

No. 200

Optimal pay-as-you-go social security
with endogenous retirement

Koichi Miyazaki
January 18, 2016

Optimal pay-as-you-go social security with endogenous retirement*

Koichi Miyazaki[†]

Faculty of Economics

Kagawa University

January 18, 2016

Abstract

This paper considers an overlapping-generations model with pay-as-you-go social security and retirement decision making by an old agent. In addition, the paper assumes that labor productivity depreciates. Under this setting, socially optimal allocations are examined. The first-best allocation is an allocation that maximizes welfare when a social planner distributes resources and forces an old agent to work and retire as she wants. The second-best allocation is an allocation that maximizes welfare when she can use only pay-as-you-go social security in a decentralized economy. The paper finds a range of an old agent's labor productivity such that the first-best allocation is achieved in the decentralized economy. This differs from the finding in Michel and Pestieau [“Social security and early retirement in an overlapping-generations growth model”, *Annals of Economics & Finance*, 2013] that the first-best allocation cannot be achieved in the decentralized economy.

Keywords: Optimal pay-as-you-go social security, endogenous retirement, depreciation of labor productivity, first-best and second-best allocation

JEL Classification: D91, H21, H55, J26

*This paper was previously circulated under the title, “Optimal pay-as-you-go social security when retirement is endogenous and labor productivity depreciates.” The author began this work when he was at National Taiwan University. The author thanks Been-Lon Chen, Hung-Ju Chen, Nan-Kuang Chen, Chih-Fang Lai, Yiting Li, Masanori Kashiwagi, Yi-Chan Tsai, and the seminar participants at National Taiwan University for their helpful comments and suggestions. The author acknowledges financial support from the Ministry of Science and Technology of Taiwan (Grant number:102-2410-H-002-020). Of course, all errors are the author's responsibility.

[†]Corresponding author, Mailing address: 2-1, Saiwaicho, Takamatsu, Kagawa 7608523, Japan, E-mail address: kmiyazaki@ec.kagawa-u.ac.jp, Tel: +81-87-832-1851, Fax: +81-87-832-1820

1 Introduction

This paper studies the optimal pay-as-you-go social security in a Diamond (1965)-type overlapping-generations model with endogenous retirement and the depreciation of an agent's labor productivity. It is known that in a standard Diamond-type overlapping-generations model, the optimal allocation or the golden rule is achieved in the decentralized economy by using pay-as-you-go social security.¹ In this standard model, an agent will retire once s/he becomes old. This paper loosens this assumption by adding an old agent's decision making on retirement age. This is a natural extension because several empirical works have shown that social security benefits affect an individual's decision making on retirement (see, for instance, Gruber and Wise (1999) and Feng and Pestieau (2007).)

Michel and Pestieau (2013) consider a similar environment to this paper's and investigate the first-best allocation and the second-best allocation. The first-best allocation is an allocation that maximizes welfare subject to the resource constraint and an old agent's time constraint.² In other words, a social planner can distribute resources and force an old agent to work and retire as she wants. In contrast, the social planner can redistribute resources across generations through taxation but cannot force an old agent to work or retire as she wants in a decentralized economy. The second-best allocation is an (equilibrium) allocation that maximizes welfare in the decentralized economy.³ The comparison between these two social optima is meaningful because in the real world, the redistribution of resources is feasible through taxation, whereas no one can force others to work or retire because of civil liberties and other factors. Michel and Pestieau (2013) show that the first-best allocation cannot be achieved in the decentralized economy.

This paper imposes the assumption that labor productivity depreciates over age in the model in Michel and Pestieau (2013). This is a natural assumption.⁴ Under this setting, this paper finds the range of an old agent's labor productivity in which the first-best allocation is achieved in the decentralized economy. As in Michel and Pestieau (2013), this paper considers an additively separable natural log utility function and a Cobb-Douglas production function because of the tractability of the model.⁵

Why can the first-best allocation be achieved in the decentralized economy when the depreciation of labor productivity is taken into account? Consider the case in which an old agent's labor productivity is low (relative

¹Optimality is defined by using an agent's lifetime utility in the steady state. For the formal definition, see Section 4.

²Welfare is defined by an agent's lifetime utility in the steady state. The formal definition is given in Section 4.

³An equilibrium concept in this paper is a standard perfect foresight competitive equilibrium.

⁴For instance, Figure 2 in Kitao (2014) shows that labor productivity reaches its peak at age 50 and starts to decrease over age.

⁵Michel and Pestieau (2013) argue that "... the reason for the omission [of the study of the interaction between social security and retirement] [is] in the analytical difficulty of studying the dynamics of overlapping-generations models with endogenous labor supply."

to a young agent's labor productivity) and the capital share in the production function is also low. Since an old agent's labor productivity is low, an old agent fully retires in the first-best allocation. Moreover, capital accumulation is set to satisfy the golden rule, which is $f'(k) = 1$, in the first-best allocation.⁶ Consider the decentralized economy with a zero payroll tax rate, which is called a laissez-faire economy. When the capital share is low enough, capital is over-accumulated in equilibrium; that is, $r = f'(k) < 1$, where r is the interest rate, and an old agent works in equilibrium because of the low return from savings. As the payroll tax rate increases, the labor supplied by an old agent decreases and savings by a young agent decreases. How capital per unit of effective labor, k , changes is ambiguous in general, but under this paper's setting, k is constant as long as an old agent's labor supply is positive. Once an old agent fully retires, k starts to decrease as the payroll tax rate increases. Since an old agent fully retires in the first-best allocation and the labor supply is bounded below by 0, the social planner will set a sufficiently high payroll tax rate that an old agent fully retires and the golden rule is achieved in the decentralized economy. Thus, the first-best allocation is achieved in the decentralized economy.

This argument does not hold once an old agent partially retires in the first-best allocation. A sufficiently high payroll tax rate can achieve the golden rule, whereas an old agent works less than the first-best level. In contrast, a low tax rate can achieve the first-best level of an old agent's labor supply, while capital is over-accumulated. Therefore, the first-best allocation is not achieved in the decentralized economy, as Michel and Pestieau (2013) suggest.

Another finding that differs from Michel and Pestieau (2013) is that the second-best payroll tax rate can achieve the golden rule level of capital with an under-supply of an old agent's labor. In Michel and Pestieau (2013), the second-best payroll tax rate is lower than the rate at which the golden rule is achieved. In other words, when labor productivity does not depreciate at all, the labor supply by an old agent is at the first-best level and capital is over-accumulated in the second-best allocation. This paper presents a case in which capital per unit of effective labor is at the first-best level and the labor supplied by an old agent is lower than the first-best level in the second-best allocation.

Before Michel and Pestieau (2013), a theoretical analysis in a dynamic general equilibrium setting was performed by Hu (1979). He analyzes the effects of pay-as-you-go social security on an old agent's retirement decision making, equilibrium, and welfare in a more general two-period overlapping-generations setting. Under this setting, he shows that more generous social security benefits make an old agent retire earlier, which is consistent with the empirical findings of Gruber and Wise (1999). An old agent's retirement decision making has

⁶ k represents capital per unit of effective labor. Formally, $k = \frac{\text{Amount of savings}}{\text{Effective labor}}$. Moreover, the population growth rate is set at 1 here.

been analyzed in various frameworks. For example, Aísa et al. (2012) focus on the effects of life expectancy on the labor supply of the elderly, while Mizuno and Yakita (2013) focus on fertility decisions. Some papers investigate how exogenously determined retirement age affects the economy. For instance, Sala-i-Martin (1996) and Miyazaki (2014) show the possibility that early retirement induced by pay-as-you-go social security increases aggregate output. In this sense, pay-as-you-go social security is a mechanism that stimulates economic growth by having less productive old people retire early. This paper considers endogenous retirement instead of exogenous retirement and shows that pay-as-you-go social security is a mechanism that improves *welfare* by having less productive old people retire early in some cases.

The remainder of this paper is organized as follows. Section 2 describes the model in full, and Section 3 analyzes the equilibrium. Sections 4 analyzes welfare, and Section 5 discusses the findings in this paper. Section 6 concludes.

2 Model

The model is based on a Diamond-type overlapping-generations model, and the paper extends it by adding an old agent's decision making on his/her labor supply. The formal description is as follows.

Time is discrete and continues forever, $t = 1, 2, \dots$

2.1 Agent

An agent lives for two periods: young and old. A young agent is endowed with one unit of time, and supplies labor inelastically. After production occurs, a young agent receives labor income, w_t . A young agent also has to pay the payroll tax, which is denoted by $\tau \in [0, 1]$. With disposable income, $(1 - \tau)w_t$, a young agent decides how much to consume, c_t^y , and how much to save, s_t . Thus, a young agent's budget constraint is

$$c_t^y + s_t = (1 - \tau)w_t. \quad (1)$$

A young agent's saving at date t , s_t , is used for production at $t + 1$ as a capital good without any transformation cost from one unit of a consumption good to a capital good. A young agent at date t saves s_t and becomes old at $t + 1$, when s/he receives the interest income, $r_{t+1}s_t$, where r_{t+1} is the interest rate. An old agent is also endowed with one unit of time. If s/he works for $1 - l_{t+1}^o$ unit of time, s/he receives labor income, $(1 - l_{t+1}^o)\theta w_{t+1}$, where $\theta \in (0, 1]$ is an old agent's labor productivity relative to that of a young agent. In addition, s/he has to pay the

payroll tax, and thus his/her disposable labor income is $(1 - \tau)(1 - l_{t+1}^o)\theta w_{t+1}$.⁷ Once an old agent retires, s/he is eligible to receive the social security benefit, $l_{t+1}^o P_{t+1}$. If an old agent works for $1 - l_{t+1}^o$ units of time, because s/he works when s/he is young, the age $1 + (1 - l_{t+1}^o) = 2 - l_{t+1}^o$ can be considered as his/her retirement age. An old agent's budget constraint in period $t + 1$ is

$$c_{t+1}^o = r_{t+1}s_t + (1 - \tau)(1 - l_{t+1}^o)\theta w_{t+1} + l_{t+1}^o P_{t+1}. \quad (2)$$

From Equations (1) and (2), an agent's lifetime budget constraint is

$$c_t^y + \frac{c_{t+1}^o}{r_{t+1}} = (1 - \tau)w_t + \frac{(1 - \tau)(1 - l_{t+1}^o)\theta w_{t+1} + l_{t+1}^o P_{t+1}}{r_{t+1}}. \quad (3)$$

An agent's lifetime utility is represented by the following function:

$$U(c_t^y, c_{t+1}^o, l_{t+1}^o) := \ln(c_t^y) + \beta \ln(c_{t+1}^o) + \gamma \ln(l_{t+1}^o),$$

where $\beta \in (0, 1]$ and $\gamma \in (0, 1]$ are the weights on the preference over consumption when old and leisure when old, respectively, relative to consumption when young. Hence, an agent's problem is

$$\begin{aligned} \max_{c_t^y, c_{t+1}^o, l_{t+1}^o} \quad & U(c_t^y, c_{t+1}^o, l_{t+1}^o) \\ \text{s.t.} \quad & \text{Equation (3)} \\ & 0 \leq l_{t+1}^o \leq 1. \end{aligned} \quad (4)$$

The last constraint is a time constraint for an old agent. Note that since the utility from leisure is represented by $\gamma \ln(l_{t+1}^o)$, in the solution to the maximization problem, $l_{t+1}^o > 0$ must hold. Therefore, Equation (4) can be replaced by $l_{t+1}^o \leq 1$.

An initially old agent has saved $s_0 > 0$ at date 0. It is assumed that the population is constant forever and is normalized to one.⁸

⁷The tax rate can depend on age and labor income, and such a tax scheme is richer than this paper's. As will be shown in the following section, however, the current tax scheme is sufficient to show that the first-best allocation can be achieved in a decentralized economy.

⁸This assumption is not qualitatively or quantitatively harmful for the following results as long as the growth rate is given exogenously.

2.2 Firm

At each date, the representative firm produces a single unit of output using labor and capital. Let L_t and K_t be the aggregate effective labor and aggregate capital, respectively. A production function is $F(K, L) = K^\alpha L^{1-\alpha}$, where $\alpha \in (0, 1)$. I assume that factor markets are perfectly competitive and that capital is assumed to be fully depreciated once production occurs. Thus, a profit-maximizing firm's problem is

$$\max_{K_t, L_t} \{F(K_t, L_t) - w_t L_t - r_t K_t\}$$

for all t , where r_t is the rental rate of capital. Let k_t be capital per unit of effective labor at date t , that is, $k_t := K_t/L_t$. Let $f(k) := k^\alpha$. Then, from the firm's problem,

$$r_t = f'(k_t) = \alpha k_t^{\alpha-1} \quad (5)$$

and

$$w_t = f(k_t) - k_t f'(k_t) = (1 - \alpha) k_t^\alpha. \quad (6)$$

2.3 Government

The government in every period balances the pay-as-you-go social security budget according to the following formula: for all $t \geq 0$,

$$l_{t+1}^o P_{t+1} = \tau w_{t+1} [1 + \theta(1 - l_{t+1}^o)]. \quad (7)$$

This equation implicitly ignores the case of $l_{t+1}^o = 0$. Since an old agent does not choose $l_{t+1}^o = 0$ in equilibrium, the government budget constraint is well defined.

2.4 Equilibrium and the steady state

An equilibrium concept is a standard *perfect foresight competitive equilibrium*. The formal definition is given below.

Definition 1. Given the initial capital stock, $K_1 = s_0$, the payroll tax rate, $\tau \in [0, 1)$, and the productivity of an old agent, $\theta \in (0, 1]$, an **equilibrium** consists of a sequence of an agent's consumption when young and when

old, savings, and leisure when old, $(c_1^o, l_1^o, (c_t^y, s_t, c_{t+1}^o, l_{t+1}^o)_{t=1}^\infty)$; a sequence of prices, $(r_t, w_t)_{t=1}^\infty$; a sequence of factors of production, $(K_t, L_t)_{t=1}^\infty$; and a sequence of social security benefits, $(P_t)_{t=1}^\infty$; such that:

1. Given the payroll tax rate, prices and social security benefits, $(c_t^y, s_t, c_{t+1}^o, l_{t+1}^o)$ solves a young agent's problem for all $t \geq 1$.
2. The initially old agent solves

$$\begin{aligned} \max_{c_1^o, l_1^o} \quad & \beta \ln(c_1^o) + \gamma \ln(l_1^o) \\ \text{s.t.} \quad & c_1^o = (1 - \tau)(1 - l_1^o)\theta w_1 + l_1 P_1 + r_1 s_0, \quad 0 \leq l_1^o \leq 1. \end{aligned}$$

3. The factor prices are determined by Equations (5) and (6).
4. A capital market and a labor market clear, that is, $K_t = s_{t-1}$ and $L_t = 1 + \theta(1 - l_t^o)$.
5. The government budget constraint is satisfied, that is, (7) holds.

A combination of time-invariant variables is called a *steady state*.

3 Equilibrium analysis

In this section, an equilibrium is characterized.

Proposition 1. *Suppose that the payroll tax rate satisfies*

$$\tau \geq \hat{\tau} := \frac{\beta(1 - \alpha)\theta - \gamma\alpha}{(1 - \alpha)[\beta(1 + \theta) + \gamma]}. \quad (8)$$

Then, in equilibrium, an old agent fully retires, $l_{t+1}^o = 1$ for all t , and the equilibrium economic growth is characterized by

$$k_{t+1} = \frac{\alpha\beta(1 - \tau)(1 - \alpha)}{(1 + \beta)\alpha + \tau(1 - \alpha)} k_t^\alpha.$$

Furthermore, there exists a unique non-zero steady state, $k_f(\tau)$, where

$$k_f(\tau) := \left[\frac{\alpha\beta(1 - \tau)(1 - \alpha)}{(1 + \beta)\alpha + \tau(1 - \alpha)} \right]^{\frac{1}{1 - \alpha}}, \quad (9)$$

and it is globally stable.⁹

Proof. See the Appendix.

Q.E.D.

If Equation (8) is satisfied, then an old agent will retire fully in equilibrium. The first implication is that if the payroll tax rate is too high, since an old agent loses the incentive to work, s/he will retire fully. However, Equation (8) holds even if τ is low.

If $\beta(1 - \alpha)\theta \leq \gamma\alpha$ or $\alpha \geq \frac{\beta\theta}{\beta\theta + \gamma}$ holds, then the RHS of Equation (8) is less than or equal to zero. This implies that for any $\tau \in [0, 1)$, an old agent fully retires in equilibrium. One case in which $\beta(1 - \alpha)\theta < \gamma\alpha$ holds is where γ is sufficiently large. That is, for an agent, leisure in the second period is more valuable relative to consumption when young and when old. Therefore, even though the payroll tax rate is low enough, an old agent will retire fully. Another case is where α is large. Since $1 - \alpha$ represents the labor share, if α is large, then the return from labor is low. Thus, an old agent will lose the incentive to work. One more case is where θ is sufficiently small. Since θ is an old agent's productivity relative to a young agent's productivity, the return from work for an old agent is sufficiently low. Therefore, an old agent will retire fully.

Let

$$g_f(\tau) := \frac{\alpha\beta(1 - \tau)(1 - \alpha)}{(1 + \beta)\alpha + \tau(1 - \alpha)}.$$

Thus, $k_{t+1} = g_f(\tau)k_t^\alpha$ (Equation (17)). It is clear that $\frac{\partial g_f}{\partial \tau} < 0$. Therefore, the following result holds.

Proposition 2. *Suppose that $\tau \geq \hat{\tau}$. Then, a slight increase in τ decreases the law of motion of capital per unit of effective labor.*

As long as τ satisfies Equation (8), all agents will retire when they become old. This situation is the same as the standard overlapping-generations model with production in Diamond (1965). Therefore, a higher τ reduces the amount of savings, which leads the law of motion of capital per unit of effective labor to rotate downward.

From this result, the following result regarding the steady state is derived.

Corollary 1. *Suppose that $\tau \geq \hat{\tau}$. Then, a slight increase in τ depresses the steady state capital per unit of effective labor.*

The next proposition considers the case in which τ is less than $\hat{\tau}$.

⁹In the following, the subscript “f” indicates “full retirement.”

Proposition 3. Suppose that the payroll tax rate satisfies $0 \leq \tau < \hat{\tau}$. Then, in equilibrium, an old agent in period $t + 1$ works for $1 - l_{t+1}^o$, where

$$l_{t+1}^o = \frac{\gamma(1 - \tau)(\alpha + \theta) + \tau(1 + \theta)[\beta(1 - \alpha) + \gamma]}{\theta[\beta(1 - \alpha) + \gamma]} < 1, \quad (10)$$

and the equilibrium economic growth is characterized by

$$k_{t+1} = \frac{\alpha[\beta(1 - \alpha) + \gamma]}{\alpha(1 + \beta + \gamma) + \theta(1 + \alpha\beta)} k_t^\alpha. \quad (11)$$

Moreover, there is a unique non-zero steady state, $k_p(\tau)$, where

$$k_p(\tau) := \left[\frac{\alpha[\beta(1 - \alpha) + \gamma]}{\alpha(1 + \beta + \gamma) + \theta(1 + \alpha\beta)} \right]^{\frac{1}{1 - \alpha}}, \quad (12)$$

and it is globally stable.¹⁰

Proof. See the Appendix. Q.E.D.

Note that for $0 < \tau \leq \hat{\tau}$ to be true, α needs to satisfy

$$\alpha < \bar{\alpha} := \frac{\beta\theta}{\beta\theta + \gamma}.$$

This fact proves the following result.

Proposition 4. Suppose that $0 \leq \tau < \hat{\tau}$. Then, a slight increase in τ increases l_{t+1}^o .

Proof. See the Appendix. Q.E.D.

If τ increases, when an old agent works, s/he has to pay more tax. This decreases the old agent's incentive to work. Therefore, working time, $1 - l_{t+1}^o$, decreases as τ increases.

Let

$$g_p(\tau) =: \frac{\alpha[\beta(1 - \alpha) + \gamma]}{\alpha(1 + \beta + \gamma) + \theta(1 + \alpha\beta)}.$$

Then, Equation (11) is written as $k_{t+1} = g_p(\tau)k_t^\alpha$. It is clear that a change in τ does not affect the law of motion of capital per unit of effective labor. Hence, a change in τ does not affect the output per unit of effective labor. From Proposition 4, if τ increases, then an old agent works less. Since Equation (11) does not change in τ , an

¹⁰In the following, the subscript “ p ” indicates “partial retirement.”

increase in τ also decreases the amount of savings. In general, this could affect economic growth through capital formation. Since an old agent supplies less labor than before, however, the capital stock per unit of effective labor is not affected at all.

Corollary 2. *Suppose that $0 \leq \tau < \hat{\tau}$. Then, τ does not change the steady state at all.*

The output per unit of effective labor is not affected by a change in τ . Aggregate output, however, will decrease as τ increases. This is because even though the capital stock per unit of effective labor does not change, the aggregate labor supplied by an old agent will decrease as τ increases.

4 Welfare

This section discusses welfare. Following the previous literature (e.g., Michel and Pestieau (2013)), welfare is defined by an agent's lifetime utility in the steady state.

4.1 First-best allocation

First, as a benchmark, I characterize the first-best allocation in the model. To define the first-best allocation, first, I define the feasible allocation. A steady-state allocation, (c^y, c^o, l^o, k) , is *feasible* if $(c^y, c^o, l^o, k) \in \mathbb{R}_+^4$ satisfies the resource constraint and the time constraint, that is,

$$c^y + c^o = [1 + \theta(1 - l^o)][f(k) - k] \text{ and } l^o \in [0, 1].$$

The *first-best* allocation is a feasible allocation that maximizes

$$\ln(c^y) + \beta \ln(c^o) + \gamma \ln(l^o).$$

Let

$$\hat{\theta} := \frac{\gamma}{1 + \beta}.$$

A high $\hat{\theta}$ implies that leisure is more valuable than lifetime consumption. If an old agent's labor productivity is higher than $\hat{\theta}$, that is, $\theta(1 + \beta) > \gamma$, then an old agent should work more and agents should consume more at the first-best allocation.

Proposition 5. *If $\theta > \widehat{\theta}$, the first-best allocation is*

$$c^{y*} = \frac{(1+\theta)(1-\alpha)}{1+\beta+\gamma} \alpha^{\frac{\alpha}{1-\alpha}}, c^{o*} = \beta c^{y*}, l^{o*} = \frac{\gamma(1+\theta)}{\theta(1+\beta+\gamma)} < 1, k^* = \alpha^{\frac{1}{1-\alpha}}.$$

If $\theta \leq \widehat{\theta}$, the first-best allocation is

$$c^{y*} = \frac{1-\alpha}{1+\beta} \alpha^{\frac{\alpha}{1-\alpha}}, c^{o*} = \beta c^{y*}, l^{o*} = 1, k^* = \alpha^{\frac{1}{1-\alpha}}.$$

Proof. See the Appendix. *Q.E.D.*

In general, these first-best allocations cannot be achieved in the decentralized economy without any intervention. Moreover, if the social planner redistributes resources across across generations and controls the retirement age, then the first-best allocation can be achieved in the decentralized economy. Inter-generational redistribution is feasible through taxation, whereas retirement age control is quite difficult to implement in the real world.¹¹ Thus, in the next section, I consider an environment in which only pay-as-you-go social security is feasible for the social planner to redistribute outputs across generations, and I characterize the second-best allocation in the decentralized economy. Michel and Pestieau (2013) investigate the same question and conclude that the first-best allocation is not achieved if pay-as-you-go social security is the only available policy for the social planner. The following section, however, shows that their result relies on their implicit assumption that labor productivity does not depreciate over age.

4.2 Second-best allocation

Let $W(\tau)$ denote an agent's lifetime utility in the steady state given τ . Formally,

$$W(\tau) := \ln(c^y) + \beta \ln(c^o) + \gamma \ln(l^o),$$

where (c^y, c^o, l^o) is a steady-state allocation to which an equilibrium allocation, $(c_t^y, c_{t+1}^o, l_{t+1}^o)_{t=1}^{\infty}$, converges given τ . Thus, the social planner's second-best problem is to maximize $W(\tau)$ by choosing $\tau \in [0, 1)$.

Proposition 6.

1. *Suppose that $\theta \leq \widehat{\theta}$. Then,*

¹¹ Although the social security systems have the eligibility conditions to receive the social security benefits, individuals have the right to determine when s/he retires. No one can (directly) force individuals to work and retire at a certain age.

(i) For $\alpha \geq \tilde{\alpha} := \frac{\beta}{1+2\beta}$, the second-best payroll tax rate is 0, an old agent fully retires, and capital is under-accumulated.

(ii) For $\alpha < \tilde{\alpha}$, the second-best payroll tax rate is $\tilde{\tau}_f \in (0, 1)$, an old agent fully retires, and capital is the golden rule, where $\tilde{\tau}_f$ maximizes $W(\tau)$ conditional on an old agent fully retiring.

2. Suppose that $\beta + \gamma \geq \theta > \hat{\theta}$. Then,

(i) For $\alpha \geq \bar{\alpha} = \frac{\beta\theta}{\beta\theta+\gamma}$, the second-best tax rate is 0, an old agent fully retires, and capital is under-accumulated.

(ii) For $\bar{\alpha} > \alpha \geq \hat{\alpha} := \frac{\beta+\gamma-\theta}{1+\gamma+\beta(2+\theta)}$, the second-best tax rate is 0, an old agent partially retires, and capital is under-accumulated.

(iii) For $\alpha < \hat{\alpha}$, the second-best payroll tax rate is $\arg \max \{W(\tilde{\tau}_p), W(\tilde{\tau}_f)\}$, and an old agent fully retires and capital is the golden rule when the second-best payroll tax rate is $\tilde{\tau}_f$, while an old agent partially retires and capital is over-accumulated when the second-best payroll tax rate is $\tilde{\tau}_p$, where $\tilde{\tau}_p$ maximizes $W(\tau)$ conditional on an old agent partially retiring.

3. If $\theta > \beta + \gamma$, 2(i) and 2(ii) hold.

Proof. See the Appendix.

Q.E.D.

Recall that when $\theta \leq \hat{\theta}$, since an old agent's labor productivity is low, in the first-best allocation, an old agent fully retires. When α is high, the labor share is low. In addition, the interest rate is high, that is, $r > 1$, in a laissez-faire equilibrium.¹² Because of these facts and an old agent's low labor productivity, an old agent fully retires in a laissez-faire equilibrium. Since $l^o \leq 1$ and l^o increases as τ increases, an old agent fully retires in equilibrium for all $\tau \in [0, 1)$. Because when capital is under-accumulated, transfer from a young agent to an old agent lowers welfare, $\tau = 0$ is the second-best payroll tax rate. When α is low enough, the labor share is high and the interest rate is low, that is, $r < 1$ in a laissez-faire equilibrium. Thus, even though an old agent's labor productivity is low, an old agent works in a laissez-faire equilibrium. As τ increases, savings by a young agent and the labor supply of an old agent decrease. Then, how capital per unit of effective labor changes is generally ambiguous, although as shown in Proposition 3, capital per unit of effective labor is constant as long as an old agent supplies a positive amount of labor. After τ reaches the level at which an old agent fully retires, capital per unit of effective labor starts decreasing in τ . Since at $\tilde{\tau}_f$, welfare is maximized conditional on an old agent

¹²A *laissez-faire* equilibrium is an equilibrium with $\tau = 0$.

fully retiring, $k_f(\tilde{\tau}_f) = k^*$. Thus, when $\tau = \tilde{\tau}_f$, $k = k^*$ and $l^o = l^*$ are achieved, and hence, $\tilde{\tau}_f$ is the second-best payroll tax rate.

Recall again that when $\theta > \hat{\theta}$, an old agent partially retires in the first-best allocation, that is, $l^{o*} < 1$. When α is high, that is, $\alpha \geq \bar{\alpha}$, the labor share is low and the interest rate in a laissez-faire equilibrium is high. Thus, even though an old agent's labor productivity is high, an old agent fully retires in a laissez-faire equilibrium. By the same argument as for Case 1(i), the second-best payroll tax rate is 0, and an old agent fully retires in the second-best allocation. When α is intermediate, that is, $\bar{\alpha} > \alpha \geq \hat{\alpha}$, capital is still under-accumulated in a laissez-faire equilibrium, while an old agent works in the laissez-faire equilibrium. Since α is intermediate, the labor share, $1 - \alpha$, is not so high. Thus, in the laissez-faire equilibrium, an old agent works less than the first-best level, that is, $1 - l^o < 1 - l^{o*}$. Because capital is under-accumulated, a low tax rate is better in terms of capital accumulation. In addition, since an old agent works less than before as τ increases, a low tax rate is also better in terms of labor. Therefore, $\tau = 0$ is the second-best payroll tax rate. In addition, leisure in the second-best allocation is not larger than the first-best level, that is, $l^o \geq l^{o*}$.¹³ When α is low, that is, $\alpha < \hat{\alpha}$, in a laissez-faire equilibrium, the capital per unit of effective labor satisfies $f'(k) < 1$. In addition, since the labor share is high and an old agent's labor productivity is high, an old agent partially retires in a laissez-faire equilibrium and $1 - l^o(0) > 1 - l^*$. The logic for this case is similar to that for Case 1(ii). As τ increases, l^o increases and k is constant until an old agent fully retires. Once an old agent fully retires, k starts decreasing as τ increases. Different from Case 1(ii), the first-best labor supply by an old agent is positive, that is, $1 - l^{o*} > 0$. Thus, there is trade-off between capital accumulation and labor supply by an old agent. At $\tilde{\tau}_f$, $l^o(\tilde{\tau}_f) = 1$ and $k_f(\tilde{\tau}_f) = k^*$. In this case, capital is the golden rule, whereas an old agent supplies less labor than the first-best level. At $\tilde{\tau}_p$, which is less than $\tilde{\tau}_f$, $l^o(\tilde{\tau}_p) = l^{o*}$ and $k_p(\tilde{\tau}_p) > k^*$. In this case, the labor supply by an old agent is first-best level, while capital is not the golden rule. Which tax rate, $\tilde{\tau}_f$ or $\tilde{\tau}_p$, is the second-best payroll tax rate depends on which welfare loss is severer. If the welfare loss from the over-accumulation of capital is severer than that from the over-supply of an old agent's labor, then $\tilde{\tau}_f$ is the second-best payroll tax rate; otherwise, $\tilde{\tau}_p$ is the second-best payroll tax rate. Unfortunately, it is not easy to derive the conditions under which $\tilde{\tau}_f$ is the second-best payroll tax rate and under which $\tilde{\tau}_p$ is the second-best payroll tax rate. In the next section, I propose sufficient conditions under which $\tilde{\tau}_f$ is the second-best payroll tax rate and under which $\tilde{\tau}_p$ is the second-best payroll tax rate.

When $\theta \leq \hat{\theta}$ and $\alpha < \tilde{\alpha}$, the second-best allocation coincides with the first-best allocation. This finding is

¹³When an old agent partially retires in the laissez-faire equilibrium, $l^o = \frac{\gamma(\alpha+\theta)}{\theta[\beta(1-\alpha)+\gamma]}$. From this and l^{o*} in Proposition 5, $l^o \geq l^{o*}$ if and only if $\alpha \geq \hat{\alpha}$.

different from the conclusion of Michel and Pestieau (2013). When $\theta > \hat{\theta}$, however, the second-best allocation and the first-best allocation are different, as argued by Michel and Pestieau (2013). Michel and Pestieau (2013) claim that to achieve the first-best allocation in the decentralized economy, the social planner needs to control both the pay-as-you-go social security tax rate and retirement age. The finding in this paper implies that their argument relies on the implicit assumption that labor productivity does not depreciate. As Proposition 6 shows, if labor productivity sufficiently depreciates and the capital share in the production function is not too high, then pay-as-you-go social security is enough to achieve the first-best allocation in the decentralized economy. In addition, from the above arguments, if an old agent partially retires in the first-best allocation, then pay-as-you-go social security only is not enough to implement the first-best allocation in the decentralized economy.

5 Discussion

5.1 The second-best payroll tax rate when $\theta > \hat{\theta}$ and $\alpha < \hat{\alpha}$

When $\theta > \hat{\theta}$ and $\alpha < \hat{\alpha}$, as Proposition 6 shows, which payroll tax rate, $\tilde{\tau}_f$ or $\tilde{\tau}_p$, is the second-best tax rate is ambiguous. In what follows, I give sufficient conditions under which $\tilde{\tau}_f$ is the second-best payroll tax rate and under which $\tilde{\tau}_p$ is the second-best payroll tax rate. First, I provide a sufficient condition under which $\tilde{\tau}_f$ is the second-best payroll tax rate.

Proposition 7. *Suppose that $\theta = \hat{\theta}$. If α , β , and γ satisfy*

$$0 < \alpha < \frac{(1 + \beta)(\beta + \gamma) - \gamma}{(1 + \beta)(1 + \gamma) + \beta(2 + 2\beta + \gamma)},$$

then $W(\tilde{\tau}_f) > W(\tilde{\tau}_p)$, that is, $\tilde{\tau}_f$ is the second-best payroll tax rate.

Proof. See the Appendix. Q.E.D.

In this case, capital per unit of effective labor is the first-best level, while an old agent provides less labor than the first-best level. This finding also differs from the finding in Michel and Pestieau (2013), who claim that the second-best payroll tax rate is lower than the payroll tax rate that achieves the golden rule. Their argument is true as long as an old agent partially retires in the decentralized economy. Once the corner solution, $l^o = 1$, is taken into account, however, the second-best payroll tax rate can coincide with the first-best payroll tax rate. Note that in this case, an old agent supplies less labor than the first-best level.

The intuition is as follows. Let c_f^y and r_f denote the consumption of a young agent under $\tilde{\tau}_f$ and the interest rate under $\tilde{\tau}_f$, respectively. Moreover, let l_f^o denote the leisure of an old agent under $\tilde{\tau}_f$. Then, welfare under $\tilde{\tau}_f$ is

$$\begin{aligned} W(\tilde{\tau}_f) &= \ln(c_f^y) + \beta \ln(\beta r_f c_f^y) + \gamma \ln(l_f^o) \\ &= (1 + \beta) \ln(c_f^y) + \beta \ln(\beta), \end{aligned}$$

where $c_f^o = \beta r_f c_f^y$, $r_f = 1$ and $l_f^o = 1$ are used. By analogy, welfare under $\tilde{\tau}_p$ is written as

$$\begin{aligned} W(\tilde{\tau}_p) &= \ln(c_p^y) + \beta \ln(\beta r_p c_p^y) + \gamma \ln(l_p^o) \\ &= (1 + \beta) \ln(c_p^y) + \beta \ln(\beta) + \beta \ln(r_p) + \gamma \ln(l_p^o). \end{aligned}$$

From these,

$$W(\tilde{\tau}_f) - W(\tilde{\tau}_p) = (1 + \beta) \ln\left(\frac{c_f^y}{c_p^y}\right) + \beta \ln\left(\frac{1}{r_p}\right) + \gamma \ln\left(\frac{1}{l_p^o}\right). \quad (13)$$

Note that the first and second terms in Equation (13) are negative and the last term in Equation (13) is positive. When θ is small, l_p^o is close to 1 because an old agent will receive a small amount even though s/he works. Therefore, the third term in Equation (13) is very close to zero. If α is small enough, capital per unit of effective labor is over-accumulated in an economy. This implies that r_p is much smaller than 1 and hence, the second term in Equation (13) becomes large, and in total, $W(\tilde{\tau}_f) - W(\tilde{\tau}_p)$ becomes positive. If the conditions in Proposition 7 hold, the above argument is true. Although θ is equal to $\hat{\theta}$ in Proposition 7, because of the continuity of W in θ , as long as θ is larger than $\hat{\theta}$ but sufficiently close to $\hat{\theta}$, $\tilde{\tau}_f$ is optimal if α is sufficiently small. This argument, however, may not hold if θ is large enough.

Proposition 8. *Suppose that $\beta + \gamma > 1$ and $\beta = \gamma$. In addition, suppose $\theta = 1$. Then, even though α is small enough, $W(\tilde{\tau}_p) > W(\tilde{\tau}_f)$, that is, $\tilde{\tau}_p$ is the second-best payroll tax rate.*

Proof. See the Appendix. Q.E.D.

As α increases, the gap between the interest rate and the population growth rate in the laissez-faire equilibrium shrinks. Therefore, it is reasonable to conjecture that if $\theta = 1$, for any $\alpha \in (0, 1)$, $\tilde{\tau}_p$ is optimal. Note that this proposition corresponds to the welfare analysis in Michel and Pestieau (2013).

Under the condition in Proposition 8, since θ is large enough, from the social planner's point of view, the gain from more labor supplied by productive old agents dominates the loss from the over-accumulated capital per unit of effective labor. Although in Proposition 8, θ is set at 1, the statement could hold if θ is smaller than 1 but sufficiently close to 1 because of the continuity of W in θ .

To fix the idea, I provide a numerical example. Set $\beta = \gamma = 1$. In this case, $\hat{\theta} = \frac{\gamma}{1+\beta} = 0.333$. Since Propositions 7 and 8 consider extreme cases of θ , consider a medium value of θ and set $\theta = 0.7$.¹⁴ Then, $\hat{\alpha} = 0.2766$, $\tilde{\alpha} = 0.3953$, and $\bar{\alpha} = 0.4118$. Consider two values of α , $\alpha = 0.2$ and $\alpha = 0.1$. When $\alpha = 0.2$, the optimal tax rate is $\tilde{\tau}_p = 0.0556$ and $W(\tilde{\tau}_p) = -2.6282$. For reference, $\tilde{\tau}_f = 0.2500$ and $W(\tilde{\tau}_f) = -2.6373$. Figure 1 shows how W changes in τ in this case. When $\alpha = 0.1$, the optimal tax rate is $\tilde{\tau}_f = 0.3889$ and $W(\tilde{\tau}_f) = -2.1087$. For reference, $\tilde{\tau}_p = 0.1139$ and $W(\tilde{\tau}_p) = -2.2037$. Figure 2 shows this case. As these numerical examples show, if α is high in $(0, \hat{\alpha})$, $\tilde{\tau}_p$ is optimal, while if α is low, then $\tilde{\tau}_f$ is optimal. In addition, note that W is not a unimodal function. This suggests that a first-order condition is not necessarily sufficient to specify the *globally* optimal tax rate.

5.2 Application to the real economy

Determining whether or not the real economy is at the (second-best) optimum is an important and interesting question. If the criterion for the judgment is simple and easy to use, it is good for society.¹⁵

Proposition 6 provides a simple criterion by which to determine the optimal payroll tax rate. For example, consider the U.S. economy. Suppose that a young agent starts working at age 20 and becomes eligible to receive social security at age 65. Then, one period in this paper corresponds to 45 years. Following Kitao (2014), since a subjective discount factor is 0.98, β is set to $(0.9815)^{45} = 0.4316$. The weight on leisure relative to consumption is 0.5123 in Kitao (2014), and so $\gamma = \beta * 0.5123 = 0.2211$. The capital share, α , is set at 0.4. Under these parameter values, $\hat{\theta} = 0.1544$ and $\tilde{\alpha} = 0.2316$. Hence, if an old agent's labor productivity depreciates sufficiently, that is, $\theta \leq 0.1544$, since $\tilde{\alpha} < \alpha = 0.4$, the second-best payroll tax rate is 0, and an old agent fully retires. Consider the case in which $\theta > \hat{\theta} = 0.1544$. Since $\hat{\alpha}$ is strictly decreasing in θ , $\hat{\alpha} \leq \frac{\beta + \gamma - \hat{\theta}}{1 + \gamma + \beta(2 + \hat{\theta})} = 0.2316$. Thus, for all $\theta \geq \hat{\theta}$, $\alpha = 0.4 > 0.2316 \geq \hat{\alpha}$; by Proposition 6, the second-best payroll tax rate is 0, and an old agent partially retires. Although the model is too simple to describe the real economy, the model suggests that for any $\theta \in (0, 1]$, the second-best payroll tax rate is 0. This result coincides with the result in Abel et al. (1989).

¹⁴This number is chosen arbitrarily.

¹⁵Abel et al. (1989) consider a standard Diamond-type overlapping-generations model and estimate the rate of return from saving, r , using U.S. data. They conclude that the U.S. economy is dynamically efficient. For a stochastic overlapping-generations model, see Chattopadhyay (2006) and Barbie et al. (2007).

However, depending on the size of θ , an old agent's working behavior at the second-best payroll tax rate will be different. If θ is low enough, then an old agent should retire fully at the second-best payroll tax rate, while if θ is large enough, then an old agent should work for some time and retire later at the second-best payroll tax rate

6 Conclusions

This paper studies a Diamond-type overlapping-generations model with pay-as-you-go social security and an old agent's retirement decision making. The paper characterizes a perfect foresight competitive equilibrium, the first-best allocation, and the second-best allocation. Moreover, the relationship between the first-best allocation and the second-best allocation is discussed. In contrast to the finding in Michel and Pestieau (2013), the paper shows the possibility that using pay-as-you-go social security alone can lead the economy to the first-best optimum.

The key parameters for deriving the results in this paper are capital share in the production function and an old agent's labor productivity relative to a young agent's labor productivity. If an old agent's labor productivity is not too low, then correcting the over-accumulation of capital can lower welfare. However, if an old agent's labor productivity is low, then inducing an old agent to retire early increases welfare. Recently, in many developed countries, people have been living longer than before and old people's productivity is also higher than before. In such a case, instead of providing generous social security benefits and having old people retire early, providing less generous social security benefits and making them work longer can improve welfare and can also solve the sustainability problem of the social security system.

A Appendix

A.1 Proof of Proposition 1

Proof. Consider an agent's problem. Let

$$L = \ln(c_t^y) + \beta \ln(c_{t+1}^o) + \gamma \ln(l_{t+1}^o) + \lambda \left\{ (1 - \tau)w_t + \frac{(1 - \tau)(1 - l_{t+1}^o)\theta w_{t+1} + l_{t+1}^o P_{t+1}}{r_{t+1}} - c_t^y - \frac{c_{t+1}^o}{r_{t+1}} \right\} + \eta(1 - l_{t+1}^o)$$

be the Lagrangian of an agent's problem. Taking the derivative of L with respect to c_t^y , c_{t+1}^o and l_{t+1}^o ,

$$c_{t+1}^o = \beta r_{t+1} c_t^y \tag{14}$$

and

$$\frac{\gamma}{l_{t+1}^o} \geq \frac{(1-\tau)\theta w_{t+1} - P_{t+1}}{r_{t+1} c_t^y} \quad (15)$$

with equality holding if $l_{t+1}^o < 1$.

Consider the case in which $l_{t+1}^o = 1$ in equilibrium. This implies that Equation (15) is

$$\gamma \geq \frac{(1-\tau)\theta w_{t+1} - P_{t+1}}{r_{t+1} c_t^y}. \quad (16)$$

In this case, from the agent's lifetime budget constraint,

$$c_t^y = \frac{1}{1+\beta} \left[(1-\tau)w_t + \frac{P_{t+1}}{r_{t+1}} \right].$$

Thus,

$$s_t = (1-\tau)w_t - c_t^y = \frac{\beta(1-\tau)}{1+\beta} w_t - \frac{1}{1+\beta} \frac{P_{t+1}}{r_{t+1}}.$$

Since $k_{t+1} = \frac{K_{t+1}}{L_{t+1}} = s_t$ in this case,

$$\begin{aligned} k_{t+1} &= \frac{\beta(1-\tau)}{1+\beta} w_t - \frac{1}{1+\beta} \frac{P_{t+1}}{r_{t+1}} \\ &= \frac{\beta(1-\tau)}{1+\beta} (1-\alpha)k_t^\alpha - \frac{1}{1+\beta} \frac{\tau w_{t+1}}{r_{t+1}} \\ &= \frac{\beta(1-\tau)}{1+\beta} (1-\alpha)k_t^\alpha - \frac{\tau}{1+\beta} \frac{1-\alpha}{\alpha} k_{t+1}, \end{aligned}$$

where $w_t = (1-\alpha)k_t^\alpha$, $r_t = \alpha k_t^{1-\alpha}$, and $P_{t+1} = \tau w_{t+1}$ are used. From this equation, the law of motion of capital per unit of effective labor is

$$k_{t+1} = \frac{\alpha\beta(1-\tau)(1-\alpha)}{(1+\beta)\alpha + \tau(1-\alpha)} k_t^\alpha. \quad (17)$$

Equation (16) can be rewritten as

$$\frac{\gamma r_{t+1}}{1+\beta} \left[(1-\tau)w_t + \frac{\tau w_{t+1}}{r_{t+1}} \right] \geq (1-\tau)\theta w_{t+1} - \tau w_{t+1}.$$

Rearranging this equation, I have

$$\frac{\gamma(1-\tau)(1-\alpha)}{1+\beta}k_t^\alpha \geq \left[(1-\tau)\theta - \tau - \frac{\gamma\tau}{1+\beta} \right] \frac{1-\alpha}{\alpha}k_{t+1}. \quad (18)$$

If the RHS of Equation (18) is less than or equal to zero, that is,

$$\tau \geq \frac{\theta(1+\beta)}{(1+\theta)(1+\beta) + \gamma},$$

then Equation (18) automatically holds. Hence, $l_{t+1}^o = 1$.

Suppose $0 \leq \tau < \frac{\theta(1+\beta)}{(1+\theta)(1+\beta) + \gamma}$. Then, from Equation (18), if k_t and k_{t+1} satisfy

$$k_{t+1} \leq \frac{\alpha\gamma(1-\tau)}{\{(1-\tau)\theta - \tau\}(1+\beta) - \gamma\tau} k_t^\alpha,$$

then $l_{t+1}^o = 1$. Since the law of motion of capital per unit of effective labor is specified by Equation (17), if

$$\frac{\alpha\beta(1-\tau)(1-\alpha)}{(1+\beta)\alpha + \tau(1-\alpha)} \leq \frac{\alpha\gamma(1-\tau)}{\{(1-\tau)\theta - \tau\}(1+\beta) - \gamma\tau},$$

then $l_{t+1}^o = 1$. This equation can be rewritten as

$$\tau \geq \frac{\beta(1-\alpha)\theta - \gamma\alpha}{(1-\alpha)[\beta(1+\theta) + \gamma]}.$$

Note that for all $\gamma > 0$,

$$\frac{\theta(1+\beta)}{(1+\theta)(1+\beta) + \gamma} > \hat{\tau} = \frac{\beta(1-\alpha)\theta - \gamma\alpha}{(1-\alpha)[\beta(1+\theta) + \gamma]}.^{16}$$

Q.E.D.

A.2 Proof of Proposition 3

Proof. Suppose $l_{t+1}^o < 1$. Then, from Equation (15),

$$[(1-\tau)\theta w_{t+1} - P_{t+1}]l_{t+1}^o = \gamma r_{t+1} c_t^y.$$

¹⁶The inequality holds if $\theta(1+\beta)(1-\alpha)\beta(1+\theta) + \theta(1+\beta)\gamma > \beta(1-\alpha)\theta(1+\theta)(1+\beta) + \beta(1-\alpha)\theta\gamma - \gamma\alpha[(1+\theta)(1+\beta) + \gamma]$. Since the first terms on both sides are the same and $\theta\gamma\beta(1-\alpha) > \theta\gamma\beta(1-\alpha)$, the LHS is greater than the RHS.

Plugging this and Equation (14) into the agent's lifetime budget constraint,

$$c_t^y = \frac{1}{1 + \beta + \gamma} \left[(1 - \tau)w_t + \frac{(1 - \tau)\theta w_{t+1}}{r_{t+1}} \right].$$

From this,

$$s_t = (1 - \tau)w_t - c_t^y = \frac{(\beta + \gamma)(1 - \tau)}{1 + \beta + \gamma} w_t - \frac{(1 - \tau)\theta}{1 + \beta + \gamma} \frac{w_{t+1}}{r_{t+1}}.$$

Now consider an agent's choice of l_{t+1}^o . From Equation (15),

$$l_{t+1}^o = \frac{\gamma r_{t+1} c_t^y}{(1 - \tau)\theta w_{t+1} - P_{t+1}}.$$

Since in equilibrium, $l_{t+1}^o P_{t+1} = \tau w_{t+1} [1 + \theta(1 - l_{t+1}^o)]$ holds,

$$l_{t+1}^o = \frac{\gamma r_{t+1} c_t^y}{(1 - \tau)\theta w_{t+1} - \frac{\tau w_{t+1} [1 + \theta(1 - l_{t+1}^o)]}{l_{t+1}^o}}.$$

From this,

$$\theta w_{t+1} l_{t+1}^o = \gamma r_{t+1} c_t^y + \tau w_{t+1} (1 + \theta)$$

holds. By using k_t and k_{t+1} , l_{t+1}^o in equilibrium is written as

$$l_{t+1}^o = \frac{\gamma \alpha (1 - \tau)}{\theta (1 + \beta + \gamma)} \frac{k_t^\alpha}{k_{t+1}} + \frac{\gamma (1 - \tau)}{1 + \beta + \gamma} + \frac{\tau (1 + \theta)}{\theta}. \quad (19)$$

Since $k_{t+1} = \frac{s_t}{1 + \theta(1 - l_{t+1}^o)}$,

$$\begin{aligned} k_{t+1} &= \frac{\frac{(\beta + \gamma)(1 - \tau)}{1 + \beta + \gamma} (1 - \alpha) k_t^\alpha - \frac{1}{1 + \beta + \gamma} \frac{(1 - \tau)\theta(1 - \alpha)}{\alpha} k_{t+1}}{1 + \theta \left[1 - \frac{\gamma \alpha (1 - \tau)}{\theta (1 + \beta + \gamma)} \frac{k_t^\alpha}{k_{t+1}} - \frac{\gamma (1 - \tau)}{1 + \beta + \gamma} - \frac{\tau (1 + \theta)}{\theta} \right]} \\ &= \frac{\frac{(\beta + \gamma)(1 - \alpha)}{1 + \beta + \gamma} k_t^\alpha - \frac{\theta}{1 + \beta + \gamma} \frac{1 - \alpha}{\alpha} k_{t+1}}{1 + \theta - \frac{\gamma \alpha}{1 + \beta + \gamma} \frac{k_t^\alpha}{k_{t+1}} - \frac{\theta \gamma}{1 + \beta + \gamma}}. \end{aligned}$$

From this equation, the law of motion of capital per unit of effective labor is

$$k_{t+1} = \frac{\alpha[\beta(1-\alpha) + \gamma]}{\alpha(1+\beta+\gamma) + \theta(1+\alpha\beta)} k_t^\alpha.$$

Plugging this into Equation (19), I have

$$l_{t+1}^o = \frac{\gamma(1-\tau)(\alpha + \theta) + \tau(1+\theta)[\beta(1-\alpha) + \gamma]}{\theta[\beta(1-\alpha) + \gamma]}.$$

Q.E.D.

A.3 Proof of Proposition 4

Proof. Taking the derivative of l_{t+1}^o with respect to τ , I obtain

$$\frac{\partial l_{t+1}^o}{\partial \tau} = \frac{-\gamma(\alpha + \theta) + (1+\theta)[\beta(1-\alpha) + \gamma]}{\theta[\beta(1-\alpha) + \gamma]} = \frac{\gamma(1-\alpha) + \beta(1+\theta)(1-\alpha)}{\theta[\beta(1-\alpha) + \gamma]} > 0. \quad (20)$$

Q.E.D.

A.4 Proof of Proposition 5

Proof. Let

$$L = \ln(c^y) + \beta \ln(c^o) + \gamma \ln(l^o) + \lambda \{ [1 + \theta(1 - l^o)][f(k) - k] - c^y - c^o \} + \eta(1 - l_{t+1}^o)$$

be the Lagrangian of the problem. Taking the derivative of L with respect to c^y , c^o , and l^o ,

$$c^o = \beta c^y, \quad (21)$$

$$f'(k) - 1 = 0, \quad (22)$$

and

$$\frac{\gamma}{l^o} \geq \frac{1}{c^y} \theta [f(k) - k] \quad (23)$$

with equality holding if $l^o < 1$. From Equation (22),

$$k^* = \alpha^{\frac{1}{1-\alpha}}.$$

Suppose $l^{o*} = 1$. Then, from the resource constraint, $c^y + c^o = f(k^*) - k^*$ holds. Since $c^o = \beta c^y$,

$$c^{y*} = \frac{1-\alpha}{1+\beta} \alpha^{\frac{\alpha}{1-\alpha}}.$$

Since these must satisfy Equation (23),

$$\theta \leq \hat{\theta}.$$

Suppose $l^{o*} < 1$. In this case, $c^y = \frac{\theta l^{o*}}{\gamma} [f(k^*) - k^*]$ holds. From the resource constraint, I obtain

$$c^{y*} = \frac{(1-\alpha)(1+\theta)}{1+\beta+\gamma} \alpha^{\frac{\alpha}{1-\alpha}}.$$

By using this, I have

$$l^{o*} = \frac{(1+\theta)\gamma}{\theta(1+\beta+\gamma)},$$

which is less than 1 when $\theta > \hat{\theta}$.

Q.E.D.

A.5 Proof of Proposition 6

In the proof, first, I characterize the second-best allocation conditional on an old agent fully retiring, and next, I characterize the second-best allocation conditional on an old agent partially retiring. After that, combining the findings from these two cases, the proof is completed.

A.5.1 The second-best allocation when an old agent fully retires

Suppose that given $\tau \in [0, 1)$, an old agent fully retires. Then, from Proposition 1, in the steady state,

$$c^y = (1-\tau)w - s = (1-\tau)(1-\alpha)k_f^\alpha - k_f,$$

$$c^o = rs + P = \alpha k_f^\alpha + \tau(1-\alpha)k_f^\alpha,$$

and

$$l^o = 1.$$

Taking the derivative of W with respect to τ , I obtain

$$\frac{\partial W}{\partial \tau}(\tau) = \frac{1}{c^y} \left[\frac{\partial c^y}{\partial \tau} + \frac{1}{r} \frac{\partial c^o}{\partial \tau} \right].$$

In the brackets,

$$\begin{aligned} \frac{\partial c^y}{\partial \tau} + \frac{1}{r} \frac{\partial c^o}{\partial \tau} &= (1-\tau)(1-\alpha)\alpha k_f^{\alpha-1} \frac{\partial k_f}{\partial \tau} - (1-\alpha)k_f^\alpha - \frac{\partial k_f}{\partial \tau} \\ &\quad + \frac{1}{\alpha k_f^{\alpha-1}} \left[\alpha^2 k_f^{\alpha-1} \frac{\partial k_f}{\partial \tau} + (1-\alpha)k_f^\alpha + \tau(1-\alpha)\alpha k_f^{\alpha-1} \frac{\partial k_f}{\partial \tau} \right] \\ &= (\alpha k_f^{\alpha-1} - 1) \left[\frac{\partial k_f}{\partial \tau} (1-\tau) - \frac{k}{\alpha} \right] (1-\alpha). \end{aligned}$$

From Corollary 1, $\frac{\partial k_f}{\partial \tau} < 0$. Since $\tau \in [0, 1)$, $c^y > 0$ and $\alpha \in (0, 1)$, $\frac{\partial k_f}{\partial \tau} (1-\tau) - \frac{k}{\alpha} < 0$. Hence, $\frac{\partial W}{\partial \tau}(\tau) > 0$ if and only if $\alpha k_f^{\alpha-1}(\tau) - 1 < 0$.

From Equation (9), k_f is strictly decreasing in τ . Therefore, if $\alpha k_f^{\alpha-1}(0) \geq 1$ holds, $\alpha k_f^{\alpha-1}(\tau) \geq 1$ for all $\tau \in [0, 1)$. In this case, $\frac{\partial W}{\partial \tau} \leq 0$ for all $\tau \in [0, 1)$. When $\tau = 0$,

$$k_f(0) = \left[\frac{\alpha\beta(1-\alpha)}{\alpha(1+\beta)} \right]^{\frac{1}{1-\alpha}}.$$

Then, $\alpha(k_f(0))^{\alpha-1} \geq 1$ if and only if $\alpha \geq \tilde{\alpha} := \frac{\beta}{1+2\beta}$.

Lemma 1. *Suppose $\alpha > \tilde{\alpha}$ and an old agent fully retires. Then, $\frac{\partial W}{\partial \tau} < 0$ for all $\tau \in [0, 1)$. In addition, if $\alpha = \tilde{\alpha}$, then $\frac{\partial W}{\partial \tau}(0) = 0$ and $\frac{\partial W}{\partial \tau} < 0$ for all $\tau \in (0, 1)$. Suppose $\alpha < \tilde{\alpha}$ and an old agent fully retires. Then, there is a unique $\tilde{\tau}_f \in (0, 1)$ that maximizes $W(\tau)$, where*

$$\tilde{\tau}_f = \frac{\beta - \alpha(1+2\beta)}{(1-\alpha)(1+\beta)}.$$

Proof. Note that the sign of $\frac{\partial W}{\partial \tau}$ is determined by $\alpha(k_f(\tau))^{\alpha-1} - 1$. If $\alpha(k_f(\tau))^{\alpha-1} < 1$, $\frac{\partial W}{\partial \tau} > 0$, and if $\alpha(k_f(\tau))^{\alpha-1} > 1$, $\frac{\partial W}{\partial \tau} < 0$. Note further that $\alpha(k_f(\tau))^{\alpha-1}$ is strictly increasing and continuous in τ . If $\alpha < \tilde{\alpha}$,

then $\alpha(k_f(0))^{\alpha-1} < 1$. As τ approaches 1, $\alpha k_f^{\alpha-1} \rightarrow +\infty$. Hence, there is a unique $\tilde{\tau}_f \in (0, 1)$ such that

$$\frac{\partial W}{\partial \tau}(\tau) \begin{cases} > \\ = \\ < \end{cases} \left. \vphantom{\frac{\partial W}{\partial \tau}} \right\} 0 \text{ if } \tau \begin{cases} < \\ = \\ > \end{cases} \left. \vphantom{\tau} \right\} \tilde{\tau}_f.$$

Moreover, since $\tilde{\tau}_f$ must satisfy $\alpha(k_f(\tilde{\tau}_f))^{\alpha-1} - 1 = 0$,

$$\tilde{\tau}_f := \frac{\beta - \alpha(1 + 2\beta)}{(1 - \alpha)(1 + \beta)}.$$

Note that when $\alpha = \tilde{\alpha}_f$, $\tilde{\tau}_f = 0$. Since the numerator of $\tilde{\tau}_f$ is decreasing in α , for all $\alpha \leq \tilde{\alpha}_f$, $\tilde{\tau}_f \geq 0$. Note also that $\tilde{\tau}_f$ is strictly decreasing in α .¹⁷ As $\alpha \rightarrow 0$, $\tilde{\tau}_f \rightarrow \frac{\beta}{1+\beta} < 1$. Since $\tilde{\tau}_f$ is continuous in α , $\tilde{\tau}_f < 1$ for all $\alpha \leq \tilde{\alpha}_f$. *Q.E.D.*

It is known that if $\tau < \hat{\tau}$, an old agent does not fully retire in equilibrium. Thus, it is useful to know when $\tilde{\tau}_f \geq \hat{\tau}$ holds.

Lemma 2. $\tilde{\tau}_f \geq \hat{\tau}$ if and only if

$$\alpha \leq \hat{\alpha} := \frac{\beta + \gamma - \theta}{1 + \gamma + \beta(2 + \theta)},$$

with equality holding if $\alpha = \hat{\alpha}$.

Proof. $\tilde{\tau}_f \geq \hat{\tau}$ if and only if

$$\beta[\beta(1 + \theta) + \gamma] - \alpha(1 + 2\beta)[\beta(1 + \theta) + \gamma] \geq \beta(1 - \alpha)\theta(1 + \beta) - \gamma\alpha(1 + \beta).$$

Rearranging this equation, I obtain

$$\beta + \gamma - \theta \geq \alpha[1 + 2\beta + \gamma + \beta\theta].$$

From this, $\hat{\alpha}$ is derived. *Q.E.D.*

¹⁷Taking the derivative of $\tilde{\tau}_f$ with respect to α , I obtain $-\frac{1+2\beta}{(1-\alpha)(1+\beta)} + \frac{[\beta-\alpha(1+2\beta)](1+\beta)}{(1-\alpha)^2(1+\beta)^2}$. Rearranging this, the sign is determined by $-(1+2\beta)(1-\alpha) + [\beta-\alpha(1+2\beta)] = -1-\beta < 0$.

Note that $\hat{\alpha} > 0$ if and only if $\theta < \beta + \gamma$. $\tilde{\tau}_f$ is well-defined when $\alpha \leq \tilde{\alpha}$ by Lemma 1. Thus, the sizes of $\hat{\alpha}$ and $\tilde{\alpha}$ are necessary information in the discussion of the second-best payroll tax rate.

Lemma 3. $\hat{\alpha} \geq \tilde{\alpha}$ if and only if

$$\theta \leq \hat{\theta} := \frac{\gamma}{1+\beta},$$

with equality holding if $\theta = \hat{\theta}$.

Proof. $\hat{\alpha} > \tilde{\alpha}$ if and only if

$$(\beta + \gamma - \theta)(1 + 2\beta) > \beta[1 + \gamma + \beta(2 + \theta)].$$

Rearranging this equation with respect to θ , I obtain

$$\gamma(1 + \beta) > \theta(1 + \beta)^2.$$

From this, $\hat{\theta}$ is derived.

Q.E.D.

Recall that if $\alpha > \bar{\alpha} = \frac{\beta\theta}{\beta\theta + \gamma}$, then an old agent will retire fully in equilibrium for all $\tau \in [0, 1)$.

Lemma 4. $\bar{\alpha} \leq \tilde{\alpha}$ if and only if $\theta \leq \hat{\theta}$, with equality holding if $\theta = \hat{\theta}$.

Proof.

$$\begin{aligned} \frac{\beta\theta}{\beta\theta + \gamma} < \frac{\beta}{1 + 2\beta} &\Leftrightarrow \frac{\theta}{\beta\theta + \gamma} < \frac{1}{1 + 2\beta} \\ &\Leftrightarrow \theta(1 + 2\beta) < \beta\theta + \gamma \\ &\Leftrightarrow \theta < \frac{\gamma}{1 + \beta}. \end{aligned}$$

Q.E.D.

A.5.2 The second-best allocation when an old agent partially retires

From Proposition 3, when an old agent partially retires, in the steady state,

$$\begin{aligned} c^y &= (1-\tau)(1-\alpha)k_p^\alpha - s, \\ s &= k_p[1+\theta(1-l^o)], \\ c^o &= \alpha k_p^{\alpha-1}s + (1-\tau)(1-l^o)\theta(1-\alpha)k_p^\alpha + l^o P, \\ l^o &= \frac{\gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\beta(1-\alpha) + \gamma]}{\theta[\beta(1-\alpha) + \gamma]}, \end{aligned}$$

and

$$l^o P = \tau(1-\alpha)k_p^\alpha[1+\theta(1-l^o)].$$

From these equations,

$$\begin{aligned} c^y &= (1-\tau)(1-\alpha)k_p^\alpha - k_p[1+\theta(1-l^o)], \\ c^o &= \alpha k_p^\alpha + \theta(1-l^o)k_p^\alpha + \tau(1-\alpha)k_p^\alpha. \end{aligned}$$

Taking the derivative of W with respect to τ , I have

$$\begin{aligned} \frac{\partial W}{\partial \tau} &= \frac{1}{c^y} \left\{ -(1-\alpha)k_p^\alpha + (1-\tau)(1-\alpha)\alpha k_p^{\alpha-1} \frac{\partial k_p}{\partial \tau} - [1+\theta(1-l^o)] \frac{\partial k_p}{\partial \tau} + k_p \theta \frac{\partial l^o}{\partial \tau} \right\} \\ &\quad + \frac{\beta}{c^o} \left\{ \alpha^2 k_p^{\alpha-1} \frac{\partial k_p}{\partial \tau} + \theta(1-l^o)\alpha k_p^{\alpha-1} \frac{\partial k_p}{\partial \tau} - \theta k_p^\alpha \frac{\partial l^o}{\partial \tau} \right. \\ &\quad \left. + (1-\alpha)k_p^\alpha + \tau(1-\alpha)\alpha k_p^{\alpha-1} \frac{\partial k_p}{\partial \tau} \right\} + \frac{\gamma}{l^o} \frac{\partial l^o}{\partial \tau}. \end{aligned}$$

Since $\frac{\partial k_p}{\partial \tau} = 0$ in the steady state,

$$\begin{aligned} \frac{\partial W}{\partial \tau} &= \frac{1}{c^y} \left\{ -(1-\alpha)k_p^\alpha + k_p \theta \frac{\partial l^o}{\partial \tau} \right\} + \frac{\beta}{c^o} \left\{ -\theta k_p^\alpha \frac{\partial l^o}{\partial \tau} + (1-\alpha)k_p^\alpha \right\} + \frac{\gamma}{l^o} \frac{\partial l^o}{\partial \tau} \\ &= \frac{1}{c^y} (1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} \right] + \frac{\partial l^o}{\partial \tau} \left[\frac{k_p \theta}{c^y} \frac{\alpha-1}{\alpha} + \frac{\gamma}{l^o} \right], \end{aligned} \tag{24}$$

where I use the equilibrium condition $c^o = \beta r c^y$. Since $\frac{\gamma}{l^o} = \frac{(1-\tau)\theta(1-\alpha)k_p^\alpha - P}{r c^y}$,

$$\begin{aligned}\frac{\partial W}{\partial \tau} &= \frac{1}{c^y} \left\{ (1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} \right] - \frac{\partial l^o}{\partial \tau} \theta \frac{1-\alpha}{\alpha} k_p + \frac{\partial l^o}{\partial \tau} \frac{(1-\tau)\theta(1-\alpha)k_p^\alpha - P}{r} \right\} \\ &= \frac{1}{c^y} \left\{ (1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} \right] - \frac{\partial l^o}{\partial \tau} \theta \frac{1-\alpha}{\alpha} \tau k_p - \frac{\partial l^o}{\partial \tau} \frac{P}{r} \right\}.\end{aligned}\quad (25)$$

Since $c^y > 0$, the sign of $\frac{\partial W}{\partial \tau}$ is equivalent to

$$\begin{aligned}(1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} \right] - \frac{\partial l^o}{\partial \tau} \theta \frac{1-\alpha}{\alpha} \tau k_p - \frac{\partial l^o}{\partial \tau} \frac{P}{r} &= (1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} \right] - \frac{\partial l^o}{\partial \tau} \frac{1-\alpha}{\alpha} \tau k_p \frac{1+\theta}{l^o} \\ &= (1-\alpha)k_p^\alpha \left[-1 + \frac{1}{r} - \frac{\partial l^o}{\partial \tau} \frac{1}{\alpha} \tau k_p^{1-\alpha} \frac{1+\theta}{l^o} \right],\end{aligned}\quad (26)$$

where $P = \frac{\tau w [1+\theta(1-l^o)]}{l^o}$ and $\frac{P}{r} = \tau \frac{1-\alpha}{\alpha} k_p \frac{1+\theta(1-l^o)}{l^o}$ are used.

Replacing k_p , l^o , and $\frac{\partial l^o}{\partial \tau}$ with Equation (12), Equation (10) and Equation (20) respectively, Equation (26) can be rewritten as

$$(1-\alpha)k_p^\alpha \left[-1 + \frac{\beta(1-\alpha) + \gamma}{\alpha(1+\beta+\gamma) + \theta(1+\alpha\beta)} \frac{\gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\alpha\gamma - \beta\theta(1-\alpha)]}{\gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\beta(1-\alpha) + \gamma]} \right].$$

Hence, $\frac{\partial W}{\partial \tau} > 0$ if and only if

$$-1 + \frac{\beta(1-\alpha) + \gamma}{\alpha(1+\beta+\gamma) + \theta(1+\alpha\beta)} \frac{\gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\alpha\gamma - \beta\theta(1-\alpha)]}{\gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\beta(1-\alpha) + \gamma]} > 0.\quad (27)$$

Proposition 9. Suppose that an old agent partially retires. If $\alpha \geq \max\{0, \hat{\alpha}\}$, then $\frac{\partial W}{\partial \tau} \leq 0$ for all $\tau \in [0, 1)$.

Assume further that $\theta < \beta + \gamma$. If $\alpha < \hat{\alpha}$, then

$$\tilde{\tau}_p := \frac{\gamma\{\beta + \gamma - \theta - \alpha[1 + \gamma + \beta(2 + \theta)]\}}{(1-\alpha)[\beta(1+\theta) + \gamma](1 + \beta + \gamma)} \in (0, 1)\quad (28)$$

such that $\frac{\partial W}{\partial \tau} > 0$ for $\tau \in [0, \tilde{\tau}_p)$, $\frac{\partial W}{\partial \tau} = 0$ for $\tau = \tilde{\tau}_p$, and $\frac{\partial W}{\partial \tau} < 0$ for $\tau > \tilde{\tau}_p$.

Proof. Equation (27) is equivalent to

$$\begin{aligned}& [\beta(1-\alpha) + \gamma] \{ \gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\alpha\gamma - \beta\theta(1-\alpha)] \} \\ & > [\alpha(1+\beta+\gamma) + \theta(1+\alpha\beta)] \{ \gamma(1-\tau)(\alpha+\theta) + \tau(1+\theta)[\beta(1-\alpha) + \gamma] \}.\end{aligned}$$

Rearranging this equation, I obtain

$$\tau < \tilde{\tau}_p := \frac{\gamma\{\beta + \gamma - \theta - \alpha[1 + \gamma + \beta(2 + \theta)]\}}{(1 - \alpha)[\beta(1 + \theta) + \gamma](1 + \beta + \gamma)}.$$

When $\alpha \geq \max\{0, \hat{\alpha}\}$, the numerator of $\tilde{\tau}_p$ is smaller than 0 with equality holding if $\alpha = \hat{\alpha} > 0$. Therefore, $\frac{\partial W}{\partial \tau} < 0$ for all τ .

Assuming $\theta < \beta + \gamma$, $\hat{\alpha} > 0$. When $\alpha < \hat{\alpha}$, then for $\tau < \tilde{\tau}_p$, $\frac{\partial W}{\partial \tau} > 0$, for $\tau = \tilde{\tau}_p$, $\frac{\partial W}{\partial \tau} = 0$, and for $\tau > \tilde{\tau}_p$, $\frac{\partial W}{\partial \tau} < 0$. Therefore, $\tau = \tilde{\tau}_p$ maximizes welfare. When $\alpha < \hat{\alpha}$, because $\beta + \gamma - \theta - \alpha[1 + \gamma + \beta(2 + \theta)] > 0$, $\tilde{\tau}_p > 0$. $\tilde{\tau}_p < 1$ if and only if $\gamma[\beta + \gamma - \theta] - \gamma\alpha[1 + \gamma + \beta(2 + \theta)] < (1 - \alpha)(1 + \beta + \gamma)\beta(1 + \theta) + (1 - \alpha)(1 + \beta + \gamma)\gamma$. Rearranging this, I have $0 < (1 - \alpha)(1 + \beta + \gamma)\beta(1 + \theta) + \gamma(1 + \theta) + \gamma\alpha(\beta + \theta)$, which is always true. *Q.E.D.*

From the equilibrium analysis, $\tilde{\tau}_p$ induces a partial-retirement equilibrium if and only if $\tilde{\tau}_p < \hat{\tau}$.

Lemma 5. $\tilde{\tau}_p \leq \hat{\tau}$ if and only if $\theta \geq \hat{\theta}$, with equality holding if $\theta = \hat{\theta}$.

Proof. $\tilde{\tau}_p < \hat{\tau}$ if and only if

$$\frac{\gamma\{\beta + \gamma - \theta - \alpha[1 + \gamma + \beta(2 + \theta)]\}}{1 + \beta + \gamma} < \beta(1 - \alpha)\theta - \gamma\alpha.$$

Rearranging this equation, I obtain

$$\alpha\beta[\theta(1 + \beta) - \gamma] < (\beta + \gamma)[\theta(1 + \beta) - \gamma]. \quad (29)$$

Hence, if $\theta > \hat{\theta}$, $\theta(1 + \beta) - \gamma > 0$, and Equation (29) is equivalent to

$$\alpha < \frac{\beta + \gamma}{\beta}.$$

Since $\alpha \in (0, 1)$, this inequality always holds. If $\theta < \hat{\theta}$, $\theta(1 + \beta) - \gamma < 0$. Then, Equation (29) is equivalent to

$$\alpha > \frac{\beta + \gamma}{\beta},$$

which does not hold at all because $\alpha \in (0, 1)$. When $\theta = \hat{\theta}$, $\tilde{\tau}_p = \hat{\tau}$.

Q.E.D.

A.5.3 Proof of Proposition 6

Proof. When $\theta \leq \hat{\theta}$, by Lemmas 3 and 4, $\bar{\alpha} < \tilde{\alpha} < \hat{\alpha}$ holds. If $\alpha \geq \hat{\alpha} > \bar{\alpha}$, for any $\tau \in [0, 1)$, an old agent fully retires in equilibrium. Moreover, if $\alpha > \tilde{\alpha}$, a lower tax results in higher welfare. Hence, the optimal payroll tax rate is 0.

Suppose that $\alpha \geq \tilde{\alpha}$ and $\alpha < \hat{\alpha}$. Since $\alpha > \bar{\alpha}$, an old agent fully retires in equilibrium for all $\tau \in [0, 1)$. Since $\alpha \geq \tilde{\alpha}$, a lower tax rate is better. Therefore, the optimal tax rate is 0.

Suppose that $\alpha \geq \bar{\alpha}$ and $\alpha < \tilde{\alpha}$. Similar to the above cases, an old agent fully retires in equilibrium for all $\tau \in [0, 1)$. Since $\alpha < \tilde{\alpha}$, by Lemma 3, $\tilde{\tau}_f \in (0, 1)$ is optimal.

Finally, suppose that $\alpha < \bar{\alpha}$. Since $\alpha < \tilde{\alpha}$, $\tilde{\tau}_f$ gives the highest welfare when an old agent fully retires, whereas $\hat{\tau}$ maximizes welfare when an old agent partially retires. Since under $\tau = \hat{\tau}$, an old agent fully retires, $W(\tilde{\tau}_f) > W(\hat{\tau})$ holds. Hence, $\tilde{\tau}_f$ is optimal.

When $\theta > \hat{\theta}$, $\hat{\alpha} \leq \tilde{\alpha} \leq \bar{\alpha}$ holds. If $\alpha \geq \bar{\alpha}$, an old agent fully retires in equilibrium for all $\tau \in [0, 1)$. Since $\alpha \geq \tilde{\alpha}$, a lower tax rate is better. Therefore, $\tau = 0$ is optimal, and an old agent fully retires.

If $\alpha \geq \tilde{\alpha}$ and $\alpha < \bar{\alpha}$, $\tau = \hat{\tau} > 0$. In this case, since $\alpha \geq \tilde{\alpha}$, if an old agent fully retires, $\tau = \hat{\tau}$ gives the highest welfare. If an old agent partially retires, from Proposition 9, $\tau = 0$ gives the highest welfare. Note that from Proposition 9, $W(\tau)$ is strictly decreasing in $\tau \in (0, \hat{\tau})$ and $W(0) > W(\tau)$ for all $\tau \in (0, \hat{\tau})$. Since W is continuous in τ ,¹⁸ $W(0) > W(\hat{\tau})$. Therefore, the optimal tax rate is 0, and an old agent partially retires.

If $\alpha \geq \hat{\alpha}$ and $\alpha < \tilde{\alpha}$, then, by Lemmas 1 and 2, $\hat{\tau}$ gives the highest welfare if an old agent fully retires in equilibrium. From Proposition 9, $\tau = 0$ gives the highest welfare if an old agent partially retires in equilibrium. The same reason in the previous case is applicable to this case. Thus, the optimal tax rate is $\tau = 0$, and an old agent partially retires.

If $\alpha < \hat{\alpha}$, by Lemmas 1 and 2, $\tilde{\tau}$ gives the highest welfare if an old agent fully retires in equilibrium. If an old agent partially retires in equilibrium, then, by Proposition 9, $\tilde{\tau}_p$ maximizes welfare. However, because of the complexity of the calculation, it is not possible to determine which tax rate is optimal. *Q.E.D.*

¹⁸Note that k is continuous in τ , all variables are determined by k , and W is continuous in k . Therefore, W is continuous in τ .

A.6 Proof of Proposition 7

Proof. When $\theta = \hat{\theta}$, Lemma 5 implies that $\tilde{\tau}_p = \hat{\tau}$. From Lemma 2, if $\alpha < \frac{\beta + \gamma - \hat{\theta}}{1 + \gamma + \beta(2 + \hat{\theta})}$, then $\tilde{\tau}_f > \hat{\tau}$. The equation $\alpha < \frac{\beta + \gamma - \hat{\theta}}{1 + \gamma + \beta(2 + \hat{\theta})}$ is equivalent to

$$\alpha < \frac{(1 + \beta)(\beta + \gamma) - \gamma}{(1 + \beta)(1 + \gamma) + \beta(2 + 2\beta + \gamma)}.$$

Hence, $W(\tilde{\tau}_f) > W(\hat{\tau}) = W(\tilde{\tau}_p)$.

Q.E.D.

A.7 Proof of Proposition 8

Proof. Since $c_f^y = (1 - \tilde{\tau}_f)(1 - \alpha)(k_f(\tilde{\tau}_f))^\alpha - k_f(\tilde{\tau}_f)$,

$$c_f^y = k_f(\tilde{\tau}) \frac{1 - \alpha}{\alpha(1 + \beta)}.$$

By analogy, since $c_p^y = (1 - \tilde{\tau}_p)(1 - \alpha)(k_p(\tilde{\tau}_p))^\alpha - k_p(\tilde{\tau}_p)[1 + \theta(1 - l_p^a)]$,

$$c_p^y = k_p(\tilde{\tau}_p) \frac{(1 - \tilde{\tau}_p)(1 - \alpha)(\alpha + \theta)}{\alpha[\beta(1 - \alpha) + \gamma]}.$$

From these,

$$\frac{c_f^y}{c_p^y} = \frac{k_f(\tilde{\tau}_f)}{k_p(\tilde{\tau}_p)} \frac{\beta(1 - \alpha) + \gamma}{(1 + \beta)(\alpha + \theta)(1 - \tilde{\tau}_p)}.$$

Using the expressions of k_f and k_p (Equations (9) and (12)), I obtain

$$\begin{aligned} \frac{c_f^y}{c_p^y} &= \left[\frac{\alpha(1 + \beta + \gamma) + \theta(1 + \alpha\beta)}{\beta(1 - \alpha) + \gamma} \right]^{\frac{1}{1-\alpha}} \frac{\beta(1 - \alpha) + \gamma}{(1 + \beta)(\alpha + \theta)} \\ &\times \frac{(1 - \alpha)(1 + \beta + \gamma)[\beta(1 + \theta) + \gamma]}{(1 + \beta + \gamma)(1 - \alpha)\beta(1 + \theta) + \gamma[1 + \alpha\beta + \theta(1 + \alpha)]}. \end{aligned} \quad (30)$$

When $\theta = 1$ and $\beta = \gamma$, Equation (30) becomes

$$\left[\frac{\alpha(1 + 2\beta) + 1 + \alpha\beta}{\beta(1 - \alpha) + \beta} \right]^{\frac{1}{1-\alpha}} \frac{\beta(1 - \alpha) + \beta}{(1 + \beta)(\alpha + 1)} \frac{(1 - \alpha)(1 + 2\beta)3\beta}{(1 + 2\beta)(1 - \alpha)2\beta + \beta[2 + \alpha\beta + \alpha]}.$$

Thus,

$$\lim_{\alpha \rightarrow 0} \frac{c_f^y}{c_p^y} = \frac{3\beta(1+2\beta)}{4\beta(1+\beta)^2}.$$

Since $r_p = \frac{\alpha(1+\beta+\gamma)+\theta(1+\alpha\beta)}{\beta(1-\alpha)+\gamma}$, when $\beta = \gamma$ and $\theta = 1$,

$$r_p = \frac{\alpha(1+2\beta)+1+\alpha\beta}{\beta(1-\alpha)+\beta}.$$

Thus,

$$\lim_{\alpha \rightarrow 0} r_p = \frac{1}{2\beta}.$$

From Equations (10) and (28), when $\beta = \gamma$ and $\theta = 1$,

$$\lim_{\alpha \rightarrow 0} l_p^o = \frac{2\beta}{1+2\beta}.$$

Hence,

$$\lim_{\alpha \rightarrow 0} [W(\tilde{\tau}_f) - W(\tilde{\tau}_p)] = \ln \left(\frac{3^{1+\beta}(1+2\beta)^{1+2\beta}}{2^{2+2\beta}(1+\beta)^{2+2\beta}} \right).$$

If $\frac{3^{1+\beta}(1+2\beta)^{1+2\beta}}{2^{2+2\beta}(1+\beta)^{2+2\beta}} < 1$, $\lim_{\alpha \rightarrow 0} [W(\tilde{\tau}_f) - W(\tilde{\tau}_p)] < 0$.

Let

$$h(\beta) := \frac{3^{1+\beta}(1+2\beta)^{1+2\beta}}{2^{2+2\beta}(1+\beta)^{2+2\beta}} = \left(\frac{1+2\beta}{1+\beta} \right)^{1+2\beta} \left(\frac{3}{4} \right)^{1+\beta} \frac{1}{1+\beta}.$$

Taking the derivative of h with respect to β , I obtain

$$\frac{dh}{d\beta} = \left(\frac{1+2\beta}{1+\beta} \right)^{2\beta} \left(\frac{3}{4} \right)^{1+\beta} \frac{1}{(1+\beta)^2} \left[2 \left\{ \ln \left(\frac{1+2\beta}{1+\beta} \right) + \frac{\beta}{(1+2\beta)(1+\beta)} \right\} (1+2\beta) + (1+2\beta) \ln \left(\frac{3}{4} \right) - \frac{2\beta}{1+\beta} \right].$$

The fourth term of $\frac{dh}{d\beta}$ is

$$2(1+2\beta) \ln \left(\frac{1+2\beta}{1+\beta} \right) + (1+2\beta) \ln \left(\frac{3}{4} \right) = (1+2\beta) \ln \left(\frac{1+4\beta+4\beta^2}{1+2\beta+\beta^2} \cdot \frac{3}{4} \right).$$

Note that $\beta + \gamma > 1$ and $\beta = \gamma$ imply that $\beta > \frac{1}{2}$. For $\beta \in (1/2, 1]$, $3(1+4\beta+4\beta^2) > 4(1+2\beta+\beta^2)$ because $8\beta^2+4\beta-1 = 8\left(\beta+\frac{1}{4}\right)^2 - \frac{3}{2}$, which is increasing in β on $(1/2, 1]$, and at $\beta = 1/2$, the value is 3. Therefore, since

$$\ln \left(\frac{1+4\beta+4\beta^2}{1+2\beta+\beta^2} \cdot \frac{3}{4} \right) > 0,$$

$h(\beta)$ is strictly increasing in β on $(1/2, 1]$. Since

$$h(1) = \left(\frac{3}{2}\right)^3 \left(\frac{3}{4}\right)^2 \frac{1}{2} = \frac{243}{256} < 1$$

even for a sufficiently small α , $W(\tilde{\tau}_p) > W(\tilde{\tau}_f)$ holds.

Q.E.D.

References

- Abel, Andrew B., Gregory M. Mankiw, Lawrence H. Summers, and Richard J. Zeckhauser (1989) Assessing dynamic efficiency: Theory and evidence. *Review of Economic Studies* 56, 1–19.
- Aísa, Rosa, Fernando Pueyo, and Marcos Sanso (2012) Life expectancy and labor supply of the elderly. *Journal of Population Economics* 25, 545–568.
- Barbie, M., Marcus Hagedorn, and Ashok Kaul (2007) On the interaction between risk sharing and capital accumulation in a stochastic OLG model with production. *Journal of Economic Theory* 137, 568–579.
- Chattopadhyay, Subir (2006) Optimality in stochastic OLG models: Theory for tests. *Journal of Economic Theory* 131, 282–294.
- Diamond, Peter A. (1965) National debt in a neoclassical growth model. *American Economic Review* 55, 1126–1150.
- Fenge, Robert and Pierre Pestieau (2007) *Social security and early retirement*. The MIT Press.
- Gruber, Jonathan and David A. Wise (1999) *Social security and retirement around the world*. University of Chicago Press.
- Hu, Sheng-Cheng (1979) Social security, the supply of labor, and capital accumulation. *American Economic Review* 69, 274–283.
- Kitao, Ssagiri (2014) Sustainable social security: Four options. *Review of Economic Dynamics* 17, 756–779.
- Michel, Philippe and Pierre Pestieau (2013) Social security and early retirement in an overlapping-generations growth model. *Annals of Economics & Finance* 14, 705–719.
- Miyazaki, Koichi (2014) The effects of the raising-the-official-pension-age policy in an overlapping generations economy. *Economics Letters* 123, 329–332.
- Mizuno, Masakatsu and Akira Yakita (2013) Elderly labor supply and fertility decisions in aging-population economies. *Economics Letters* 121, 395–399.
- Sala-i-Martin, Xavier X (1996) A positive theory of social security. *Journal of Economic Growth* 1, 277–304.

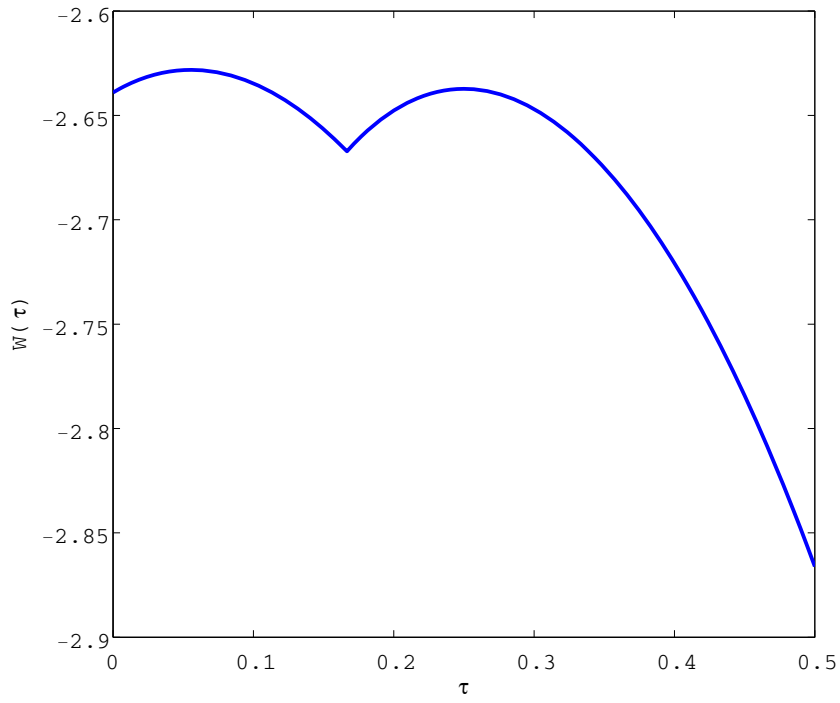


Figure 1: $W(\tau)$ when $\alpha = 0.2$

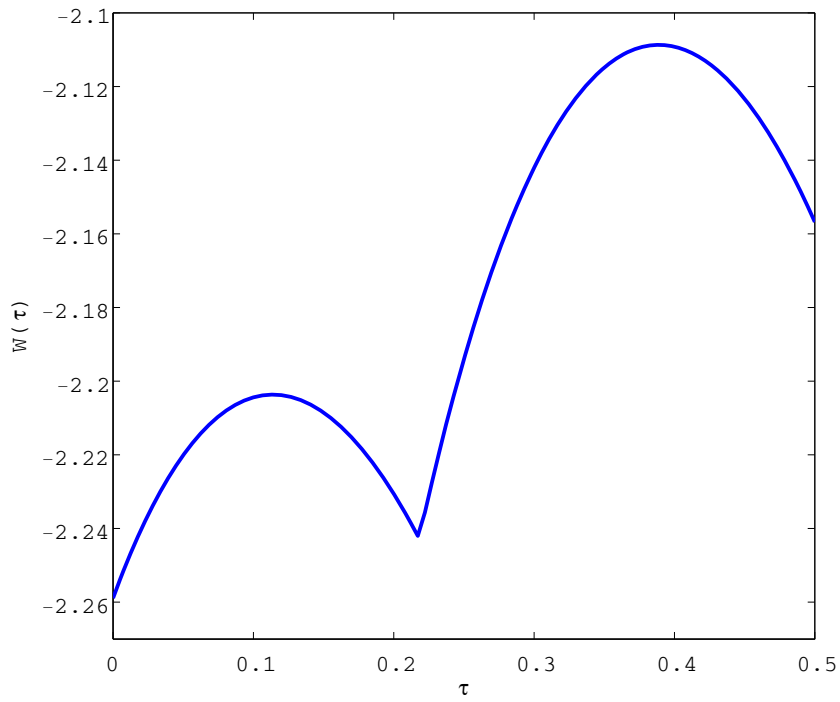


Figure 2: $W(\tau)$ when $\alpha = 0.1$