Discussion Papers In Economics And Business

Aging, Inequality and Social Security

Ryo Arawatari and Tetsuo Ono

Discussion Paper 08-19

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

Aging, Inequality and Social Security

Ryo Arawatari and Tetsuo Ono

Discussion Paper 08-19

April 2008

この研究は「大学院経済学研究科・経済学部記念事業」
基金より援助を受けた、記して感謝する。

Graduate School of Economics and
Osaka School of International Public Policy (OSIPP)
Osaka University, Toyonaka, Osaka 560-0043, JAPAN

Aging, Inequality and Social Security*

Ryo Arawatari[†] and Tetsuo Ono[‡]

Graduate School of Economics, Osaka University

April 2008

Abstract

This paper develops an overlapping-generations model including wage inequality within a generation and intra- and intergenerational resource reallocation via social security. Based on the concept of a stationary Markov perfect equilibrium, the paper focuses on the feedback mechanism between current individuals' decisions on saving and future voting on social security. The paper demonstrates the determination of social security via probabilistic voting and its consequence for consumption inequality within a generation. It is shown that when the elderly are politically powerful, (i) the economy attains an oscillatory path of inequality and social security, and (ii) aging may reduce consumption inequality.

Keywords: Aging; Inequality; Social security; Political Economy; Stationary Markov Perfect Equilibrium.

JEL Classification: D72, H55, J10.

*Arawatari acknowledges financial support from the Research Fellowships for Young Scientists of the Japan Society for the Promotion of Science (JSPS). Ono acknowledges financial support from the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for Young Scientists (B) (No. 20730208).

[†]Graduate School of Economics, Osaka University, 1-7, Machikaneyama, Toyonaka, Osaka 560-0043, Japan. E-mail address: ege001ar@mail2.econ.osaka-u.ac.jp

[‡]Corresponding Author. Graduate School of Economics, Osaka University, 1-7, Machikaneyama, Toyonaka, Osaka 560-0043, Japan. E-mail address: tono@econ.osaka-u.ac.jp

1 Introduction

Pay-as-you-go (PAYG) social security systems are at the center of public sector expenditure in many OECD countries. Benefit levels and contribution rates are determined in the political process and thus affected by political conflict between the beneficiaries (the elderly) and the contributors (the young). In order to assess the efficiency of existing social security and to obtain a prediction of the future of social security, there is a need to develop a politico-economic framework that incorporates this political conflict between generations.

A key to the development of the framework is a feedback mechanism between social security policy and individual decision making. Expectations about future policy affect current individuals' decision making on saving or education, which in turn has an effect on the future distribution of income and thus on the future voting behavior on the policy. Several researchers have provided politico-economic frameworks that incorporate this feedback mechanism based on the concept of a stationary Markov perfect equilibrium (see, for example, Grossman and Helpman, 1998; Azariadis and Galasso, 2002; Hassler et al., 2003; Forni, 2005; Hassler et al., 2005; Bassetto, 2007; Gonzalez-Eiras and Niepelt, 2007; Song, 2007).

Social security has a role of intragenerational redistribution via its Beveridgean part in addition to intergenerational redistribution. However, the above-mentioned studies abstracted the intragenerational redistribution aspect by assuming identical agents within a generation. Some researchers have considered the politics of intragenerational redistribution, but abstracted the feedback mechanism by assuming once-and-for-all voting (Casamatta, Cremer and Pestieau, 2000, 2001; Conde-Ruiz and Profeta, 2007; Koethenbuerger, Poutvaara and Profeta, 2008). They did not mention the political effect of social security on inequality within a generation. Therefore, the question arises of whether or not the politics of social security has an effect on intragenerational redistribution, and thus on inequality, within a generation in the presence of the feedback mechanism. The answer is no, but it is only for the case of wealth inequality under a pure Bismarckian scheme (Song, 2007).[1]

The aim of this paper is to characterize the dynamic political economy of social security and consumption inequality under a social security system including the Beveridgean and the Bismarckian parts. In addition, we consider population aging resulting from a decline in fertility rates, and investigate how aging affects social security and consumption

[1]Hassler et al. (2003) and Hassler et al. (2005) also analyzed income inequality within a generation, but there is no capital accumulation in their framework. In addition, their focus is different from Song (2007) and ours. They focus on the endogenous determination of income inequality affected by politics; Song (2007) and our paper assume income inequality as given and focus on endogenous determination of wealth inequality (Song, 2007) or consumption inequality (the current paper) affected by politics.

inequality through the political process. Therefore, the paper provides dynamic political-economy implications of aging, inequality and social security.[2]

For the purpose of analysis, we develop a simple two-period overlapping-generations model with a linear production function and a quasi-linear utility function. Agents differ in their working ability and thus in their labor income in youth; they retire in old age and receive social security benefits. The pension benefit consists of two parts: a contributory part (Bismarckian) and a noncontributory part (Beveridgean). This pension benefit is determined in a probabilistic voting framework (Lindbeck and Weibull, 1987). We utilize the concept of a stationary Markov perfect equilibrium where the social security tax rate is conditioned on capital as a state variable.

Under the framework presented above, the current paper shows the following two results. First, the political power of the elderly related to aging plays a key role in determining the dynamic patterns of inequality and social security. When the political power of the elderly is weak, the economy displays a monotone path of inequality and social security toward the steady state. On the other hand, when their power is strong, the economy displays an oscillatory convergence toward, or a period-two cycle around, the steady state.

Second, the model presents the opposite economic and political effects of aging on the size of the social security similar to Galasso and Profeta (2004, 2007). Given this result, we examine the effect of aging on consumption inequality (evaluated in terms of utility). It is shown that under a pure Beveridgean scheme, aging results in a lower level of inequality when the political power of the elderly is strong. Under a pure Bismarckian scheme, aging results in a lower level of inequality when the initial population growth rate is low. The result implies that aging may reduce consumption inequality through the political process.

The organization of this paper is as follows. Section 2 develops the model and characterizes economic equilibrium. Section 3 characterizes a Ramsey allocation defined as a feasible plan chosen by a benevolent social planner. Section 4 characterizes a political equilibrium, shows its existence and stability, compares it to the Ramsey allocation and analyzes the effect of aging on the amount of social security. Section 5 presents the dynamics of inequality and investigates the effect of aging on consumption inequality. Section 6 discusses the assumptions of this paper. Section 7 provides concluding remarks. The proofs are provided in the appendix.

[2]There are several studies analyzing the political economy of social security focusing on the role of social security systems (see, for example, Casamatta et al., 2000, 2001; Cremer et al., 2007; Conde-Ruiz and Profeta, 2007; Koethenbuerger, Poutvaara and Profeta, 2008). However, these studies focus on the steady state by assuming once-and-for-all voting; the dynamic motion of social security and inequality is abstracted from their analysis.

2 The Model and Economic Equilibrium

Consider an overlapping-generations model where the economy is composed of individuals and perfectly competitive firms. Time is discrete and denoted as $t = 0, 1, 2, \cdots$. A new generation is born in each period $t = 0, 1, 2, \cdots$. The size of a generation born in period t (generation t) is given by N^t. The population grows at a constant rate $n > -1$: $N^t = (1 + n)N^{t-1}$.

Individuals in each generation live for two periods: youth and old age. Each individual is endowed with one unit of labor in youth, supplies it inelastically to firms, and retires in old age. An individual has a probability $\lambda \in (0, 1)$ to be endowed with the high productivity $w^h > 0$ and a probability $1-\lambda$ to be endowed with the low productivity $w^l \in (0, w^h)$. Individuals with w^i are referred to as type $i(= h, l)$. The average productivity is assumed to be fixed and given by $\bar{w} = \lambda w^h + (1 - \lambda)w^l$. The aggregate effective labor supply in period t is given by $L_t = (\lambda w^h + (1 - \lambda)w^l)N^t$.

To obtain analytical results, we assume a linear production function in which gross output, Y, is produced using labor, L, and capital, K : $Y_t = L_t + RK_t$. The wage rate, 1, and the gross rental price of capital, R, are determined by the marginal productivity conditions for factor prices and are already substituted into the production function. For simplicity, the two types of labor are assumed to be perfect substitutes in production in terms of efficiency units of labor input, and capital is assumed to fully depreciate at the end of the production process.

A type-i individual in generation t evaluates consumption bundles (c_t^{yi}, c_{t+1}^{oi}) by the utility function $U^{it} = u(c_t^{yi}) + \beta c_{t+1}^{oi}$ where c_t^{yi} represents the consumption at time t of type i young individual and c_{t+1}^{oi} represents the consumption at time $t + 1$ of type i old individual. $\beta \in (0, 1)$ is the time discount factor. The function u is concave and twice continuously differentiable with $\lim_{c\to 0} u'(c) = +\infty$. Quasi-linearity of U^{it} implies that all income effects are absorbed by old age consumption. The role of this assumption will be discussed in Section 6.

A type-i individual is subject to a type-independent payroll tax $\tau_t \in [0, 1]$ and expects the type-dependent old age pension b_{t+1}^i. He/she can allocate his/her disposable labor income between consumption c_t^{yi} and saving z_t^i. When he/she retires, his/her consumption c_{t+1}^{oi} is equal to the gross return on saving, Rz_t^i, plus the pension benefit, b_{t+1}^i. Formally, type i individual solves the following problem:

$$\max_{z_t^i \geqq 0} u(c_t^{yi}) + \beta c_{t+1}^{oi}$$
$$\text{s.t. } c_t^{yi} + z_t^i = (1 - \tau_t)w^i,$$
$$c_{t+1}^{oi} = Rz_t^i + b_{t+1}^i.$$

The first-order condition associated with an interior solution of z_t^i is:

$$u'(c_t^{yi}) = \beta R \text{ for } i = h, l, \quad (1)$$

or:

$$c_t^{yi} = \bar{c}^y \equiv u'^{-1}(\beta R) \text{ for } i = h, l. \quad (2)$$

With (2) and the budget constraints, the saving and the old-age consumption of type i are given by, respectively:

$$z_t^i = (1 - \tau_t)w^i - \bar{c}^y, \quad (3)$$
$$c_{t+1}^{oi} = R \cdot \{(1 - \tau_t)w^i - \bar{c}^y\} + b_{t+1}^i. \quad (4)$$

Because of the assumption of the quasi-linear utility function, the consumption in youth is type-independent and constant over time, and the saving is independent of the pension benefit b_{t+1}^i.[3] We assume that pension benefit of type-i, b_{t+1}^i, consists of two parts: a contributory part that is directly related to individual earning, w^i, and a noncontributory part that depends on average earnings, $\bar{w}$. With a pay-as-you-go (PAYG) scheme, the average return of the social security system is given by the population growth rate. These properties yield the following expression for b_{t+1}^i:

$$b_{t+1}^i = (1 + n) \cdot \tau_{t+1} \cdot (\alpha w^i + (1 - \alpha)\bar{w}). \quad (5)$$

The parameter $\alpha \in [0, 1]$ is the Bismarckian factor that is the fraction of the pension benefits that is related to contributions. When $\alpha = 1$, the pension scheme is pure Bismarckian or contributory; when $\alpha = 0$, pension benefits are uniform and the scheme is Beveridgean. The case of an endogenous choice of α by voting is discussed in Section 6.

At the start of the economy, time 0, there are young agents in generation 0 as well as the initial old agents. A type-i initial old agent is endowed with z_{-1}^i units of saving at the beginning of period 0 and obtains the return Rz_{-1}^i and the pension benefit b_0^i, which are consumed.

Definition 1: For a given sequence of taxes, $\{\tau_t\}_{t=0}^{\infty}$, the initial condition $z_{-1}^i > 0 (i = h, l)$ and a set of parameters, $(n, \alpha, w^h, w^l, \lambda, R)$, an *economic equilibrium* is a se-

[3] The saving function indicates that for type-l individuals, there is a possibility of $z_t^l < 0$. This implies that type-l individuals want to borrow from type-h individuals in the credit market in order to finance their consumption in youth. This paper assumes that such borrowing is available for type-l individuals: there is no limit of access to the credit market. Although a borrowing constraint may provide interesting results and implications for the political economy (see, for example, Bellettini and Berti Ceroni, 2007), this paper leaves this for future research and instead focuses on the interaction between aging, inequality and social security under the condition of no credit constraint.

quence of allocations, $\left\{c_t^{yi}, c_{t+1}^{oi},\ z_t^i, b_t^i, K_t\right\}_{t=0}^{\infty}$, such that (i) the utility maximization problem is solved for each type i; (ii) the social security budget constraint is balanced in every period, $N^t(\lambda b_{t+1}^h + (1-\lambda)b_{t+1}^l) = N^{t+1}(\lambda w^h + (1-\lambda)w^l)\tau_{t+1}$, under the benefit rule (5); and (iii) the capital market clears: $K_{t+1} = N^t(\lambda z_t^h + (1-\lambda)z_t^l)$.

Let $k_t \equiv K_t/N^t$ denote capital per young person in period t. Given a sequence of taxes, $\{\tau_t\}_{t=0}^{\infty}$, and the initial condition $k_0 > 0$, an equilibrium sequence of k_t, $\{k_t\}_{t=0}^{\infty}$, is characterized by:

$$(1+n)k_{t+1} = (1-\tau_t)\bar{w} - \bar{c}^y. \tag{6}$$

3 Ramsey Allocation

In this section, we characterize a *Ramsey allocation* defined as a feasible plan chosen by a benevolent social planner who can commit to a policy sequence at time zero. The allocation maximizes the following objective function:

$$W = \beta\left\{\lambda c_0^{oh} + (1-\lambda)c_0^{ol}\right\} + \textstyle\sum_{t=0}^{\infty}\eta^{t+1}\left[\lambda\left\{u(c_t^{yh}) + \beta c_{t+1}^{oh}\right\} + (1-\lambda)\left\{u(c_t^{yl}) + \beta c_{t+1}^{ol}\right\}\right]$$

subject to

$$\begin{aligned} &c_t^{yi} = \bar{c}^y, c_{t+1}^{oi} = R\cdot\{(1-\tau_t)w^i - \bar{c}^y\} + b_{t+1}^i, \\ &\lambda z_t^h + (1-\lambda)z_t^l \geq 0, \end{aligned}$$

where c_t^{yi} and c_{t+1}^{oi} are households' consumption functions and the third inequality is the nonnegative constraint of aggregate saving. The planner discounts future generations' utilities with a discount factor of $\eta \in (0,1)$.

The Ramsey problem described above is dynamic in nature, but it admits a simple static representation. To show this, we substitute households' budget constraints into the objective function to obtain:

$$W = \beta R\{\lambda z_{-1}^h + (1-\lambda)z_{-1}^l\} + \tilde{W},$$

where

$$\tilde{W} \equiv \textstyle\sum_{t=0}^{\infty}\eta^t\left[\beta\bar{w}\{(1+n) - \eta R\}\tau + \eta\{u(\bar{c}^y) + \beta R(\bar{w} - \bar{c}^y)\}\right].$$

The problem reduces to a sequence of identical static optimization problems over τ.

Let τ^R denote the solution to the Ramsey problem described above. We establish the following result.

Proposition 1: *The Ramsey allocation has*

$$\tau^R \begin{cases} = \frac{\bar{w}-\bar{c}^y}{\bar{w}} & \text{if} \quad 1+n > \eta R, \\ \in \left[0, \frac{\bar{w}-\bar{c}^y}{\bar{w}}\right] & \text{if} \quad 1+n = \eta R, \\ = 0 & \text{if} \quad 1+n < \eta R. \end{cases}$$

Figure 1 shows the tax rate in the Ramsey allocation as a function of the population growth rate, n. For $1+n < (>)\eta R$, the return from intergenerational transfer $1+n$ is lower (higher) than the return from saving, R, multiplied by the discount factor η from the viewpoint of the social planner. Therefore, the planner chooses $\tau^R = 0 (= 1)$ if $1+n < (>)\eta R$. For $1+n = \eta R$, the planner is indifferent between the two ways of resource reallocation.

[Figure 1 about here.]

4 Political Equilibrium

This section characterizes a political equilibrium. Section 3.1 provides the definition of political equilibrium based on the concept of a stationary Markov perfect equilibrium with probabilistic voting. Section 3.2 provides the characterization of the political equilibrium path of capital and tax rate. We then focus on the steady-state political equilibrium and show the existence and stability of the steady-state equilibrium. In Section 3.3, based on the characterization of the steady-state equilibrium, we examine how the size of the welfare state, represented by the tax rate, is affected by population aging resulting from a decline in the fertility rate.

4.1 The Political Environment

The tax rate τ_t is chosen by some repeated political process at the beginning of each period. This paper assumes that τ_t is determined in a probabilistic voting framework (see, for example, Lindbeck and Weibull, 1987). In this framework, political decision makers' platforms in period t simply maximize a weighted average utility of all voters alive in that period. Thus, the equilibrium policy maximizes a political objective function given by:

$$\begin{aligned} G_t &= \gamma \cdot \left[\lambda c_t^{oh} + (1-\lambda) c_t^{ol}\right] \\ &\quad + (1+n) \cdot \left[\lambda \left\{u(c_t^{yh}) + \beta c_{t+1}^{oh}\right\} + (1-\lambda) \left\{u(c_t^{yl}) + \beta c_{t+1}^{ol}\right\}\right], \end{aligned}$$

where the political weight on the old is given by $\gamma \in [0, \infty)$ and the weight on the young is given by the gross rate of population growth. When $\gamma = 0$, only the young are entitled to vote; when $\gamma > 0$, both the young and the old vote on the tax rate. A greater weight on the old implies that they have greater political power.

We focus on a stationary Markov perfect equilibrium, in which the state of the economy is summarized by the stock of capital per young person in the economy. We restrict the choice of political decision makers to the Markovian policy function:

$$\tau_t = \theta(k_t),$$

where $\theta(\cdot)$ is a stationary, differentiable function. Let the political decision maker in period t, with expectations $\theta^e(k_{t+1})$ about the period $t+1$ tax rate find it optimal to set the transfer $\theta(k_t)$. The decision maker's expectations about the future are consistent with the incentives that will actually face the future decision maker if and only if $\theta = \theta^e$. Formally, the Markov perfect equilibrium is defined as follows.

Definition 2: A *Markov perfect political equilibrium* is a sequence of allocations, $\{c_t^{yi}, c_{t+1}^{oi}, z_t^i, b_t^i, k_t\}_{t=0}^{\infty}$ and a differentiable policy function $\theta(\cdot)$ such that (i) conditions at Definition 1 (economic equilibrium) are satisfied; and (ii) $\theta(\cdot)$ is a fixed point of the mapping from $\theta^e(\cdot)$ to $\theta(\cdot)$, where $\theta(\cdot)$ is the solution to the policy maker's problem and $\theta^e(\cdot)$ is the expected policy function.

4.2 Existence and Stability of Political Equilibrium

In order to find the solution to the problem of maximization G_t, we take the derivative of the objective function G_t with respect to $\theta(k_t)$, which yields:

$$\begin{aligned}
\frac{\partial G_t}{\partial \theta(k_t)} = \; & \underbrace{\gamma\lambda(1+n)\left(\alpha w^h + (1-\alpha)\bar{w}\right)}_{(*1)} + \underbrace{\gamma(1-\lambda)(1+n)\left(\alpha w^l + (1-\alpha)\bar{w}\right)}_{(*2)} \\
& \underbrace{+(1+n)\lambda\left\{-\beta R w^h + \beta(1+n)\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^h + (1-\alpha)\bar{w}\right)\right\}}_{(*3)} \\
& \underbrace{+(1+n)(1-\lambda)\left\{-\beta R w^l + \beta(1+n)\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^l + (1-\alpha)\bar{w}\right)\right\}}_{(*4)},
\end{aligned}$$

where the term (*1) ((*2)) represents the marginal benefit for high-productivity (low-productivity) old people, and comes from the increase in $\theta(k_t)$. The term (*3)((*4)) represents the sum of the marginal cost and benefit for high-productivity (low-productivity) young people, and comes from the increase in $\theta(k_t)$ and the increase in $\theta(k_{t+1})$ through a change in $\theta(k_t)$, respectively. The derivation of the above equation is provided in the appendix.

Because we have $(1+n)k_{t+1} = (1-\theta(k_t))\bar{w} - \bar{c}^y$ in equilibrium, it follows that $\partial k_{t+1}/\partial\theta(k_t) = -\bar{w}/(1+n)$. Substituting this derivative into the condition of $\partial G_t/\partial\theta(k_t) = 0$ and rearranging the terms, we obtain the differential equation $\partial\theta(k_{t+1})/\partial k_{t+1} = -\mu$ where:

$$\mu \equiv \frac{\beta R - \gamma}{\beta \bar{w}}.$$

The parameter μ takes a positive or negative value depending on the sizes of β, γ and R.[4]

Integrating the differential equation, $\partial\theta(k_{t+1})/\partial k_{t+1} = -\mu$, we obtain:

$$\theta(k_{t+1}) = A - \mu k_{t+1}. \tag{7}$$

The constant of integration, A, is a free parameter that is determined by the first decision maker's expectation of future policies. Any period-t decision maker, who expects the next decision maker to choose the policy according to (7), is indifferent between all feasible taxes. That is, given the expectation that $\theta^e(k_{t+1}) = A - \mu k_{t+1}$, $\partial G_t/\partial\theta(k_t) = 0$ holds for all $\theta(k_t) \in [0,1]$. This is because a change in $\theta(k_t)$ would be exactly compensated by a corresponding change in $\theta^e(k_{t+1})$, driven by a variation in k_{t+1}. Therefore, every decision maker is willing to act according to (7), lagged one period, in order to validate the previous decision maker's expectations.

The equilibrium policy function $\theta(k_t) = A - \mu k_t$ can be combined with the accumulation equation (6) to derive the following law of motion for the stock of capital:

$$k_{t+1} = \frac{\bar{w}(1-A) - \bar{c}^y}{1+n} + \frac{\mu\bar{w}}{1+n}k_t. \tag{8}$$

Using this law of motion and the policy function, we can write the law of motion of the equilibrium tax as:

$$\theta_{t+1} = A - \frac{\mu(\bar{w} - \bar{c}^y)}{1+n} + \frac{\mu\bar{w}}{1+n}\theta_t. \tag{9}$$

We impose the following assumption (Assumption 1) to ensure that given $k_t > 0$, $k_{t+1} > 0$ and $\theta_t \in [0,1)$ hold.[5]

Assumption 1: $\bar{w} > \bar{c}^y = u'^{-1}(\beta R)$.

[4]For example, assume that a generation is 30 years in length and set $\beta = (0.98)^{30}$ and $\gamma = 1$. The value of β is based on the single-period discount rate found by Auerbach and Kotlikoff (1987) of 0.98. The value of γ implies that the old have equal political power with the young. If the single-period gross interest rate is given by 1.02, then $\beta R = (0.98 \times 1.02)^{30} < \gamma = 1 : \mu$ takes a negative value. Alternatively, if it is given by 1.03, then $\beta R = (0.98 \times 1.03)^{30} > \gamma = 1 : \mu$ takes a positive value.

[5]In the Appendix (Section 8.4), it is shown that under Assumption 1, there exists a nonempty set of k_t that leads to $k_{t+1} > 0$ and $\theta_t \in [0,1)$.

The political equilibrium attains a steady state if a sequence of capital and tax is stationary over time: $k_t = k_{t+1} = \bar{k}$ and $\theta_t = \theta_{t+1} = \bar{\theta}$. Suppose that there is a steady state with $\bar{k} > 0$ and $\bar{\theta} \in [0, 1)$. Equations (8) and (9) imply that the equilibrium sequence of $\{k_t, \theta_t\}$ displays monotone convergence toward the steady state if $\mu\bar{w}/(1 + n) \in (0, 1)$; constant levels over time if $\mu\bar{w}/(1+n) = 0$; oscillatory convergence toward the steady state if $\mu\bar{w}/(1+n) \in (-1, 0)$; and a period-two cycle around the steady state if $\mu\bar{w}/(1+n) = -1$. Therefore, we impose the following assumption (Assumption 2) on the set of parameters to ensure the stability of the interior steady state. Figure 2 illustrates the set of parameters that satisfies Assumption 2 in a $R - \gamma$ space.

Assumption 2: $0 < (1+n) - \mu\bar{w}$ and $0 \leq (1+n) + \mu\bar{w}$, i.e., $\gamma \in (\beta(R - (1+n)), \beta(R + (1+n))]$.

[Figure 2 about here.]

Define two critical values of the free parameter A as:

$$A_{low} \equiv \frac{\mu(\bar{w} - \bar{c}^y)}{1+n}, \; A_{high} \equiv \frac{\bar{w} - \bar{c}^y}{\bar{w}} \; . \tag{10}$$

Using these definitions, we establish the existence of a steady-state equilibrium.

Proposition 2: *If $A \in [A_{low}, A_{high})$, there exists a steady-state equilibrium with*

$$\bar{k} = \frac{\bar{w}(1-A) - \bar{c}^y}{(1+n) - \mu\bar{w}} > 0 \text{ and } \bar{\theta} = \frac{(1+n)A - \mu(\bar{w} - \bar{c}^y)}{(1+n) - \mu\bar{w}} \in [0, 1).$$

If $A = A_{low}$, then $\bar{\theta} = 0$; if $A \in (A_{low}, A_{high})$, then $\bar{\theta} \in (0, 1)$.

To understand the result shown in Proposition 2, consider the case of $\mu > 0$(that is, $\beta R > \gamma$) as an example. First, $\bar{k} > 0$ is rewritten as $A < A_{high}$. This implies that if the expectation parameter A is high such that $A \geq A_{high}$, the tax rate is too high to save: there is no resource remaining for saving and thus no accumulation of capital. Second, $\bar{\theta} < 1$ always holds under the condition of $A < A_{high}$ that requires $\bar{k} > 0$. Finally, $\bar{\theta} \geq 0$ is rewritten as $A \geq A_{low}$. If $A < A_{low}$, the tax rate becomes negative. Therefore, there exists a steady-state political equilibrium with $\bar{k} > 0$ and $\bar{\theta} \in [0, 1)$ if A is set within the range $[A_{low}, A_{high})$.

The result established in Proposition 2 indicates that the steady-state levels of capital and tax are independent of the parameter α representing the degree of redistribution via the pension system. This result depends on the following two assumptions: (i) the quasi-linear utility function of individuals; and (ii) the political objective function defined as the weighted sum of utilities of individuals. Because of these assumptions, the terms related to the parameter α are displayed in a linear form and thus redistribution effects on types l and h offset each other. Therefore, the parameter α has no effect on the determination of capital and tax. However, as we will show in Section 5, the redistribution effect on consumption still remains under these assumptions, implying that the pension system has an effect on consumption inequality.

Having shown the existence of a steady-state equilibrium with $\bar{k} > 0$, we next examine stability of the steady state. Equations (8) and (9) indicate that if $\mu = 0 (\beta R = \gamma)$, the levels of capital and tax are constant over time. On the other hand, if $\mu \neq 0 (\beta R \neq \gamma)$, the time paths of capital and tax display dynamics tending toward or around the steady state. The stability property for the latter case is formally summarized in the following proposition.

Proposition 3: *Assume that $\mu \neq 0(\beta R \neq \gamma)$. Given $k_0 \in ((\bar{c}^y - \bar{w})/\mu\bar{w} + A/\mu, A/\mu]$ if $\mu > 0$ (i.e., $\gamma < \beta R$) or $k_0 \in [A/\mu, (\bar{c}^y - \bar{w})/\mu\bar{w} + A/\mu)$ if $\mu < 0$ (i.e., $\gamma > \beta R$), there exists an equilibrium sequence of $(k_t, \theta_t)_{t=0}^{\infty}$ displaying* (i) *monotone convergence toward the interior steady-state equilibrium if $\gamma \in (\beta(R-(1+n)), \beta R)$*; (ii) *oscillatory convergence toward the interior steady-state equilibrium if $\gamma \in (\beta R, \beta(R+(1+n)))$; and* (iii) *a period-two cycle around the interior steady-state equilibrium if $\gamma = \beta(R+(1+n))$.*

The conditions of the parameters β, R, γ and n for the stability property are immediately found by equations (8) and (9). For the result established in Proposition 3, two remarks are in order. First, the relative magnitude between the political weight on the old, γ, and the gross interest rate multiplied by the discount factor, βR, plays a key role in determining the stability property of the equilibrium path. A larger γ leads to a higher tax rate because the old have greater political power in policy decision making, while a larger βR leads to a lower tax rate because a higher return from saving ensures a higher level of consumption in the second period of the life cycle, thereby giving the policy maker an incentive to set a lower current tax rate. Given these competing effects, if $\gamma < (>) \beta R$, i.e., if $\mu > (<) 0$, the former effect is smaller (larger) than the latter effect, implying that an increase in the current capital stock, k_t, results in a lower (higher) tax burden and thus an increase (a decrease) in the capital stock in the next period. Therefore, the political economy displays a monotone (an oscillatory) sequence of capital if $\gamma < (>) \beta R$, i.e., if $\mu > (<) 0$.

Second, the property of the equilibrium path depends on the expectation parameter A. To see this, suppose first that $\gamma < \beta R$: the economy displays a monotone convergence toward the steady state. If A is set within the range (A_{low}, A_{high}), the economy displays an increasing or decreasing sequence of the tax depending on the initial condition. However, if $A = A_{low}$, there is no increasing sequence of the tax because the steady-state tax is given by $\bar{\theta} = 0$; the economy displays only a decreasing sequence of the tax. Alternatively, suppose that $\gamma > \beta R$: the economy displays an oscillatory convergence toward, or a period-two cycle around, the steady state. In this case, the policy parameter A must be set within the range (A_{low}, A_{high}). If $A = A_{low}$, the sequence of tax is unfeasible because $\theta_1 < 0$ holds for a given $\theta_0 > 0$.

Figure 1 compares the politically determined tax rate (the political tax rate) to the tax rate in the Ramsey allocation (the Ramsey tax rate). Comparison is made for the following three cases. First, if the gross rate of population growth is low such that $1 + n < \eta R$, the political tax rate is generally higher than the Ramsey tax rate. The political equilibrium resembles the Ramsey allocation if and only if the expectation parameter A is unexpectedly set at $A = A_{low}$. Second, if the gross rate of population growth is high such that $1 + n > \eta R$, the political tax rate is lower than the Ramsey tax rate. The Ramsey allocation requires 100% taxation because resource reallocation via intergenerational transfer yields a higher return than that via private saving. However, political equilibrium attains a lower tax rate than that required by the Ramsey allocation because of the political power of the young, who want to avoid a tax burden. Finally, if the gross rate of population growth is by chance equal to ηR, any political tax rate is efficient because the Ramsey allocation attains indeterminate tax rates given by the set $[0, 1]$. Therefore, the comparison provides the following implication for political equilibrium in terms of efficiency: the political economy attains an antiwelfare state when it has a high rate of population growth whereas it attains a prowelfare state when it has a low rate of population growth.

4.3 Population Aging and the Size of the Welfare State

Based on the characterization of the political equilibrium, we now consider how the size of the welfare state, represented by the tax rate θ, is affected by population aging represented by a decrease in population growth (n).

Proposition 4: *Suppose that $\gamma \in (\beta(R - (1 + n)), \beta(R + (1 + n)))$: the equilibrium path displays a stable convergence toward the steady state with $\bar{\theta} \in (0, 1)$. A lower population growth rate leads to a smaller (constant, larger) size of the welfare state if and only if $\mu > (=, <) 0$: $\partial\bar{\theta}/\partial n \gtreqless 0$ if and only if $\mu \gtreqless 0$, i.e., $\beta R \gtreqless \gamma$.*

The result for the case of $\mu = 0$ is immediately apparent from the fact of $\bar{\theta} = A$. To understand the results for the case of $\mu \neq 0$, we first focus our attention on the capital market-clearing condition in the steady state derived from (8): $(1 + n)\bar{k} = (1 - (A - \mu\bar{k}))\bar{w} - \bar{c}^y$ and consider the effects of the parameters n and γ on the steady-state level of capital. We then turn our focus to the policy function $\theta = A - \mu k$ and consider the effects of n and γ on the tax rate via the steady-state level of capital.

Figure 3 illustrates the steady-state level of capital given by the point at the intersection of $(1+n)k$ with $(1-(A-\mu k))\bar{w} - \bar{c}^y$. As illustrated in Figure 3, a lower population growth leads to a higher level of per capita capital for both cases of $\mu > 0$ and $\mu < 0$. This is because, given the aggregate level of capital, a smaller size of population leads to a higher level of per capita capital. Given the policy function $\bar{\theta} = A - \mu\bar{k}$, a higher level of per capita capital results in a lower (higher) level of tax rate if and only if $\mu > (<) 0$, i.e., $\beta R > (<) \gamma$. The relative magnitude between βR and γ is important for the determination of the effect of aging on the size of the welfare system.

[Figure 3 about here.]

The terms βR and γ indicate that population aging has two opposite economic and political effects on the size of the welfare state (see, for example, Razin, Sadka and Swagel, 2002; Galasso and Profeta, 2004, 2007). The economic effect is captured by the increase in the share of retirees to workers, which reduces the profitability of a pay-as-you-go social security system compared to saving. This effect is represented by the term βR. A larger R implies that the pay-as-you-go social security system is less profitable than saving, which gives individuals an incentive to downsize the social security system. The political effect is captured by a decrease in the political weight for the young, that is, an increase in the political weight for the old. This effect is represented by the term γ. A larger γ implies that the political power of the old becomes stronger and thus a larger social security system is supported by voting. The overall effect of aging depends on the relative magnitude between the two terms βR and γ.

5 Inequality

This section introduces an index that measures an inequality between the two types of agents in terms of consumption, i.e., lifetime utility. We focus on inequality of economic welfare by consumption rather than income because consumption reflects not only current income but also lifetime resources. For the purpose of analysis, we define the index of the lifetime utility inequality as:

$$I_t = \frac{u(\bar{c}^y) + \beta\left[R\{(1-\theta(k_t))w^h - \bar{c}^y\} + (1+n)\theta(k_{t+1})(\alpha w^h + (1-\alpha)\bar{w})\right]}{u(\bar{c}^y) + \beta\left[R\{(1-\theta(k_t))w^l - \bar{c}^y\} + (1+n)\theta(k_{t+1})(\alpha w^l + (1-\alpha)\bar{w})\right]},$$

where the numerator and the denominator represent the lifetime utilities of types h and l, respectively. Based on this measure, we investigate how inequality changes over time

(Section 5.1). We also consider how inequality is affected by population aging (Section 5.2). For the tractability of analysis, we make the following assumption.

Assumption 3: $u(\bar{c}^y) \geq \beta R \bar{c}^y$.

Assumption 3 guarantees that the level of the lifetime utility for each type of agent is greater than zero. Since the utility level of type h is greater than the utility level of type l, the index takes a value between one and infinity under Assumption 3: $I_t \in (1, \infty)$. A higher I_t implies greater inequality in terms of lifetime utility. If $u(\bar{c}^y) = \beta R \bar{c}^y$, the index is identical to that measuring the lifetime income inequality.

5.1 Dynamics of Inequality

This section investigates the time path of inequality. The inequality index I_t is rewritten as a function of k_t by using the law of motion for capital (8). Thus, we can display the time path of the lifetime utility (or income) inequality based on the dynamics of capital accumulation presented in Proposition 3.

Proposition 5

(i) *Suppose that $\mu > 0$, i.e., $\gamma < \beta R$. The inequality is increased by capital accumulation:* $\partial I_t / \partial k_t > 0$.

(ii) *Suppose that $\mu < 0$, i.e., $\gamma > \beta R$. The inequality is nonincreased by capital accumulation, i.e., $\partial I_t / \partial k_t \leq 0$, if $\bar{c}^y < (1+n)\bar{w}/\{(1+n) - \mu \bar{w}\}$ and $A \in (-\mu \bar{c}^y/(1+n), A_{high})$.*

In order to interpret the result established in Proposition 5, we differentiate I_t with respect to k_t and rearrange the terms to obtain:

$$\begin{aligned} &\frac{1}{\beta}[u(\bar{c}^y) + \beta c^{ol}_{t+1}]^2 \frac{\partial I_t}{\partial k_t} \\ = \ &\underbrace{R \frac{\partial \theta(k_t)}{\partial k_t}(-w^h + w^l)\left[(u(\bar{c}^y) - \beta R \bar{c}^y) + (1+n)\theta(k_{t+1})(1-\alpha)\bar{w}\right]}_{\text{Labor income effect}} \\ &\underbrace{+(1+n)\frac{\partial \theta(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t}(w^h - w^l)\left[\overbrace{\alpha(u(\bar{c}^y) - \beta R \bar{c}^y)}^{\text{Bismarckian effect}} \overbrace{-(1-\alpha)\bar{w}\beta R(1-\theta(k_t))}^{\text{Beveridgean effect}}\right]}_{\text{Pension effect}}, \end{aligned} \quad (11)$$

where the first term shows the effect of capital on inequality through labor income and the second term shows the effect of capital on inequality through the pension benefit. The second term includes the two effects via the Bismarckian and Beveridgean parts of the pension benefit.

Suppose that $\mu > 0$: a higher level of capital results in a lower tax rate and a higher level of the following period's capital: $\partial\theta(k_t)/\partial k_t = \partial\theta(k_t)/\partial k_t = -\mu < 0$ and $\partial k_{t+1}/\partial k_t = \mu\bar{w}/(1+n) > 0$. Given these results, there are three effects of capital on inequality. First, both types of agents benefit from a decrease in the tax burden, but the marginal benefit for type h, w^h, is greater than the marginal benefit for type l, w^l. The effect via labor income is negative in terms of equality. Second, because of a decrease in the tax rate, both types of agents are harmed by a decrease in the Bismarckian part of the pension benefit, but the marginal loss for type h, w^h, is greater than that for type l, w^l. The effect via the Bismarckian part is thus positive in terms of equality. Third, both types of agents suffer the same degree of loss in the Beveridgean part of the pension benefit. This implies that the effect via the Beveridgean part is negative in terms of equality. In equilibrium, the sum of the first and the third negative effects overcomes the second positive one. Therefore, if $\mu > 0$, inequality is increased by capital accumulation.

Alternatively, suppose that $\mu < 0$: a higher level of capital results in a higher tax rate and a lower level of the next period's capital: $\partial\theta(k_t)/\partial k_t = \partial\theta(k_t)/\partial k_t = -\mu > 0$ and $\partial k_{t+1}/\partial k_t = \mu\bar{w}/(1+n) < 0$. The three effects observed for the case of $\mu > 0$ are reversed: the effect via the labor income effect is positive; the effect via the Bismarckian part is negative; and the effect via the Beveridgean part is positive. The sum of the positive effects is equal to or greater than the negative one if $A \in [-\mu\bar{c}^y/(1+n), A_{high})$. This set is nonempty if $\bar{c}^y < (1+n)/\{(1+n) - \mu\bar{w}\}$.

The result established in Proposition 5 suggests that the dynamic pattern of inequality depends on the political power of the elderly. When the political power of the elderly is weak such that $\mu > 0$ (i.e., $\gamma < \beta R$), the economy experiences an expanding (shrinking) inequality along the transition toward the steady state if the initial level of capital is below (above) the steady-state level. However, when the political power is strong such that $\mu < 0$ (i.e., $\gamma > \beta R$), it may experience fluctuation of inequality along the transition toward the steady state. The result indicates that the property of inequality dynamics depends heavily on the political power of the elderly.

5.2 Population Aging and Inequality

This section focuses on the steady-state level of inequality and investigates how inequality is affected by population aging via the political determination of social security. For the tractability of analysis, we focus on two extreme social security systems, the pure Beveridgean ($\alpha = 0$) and the pure Bismarckian ($\alpha = 1$), and consider how population growth rate (n) affects the steady-state level of inequality in each case.

The steady-state level of inequality is given by:

$$\bar{I} = \frac{u(\bar{c}^y) + \beta\left[R\{(1-\theta(\bar{k}))w^h - \bar{c}^y\} + (1+n)\theta(\bar{k})(\alpha w^h + (1-\alpha)\bar{w})\right]}{u(\bar{c}^y) + \beta\left[R\{(1-\theta(\bar{k}))w^l - \bar{c}^y\} + (1+n)\theta(\bar{k})(\alpha w^l + (1-\alpha)\bar{w})\right]}.$$

By using this, we first consider the effect of a decrease in population growth on inequality.[6]

Proposition 6

(i) *Suppose that the pension system is purely Beveridgean:* $\alpha = 0$. *An inequality is* (a) *increased by a lower population growth, i.e.,* $\partial\bar{I}/\partial n < 0$ *if* $\mu \geq 0$; *and* (b) *decreased by a lower population growth, i.e.,* $\partial\bar{I}/\partial n > 0$ *if* $\mu < 0$ *and* $A \in (A_{low}, \hat{A}]$ *where:*

$$\hat{A} \equiv \frac{-\mu(\bar{w} - \bar{c}^y)\mu\bar{w}}{(1+n)\{(1+n) - 2\mu\bar{w}\}}.$$

(ii) *Suppose that the pension system is purely Bismarckian:* $\alpha = 1$. *An inequality is not increased by lower population growth, i.e.,* $\partial\bar{I}/\partial n \geq 0$ *if* (a) $\mu > 0$ *and* $1 + n \geq R$, (b) $\mu = 0$, *or* (c) $\mu < 0$ *and* $1 + n \leq R$.

We differentiate $\bar{I}$ with respect to n to obtain:

$$\begin{aligned}
&\frac{1}{\beta}[u(\bar{c}^y) + \beta\bar{c}^{ol}]^2\frac{\partial\bar{I}}{\partial n} \\
=\; &\underbrace{\left(-R\frac{\partial\theta(\bar{k})}{\partial n}\right)(w^h - w^l)\left[(u(\bar{c}^y) - \beta R\bar{c}^y) + \beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right]}_{\text{Effect via labor income}} \\
&\underbrace{+\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n} + \theta(\bar{k})\right\}\alpha(w^h - w^l)\left[(u(\bar{c}^y) - \beta R\bar{c}^y) + \beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right]}_{\text{Effect via Bismarckian part of the pension benefit}} \\
&\underbrace{-\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n} + \theta(\bar{k})\right\}(1-\alpha)\bar{w}\beta(w^h - w^l)\left[R(1-\theta(\bar{k})) + (1+n)\theta(\bar{k})\alpha\right]}_{\text{Effect via Beveridgean part of the pension benefit}}.
\end{aligned} \tag{12}$$

This shows that population growth has effects on inequality through the three factors of labor income and the Bismarckian and Beveridgean parts of the pension benefit. In what follows, we consider how these three factors may interact to affect the steady-state level of inequality.

[6] The effect of α on inequality is straightforward. Given that the tax rate $\bar{\theta}$ is independent of α (Proposition 2), the numerator of the index is increasing in α and the denominator of the index is decreasing in α. Therefore, we obtain $\partial\bar{I}/\partial\alpha > 0$: a shift toward the Bismarckian part leads to an expansion of the steady-state inequality.

Let us consider an economy where the pension system is purely Beveridgean ($\alpha = 0$): the effect via the Bismarckian part disappears in this case. First, suppose that $\mu > 0$. A lower population growth rate leads to a lower tax rate (Proposition 4(i)) and thus produces benefits for both types of young agents via labor income and losses for both types of old agents via the Beveridgean part of the pension benefits. However, the marginal benefit for type l young agents, w^l, is lower than the marginal benefit for type h young agents, w^h, implying that the effect of a lower population growth via labor income is negative in terms of equality. In addition, both types of old agents suffer the same degree of loss in the Beveridgean part of the pension benefits, implying that the effect via the Beveridgean part is negative in terms of equality. Given these two negative effects on equality, a lower population growth rate leads to a higher level of inequality if $\alpha = 0$ and $\mu > 0$.

Second, suppose that $\mu = 0$. In this case, population growth has no effect on the politically determined tax rate (Proposition 4(i)), implying that the effect via labor income disappears. There remains only the effect via the Beveridgean part of the pension benefit. Therefore, a lower population growth rate leads to a higher level of inequality if $\alpha = 0$ and $\mu = 0$. Third, suppose that $\mu < 0$. A lower population growth rate leads to a higher tax rate (Proposition 4(i)) and thus produces losses for both types of young agents via labor income. However, the marginal loss for a type l agent, w^l, is smaller than the marginal loss for a type h agent, w^h, implying that the effect of a lower population growth rate via labor income is positive in terms of equality. The effect via the Beveridgean part of the pension benefit is ambiguous because it contains two competing effects: a positive effect caused by a decrease in the pension benefit because of a lower population growth rate, and a negative effect caused by a decrease in the tax rate. The condition $A \in (A_{low}, \hat{A}]$ ensures that the former overcomes the latter: inequality is decreased by a lower population growth rate.

Next, consider the economy where the pension system is Bismarckian ($\alpha = 1$). The effect via the Beveridgean part disappears in this case. First, suppose that $\mu > 0$. As in the case when $\alpha = 0$, a lower population growth rate yields a negative effect via labor income in terms of equality. However, it yields a positive effect via the Bismarckian part of the pension benefit in terms of equality. The condition $1 + n \geq R$ ensures that the positive effect overcomes the negative one. Second, suppose that $\mu = 0$. In this case, population growth has no effect on the politically determined tax rate (Proposition 4(i)), implying that the effect via labor income disappears in this case. There remains only the positive effect via the Bismarckian part of the pension benefit. Finally, suppose that $\mu < 0$. As in the case when $\alpha = 0$, a lower population growth rate yields a positive effect via labor income in terms of equality. However, the effect via the Bismarckian part is ambiguous because it contains the two competing effects as explained in the case when

$\alpha = 0$. If $1 + n \leq R$, the sum of the positive effects overcomes the negative one: the inequality level is decreased by a lower population growth rate.

The result established in Proposition 6 has the following policy implication. For countries with an aging population, the condition $\mu < 0$ (i.e. $\beta R < \gamma$) is more likely to hold because they feature greater political power of the elderly. Under this condition, the size of the social security is increased by aging (Proposition 4). Given this result and the condition that the population growth rate is low such that $1 + n \leq R$, aging leads to a reduction of inequality via the increasing size of redistribution (Proposition 6). This result is in sharp contrast to the effect of a decrease in mortality rates on consumption inequality. When the decrease in mortality rates is responsible for population aging, consumption inequality is expanded by aging because uninsured idiosyncratic events have persistent impacts on individual income (Deaton and Paxson, 1994; Ohtake and Saito, 1998). Although our result is based on theoretical analysis with some restrictive assumptions, it indicates that the source of population aging is important in determining the effect of aging on inequality.

6 Discussions

This section discusses the roles of two assumptions in the current analysis: the quasi-linear utility function and an exogenously given pension scheme.

6.1 The Role of a Quasi-linear Utility Function

So far, we have assumed that the marginal utility of consumption in old age is constant over time: $U^{it} = u(c_t^{yi}) + \beta c_{t+1}^{oi}$. To consider the role of this assumption, let us make the alternative assumption that the marginal utility of consumption in youth is constant over time: $U^{it} = c_t^{yi} + \beta v(c_{t+1}^{oi})$. Under this assumption, it is shown below that the political equilibrium generally fails to attain an interior solution.

Under the assumption of $U^{it} = c_t^{yi} + \beta v(c_{t+1}^{oi})$, the first-order condition of utility maximization is given by $\beta R v'(c_{t+1}^{oi}) = 1$, or $c_{t+1}^{oi} = \bar{c}^o \equiv v'^{-1}(1/\beta R)$. This implies that all income effects are absorbed by the consumption in youth. With this condition, the saving of type i agent in period t is:

$$z_t^i = (\bar{c}^o - b_{t+1}^i)/R = \left(\bar{c}^o - (1+n)\theta(k_{t+1})\left(\alpha w^i + (1-\alpha)\bar{w}\right)\right)/R.$$

With this saving function, the capital market clearing condition is written as:

$$(1+n)k_{t+1} = \bar{c}^o/R - (1+n)\theta(k_{t+1})\bar{w}/R,$$

which leads to $\partial k_{t+1}/\partial\theta(k_t) = 0$.

Given the $\beta R v'(c^{oi}_{t+1}) = 1$, the derivative of G_t with respect to $\theta(k_t)$ is:

$$\frac{\partial G_t}{\partial\theta(k_t)} = (1+n)\bar{w}\left[\frac{\gamma}{\beta R} - 1 + \frac{1+n}{R}\cdot\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\cdot\frac{\partial k_{t+1}}{\partial\theta(k_t)}\right].$$

We substitute $\partial k_{t+1}/\partial\theta(k_t) = 0$ into this equation to obtain:

$$\frac{\partial G_t}{\partial\theta(k_t)} = (1+n)\bar{w}\left(\frac{\gamma}{\beta R} - 1\right).$$

This implies that the solution to the problem is generally given by $\theta(k_t) = 0$ or 1. Therefore, the result implies that as long as we adopt a quasi-linear utility function for the tractability of analysis, we must assume that the utility of consumption in old age is linear in order to attain an interior solution.

6.2 An Endogenous Choice of a Pension System

We have considered the case in which the pension system, represented by the parameter α, is exogenously given. We now consider an alternative case where α is chosen by voting, and show that voters are indifferent between any levels of α under the current economic environment.

Because we focus on Markov perfect equilibrium, we restrict the choice of a pension system to the following Markovian policy function:

$$\alpha_t = \psi(k_t),$$

where ψ is a stationary, differentiable function. The derivative of G_t with respect to $\psi(k_t)$ leads to:

$$\begin{aligned}
\frac{\partial G_t}{\partial\psi(k_t)} &= \gamma\cdot\underbrace{\left[\lambda(1+n)\theta(k_t)(w^h-\bar{w}) + (1-\lambda)(1+n)\theta(k_t)(w^l-\bar{w})\right]}_{(\#1)} \\
&\quad +(1+n)\cdot\underbrace{\left[\lambda\beta(1+n)\theta(k_{t+1})(w^h-\bar{w}) + (1-\lambda)\beta(1+n)\theta(k_{t+1})(w^l-\bar{w})\right]}_{(\#2)} \\
&\quad \times\frac{\partial\psi(k_{t+1})}{\partial k_{t+1}}\cdot\frac{\partial k_{t+1}}{\partial\psi(k_t)} \\
&= 0 \ \forall\psi(k_t)\in[0,1],
\end{aligned}$$

where the last equality holds because the terms (#1) and (#2) are zero. Therefore, there is a continuum of solutions that maximize the objective function.

The reason for this result is that in the current environment, there is no distortion produced by α. In order to obtain a unique, interior solution of α, there is a need to introduce a distortionary effect associated with the social security system. An example is a tax base net of distortion given by $\phi(\alpha) = [1-\eta(1-\alpha)]$ where η identifies a distortionary effect associated with the noncontributory part of the social security system (Conde-Ruiz and Profeta, 2007). Given this distortion, the pension benefit for type i is $b^i_{t+1} = (1+n)\cdot\tau_{t+1}\cdot(\alpha_{t+1}w^i+(1-\alpha_{t+1})\bar{w})\cdot\phi(\alpha_{t+1})$. This extension may provide a richer political implication, but definitely makes it impossible to solve the model analytically. Because of this limitation, we have assumed α to be constant over time, and focused our attention on the determination of the tax rate and its impact on inequality.

7 Conclusion

This paper presents a simple political-economy model that includes intergenerational as well as intragenerational distribution parts of social security. In particular, based on the concept of a stationary Markov perfect equilibrium, the paper focuses on the feedback mechanism between current individuals' decision on saving and future voting over social security. In the presence of this mechanism, it considers how politics affects dynamic paths of inequality and social security. In addition, it considers population aging resulting from a decline in fertility rates, and analyzes the effects of aging on inequality and social security.

The main contribution of this paper is twofold. First, the feedback mechanism creates various dynamic patterns of inequality and social security. In particular, these are affected by the political power of the elderly. When their power is weak, the economy displays a monotone increase or decrease in the size of social security and the level of inequality. However, when the elderly's political power is strong, the economy displays an oscillatory path of inequality and social security. This result suggests that an aging economy associated with great political power of the elderly will tend to fluctuate in inequality and social security over time.

Second, the effects of aging on inequality differ between the two pension systems. Under a pure Beveridgean scheme, aging increases (decreases) inequality if the political power of the elderly is weak (strong). Under a pure Bismarckian scheme, aging decreases inequality when the political power of the elderly is strong and the population growth rate is low. These two results indicate that regardless of the social security system, an aging economy may experience a decrease in inequality when a fertility decline is responsible for population aging.

Our analysis is subject to a number of caveats. First, we ignore population aging resulting from a decrease in mortality rates. Joint analysis of decreases in mortality and fertility rates would produce a richer set of policy implications. Second, we abstract the political determination of the social security scheme. An interesting extension would be to consider the joint determination of the size of social security and the degree of redistributiveness as discussed in Section 6.2. Third, for analytical convenience, we assume a quasi-linear utility function and a linear production function. A natural extension of the model is to relax this assumption. In a related work, Song (2007) presents a method to derive an analytical solution under the assumption of a log-linear utility function and a Cobb–Douglas production function. Utilizing his method would produce a wider set of results and policy implications for aging, inequality and social security.

8 Appendix

8.1 The derivative of G_t with respect to $\theta(k_t)$

The derivative of G_t with respect to $\theta(k_t)$ yields:

$$
\begin{aligned}
\frac{\partial G_t}{\partial \theta(k_t)} &= \gamma\left[\lambda(1+n)\left(\alpha w^h+(1-\alpha)\bar{w}\right)+(1-\lambda)(1+n)\left(\alpha w^l+(1-\alpha)\bar{w}\right)\right] \\
&\quad +(1+n)\Bigg[\lambda\left\{u'(c_t^{yh})\left(-w^h-\tfrac{\partial z_t^h}{\partial\theta(k_t)}\right)\right. \\
&\quad \left.+\ \beta\left(R\tfrac{\partial z_t^h}{\partial\theta(k_t)}+(1+n)\tfrac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\tfrac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^h+(1-\alpha)\bar{w}\right)\right)\right\} \\
&\quad +(1-\lambda)\left\{u'(c_t^{yl})\left(-w^l-\tfrac{\partial z_t^l}{\partial\theta(k_t)}\right)\right. \\
&\quad \left.+\ \beta\left(R\tfrac{\partial z_t^l}{\partial\theta(k_t)}+(1+n)\tfrac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\tfrac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^l+(1-\alpha)\bar{w}\right)\right)\right\}\Bigg] \\
&= \gamma(1+n)\bar{w}+(1+n)\Bigg[\lambda\left\{-\beta R w^h+\beta(1+n)\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^h+(1-\alpha)\bar{w}\right)\right\} \\
&\quad +(1-\lambda)\left\{-\beta R w^l+\beta(1+n)\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial\theta(k_t)}\left(\alpha w^l+(1-\alpha)\bar{w}\right)\right\}\Bigg],
\end{aligned}
$$

where the second equality comes from $u'(c_t^{yh})=\beta R$ and $\lambda(1+n)\left(\alpha w^h+(1-\alpha)\bar{w}\right)+(1-\lambda)(1+n)\left(\alpha w^l+(1-\alpha)\bar{w}\right)=(1+n)\bar{w}$. Because we have $(1+n)k_{t+1}=(1-\theta(k_t))\bar{w}-\bar{c}^y$ in equilibrium, it follows that $\partial k_{t+1}/\partial\theta(k_t)=-\bar{w}/(1+n)$. We substitute this into the above equation to obtain:

$$
\begin{aligned}
\frac{\partial G_t}{\partial \theta(k_t)} &= \gamma(1+n)\bar{w}+(1+n)\Bigg[\lambda\left\{-\beta R w^h-\beta\bar{w}\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\left(\alpha w^h+(1-\alpha)\bar{w}\right)\right\} \\
&\quad +(1-\lambda)\left\{-\beta R w^l-\beta\bar{w}\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\left(\alpha w^l+(1-\alpha)\bar{w}\right)\right\}\Bigg].
\end{aligned}
$$

Under the assumption of an interior solution, $\partial G_t/\partial\theta(k_t)=0$ leads to:

$$\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}=-\mu\equiv-\frac{\beta R-\gamma}{\beta\bar{w}}.$$

8.2 Proof of Proposition 1

We choose τ to maximize $\tilde{W}$ under the constraint of $\lambda z^h+(1-\lambda)z^l\geq 0$. The solution must satisfy the following first-order condition:

$$\eta^t\beta\bar{w}\{(1+n)-\eta R\}-\xi_t=0, \tag{13}$$

where ξ_t is the Kuhn–Tucker multiplier associated with the constraint of $\lambda z^h+(1-\lambda)z^l\geq 0$. If $1+n>\eta R$, (13) holds if and only if $\xi_t>0$. This implies that $\lambda z^h+(1-\lambda)z^l=0$, that is, $\tau^R = (\bar{w} - \bar{c}^y)/\bar{w}$. If $1 + n = \eta R$, (13) holds if and only if $\xi_t = 0$. This implies that $\lambda z^h + (1 - \lambda) z^l \geq 0$, that is, $\tau^R \in [0, (\bar{w} - \bar{c}^y)/\bar{w}]$. Finally, if $1 + n < \eta R$, (13) implies that $\tau^R = 0$.

8.3 Proof of Proposition 2

From the capital market clearing condition (8), the steady-state level of capital is given by:

$$\bar{k} = \frac{\bar{w}(1 - A) - \bar{c}^y}{(1 + n) - \mu \bar{w}}.$$

$\bar{k} > 0$ is rewritten as:

$$A < A_{high} \equiv \frac{\bar{w} - \bar{c}^y}{\bar{w}}.$$

From the law of motion of the equilibrium tax (9), the steady-state level of tax is given by:

$$\bar{\theta} = \frac{(1 + n)A - \mu(\bar{w} - \bar{c}^y)}{(1 + n) - \mu \bar{w}}.$$

Since the denominator is positive under Assumption 2, the inequality $\bar{\theta} \geq 0$ is rewritten as $(1 + n)A - \mu(\bar{w} - \bar{c}^y) \geq 0$, that is,

$$A \geq A_{low} \equiv \frac{\mu(\bar{w} - \bar{c}^y)}{1 + n}.$$

Under Assumption 2, $\bar{\theta} \leq 1$ is rewritten as:

$$A \leq \frac{(1 + n) - \mu \bar{c}^y}{1 + n}.$$

By direct calculation, we can find that $A_{high} < \{(1 + n) - \mu \bar{c}^y\}/(1 + n)$ holds under Assumption 2. Therefore, we have $\bar{k} > 0$ and $\bar{\theta} \in [0, 1)$ if $A \in [A_{low}, A_{high})$.

8.4 Proof of Proposition 3

The conditions of the parameters β, R, γ, and n for the stability property are immediately found by equations (8) and (9). The remaining task is to provide the range of k_0 that ensures $k_{t+1} > 0$ and $\theta_t \in [0, 1)$ along the equilibrium path.

(i) The case of $\mu > 0$

The law of motion for the stock of capital (8) implies that $k_{t+1} > 0$ is equivalent to:

$$k_t > \frac{\bar{c}^y - \bar{w}}{\mu \bar{w}} + \frac{A}{\mu}. \tag{14}$$

The policy function $\theta_t = A - \mu k_t$ implies that:

$$\theta_t < A - \mu \left\{ \frac{\bar{c}^y - \bar{w}}{\mu \bar{w}} + \frac{A}{\mu} \right\} = \frac{\bar{w} - \bar{c}^y}{\bar{w}} < 1,$$

where the first inequality comes from (14) and the second inequality comes from Assumption 1. Given that $\mu > 0, \theta_t \geq 0$ is rewritten as:

$$k_t \leq \frac{A}{\mu}. \tag{15}$$

(14) and (15) lead to the set of k_t that ensures $k_{t+1} > 0$ and $\theta_t \in [0, 1)$ along the equilibrium path:

$$\frac{\bar{c}^y - \bar{w}}{\mu \bar{w}} + \frac{A}{\mu} < k_t \leq \frac{A}{\mu}. \tag{16}$$

Therefore, setting $k_t = k_0$ in (16) leads to the set of the initial condition that ensures $k_{t+1} > 0$ and $\theta_t \in [0, 1)$ for $t \geq 0$.

(ii) The case of $\mu < 0$

Taking the same procedure as in the case of $\mu > 0$, we can show that $k_{t+1} > 0$ holds if:

$$k_t < \frac{\bar{c}^y - \bar{w}}{\mu \bar{w}} + \frac{A}{\mu},$$

$\theta_t < 1$ always holds under Assumption 1; and $\theta_t \geq 0$ holds if $A/\mu \leq k_t$. We have $k_{t+1} > 0$ and $\theta_t \in [0, 1)$ along the equilibrium path if:

$$\frac{A}{\mu} \leq k_t < \frac{\bar{c}^y - \bar{w}}{\mu \bar{w}} + \frac{A}{\mu}. \tag{17}$$

Therefore, setting $k_t = k_0$ in (17) leads to the set of the initial condition that ensures $k_{t+1} > 0$ and $\theta_t \in [0, 1)$ for $t \geq 0$. ■

8.5 Proof of Proposition 4

When $\mu = 0$, the steady-state tax rate is $\bar{\theta} = A$ and is thereby independent of n and γ. In what follows, we calculate $\partial \bar{\theta}/\partial n$ and $\partial \bar{\theta}/\partial \gamma$ for the case of $\mu \neq 0$.

(i) The differentiation of $\bar{\theta}$ with respect to n leads to:

$$\frac{\partial \bar{\theta}}{\partial n} = \frac{\mu \bar{w}}{\{(1+n) - \mu \bar{w}\}^2} \left[-A + \frac{\bar{w} - \bar{c}^y}{\bar{w}} \right] = \frac{\mu \bar{w}}{\{(1+n) - \mu \bar{w}\}^2} (-A + A_{high}).$$

Given that $A < A_{high}$, $\partial \bar{\theta}/\partial n \gtreqless 0$ if and only if $\mu \gtreqless 0$.

(ii) The differentiation of $\bar{\theta}$ with respect to γ leads to:

$$\frac{\partial \bar{\theta}}{\partial \gamma} = \frac{(1+n)\bar{w}}{\{(1+n) - \mu \bar{w}\}^2} \frac{\partial \mu}{\partial \gamma} (A - A_{high}).$$

Because $A - A_{high} < 0$ and $\partial \mu / \partial \gamma < 0$ hold, we obtain $\partial \bar{\theta} / \partial \gamma > 0$. ■

8.6 Proof of Proposition 5

We differentiate I_t with respect to k_t and rearrange the terms to obtain:

$$\begin{aligned} & \frac{1}{\beta}[u(\bar{c}^y) + \beta c_{t+1}^{ol}]^2 \frac{\partial I_t}{\partial k_t} \\ = \; & R \frac{\partial \theta(k_t)}{\partial k_t}(-w^h + w^l)\left[(u(\bar{c}^y) - \beta R \bar{c}^y) + (1+n)\theta(k_{t+1})(1-\alpha)\bar{w}\right] \\ & +(1+n)\frac{\partial \theta(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t}(w^h - w^l)\left[\alpha(u(\bar{c}^y) - \beta R \bar{c}^y) - (1-\alpha)\bar{w}\beta R(1 - \theta(k_t))\right]. \end{aligned} \tag{18}$$

From the policy function and the capital market clearing condition, $\partial \theta(k_t)/\partial k_t$ and $\partial k_{t+1}/\partial k_t$ are given by

$$\begin{aligned} \partial \theta(k_t)/\partial k_t &= \partial \theta(k_t)/\partial k_t = -\mu, \\ \partial k_{t+1}/\partial k_t &= \mu \bar{w}/(1+n). \end{aligned}$$

(i) Suppose that $\mu > 0$. We can rewrite (18) as

$$\begin{aligned} \frac{1}{\beta}[u(\bar{c}^y) + \beta c_{t+1}^{ol}]^2 \frac{\partial I_t}{\partial k_t} = \; & R \frac{\partial \theta(k_t)}{\partial k_t}(-w^h + w^l)(u(\bar{c}^y) - \beta R \bar{c}^y) \\ & + R \frac{\partial \theta(k_t)}{\partial k_t}(-w^h + w^l)(1+n)\theta(k_{t+1})(1-\alpha)\bar{w} \\ & +(1+n)\frac{\partial \theta(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t}(w^h - w^l)\alpha(u(\bar{c}^y) - \beta R \bar{c}^y) \\ & +(1+n)\frac{\partial \theta(k_{t+1})}{\partial k_{t+1}} \frac{\partial k_{t+1}}{\partial k_t}(w^h - w^l)(-1)(1-\alpha)\bar{w}\beta R(1 - \theta(k_t)), \end{aligned}$$

where the first and second terms are nonnegative, the third term is nonpositive, and the fourth term is positive. Therefore, $\partial I_t / \partial k_t > 0$ holds if the sum of the first and the third terms are nonnegative. The sum is written as:

$$\begin{aligned}
& R\frac{\partial\theta(k_t)}{\partial k_t}(-w^h+w^l)(u(\bar{c}^y)-\beta R\bar{c}^y)+(1+n)\frac{\partial\theta(k_{t+1})}{\partial k_{t+1}}\frac{\partial k_{t+1}}{\partial k_t}(w^h-w^l)\alpha(u(\bar{c}^y)-\beta R\bar{c}^y) \\
= \;& (-\mu)(w^h-w^l)(u(\bar{c}^y)-\beta R\bar{c}^y)\left[R-(1+n)\frac{\mu\bar{w}}{1+n}\alpha\right] \\
= \;& (-\mu)(w^h-w^l)(u(\bar{c}^y)-\beta R\bar{c}^y)\left[(1-\alpha)R+\alpha\frac{\gamma}{\beta}\right] \\
\geq \;& 0,
\end{aligned}$$

where the inequality comes from $-\mu<0, w^h-w^l>0$ and $u(\bar{c}^y)-\beta R\bar{c}^y\geq 0$. Thus, $\partial I_t/\partial k_t>0$ if $\mu>0$.

(ii) Suppose that $\mu<0$. Dividing both sides of (18) by $(-\mu)(w^h-w^l)>0$, we have:

$$\begin{aligned}
& \frac{1}{(-\mu)(w^h-w^l)}\frac{1}{\beta}[u(\bar{c}^y)+\beta c^{ol}_{t+1}]^2\frac{\partial I_t}{\partial k_t} \\
= \;& (-R)\left[(u(\bar{c}^y)-\beta R\bar{c}^y)+(1+n)\theta(k_{t+1})(1-\alpha)\bar{w}\right] \\
& +(1+n)\frac{\mu\bar{w}}{1+n}\left[\alpha(u(\bar{c}^y)-\beta R\bar{c}^y)-(1-\alpha)\bar{w}\beta R(1-\theta(k_t)\right] \\
= \;& (-1)(R-\mu\bar{w}\alpha)(u(\bar{c}^y)-\beta R\bar{c}^y)-R(1-\alpha)\bar{w}\left[(1+n)\theta(k_{t+1})+\mu\bar{w}\beta(1-\theta(k_t))\right].
\end{aligned}$$

Under the assumption of $\mu<0$, we have:

$$0>\mu\bar{w}\beta(1-\theta(k_t))>\mu\bar{w}(1-\theta(k_t)),$$

implying that

$$\begin{aligned}
& -R(1-\alpha)\bar{w}\left[(1+n)\theta(k_{t+1})+\mu\bar{w}\beta(1-\theta(k_t))\right] \\
< \;& -R(1-\alpha)\bar{w}\left[(1+n)\theta(k_{t+1})+\mu\bar{w}(1-\theta(k_t))\right].
\end{aligned}$$

Therefore, we obtain:

$$\begin{aligned}
& \frac{1}{(-\mu)(w^h-w^l)}\frac{1}{\beta}[u(\bar{c}^y)+\beta c^{ol}_{t+1}]^2\frac{\partial I_t}{\partial k_t}\frac{1}{(-\mu)(w^h-w^l)}\frac{1}{\beta}[u(\bar{c}^y)+\beta c^{ol}_{t+1}]^2\frac{\partial I_t}{\partial k_t} \\
< \;& (-1)(R-\mu\bar{w}\alpha)(u(\bar{c}^y)-\beta R\bar{c}^y)-R(1-\alpha)\bar{w}\left[(1+n)\theta(k_{t+1})+\mu\bar{w}\beta(1-\theta(k_t))\right] \\
= \;& (-1)(R-\mu\bar{w}\alpha)(u(\bar{c}^y)-\beta R\bar{c}^y) \\
& -R(1-\alpha)\bar{w}\left[(1+n)\left\{A-\frac{\mu(\bar{w}-\bar{c}^y)}{1+n}+\frac{\mu\bar{w}}{1+n}\theta(k_t)\right\}+\mu\bar{w}\beta(1-\theta(k_t))\right],
\end{aligned}$$

where the equality in the last line comes from the law of motion of the equilibrium tax.

Rearranging the terms of this equation, we obtain:

$$\begin{aligned}
& \frac{1}{(-\mu)(w^h-w^l)}\frac{1}{\beta}[u(\bar{c}^y)+\beta c^{ol}_{t+1}]^2\frac{\partial I_t}{\partial k_t}\frac{1}{(-\mu)(w^h-w^l)}\frac{1}{\beta}[u(\bar{c}^y)+\beta c^{ol}_{t+1}]^2\frac{\partial I_t}{\partial k_t} \\
< \; & (-1)(R-\mu\bar{w}\alpha)(u(\bar{c}^y)-\beta R\bar{c}^y)-R(1-\alpha)\bar{w}\{(1+n)A+\mu\bar{c}^y\},
\end{aligned}$$

where the first term on the right-hand side is nonpositive. This implies that $\partial I_t/\partial k_t \leq 0$ holds if $(1+n)A+\mu\bar{c}^y \geq 0$, i.e.,

$$A \geq \frac{-\mu\bar{c}^y}{1+n}. \tag{19}$$

Given the definitions of A_{low} and A_{high}, we always have

$$A_{low} \equiv \frac{\mu(\bar{w}-\bar{c}^y)}{1+n} < \frac{-\mu\bar{c}^y}{1+n},$$

and we have $-\mu\bar{c}^y/(1+n) < A_{high} \equiv (\bar{w}-\bar{c}^y)/\bar{w}$ if and only if:

$$\bar{c}^y < \frac{1+n}{(1+n)-\mu\bar{w}}\bar{w}. \tag{20}$$

Therefore, from (19) and (20), we can conclude that:

$$\partial I_t/\partial k_t \leq 0 \text{ if } \bar{c}^y < \frac{1+n}{(1+n)-\mu\bar{w}}\bar{w} \text{ and } A \in \left[\frac{-\mu\bar{c}^y}{1+n}, A_{high}\right).$$

■

8.7 Proof of Proposition 6

We differentiate $\bar{I}$ with respect to n and rearrange the terms to obtain:

$$\begin{aligned}
& \frac{1}{\beta}[u(\bar{c}^y)+\beta\bar{c}^{ol}]^2\frac{\partial\bar{I}}{\partial n} \\
= \; & \left[-R\frac{\partial\theta(\bar{k})}{\partial n}w^h+\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}\alpha w^h\right] \\
& \times\left[(u(\bar{c}^y)-\beta R\bar{c}^y)+\beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right] \\
& +\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}(1-\alpha)\bar{w}\beta\left[R(1-\theta(\bar{k}))w^l+(1+n)\theta(\bar{k})\alpha w^l\right] \\
& -\left[-R\frac{\partial\theta(\bar{k})}{\partial n}w^l+\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}\alpha w^l\right] \\
& \times\left[(u(\bar{c}^y)-\beta R\bar{c}^y)+\beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right] \\
& -\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}(1-\alpha)\bar{w}\beta\left[R(1-\theta(\bar{k}))w^h+(1+n)\theta(\bar{k})\alpha w^h\right],
\end{aligned}$$

which is reduced to:

$$
\begin{aligned}
&\frac{1}{\beta}[u(\bar{c}^y)+\beta\bar{c}^{ol}]^2\frac{\partial \bar{I}}{\partial n} \\
= &\left(-R\frac{\partial\theta(\bar{k})}{\partial n}\right)(w^h-w^l)\left[(u(\bar{c}^y)-\beta R\bar{c}^y)+\beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right] \\
&+\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}\alpha(w^h-w^l)\left[(u(\bar{c}^y)-\beta R\bar{c}^y)+\beta(1+n)\theta(\bar{k})(1-\alpha)\bar{w}\right] \\
&-\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}(1-\alpha)\bar{w}\beta(w^h-w^l)\left[R(1-\theta(\bar{k}))+(1+n)\theta(\bar{k})\alpha\right]. \quad (21)
\end{aligned}
$$

(i) Suppose that $\alpha=0$. (21) is reduced to:

$$
\begin{aligned}
\frac{1}{\beta}[u(\bar{c}^y)+\beta\bar{c}^{ol}]^2\frac{\partial \bar{I}}{\partial n} = (w^h-w^l)\Bigg[&\underbrace{\left(-R\frac{\partial\theta(\bar{k})}{\partial n}\right)\{(u(\bar{c}^y)-\beta R\bar{c}^y)+\beta(1+n)\theta(\bar{k})\bar{w}\}}_{(*1)} \\
&\underbrace{-\left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\right\}\bar{w}\beta R(1-\theta(\bar{k}))}_{(*2)}\Bigg]. \quad (22)
\end{aligned}
$$

If $\mu>0, \partial\theta(\bar{k})/\partial n>0$ (Proposition 4(i)). This implies that the sign of each term in the square brackets is negative and thus $\partial\bar{I}/\partial n<0$. If $\mu=0$, $\partial\theta(\bar{k})/\partial n=0$ (Proposition 4(i)). Given this, (22) is rewritten as:

$$\frac{1}{\beta}[u(\bar{c}^y)+\beta\bar{c}^{ol}]^2\frac{\partial \bar{I}}{\partial n}=(w^h-w^l)(-1)\theta(\bar{k})\bar{w}\beta R(1-\theta(\bar{k}))<0.$$

If $\mu<0$, the sign of the term (*1) in (22) is positive; the sign of the term (*2) in (22) is nonnegative if $(1+n)(\partial\theta(\bar{k})/\partial n)+\theta(\bar{k})\le 0$. Therefore, $\partial\bar{I}/\partial n>0$ if $(1+n)(\partial\theta(\bar{k})/\partial n)+\theta(\bar{k})\le 0$. The direct calculation leads to:

$$
\begin{aligned}
(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k}) &= \frac{(1+n)A-\mu(\bar{w}-\bar{c}^y)}{(1+n)-\mu\bar{w}}+(1+n)\frac{-A\mu\bar{w}+\mu(\bar{w}-\bar{c}^y)}{\{(1+n)-\mu\bar{w}\}^2} \\
&= \frac{(1+n)A\{(1+n)-2\mu\bar{w}\}+\mu(\bar{w}-\bar{c}^y)\mu\bar{w}}{\{(1+n)-\mu\bar{w}\}^2}.
\end{aligned}
$$

This implies that:

$$(1+n)\frac{\partial\theta(\bar{k})}{\partial n}+\theta(\bar{k})\gtreqless 0 \Leftrightarrow A\gtreqless \hat{A}\equiv\frac{-\mu(\bar{w}-\bar{c}^y)\mu\bar{w}}{(1+n)\{(1+n)-2\mu\bar{w}\}}. \quad (23)$$

The critical level of A, $\hat{A}$, is larger than A_{low} under the assumption of $(1+n) - \mu\bar{w} > 0$ (Assumption 2), and is smaller than A_{high} since $\hat{A} < 0 < A_{high}$ under the assumption of $\mu < 0$. Therefore, $(1+n)(\partial\theta(\bar{k})/\partial n) + \theta(\bar{k}) \leq 0$ holds if $\mu < 0$ and $A \in (A_{low}, \hat{A}]$.

(ii) Suppose that $\alpha = 1$. (21) is reduced to:

$$\begin{aligned}\frac{1}{\beta}[u(\bar{c}^y) + \beta\bar{c}^{ol}]^2\frac{\partial \bar{I}}{\partial n} &= (w^h - w^l)\left[\left(-R\frac{\partial\theta(\bar{k})}{\partial n}\right)(u(\bar{c}^y) - \beta R\bar{c}^y)\right. \\ &\quad \left. + \left\{(1+n)\frac{\partial\theta(\bar{k})}{\partial n} + \theta(\bar{k})\right\}(u(\bar{c}^y) - \beta R\bar{c}^y)\right] \\ &= (w^h - w^l)(u(\bar{c}^y) - \beta R\bar{c}^y)\left[(-R + (1+n))\frac{\partial\theta(\bar{k})}{\partial n} + \theta(\bar{k})\right].\end{aligned}$$

Because $(w^h - w^l)(u(\bar{c}^y) - \beta R\bar{c}^y) \geq 0$ holds (Assumption 3), the sign of the term $(-R + (1+n))(\partial\theta(\bar{k})/\partial n) + \theta(\bar{k})$ is a key to determine the sign of $\partial\bar{I}/\partial n$.

Recall that $\partial\theta(\bar{k})/\partial n \gtreqless 0$ if and only if $\mu \gtreqless 0$ (Proposition 4 (i)). Given this result, we have $(-R + (1+n))(\partial\theta(\bar{k})/\partial n) + \theta(\bar{k}) > 0$ if $\mu > 0$ and $1+n \geq R$; or if $\mu < 0$ and $1+n \leq R$. If $\mu = 0$, (21) is further reduced to:

$$\frac{1}{\beta}[u(\bar{c}^y) + \beta\bar{c}^{ol}]^2\frac{\partial \bar{I}}{\partial n} = (w^h - w^l)(u(\bar{c}^y) - \beta R\bar{c}^y)\theta(\bar{k}) \geq 0,$$

where the inequality comes from $w^h - w^l > 0$ and $u(\bar{c}^y) - \beta R\bar{c}^y \geq 0$ (Assumption 3). ■

References

[1] Auerbach, A.J. and Kotlikoff, L.J., 1987. *Dynamic Fiscal Policy.* Cambridge, New York and Melbourne: Cambridge University Press.

[2] Azariadis, C. and Galasso, V., 2002. Fiscal constitutions. *Journal of Economic Theory* 103, 255–281.

[3] Bassetto, M., 2008. Political economy of taxation in an overlapping-generations economy. *Review of Economic Dynamics* 11, 18–43.

[4] Bellettini, G. and Berti Ceroni, C. 2007. Income distribution, borrowing constraints and redistributive policies. *European Economic Review* 51, 625–645.

[5] Casamatta, G., Cremer, H. and Pestieau, P., 2000. The Political economy of social security. *Scandinavian Journal of Economics* 102, 503–522.

[6] Casamatta, G., Cremer, H. and Pestieau, P., 2001. Demographic shock and social security: a political economy perspective. *International Tax and Public Finance* 8, 417–431.

[7] Conde-Ruiz, J.I. and Profeta, P., 2007. The redistributive design of social security systems. *Economic Journal*11, 686–712.

[8] Cremer, H., De Donder, P., Maldonado, D. and Pestieau, P., 2007. Voting over type and generosity of a pension system when some individuals are myopic. *Journal of Public Economics* 91, 2041–2061.

[9] Deaton, A. and Paxon, C., 1994. Intertemporal choice and inequality. *Journal of Political Economy* 102, 437–467.

[10] Forni, L., 2005. Social security as Markov equilibrium in OLG models. *Review of Economic Dynamics* 8, 178–194.

[11] Galasso, V. and Profeta, P., 2004. Lessons for an aging society: the political sustainability of social security systems. *Economic Policy* 19, 1–29.

[12] Galasso, V. and Profeta, P., 2007. How does ageing affect the welfare state? *European Journal of Political Economy* 23, 554–563

[13] Gonzalez-Eiras, M. and Niepelt, D., 2008. The future of social security. *Journal of Monetary Economics* 55, 197–218

[14] Grossman, G. and Helpman, E., 1998. Intergenerational redistribution with short-lived governments. *Economic Journal* 108, 1299–1329.

[15] Hassler, J., Storesletten, K., Rodriguez Mora, J.V. and Zilibotti, F., 2003. The survival of the welfare state. *American Economic Review* 93, 87–112.

[16] Koethenbuerger, M., Poutvaara, P. and Profeta, P., 2008. Why are more redistributive social security systems smaller? A median voter approach. *Oxford Economic Papers* 60, 275–292.

[17] Hassler, J., Krusell, P., Storesletten, K. and Zilibotti, F., 2005. The dynamics of government. *Journal of Monetary Economics* 52, 1331–1358.

[18] Lindbeck, A. and Weibull, J., 1987. Balanced-budget redistribution as the outcome of political competition. *Public Choice* 52, 273–297.

[19] Ohtake, F. and Saito, M., 1998. Population aging and consumption inequality in Japan. *Review of Income and Wealth* 44, 361–381.

[20] Razin, A., Sadka, E. and Swagel, P., 2002. The aging population and the size of the welfare state. *Journal of Political Economy* 110, 900–918.

[21] Song, Z., 2007. The dynamics of inequality and social security. Unpublished manuscript, Fudan University.

http://homepage.fudan.edu.cn/~zsong/Michael_files/Research/Inequality.pdf

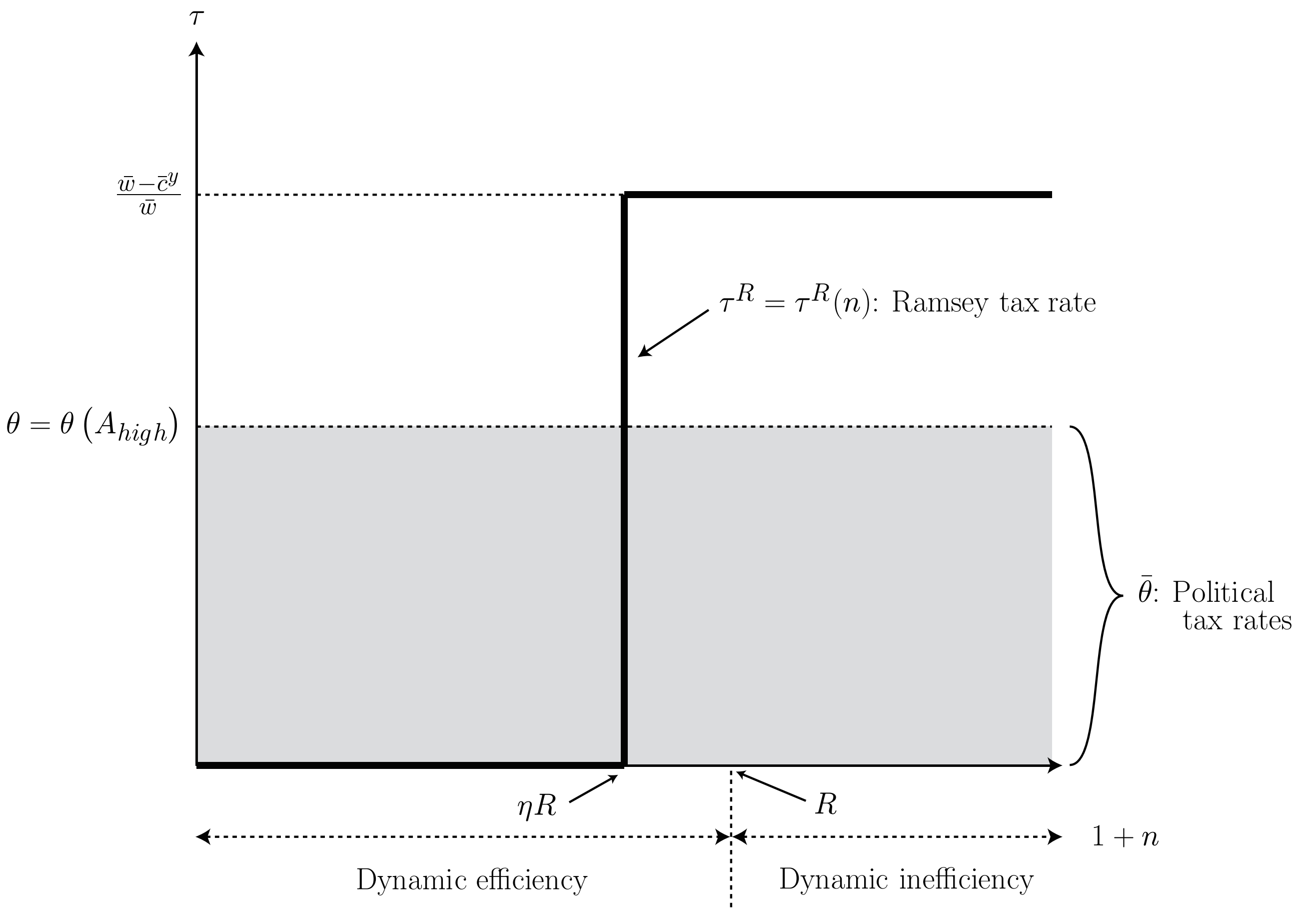

Figure 1: The figure illustrates the Ramsey tax rate and compares it to the politically determined tax rates.

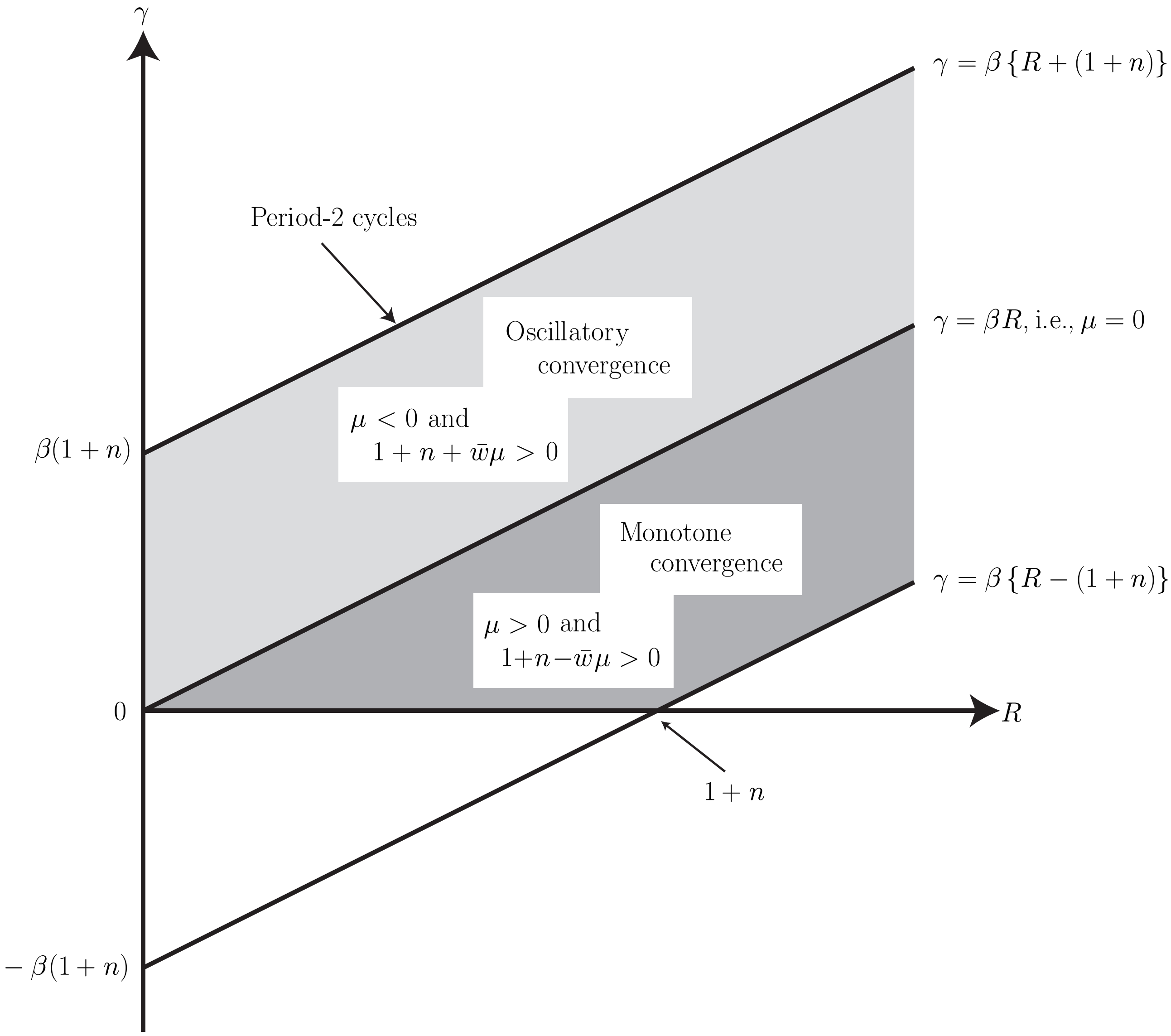

Figure 2: Stability property of political equilibrium paths.

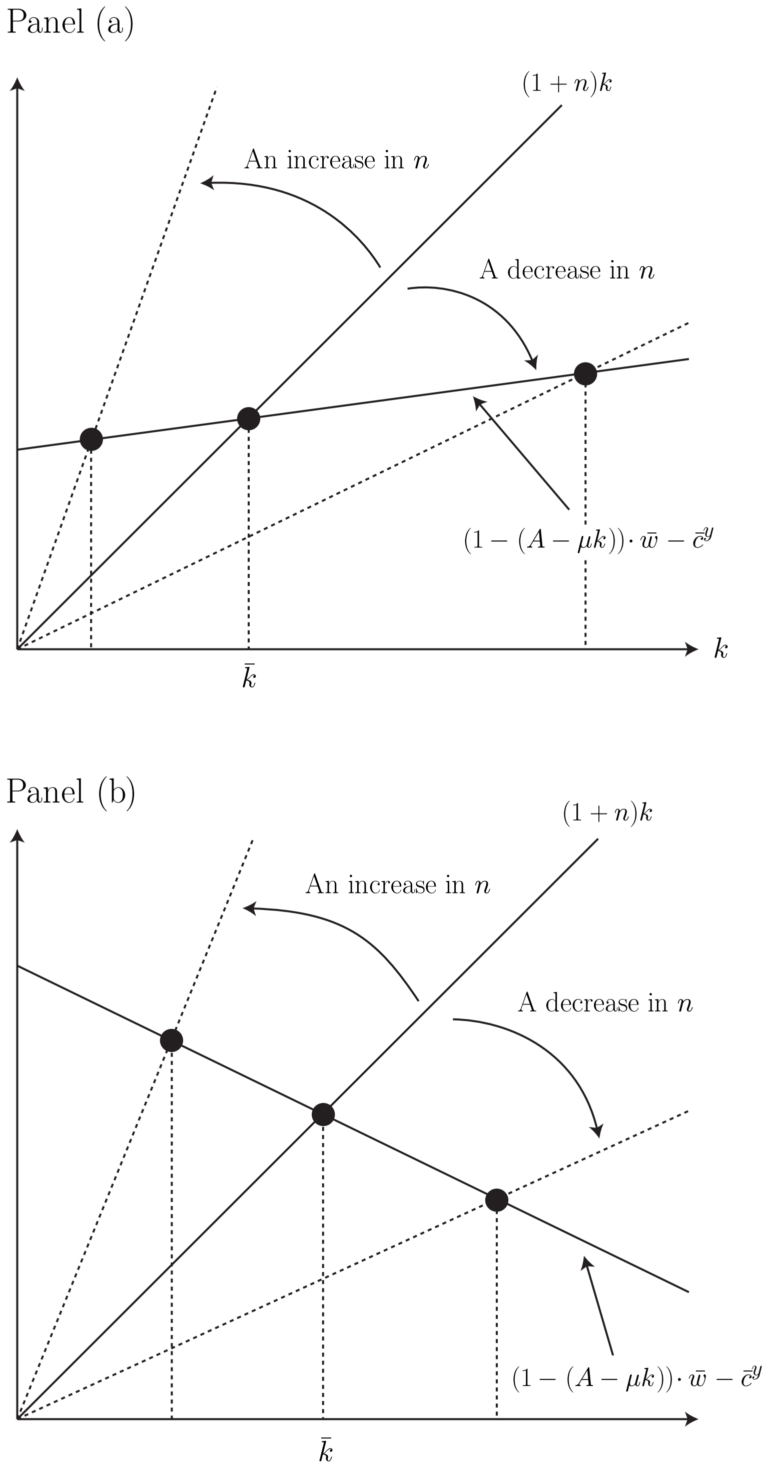

Figure 3: The figure displays the effect of a change in a population growth rate on the steady-state level of capital. Panel (a) illustrates the case of $\mu > 0$. Panel (b) illustrates the case of $\mu < 0$.