j}}{\partial \tau_t} = \frac{\partial V_t^j}{\partial \tau_t} + \frac{\partial V_t^j}{\partial k_t}\frac{\partial k_t}{\partial \tau_t} = (1+\beta\alpha)\left(-\frac{1}{1-\tau_t} + \alpha\frac{\phi'\left(\tau_t\right)}{\phi\left(\tau_t\right)}\right). \tag{62}$$

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

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} = (1+\beta\alpha)\,\alpha^i\left(-\frac{1}{1-\tau_t} + \alpha\frac{\phi'\left(\tau_t\right)}{\phi\left(\tau_t\right)}\right) \tag{63}$$

for $i = 1, 2, \cdots$. The second equality in (63) comes from the first line in (61). (62) and (63) give the first line of (36). Finally, for $t = 1$, we have

$$\frac{\partial U_1^{y,j}}{\partial \tau_1} = \frac{\partial V_1^j}{\partial \tau_1} = -\frac{1+\beta\alpha}{1-\tau_1} \tag{64}$$

and

$$\frac{\partial U_{i+1}^{y,j}}{\partial \tau_1} = \frac{\partial V_{i+1}^j}{\partial k_{i+1}} I_{1,i+1} = -\alpha^i\frac{1+\beta\alpha}{1-\tau_1} \tag{65}$$

for $i = 1, 2, \cdots$. The second equality in (65) comes from the second line in (61). (64) and (65) give the second line of (36). $\square$

7.4 Proof of Proposition 3

The first-order condition of (30) with respect to τ_1 is

$$\beta\sum_{j=h,l}\frac{\partial U_1^{o,j}}{\partial \tau_1} + \sum_{i=0}^{\infty}\rho^{i+1}\left(\sum_{j=h,l}\frac{\partial U_{i+1}^{y,j}}{\partial \tau_1}\right) = 0, \tag{66}$$

where $\partial U_i^{y,j}/\partial \tau_1$ follows from (64) and (65). This leads to

$$\beta\sum_{j=h,l}\frac{1-\alpha}{2\alpha\frac{k_1^j/k_1^i}{1+k_1^j/k_1^i} + (1-\alpha)\,\tau_1} - 2\rho\frac{1+\beta\alpha}{1-\rho\alpha}\frac{1}{1-\tau_1} = 0, \tag{67}$$

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 (59), (64) and (65), the second order condition is always satisfied.

Comparing (67) with (52), it is immediate that the closed-form solution of τ_1 follows (38) with $\upsilon \equiv \rho\left(1+\alpha\beta\right)/\left(\beta\left(1-\rho\alpha\right)\left(1-\alpha\right)\right)$. A comparison of the two first-order conditions shows that $\tau_1^R \gtreqless \tau_1^M$ if and only if $\rho \lesseqgtr \beta n/\left(1+\alpha\beta n\right)$. $\square$

7.5 Proof of Proposition 4

The first-order conditions of (30) with respect to τ_t for $t > 1$ are

$$\beta \sum_{j=h,l} \frac{\partial U_t^{o,j}}{\partial \tau_t} + \sum_{i=0}^{\infty} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} \right) = 0. \tag{68}$$

Substituting (60), (62) and (63) for $\partial U_t^{o,j}/\partial \tau_t$ and $\partial U_{t+i}^{y,j}/\partial \tau_t$, respectively, (68) leads to

$$\sum_{j=h,l} \left(\frac{(1+\beta)\left(\phi(\tau_t) + \tau_t \phi'(\tau_t)\right)}{\alpha \gamma^j + \tau_t \phi(\tau_t)} \right) - 2\rho \frac{1+\beta\alpha}{1-\rho\alpha} \frac{1}{1-\tau_t} + 2\left(\rho \frac{(1+\beta\alpha)\alpha}{1-\rho\alpha} - \beta(1-\alpha) \right) \frac{\phi'(\tau_t)}{\phi(\tau_t)} = 0. \tag{69}$$

(69) solves a constant τ_t for $t > 1$. Hence, the Ramsey tax rates converge to the steady state in two periods.

Note that the second order conditions are always satisfied. To see this, (60) shows that $\partial^2 U_t^{o,j}/\partial \tau_t^2 < 0$. Differentiating (62) and (63) with respect to τ_t establishes

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

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 (69) is unique.

Now turn to the second part of the proposition. Compare (69) and (67), we only need to show that for any τ,

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

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

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

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

$$(1+\beta)\left(\phi(\tau) + \tau\phi'(\tau)\right) \leq \beta(1-\alpha).$$

Some algebra establishes that the inequality always holds.

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

Finally we prove the third part of the proposition. Denote $L(\tau_t)$ the LHS of (69). After some algebra manipulations, $L(\tau_t)$ can be written as

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

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 (40).

7.6 CRRA Utility

Given (41), households' problem (1) becomes

$$\max_{k_{t+1}^j}\frac{\left(c_t^{y,j}\right)^{1-\sigma}-1}{1-\sigma}+\beta\frac{\left(c_{t+1}^{o,j}\right)^{1-\sigma}-1}{1-\sigma}, \tag{70}$$

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_{t+1}^j follows

$$\begin{aligned}
k_{t+1}^j &= G^j(k_t,\tau_t,\tau_{t+1},k_{t+1}) && (71)\\
&\equiv \frac{\gamma^j\left(A\alpha(k_{t+1}/n)^{\alpha-1}\beta\right)^{1/\sigma}A(1-\alpha)(1-\tau_t)(k_t/n)^\alpha-A(1-\alpha)\tau_{t+1}k_{t+1}^\alpha n^{1-\alpha}}{\left(A\alpha(k_{t+1}/n)^{\alpha-1}\beta\right)^{1/\sigma}+A\alpha(k_{t+1}/n)^{\alpha-1}}.
\end{aligned}$$

By $k_{t+1}=\sum_{j=h,l}k_{t+1}^j/2$, (71) solves private saving functions

$$k_{t+1}^j=S^j(k_t,\tau_t,\tau_{t+1}), \tag{72}$$

with

$$S_i^j=\frac{G_i^j+\left(G_i^lG_4^h-G_i^hG_4^l\right)/2}{1-\sum_{j=h,l}G_4^j/2}, \tag{73}$$

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

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

with

$$S_i=\frac{\sum_{j=h,l}S_i^j}{2}. \tag{75}$$

Given the Markovian policy rule (14), a recursive form of private and aggregate saving functions can be solved.

$$k_{t+1}^{j} = \hat{S}^{j}(k_t, \tau_t), \tag{76}$$

$$k_{t+1} = \hat{S}(k_t, \tau_t), \tag{77}$$

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 (13) still follow (19), (21) and (23), respectively.

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

$$W^{j}(k_t, \tau_t, \tau_{t+1}, k_{t+1}) \equiv \left(\gamma^j A(1-\tau_t)(k_t/n)^{\alpha} + \tau_{t+1}k_{t+1}/\alpha\right)^{1-\sigma}\left(1+\beta^{1/\sigma}\left(A\alpha(k_{t+1}/n)^{\alpha-1}\right)^{1/\sigma-1}\right)^{\sigma}. \tag{78}$$

(72) and (78) give the indirect utility function $V_t^j(k_t, \tau_t, \tau_{t+1})$, with

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

The welfare effects can be written as follows.

$$\beta\frac{\partial U_t^{o,j}}{\partial \tau_t} = \frac{\partial V_{t-1}^j}{\partial \tau_t}, \tag{79}$$

and

$$\frac{\partial U_{t+i}^{y,j}}{\partial \tau_t} = \begin{cases} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} & \text{if } i \geq 1 \\ \frac{\partial V_t^j}{\partial \tau_t} + \frac{\partial V_t^j}{\partial k_t} I_{t,t} & \text{if } i = 0 \end{cases}. \tag{80}$$

The first-order conditions of the Ramsey problem still follow (37).

7.7 Numerical Method for the Markovian Political Equilibrium

A direct application of the projection method for the present problem with heterogeneous agents is to approximate 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 (23). That is to say, the computation will be involved in solving $3 \times n^2$ nonlinear 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 $\digamma$. The following strategy substantially reduces the computational cost: the number of nonlinear equations drops from $3 \times n^2$ to n^2. First, we approximate $\digamma$ by

$$\digamma\left(k^h, k^l\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\phi_{ij}\left(k^h, k^l\right), \tag{81}$$

where $\phi_{ij}\left(k^h, k^l\right)$ 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 (23). S_i^j is easy to compute. Plugging $\digamma_1$, $\digamma_2$ and S_i^j into (**??**) and (**??**), $\hat{S}_i^j$ can be solved. Finally, choose n points in the state space $\left[k^{h,\min}, k^{h,\max}\right]$ and $\left[k^{l,\min}, k^{l,\max}\right]$, respectively, by Chebyshev collocation. The first-order condition (23) has to be satisfied for each point. Thus, the functional equation is transformed into n^2 nonlinear equations, which solve n^2 unknown coefficients a_{ij} in (81).

Following Judd (1992), the accuracy of the approximation can be indirectly assessed by the Euler equation error. Let $\tilde{\digamma}$ be the approximated $\digamma$. The Euler equation error on any given pair $\left(k^h, k^l\right)$ is measured by the percentage deviation from τ_t implied by the approximated equilibrium policy rule $\digamma\left(k^h, k^l\right)$ to the "true" optimal τ_t that solves (23) as if $\digamma = \tilde{\digamma}$. 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. The Euler equation errors over 900 points that are uniformly collected in the state space are computed. 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 $\digamma$ under logarithm utility. So we adopt a simple continuation method, i.e., using the analytical solution $\digamma$ 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 $\digamma$ turns out to be unique in the numerical experiments so far.

7.8 Numerical Method for the Ramsey Solution

Given the indirect utility V_t^j, the Ramsey problem (30) can be rewritten as

$$\max_{\{\tau_t,k_{t+1}\}_{t=1}^{\infty}} \beta \sum_{j=h,l} U_1^{o,j}\left(k_1^j, k_1, \tau_1\right) + \sum_{t=1}^{\infty} \rho^t \left(\sum_{j=h,l} V_t^j\left(k_t, \tau_t, \tau_{t+1}\right) \right), \tag{82}$$

subject to the law of motion of aggregate capital (74). The first-order conditions with respect to τ_t and k_t for $t > 1$ are

$$\sum_{j=h,l} \frac{\partial V_{t-1}^j}{\partial \tau_t} + \rho \sum_{j=h,l} \frac{\partial V_t^j}{\partial \tau_t} = \mu_{t-1} \frac{\partial k_t}{\partial \tau_t} + \rho \mu_t \frac{\partial k_{t+1}}{\partial \tau_t}, \tag{83}$$

$$\rho \sum_{j=h,l} \frac{\partial V_t^j}{\partial k_t} = -\mu_{t-1} + \rho \mu_t \frac{\partial k_{t+1}}{\partial k_t}, \tag{84}$$

where μ_t is the Lagrangian multiplier. Let $\bar{x}$ be the steady state value of variable x. Denote $\bar{V}_1^j$, $\bar{V}_2^j$ and $\bar{V}_3^j$ as the steady states of $\partial V_t^j / \partial k_t$, $\partial V_t^j / \partial \tau_t$ and $\partial V_t^j / \partial \tau_{t+1}$, respectively. Similarly, $\bar{S}_1$, $\bar{S}_2$ and $\bar{S}_3$ are referred to as the steady states of $\partial S_t / \partial k_t$, $\partial S_t / \partial \tau_t$ and $\partial S_t / \partial \tau_{t+1}$, respectively. Then (84) leads to

$$\bar{\mu} = -\frac{\rho \sum_{j=h,l} \bar{V}_1^j}{1 - \rho \bar{S}_1}. \tag{85}$$

Using (85), (83) implies

$$\sum_{j=h,l} \bar{V}_3^j + \rho \sum_{j=h,l} \bar{V}_2^j + \frac{\rho\left(\rho \bar{S}_2 + \bar{S}_3\right)}{1 - \rho \bar{S}_1} \sum_{j=h,l} \bar{V}_1^j = 0. \tag{86}$$

Moreover, (74) gives

$$\bar{k} = S\left(\bar{k}, \bar{\tau}, \bar{\tau}\right). \tag{87}$$

(86) and (87) thus solve the steady state capital stock $\bar{k}$ and the steady state social security tax rate $\bar{\tau}$.

Following Jones, Manuelli and Rossi (1993), we adopt the truncated method to solve the dynamics of the Ramsey allocation. Assume that the economy reaches the steady state after period T. Then the infinite-horizon problem (82) can be approximated by a finite-horizon one

$$\max_{\{\tau_t,k_{t+1}\}_{t=1}^{T-1}} \beta \sum_{j=h,l} U_1^{o,j}\left(k_1^j, k_1, \tau_1\right) + \sum_{t=1}^{T-1} \rho^t \left(\sum_{j=h,l} V_t^j\left(k_t, \tau_t, \tau_{t+1}\right) \right) + \Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right), \tag{88}$$

subject to the law of motion of aggregate capital (74). The value of continuation $\Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right)$ is equal to

$$\Gamma\left(k_T, \bar{\tau}, \bar{\tau}\right) = \sum_{t=T}^{\infty} \rho^t \left(\sum_{j=h,l} V_t^j\left(k_T, \bar{\tau}, \bar{\tau}\right) \right), \tag{89}$$

which corrects the error caused by "end effects". Therefore, standard nonlinear programming techniques can be applied to solve (88). For interior solutions, $\{\tau_t\}_{t=1}^{T-1}$ may be directly solved by the first-order conditions. Specifically, the effect of τ_t on $\Gamma(k_T, \bar{\tau}, \bar{\tau})$ is

$$\frac{\partial \Gamma(k_T, \bar{\tau}, \bar{\tau})}{\partial \tau_t} = \sum_{i=T-t}^{\infty} \rho^{t+i} \left(\sum_{j=h,l} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} \right) = \frac{\rho^T \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1} I_{t,T}. \tag{90}$$

Using (80) and (90), the first-order conditions of (82) with respect to τ_t for $t > 1$ can be written as

$$\rho^{t-1} \sum_{j=h,l} \frac{\partial V_{t-1}^j}{\partial \tau_t} + \rho^t \sum_{j=h,l} \frac{\partial V_t^j}{\partial \tau_t} + \sum_{i=0}^{T-t-1} \rho^{t+i} \left(\sum_{j=h,l} \frac{\partial V_{t+i}^j}{\partial k_{t+i}} I_{t,t+i} \right) + \frac{\rho^T \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1} I_{t,T} = 0. \tag{91}$$

Similarly, we have the first-order condition of (82) with respect to τ_1

$$\beta \sum_{j=h,l} \frac{\partial U_1^{o,j}}{\partial \tau_1} + \rho \sum_{j=h,l} \frac{\partial V_1^j}{\partial \tau_1} + \sum_{i=0}^{T-2} \rho^{i+1} \left(\sum_{j=h,l} \frac{\partial V_{1+i}^j}{\partial k_{1+i}} I_{1,1+i} \right) + \frac{\rho^T \sum_{j=h,l} \bar{V}_1^j}{1-\rho \bar{S}_1} I_{1,T} = 0. \tag{92}$$

(91) and (92) constitute a nonlinear equation system which solves $\{\tau_t\}_{t=1}^{T-1}$.

Table 1: Politico-Economic Equilibrium under Logarithm Utility

	t=0	t=1	t=2	t=∞
k^h/k^l	3.0000	3.0000	3.9558	3.9558
τ	0.0000	0.1196	0.1457	0.1457
$c^{o,h}$	4.5683	5.2157	5.0656	4.6834
$c^{o,l}$	1.5227	2.1703	1.8091	1.6726
$c^{y,h}$	4.3679	3.9878	3.4686	3.2069
$c^{y,l}$	1.4560	1.4242	1.2387	1.1453

Table 2: Ramsey Solution under Logarithm Utility

	t=0	t=1	t=2	t=∞
k^h/k^l	3.0000	3.0000	3.2800	3.2800
	(3.0000)	(3.0000)	(3.9558)	(3.9558)
τ	0.0000	0.4193	0.0494	0.0494
	(0.0000)	(0.1196)	(0.1457)	(0.1457)
$c^{o,h}$	4.5683	6.8384	3.9667	4.6285
	(4.5683)	(5.2157)	(5.0656)	(4.6834)
$c^{o,l}$	1.5227	3.7930	1.3589	1.5856
	(1.5227)	(2.1703)	(1.8091)	(1.6726)
$c^{y,h}$	4.3679	2.5655	3.3741	3.9361
	(4.3679)	(3.9878)	(3.4686)	(3.2069)
$c^{y,l}$	1.4560	0.8746	1.1494	1.3403
	(1.4560)	(1.4242)	(1.2387)	(1.1453)

Note: The politico-economic equilibrium outcomes are in parentheses.

Figure 1: Regions for Positive and Zero Long-Run Ramsey Social Security Taxes

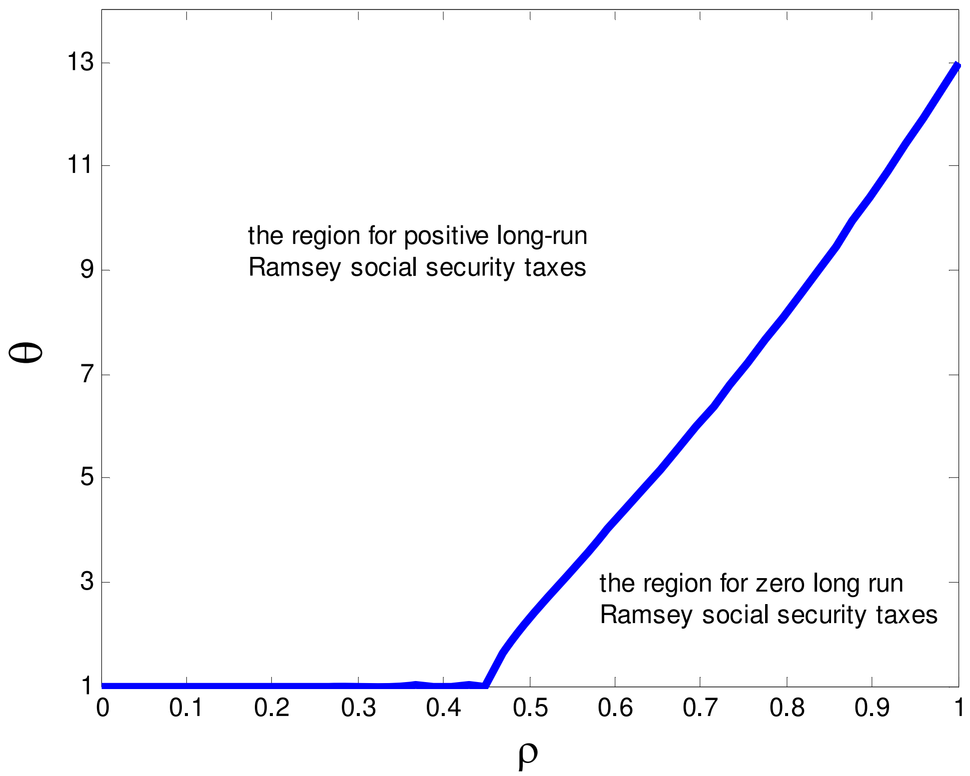

Figure 1: The other parameter values are set equal to the values in Section 3.1.

Figure 2: The Markovian Equilibrium Policy Rule under σ=2

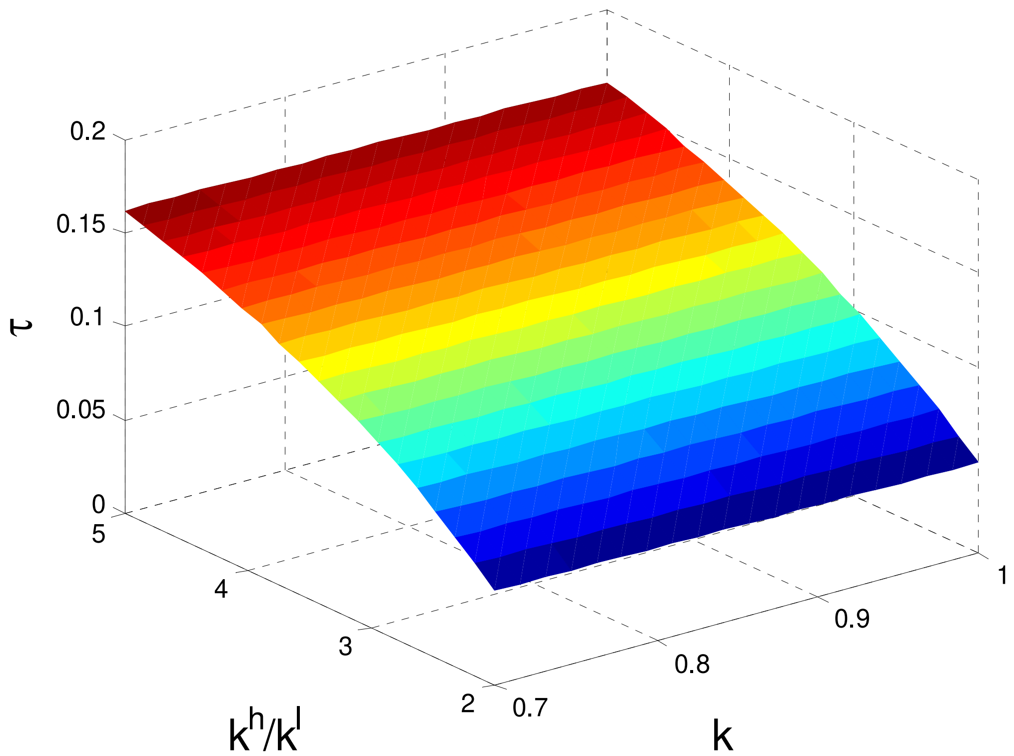

Figure 2: σ=2 and the other parameter values are set equal to the values in Section 3.1.

Figure 3: The Dynamics of Social Security under σ=2

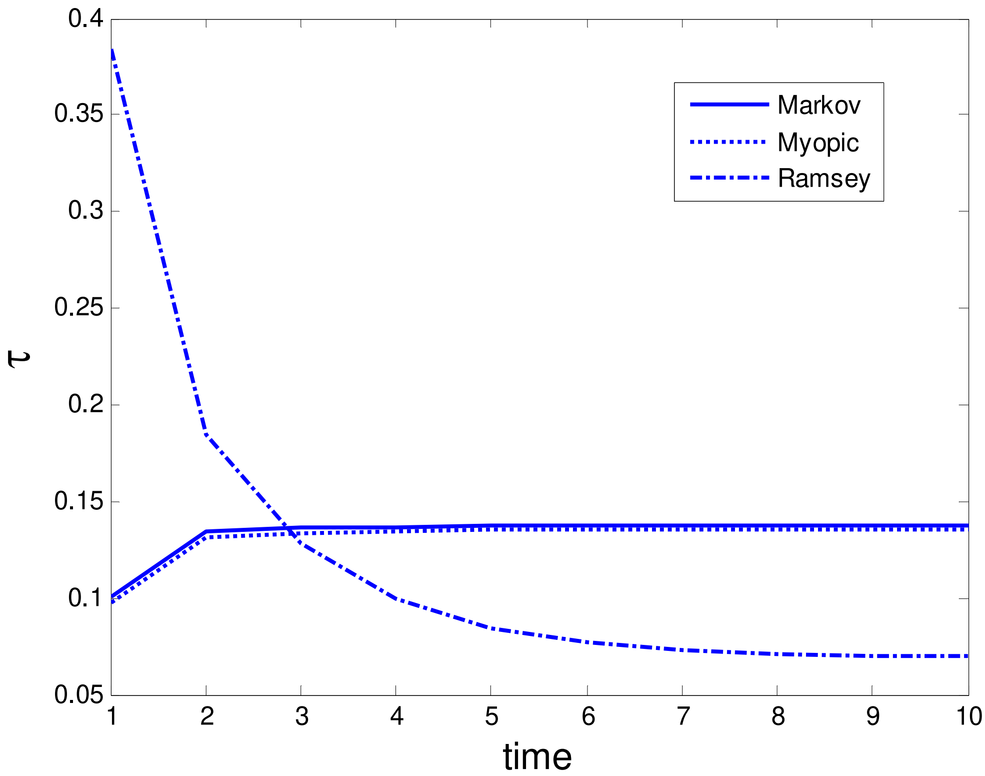

Figure 3: σ=2, ρ=βn and the other parameter values are set equal to the values in Section 3.1. The solid, dotted and dashed lines refer to social security tax rates in the Markov political equilibrium, myopic voting equilibrium and Ramsey allocation, respectively.