Discussion Papers In Economics And Business

Retirement and Social Security: A Political Economy Perspective

Ryo Arawatari Tetsuo Ono

Discussion Paper 10-04-Rev.

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

Retirement and Social Security: A Political Economy Perspective

Ryo Arawatari　Tetsuo Ono

Discussion Paper 10-04-Rev.

September 2010

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

Retirement and Social Security: A Political Economy Perspective*

Ryo Arawatari[†] Tetsuo Ono[‡]

Shinshu University Osaka University

September 4, 2010

Abstract

Countries with higher implicit taxes on continued work are associated with lower labor force participation rates of the elderly. This paper constructs a politico-economic model that accounts for this feature based on multiple, self-fulfilling expectations of agents. In this model, agents are identical at birth and can become skilled (or remain unskilled) through educational investment. When agents hold expectations of larger social security benefits, it provides a disincentive to engage in educational investment, thereby resulting in an unskilled majority. In turn, this unskilled majority supports larger social security benefits, which induces the retirement of the elderly and thus results in a lower labor force participation rate. The opposite applies when agents have expectations of smaller social security benefits in their old age.

Key words Political equilibrium; Retirement; Self-fulfilling expectations; Social security

JEL classification H55, D72, J26

*Arawatari acknowledges the financial support of the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for Young Scientists (Startup) (No. 21830046). Ono also acknowledges the financial support of the Japan Society for the Promotion of Science (JSPS) through a Grant-in-Aid for Young Scientists (B) (No. 20730208).

[†]Faculty of Economics, Shinshu University, 3-1-1, Asahi, Matsumoto, Nagano, 390-8621 Japan. E-mail: ryo_a@shinshu-u.ac.jp

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

1 Introduction

The labor force participation rate of the elderly in the OECD is negatively correlated with taxes on their continued work. For instance, Daval (2003) and Fenge and Pestieau (2005) argue that countries in the OECD can be classified into two groups: one characterized by a higher tax burden and lower labor force participation rate of the elderly, and the other characterized by a lower tax burden and higher labor force participation rate of the elderly. The former includes Austria, France, Germany and Italy; the latter includes Canada, Norway, Sweden, the United Kingdom, and the United States.

Several studies have attempted to reveal the mechanism underlying this interesting observation (see, for example, Sheshinski, 1978; Crawford and Lilien, 1981; and the recent survey by Feldstein and Liebman, 2002). In general, continued work in older age may be subject to the burden of payroll tax and forgone pension benefits. This double burden, regarded as an implicit tax on the elderly (Gruber and Wise, 1999), provides the elderly with an incentive to retire earlier. That is, a higher tax burden results in earlier retirement and thus a lower labor force participation rate of the elderly. Indeed, Gruber and Wise (1999) argue that this implicit tax on the elderly explains more than 80% of the cross-country variation in unused labor capacity for those 55 to 65 years of age.

All of the above-mentioned studies focus on the effect of the tax burden on retirement decisions. However, little attention has been given to the question why some countries choose a higher tax burden on the elderly while others select a lower tax burden, even though they share a similar political and economic background. This paper develops a politico-economic framework that responds to this important question, and demonstrates how the retirement behavior of the elderly is affected by voting on taxation. We then show the mechanism underlying the differences in social security and retirement behavior across the OECD countries.

Our framework is based on that developed in Hassler, Storesletten and Zilibotti (2007). In this approach, agents identical at birth can affect their prospects in life with educational investment. In particular, agents become either skilled or unskilled, and by undertaking costly investment can increase the probability of becoming skilled in their youth. This framework is further enriched by allowing for endogenous retirement decisions in old age. Here, each type of old agent obtains the opportunity to continue to work with some probability and makes his/her decision on retirement from the viewpoint of utility maximization. The government is assumed to provide old-age social security financed by taxes levied on the young and the working old.

Under the above-mentioned framework, there are multiple, self-fulfilling expectations of agents. In undertaking educational investment, young agents have expectations of social security in their old age. When young agents hold expectations of higher social security benefits, it provides a disincentive to engage in educational investment, thereby resulting in a lower proportion of the skilled. This implies a greater number of unskilled young individuals, which in turn increases the demand for old-age social security and thus enhances retirement. The opposite applies when young agents have expectations of lower social security in old age. The economy may then attain multiple equilibria, including a skilled-majority equilibrium featuring a low tax burden and late retirement, and an unskilled-majority equilibrium featuring a high tax burden and early retirement. Ultimately, the particular outcome attained depends on the expectations of agents.

The finding of multiple equilibria well fits the empirical evidence. In the skilled-majority equilibrium, the skilled old choose a lower level of social security and thus a lower tax burden on the working old. This gives the old an incentive to continue to work, thereby resulting in the higher labor force participation rate of the elderly. The opposite applies in the unskilled-majority equilibrium. In terms of real-world observations, the skilled-majority equilibrium represents a group of countries including Canada, Norway, Sweden, the United Kingdom, and the United States; and the unskilled-majority equilibrium comprises a group of countries including Austria, France, Germany and Italy (Daval, 2003; Fenge and Pestieau, 2005).

Interestingly, Norway and Sweden are much less unequal than Canada, the United Kingdom and the United States, even though they present a similar level of taxes and comparable labor force participation of the elderly. However, while the evidence of similarity at first appears counterintuitive, this can be explained by self-fulfilling expectations. That is, people in Norway and Sweden happen to have similar expectations, as do those in Canada, the United Kingdom and the United States, regarding future social security. This similarity of expectations leads to an identical pattern of taxation and labor force participation in each of the two groups of countries.

We undertake a numerical analysis of the pension–GDP ratio to enable further investigation of these cross-country differences. We find that the unskilled-majority equilibrium attains a higher pension–GDP ratio than the skilled-majority equilibrium, and a more equal society realizes a lower pension–GDP ratio. This numerical result appears to fit the available empirical evidence. In fact, the ratio in Austria, France, Germany and Italy, representing the unskilled-majority equilibrium, is about twice that of Canada, Norway, Sweden, the United Kingdom and the United States, representing the skilled-majority equilibrium. In addition, the evidence indicates a negative correlation between wage equality and the size of pensions across the OECD countries.

We also pursue a numerical investigation of the expected utility of the young in order to consider the welfare implications of wage inequality and the multiple expectations of agents. We find that the young obtain a higher expected utility in the skilled-majority equilibrium than in the unskilled-majority equilibrium. This is mainly because the former equilibrium requires a lower tax burden on the old and thus realizes a higher probability of being successful in education. We also find that more wage equality results in a higher expected utility of the young in the skilled-majority equilibrium, and a lower expected utility of the young in the unskilled-majority equilibrium. These dissimilar results appear to arise from differences in voting behavior in the two equilibria.

The current paper relates to the following two strands of literature. The first strand is the literature on the political economy of social security and retirement, including studies focusing on early retirement provisions (Conde-Ruiz and Galasso, 2003, 2004), old-age social security provisions (Profeta, 2002; Casamatta, Cremer and Pestieau, 2005, 2006), and the legal retirement age (Lacomba and Lagos, 2006, 2007; Galasso, 2008). The current paper is similar to that of Casamatta, Cremer and Pestieau (2005, 2006) in that we consider the political determination of social security and its impact on retirement. However, our work differs in that we introduce a link between current economic decisions and the expectations of future social security, and this creates the multiple, self-fulfilling expectations of agents that help to explain the empirical evidence.

The second strand is the literature on the dynamic political economy of social security based on the concept of Markov-perfect equilibrium. This body of work includes studies demonstrating a unique equilibrium pinned down by the initial (Grossman and Helpman, 1998; Azariadis and Galasso, 2002) and multiple, self-fulfilling expectations of agents (Hassler et al., 2003; Hassler, Storesletten and Zilibotti, 2007). Conde-Ruiz, Galasso and Profeta (2005) extend the framework in the first two studies by introducing the retirement decisions of old agents. In contrast, the current paper extends the framework in the latter two studies to demonstrate multiple equilibria as regards social security and retirement. The current paper also shows that the self-fulfilling expectations of agents explain why two different countries in terms of wage inequality display similar levels of taxes and labor force participation of the elderly.

The organization of this paper is as follows. Section 2 develops the model. Section 3 demonstrates the multiple political equilibria. Section 4 undertakes the numerical analysis, and Section 5 discusses two extensions. Section 6 provides some concluding remarks.

2 The Model

The model is a two-period-lived, overlapping-generations model based on that discussed in Hassler, Storesletten and Zilibotti (2007). Time is discrete and denoted by $t = 0, 1, 2, \cdots$. The economy comprises a continuum of agents living for two periods, youth and old age. Each generation has a unit mass. Agents are identical at birth.

Consider the young agents born in period t. They can affect their prospects in life with educational investment. In particular, they become either skilled or unskilled (denoted by s and u, respectively), and by undertaking costly investment, they can increase the probability e_t of becoming skilled in their youth. The cost of investment, measured in terms of disutility, is given by $(e_t)^2/2$. Skilled agents earn a high wage, normalized to unity, whereas unskilled agents earn a low wage, normalized to $w \in [0, 1)$ over their life cycle. A lower w implies higher wage inequality between the two types of agents.

At the beginning of period $t + 1$, there are two types of old agents: the skilled and the unskilled. Our model departs from Hassler, Storesletten and Zilibotti (2007) in the following respects. First, in old age, each agent faces job loss with a probability of $1 - \mu \in [0, 1]$ because of job replacement or deterioration in health. However, they may obtain opportunities to continue to work with a probability of μ. The situation where these opportunities differ between skilled and unskilled old agents will be examined in Section 4. Second, the cost of continuing to work is given by $(l^j_{t+1})^2/2$ $(j = s, u)$, where $l^j_{t+1} \in [0, 1]$ is the amount of labor supplied by a type-j old agent. Figure 1 illustrates the timing of events.

[Figure 1 about here.]

There is no storage technology in this economy: each individual uses his/her endowments within a given period. The government provides old-age social security, b, financed by taxes levied on the young and the working old. The tax rates are age dependent: τ^o for the old and τ^y for the young. In turn, the tax rates are determined before the young agents decide on their investment and before the old agents decide on their retirement.

Therefore, the expected utility functions of the agents alive at time t are given as follows:

$$V_t^{os} = \mu \left\{ l_t^s(1-\tau_t^o) - \frac{(l_t^s)^2}{2} \right\} + b_t, \tag{1}$$

$$V_t^{ou} = \mu \left\{ w l_t^u(1-\tau_t^o) - \frac{(l_t^u)^2}{2} \right\} + b_t, \tag{2}$$

$$\begin{aligned} V_t^y &= e_t(1-\tau_t^y) + (1-e_t)w(1-\tau_t^y) - \frac{(e_t^y)^2}{2} \\ &\quad + \beta \left[e_t \left\{ \mu \left(l_{t+1}^s (1-\tau_{t+1}^o) - \frac{(l_{t+1}^s)^2}{2} \right) + b_{t+1} \right\} \right. \\ &\quad \left. +(1-e_t^y) \left\{ \mu \left(w l_{t+1}^u (1-\tau_{t+1}^o) - \frac{(l_{t+1}^u)^2}{2} \right) + b_{t+1} \right\} \right], \end{aligned} \tag{3}$$

where V_t^{os}, V_t^{ou} and V_t^y denote the utility of the skilled old, the utility of the unskilled old, and the expected utility of the young, respectively. The utility levels of V_t^{os}, V_t^{ou} and V_t^y are computed prior to individual success or failure in education. The parameter $\beta \in (0,1)$ is a discount factor.

Given these preferences, the skilled old agent chooses l_t^s to maximize V_t^{os}; the unskilled old agent chooses l_t^u to maximize V_t^{ou}; and the young agent in period t chooses e_t to maximize V_t^y by taking account of the optimal labor supply in his/her old age, l_{t+1}^s and l_{t+1}^u. The optimal choices of the old and the young are respectively given by:

$$l^{s*}(\tau_t^o) = 1-\tau_t^o;\ l^{u*}(\tau_t^o) = w(1-\tau_t^o); \tag{4}$$

$$e^{y*}(\tau_t^y, \tau_{t+1}^o) = (1-w)\left[(1-\tau_t^y) + \beta\mu(1+w)(1-\tau_{t+1}^o)^2/2 \right]. \tag{5}$$

Young agents are ex ante identical, and therefore agents in the same cohort choose the same investment. This implies that at the beginning of period $t+1$, the proportion of the unskilled old, u_{t+1}, is equal to the proportion of the unsuccessful period-t young agents, $1-e^{y*}(\tau_t^y, \tau_{t+1}^o)$:

$$u_{t+1} \equiv 1 - e^{y*}(\tau_t^y, \tau_{t+1}^o) = 1 - (1-w)\left[(1-\tau_t^y) + \beta\mu(1+w)(1-\tau_{t+1}^o)^2/2 \right].$$

The proportion of the unskilled old at the time of voting, u_{t+1}, depends on the tax levied on the young in period t, τ_t^y, and the tax levied on working old agents in period $t+1$, τ_{t+1}^o.

The tax revenues from the young and the working old are transferred to every old agent in a lump-sum fashion. The government budget is balanced in each period so it can be expressed as:

$$b_t = W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o),$$

where $W(\tau_t^o, u_t)$ is the tax revenue from the old and $Z(\tau_t^y, \tau_{t+1}^o)$ is the tax revenue from the young, defined by:

$$\begin{aligned} W(\tau_t^o, u_t) &\equiv \{(1-u_t) + u_t w^2\}\mu(1-\tau_t^o)\tau_t^o; \\ Z(\tau_t^y, \tau_{t+1}^o) &\equiv \left[w + (1-w)^2 \left\{ (1-\tau_t^y) + \beta\mu(1+w)(1-\tau_{t+1}^o)^2/2 \right\} \right] \tau_t^y. \end{aligned}$$

3 Political Equilibria

This section characterizes the political equilibria where agents vote on taxation period by period. Section 3.1 provides the definition of a political equilibrium based on the concept of a stationary Markov-perfect equilibrium with majority voting. Sections 3.2 and 3.3 provide the characterization of political equilibria classified according to the type of majority.

3.1 Definition of Political Equilibrium

Following Hassler, Storesletten and Zilibotti (2007), we assume that elections are held at the end of each period and the elected politicians set the tax rates for the following period. The old abstain from voting because they have no interest in the following period; only the young participate in voting. This is observationally equivalent to assuming that only old agents vote over current taxes at the beginning of each period. We adopt the latter interpretation in the following analysis. Our assumption about voting implies that we focus on the intragenerational, rather than intergenerational, conflict over redistribution. With optimal choices $l^{s*}(\tau_t^o), l^{u*}(\tau_t^o)$ and $e^*(\tau_t^y, \tau_{t+1}^o)$ and the government budget constraint, the indirect utility functions of the skilled and the unskilled old are respectively given by:

$$
\begin{aligned}
V_t^{os} &= \frac{\mu}{2}(1-\tau_t^o)^2 + W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o), \\
V_t^{ou} &= \frac{\mu}{2}(w)^2(1-\tau_t^o)^2 + W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o),
\end{aligned}
$$

where the term $\mu(1-\tau_t^o)^2/2$ in the first line is the expected after-tax income of the skilled old, the term $\mu(w)^2(1-\tau_t^o)^2/2$ in the second line is the expected after-tax income of the unskilled old, and the term $W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o)$ is the old-age social security benefit.

This paper focuses on stationary Markov-perfect equilibria with majority voting. The proportion of the unskilled old (u_t) summarizes the state of the economy; the identity of a decisive voter depends on this proportion. An office-seeking politician elected by voters sets policies to maximize the utility of the larger group. Given these features, we now provide the definition of the political equilibrium as follows.

Definition: A (*stationary Markov-perfect*) *political equilibrium* is defined as a triplet of functions $\{T^o, T^y, U\}$, where $T^o : [0,1] \to [0,1]$ and T^y are two public policy rules, $\tau_t^o = T^o(u_t)$ and $\tau_t^y = T^y$, and $U : [0,1] \to [0,1]$ is a private decision rule, $u_{t+1} = U(\tau_t^y)$, such that given u_0, the following functional equations hold.

1. $T^o(u_t) = \arg\max_{\tau_t^o \in [0,1]} W^{dec}(\tau_t^o, u_t)$ $(dec = os, ou)$, where:

$$
W^{dec}(\tau_t^o, u_t) = \begin{cases} W^{os}(\tau_t^o, u_t) \equiv \frac{\mu}{2}(1-\tau_t^o)^2 + W(\tau_t^o, u_t) & \text{if } u_t \leq 1/2, \\ W^{ou}(\tau_t^o, u_t) \equiv \frac{\mu}{2}(w)^2(1-\tau_t^o)^2 + W(\tau_t^o, u_t) & \text{if } u_t > 1/2 \end{cases}
$$

2. $U(\tau_t^y) = 1 - e^*(\tau_t^y, \tau_{t+1}^o)$, with $\tau_{t+1}^o = T^o(U(\tau_t^y))$

3. $T^y = \arg\max_{\tau_t^o \in [0,1]} Z(\tau_t^y, \tau_{t+1}^o)$ subject to $\tau_{t+1}^o = T^o(U(\tau_t^y))$

The first equilibrium condition requires that the decisive voter chooses τ_t^o to maximize the utility of the skilled old (if $u_t < 1/2$) or the unskilled old (if $u_t > 1/2$). In the case of equal numbers of skilled and unskilled old agents (i.e., $u_t = 1/2$), the skilled old are assumed to be decisive. The second equilibrium condition implies that all young individuals choose their investment optimally, given τ_t^y and τ_{t+1}^o, under rational expectations about future taxes and the distribution of types. The third equilibrium condition requires that the decisive old voter chooses τ_t^y to maximize revenue from the young. Rational voters understand that their choice over current redistribution affects future redistribution via the private decision rule and public policy.

3.2 The Determination of T^o and U

We now solve the equilibrium conditions recursively. Condition 1 defines a one-to-one mapping from the state variable to the equilibrium choice of taxation of the old: $\tau_t^o = T^o(u_t)$. Suppose that the skilled old form the majority: $u_t \leq 1/2$. The objective function of the majority is given by $W^{os}(\tau_t^o, u_t) \equiv \mu(1-\tau_t^o)^2/2 + W(\tau_t^o, u_t)$, which is strictly decreasing in τ_t^o : $\partial W^{os}(\tau_t^o, u_t)/\partial \tau_t^o < 0$. The skilled old pay more than they receive because the unskilled agents pay less than the skilled agents, but the revenue is distributed equally between the skilled and unskilled old agents. Therefore, the skilled old prefer $\tau_t^o = 0$:

$$T^o(u_t) = 0 \text{ if } u_t \in \left[0, \frac{1}{2}\right]. \tag{6}$$

Alternatively, suppose that the unskilled old form the majority: $u_t > 1/2$. The objective function of the majority is given by $W^{ou}(\tau_t^o, u_t) \equiv \mu(w)^2(1-\tau_t^o)^2/2 + W(\tau_t^o, u_t)$, which has the following properties:

$$\left.\frac{\partial W^{ou}(\tau_t^o, u_t)}{\partial \tau_t^o}\right|_{\tau_t^o=0} = \mu(1-u_t)(1-(w)^2) > 0,$$

$$\left.\frac{\partial W^{ou}(\tau_t^o, u_t)}{\partial \tau_t^o}\right|_{\tau_t^o=1/2} = -\frac{1}{2}\mu(w)^2 < 0,$$

$$\left.\frac{\partial^2 W^{ou}(\tau_t^o, u_t)}{\partial \tau_t^{o2}}\right| = \mu\left[-2(1-u_t) + (1-2u_t)(w)^2\right] < 0.$$

Given these properties, there is a unique tax rate τ_t^o that maximizes $W^{ou}(\tau_t^o, u_t)$:

$$\arg\max_{\tau_t^o} W^{ou}(\tau_t^o, u_t) = \frac{(1-u_t)(1-(w)^2)}{2(1-u_t) + (w)^2(2u_t - 1)} \in \left(0, \frac{1}{2}\right).$$

The optimal tax rate chosen by the unskilled old depends on the population of the unskilled old agents, u_t. This makes it impossible to obtain an analytic solution. In fact, the model without retirement decisions (Hassler, Storesletten and Zilibotti, 2007) could derive an analytic solution because the unskilled old choose $\tau_t^o = 1$ irrespective of the state variable u_t. The state-independent tax rates on the old enable us to attain analytically tractable solutions of multiple, self-fulfilling expectations of agents when employing the Hassler, Storesletten and Zilibotti (2007) framework. In order to retain analytical tractability, we impose the following assumption regarding the choice of the tax rate by the unskilled old.

Assumption 1. The politically available tax rates for the old are limited within the range $[0, \tau^o]$, where $\tau^o \in (0, 1)$ satisfies $\partial W^{ou}(\tau_t^o, u_t)/\partial \tau_t^o|_{\tau_t^o=\tau^o} > 0$.

Figure 2 illustrates a situation where Assumption 1 holds. The unskilled old wish to choose the tax rate that attains the top of the Laffer curve $W^{ou}(\tau_t^o, u_t)$: $\tau_t^o = \arg\max W^{ou}(\tau_t^o, u_t)$. However, Assumption 1 restricts their available choices to $[0, \tau^o]$ because of some institutional or political constraint. Given this restriction, the unskilled old choose $\tau_t^o = \tau^o$ from the viewpoint of utility maximization. Under Assumption 1, the mapping that satisfies Equilibrium Condition 1 is summarized as follows:

$$T^o(u_t) = \begin{cases} 0 & \text{if} \quad u_t \le \frac{1}{2}, \\ \tau^o & \text{if} \quad u_t > \frac{1}{2}. \end{cases} \tag{7}$$

[Figure 2 about here.]

Next, we rewrite Equilibrium Condition 2 by substituting in the optimal investment $e^{y*}(\tau_t^y, \tau_{t+1}^o)$. This yields the following functional equation:

$$U(\tau_t^y) = 1 - (1-w)\left\{(1-\tau_t^y) + \frac{\beta\mu}{2}(1+w)\left(1 - T^o(U(\tau_t^y))\right)^2\right\}, \tag{8}$$

where $T^o(\cdot)$ is given by (7). We derive the solution to the functional equation (8) by assuming rational expectations. The solution to the functional equation (8) is given by:

$$U(\tau_t^y) = \begin{cases} U(\tau_t^y, 0) \equiv 1 - (1-w)\left\{(1-\tau_t^y) + \frac{\beta\mu}{2}(1+w)\right\} & \text{if} \quad \tau_t^y \le \bar{\tau}^y \\ U(\tau_t^y, \tau^o) \equiv 1 - (1-w)\left\{(1-\tau_t^y) + \frac{\beta\mu}{2}(1+w)\left(1-\tau^o\right)^2\right\} & \text{if} \quad \tau_t^y > \underline{\tau}^y \end{cases} \tag{9}$$

where $\bar{\tau}^y$ and $\underline{\tau}^y$ are defined as:

$$\bar{\tau}^y \equiv 1 + \frac{-1 + \beta\mu(1-w)(1+w)}{2(1-w)},$$

$$\underline{\tau}^y \equiv 1 + \frac{-1 + \beta\mu(1-w)(1+w)\left(1-\tau^o\right)^2}{2(1-w)}.$$

The interpretation of (9) is as follows. Suppose that the agents in period t expect $\tau_{t+1}^o = 0$. Under this expectation, young agents choose their investment as $e^{y*}(\tau_t^y, 0) = (1-w)\left[(1-\tau_t^y) + \beta\mu(1+w)/2\right]$. By (7), this expectation is rational if $1 - e^{y*}(\tau_t^y, 0) \le 1/2$; that is, if $\tau_t^y \le \bar{\tau}^y$. Next, suppose that the young agents in period t expect $\tau_{t+1}^o = \tau^o$. Under this expectation, young agents choose their investment as $e^{y*}(\tau_t^y, \tau^o) = (1-w)\left[(1-\tau_t^y) + \beta\mu(1+w)(1-\tau^o)^2/2\right]$. By (7), their expectation is rational if $1 - e^{y*}(\tau_t^y, \tau^o) > 1/2$; that is, if $\tau_t^y > \underline{\tau}^y$. Figure 3 illustrates an example of the solutions to the functional equation (9).

[Figure 3 about here.]

There are multiple, self-fulfilling expectations of U for the set of $\tau_t^y \in [\underline{\tau}^y, \bar{\tau}^y]$. The particular U that arises in equilibrium depends on the expectations of agents. To illustrate U in equilibrium, we follow the method in Hassler, Storesletten and Zilibotti (2007) and introduce the critical rate of τ_t^y : $\theta \in (\max(0, \underline{\tau}^y), \bar{\tau}^y]$. The rate θ, which depends on the expectations of agents, is the highest tax rate that can yield a majority of the unskilled old. For $\tau_t^y > \theta$, the majority is the unskilled old; however, for $\tau_t^y \in (\max(0, \underline{\tau}^y), \theta]$, the majority can be either the skilled or the unskilled old depending on agents' expectations.

Given the definition of θ, the solution is now given by:

$$U(\tau_t^y) = \begin{cases} U(\tau_t^y, 0) & \text{if} \quad \tau_t^y \le \max(0, \underline{\tau}^y) \\ \{U(\tau_t^y, 0), U(\tau_t^y, \tau^o)\} & \text{if} \quad \max(0, \underline{\tau}^y) < \tau_t^y \le \theta \\ U(\tau_t^y, \tau^o) & \text{if} \quad \theta < \tau_t^y. \end{cases} \tag{10}$$

Figure 4 illustrates three possible cases of the solution in (10).

[Figure 4 about here.]

3.3 The Determination of T^y and Characterization of the Political Equilibria

Given the characterization of T^o and U satisfying Equilibrium Conditions 1 and 2, respectively, we now consider the political determination of τ_t^y that satisfies Equilibrium Condition 3. Because there are two possible cases of a majority, we introduce corresponding definitions of the political equilibria: a *skilled-majority equilibrium* and an *unskilled-majority equilibrium*. In particular, the first equilibrium condition given by (7) implies that when the majority are skilled, there is a skilled-majority equilibrium where agents expect no taxation on the old in the future ($\tau_{t+1}^o = 0$) and choose τ_t^y to induce a majority of skilled at time $t+1$ ($u_{t+1} \le 1/2$). In contrast, when the majority are unskilled, there is an unskilled-majority equilibrium where agents expect taxation on the old in the future ($\tau_{t+1}^o = \tau^o$) and choose τ_t^y to induce a majority of unskilled at time $t+1$ ($u_{t+1} > 1/2$).

In order to find a τ_t^y that satisfies Equilibrium Condition 3, we note the following properties: (i) $Z(\tau_t^y, 0) > Z(\tau_t^y, \tau^o) \forall \tau_t^y \in [0, 1]$; (ii) $Z(\tau_t^y, 0)$ and $Z(\tau_t^y, \tau^o)$ attain the tops of the Laffer curves at $\tau_t^y = \tau^{ys}$ and τ^{yu}, respectively, where:

$$\tau^{ys} \equiv \arg\max Z(\tau_t^y, 0) = \frac{1}{2(1-w)^2}\left[w + (1-w)^2\left\{1 + \frac{\beta\mu}{2}(1+w)\right\}\right],$$
$$\tau^{yu} \equiv \arg\max Z(\tau_t^y, \tau^o) = \frac{1}{2(1-w)^2}\left[w + (1-w)^2\left\{1 + \frac{\beta\mu}{2}(1+w)(1-\tau^o)^2\right\}\right];$$

and (iii) the expectations of $\tau_{t+1}^o = 0$ and τ^o are rational if and only if $\tau_t^y \le \theta (\le \bar{\tau}^y)$ and $\tau_t^y > \underline{\tau}^y$, respectively. Given these properties, the revenue from the young is illustrated in Figure 5.

[Figure 5 about here.]

Panels (a) and (b) in Figure 5 depict the skilled-majority equilibrium; that is, the revenue from the young is maximized under the expectation of $\tau_{t+1}^o = 0$. In particular, Panel (a) demonstrates the case where $\tau^{ys} \le \underline{\tau}^y$ holds. Under this condition, agents can choose $\tau_t^y = \tau^{ys}$ irrespective of the expectation parameter θ. Revenue is maximized by choosing $\tau_t^y = \tau^{ys}$. In contrast, Panel (b) illustrates the case where $\tau^{ys} > \underline{\tau}^y$ holds. Agents may have an opportunity of choosing $\tau_t^y = \tau^{ys}$ under the expectation of $\tau_{t+1}^o = 0$, but the choice depends on the expectation of θ. The choice of $\tau_t^y = \tau^{ys}$ is feasible if θ is high such that $\theta \geq \tau^{ys}$; however, it is unfeasible if θ is low such that $\theta < \tau^{ys}$. Under the expectation of $\tau^o_{t+1} = 0$, revenue is maximized by setting $\tau^y_t = \min(\theta, \tau^{ys})$, where $\tau^y_t = \tau^{ys}$ is included as a special case.

Panels (c) and (d) illustrate the unskilled-majority equilibrium; that is, revenue from the young is maximized under the expectation of $\tau^o_{t+1} = \tau^o$. In particular, Panel (c) demonstrates the case where $\tau^{yu} \leq 1$, that is, it is feasible to attain the top of the Laffer curve $Z(\tau^y_t, \tau^o)$, while Panel (d) demonstrates the case where $\tau^{yu} > 1$, that is, it is unfeasible to attain the top of the Laffer curve. In both cases, the expectation θ is assumed to be low, implying that the concerned choice that produces the unskilled-majority equilibrium dominates the option of $(\tau^y_t, \tau^o_{t+1}) = (\theta, 0)$ that produces the skilled-majority equilibrium in terms of revenue from the young.

Before proceeding to the analysis, we define the following four equations that help us characterize the political equilibria:

$$\beta\mu = f^1(w) \equiv \frac{w(2-w)}{(1+w)(1-w)^2\left\{(1-\tau^o)^2 - 1/2\right\}},$$
$$\beta\mu = f^2(w) \equiv \frac{2w^2 - 5w + 1}{(1-w)(1+w)[-(1+w) + 2(1-w)^2(1-\tau^o)^2]},$$
$$\beta\mu = f^3(w) \equiv \frac{(1+w) - \{\{w + (1-w)^2\}(1-\tau^o)^2\} - \sqrt{y}}{(1-\tau^o)^4(1-w)^2(1+w)/2},$$
$$\beta\mu = f^4(w) \equiv \frac{2(w^2 - 3w + 1)}{(1+w)(1-w)^2(1-\tau^o)^2},$$

where y in $f^3(w)$ is given by:

$$y \equiv \left[\{w + (1-w)^2\}(1-\tau^o)^2 - (1+w)\right]^2 - (1-\tau^o)^4(w(2-w))^2.$$

The implications of these equations are described below. The graphs of these equations are illustrated in Figure 6. As depicted in the figure, there is a unique level of w, denoted by $\hat{w}$, where f^2, f^3 and f^4 cross: $f^2(w) = f^3(w) = f^4(w)$ at $w = \hat{w}$.

[Figure 6 about here.]

Using these equations, we provide a characterization of each type of equilibrium in turn. We first provide the characterization of the skilled-majority equilibria.

Proposition 1

(i) *Suppose the following conditions hold:*

$$\tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \text{ and } \beta\mu \geq f^1(w).$$

There exists a set of skilled-majority equilibria such that $\forall t$, T^o is given by (7), U is given by (10), and $T^y = \tau^{ys}$. The equilibrium outcome is unique and such that $\forall t$, $\tau^y_t = \tau^{ys}, \tau^o_t = 0$, and $u_t = 1 - (1-w)\{(1-\tau^{ys}) + \beta\mu(1+w)/2\}$.

(ii) *Suppose the following three conditions hold:*

$$\begin{cases} \tau^o \in \left[\frac{2-\sqrt{2}}{2}, \frac{1}{2}\right), \text{ or } \tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \text{ and } \beta\mu < f^1(w); \\ \beta\mu \geq f^2(w); \text{ and} \\ \beta\mu \geq f^3(w). \end{cases}$$

There exists a set of skilled-majority equilibria such that $\forall t$, T^o *is given by (7),* U *is given by (10), and* $T^y = \min(\theta, \tau^{ys})$. *The equilibrium outcome is indeterminate such that* $\forall t$, $\tau_t^y = \min(\theta, \tau^{ys}), \tau_t^o = 0$, *and* $u_t = 1 - (1-w)\{(1 - \min(\theta, \tau^{ys})) + \beta\mu(1+w)/2\}$.

Proof. See Appendix 7.1.

Areas P.1(i) and P.1(ii) in Figure 6 indicate the set of parameters $(w, \beta\mu)$ satisfying the equilibrium conditions in Statements (i) and (ii) in Proposition 1, respectively. The assumption in Statement (i) ensures that it is feasible to choose $\tau_t^y = \tau^{ys}$ under the expectation of $\tau_{t+1}^o = 0$ irrespective of the expectation parameter θ. Because this choice attains the top of the Laffer curve $Z(\tau_t^y, 0)$, the choice dominates any other pair of tax rates from the viewpoint of tax revenue maximization (see Panel (a) in Figure 5). A unique equilibrium with $(\tau_t^y, \tau_{t+1}^o) = (\tau^{ys}, 0)$ then exists when the assumption in Statement (i) holds.

The first assumption in Statement (ii), which is the exact opposite of the first assumption in Statement (i), implies that the choice of τ_t^y depends on the expectations parameter θ when agents expect $\tau_{t+1}^o = 0$. In other words, it is unfeasible to set $\tau_t^y = \tau^{ys}$ irrespective of θ. For example, the decisive voter cannot choose $\tau_t^y = \tau^{ys}$ if θ is low such that $\theta < \tau^{ys}$; however, he/she can choose $\tau_t^y = \tau^{ys}$ if θ is high such that $\theta \geq \tau^{ys}$. The choice of the decisive voter depends on the expectations of agents represented by the parameter θ. The revenue from the young is maximized by setting $\tau_t^y = \min(\theta, \tau^{ys})$, and thus the equilibrium outcome becomes indeterminate, as illustrated by Panel (b) in Figure 5.

Under the current situation, there is either of the following alternative choices for the decisive voter: setting $\tau_t^y = 1$ under the expectation of $\tau_{t+1}^o = \tau^o$ if $\tau^{yu} > 1$, and setting $\tau_t^y = \tau^{yu}$ under the expectation of $\tau_{t+1}^o = \tau^o$ if $\tau^{yu} \leq 1$. The concerned choice is sustained against these two alternatives if $Z(\theta, 0) \geq Z(1, \tau^o)$ and $Z(\theta, 0) \geq Z(\tau^{yu}, \tau^o)$; that is, if $\theta \geq \tilde{\theta}$ and $\theta \geq \hat{\theta}$ where:

$$\tilde{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys} - 1)^2 + 2(\tau^{ys} - \tau^{yu})};$$
$$\hat{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys})^2 - (\tau^{yu})^2}.$$

Figure 7 illustrates the determination of $\tilde{\theta}$ and $\hat{\theta}$. Details of the derivation of $\tilde{\theta}$ and $\hat{\theta}$ are given in the Appendix. Because θ is bounded from above $\bar{\tau}^y$, $Z(\theta, 0) \geq Z(1, \tau^o)$ and $Z(\theta, 0) \geq Z(\tau^{yu}, \tau^o)$ hold if $\theta \in [\tilde{\theta}, \bar{\tau}^y]$ and $\theta \in [\hat{\theta}, \bar{\tau}^y]$. The second and third assumptions in Statement (ii) ensure that $[\tilde{\theta}, \bar{\tau}^y]$ and $[\hat{\theta}, \bar{\tau}^y]$ are nonempty.

[Figure 7 about here.]

The next proposition establishes the conditions for the existence of the unskilled-majority equilibria.

Proposition 2

(i) *Suppose the following conditions hold:*

$$\begin{cases} \tau^o \in \left[\frac{2-\sqrt{2}}{2}, \frac{1}{2}\right), \text{ or } \tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \text{ and } \beta\mu < f^1(w); \text{ and} \\ \beta\mu \leq f^4(w). \end{cases}$$

There exists a set of unskilled-majority equilibria such that $\forall t$, T^o *is given by (7),* U *is given by (10), and* $T^y = \tau^{yu}$. *The equilibrium outcome is unique such that* $\forall t$, $\tau_t^y = \tau^{yu}, \tau_t^o = \tau^o$, *and* $u_t = 1 - (1-w)\{(1-\tau^{yu}) + \beta\mu(1+w)(1-\tau^o)^2/2\}$.

(ii) *Suppose the following conditions hold:*

$$\begin{cases} \tau^o \in \left[\frac{2-\sqrt{2}}{2}, \frac{1}{2}\right), \text{ or } \tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \text{ and } \beta\mu < f^1(w); \text{ and} \\ \beta\mu > f^4(w). \end{cases}$$

There exists a set of unskilled-majority equilibria such that $\forall t$, T^o *is given by (7),* U *is given by (10), and* $T^y = 1$. *The equilibrium outcome is unique such that* $\forall t$, $\tau_t^y = 1, \tau_t^o = \tau^o$, *and* $u_t = 1 - (1-w)\beta\mu(1+w)(1-\tau^o)^2/2$.

Proof. See Appendix 7.2.

Areas P.2(i) and P.2(ii) in Figure 6 indicate the set of parameters $(w, \beta\mu)$ satisfying the equilibrium conditions in Statements (i) and (ii) in Proposition 2, respectively. The first assumptions in Statements (i) and (ii) are identical to the first assumption in Statement (ii) of Proposition 1. These mean that it is not always feasible to choose $\tau_t^y = \tau^{ys}$ under the expectation of $\tau_{t+1}^o = 0$. When this choice is always feasible, the choice dominates any pair of tax rates that induces the unskilled-majority equilibrium.

The second assumption in each statement determines the tax rate on the young τ_t^y that maximizes the revenue from the young under the expectation of $\tau_{t+1}^o = \tau^o$. The second assumption in Statement (i) is equivalent to $\tau^{yu} \leq 1$, implying that it is feasible to choose a $\tau_t^y = \tau^{yu}$ that attains the top of the Laffer curve under the expectation of $\tau_{t+1}^o = \tau^o$. In contrast, the second assumption in Statement (ii) is equivalent to $\tau^{yu} > 1$, implying that it is unfeasible to choose $\tau_t^y = \tau^{yu}$.

Given the argument thus far, we can conclude that the revenue from the young is maximized by choosing $\tau_t^y = \min(\theta, \tau^{ys})$ under the expectation of $\tau_{t+1}^o = 0$, or by choosing $\tau_t^y = \min(\tau^{yu}, 1)$ under the expectation of $\tau_{t+1}^o = \tau^o$. The latter dominates the former if $Z(\min(1, \tau^{yu}), \tau^o) > Z(\theta, 0)$; that is, if $\theta < \min(\tilde{\theta}, \hat{\theta})$. There then exists an unskilled-majority equilibrium under the conditions presented in Proposition 2 as long as agents attach a low value to the expectation parameter θ.

The following proposition summarizes the results established thus far. The proof of the proposition is immediately apparent from Figure 6.

Proposition 3

(i) *Suppose the following conditions hold:*

$$\tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \text{ and } \beta\mu \geq f^1(w).$$

There exists a unique skilled-majority equilibrium as in Proposition 1(i).

(ii) *Suppose the following conditions hold:*

$$\begin{cases} \tau^o \in \left[\frac{2-\sqrt{2}}{2}, \frac{1}{2}\right), \ \textit{or}\ \tau^o \in \left(0, \frac{2-\sqrt{2}}{2}\right) \ \textit{and}\ \beta\mu < f^1(w); \\ \beta\mu \geq f^3(w)\ \textit{for}\ w \in [0, \hat{w}]; \\ \beta\mu \geq f^2(w)\ \textit{for}\ w \in (\hat{w}, 1). \end{cases}$$

Then the equilibrium is indeterminate. There exists both a skilled-majority equilibrium, as in Proposition 1(ii), and an unskilled-majority equilibrium, as in Proposition 2(i), if $\beta\mu \leq f^4(w)$*; there exist both a skilled-majority equilibrium as in Proposition 1(ii) and an unskilled-majority equilibrium, as in Proposition 2(ii), if* $\beta\mu > f^4(w)$.

(iii) *Suppose the following conditions hold:*

$$\begin{cases} \beta\mu < f^3(w)\ \textit{for}\ w \in [0, \hat{w}]; \\ \beta\mu < f^2(w)\ \textit{for}\ w \in (\hat{w}, 1). \end{cases}$$

There exists a unique unskilled-majority equilibrium, as in Proposition 2(i), if $\beta\mu \leq f^4(w)$*; there exists a unique unskilled-majority equilibrium as in Proposition 2(ii) if* $\beta\mu > f^4(w)$.

Statement (i) in Proposition 3 imposes two assumptions, $\tau^o \in \left(0, (2-\sqrt{2})/2\right)$ and $\beta\mu \geq f^1(w)$, for the existence of a unique skilled-majority equilibria with $\tau^y = \tau^{ys}$ and $\tau^o = 0$. The first assumption implies a low tax burden in old age; the second assumption implies a low wage for unskilled workers. These assumptions jointly create an incentive for young agents to invest in education, thereby resulting in a high probability of being successful in education and thus a high ratio of skilled to unskilled workers within a generation. Unskilled-majority equilibrium cannot arise under these assumptions because they strongly induce a skilled-majority equilibrium.

If one of the above assumptions is relaxed, there arises an unskilled-majority equilibrium, as in Statements (ii) and (iii). The incentive for educational investment is then reduced because of a higher tax burden on the old and/or a higher wage for unskilled workers. In addition, the skilled workers may fail to attain the top of the Laffer curve. That is, under the expectation of $\tau^o_{t+1} = 0$, the revenue from the young is maximized by setting $\tau^y_t = \theta$ rather than $\tau^y_t = \tau^{ys}$. This equilibrium outcome depends on the expectations of agents and thus becomes indeterminate in the skilled-majority equilibrium.

A skilled-majority equilibrium and an unskilled-majority equilibrium coexist if the wage for unskilled workers is low such that the assumptions in Statement (ii) hold. Which particular equilibrium arises depends on the expectation of agents: the skilled-majority (the unskilled-majority) equilibrium arises when θ is high (low) such that $Z(\theta, 0) \geq (<) Z(\min(\tau^{yu}, 1), \tau^o)$ holds. In contrast, when the wage for the unskilled is high such that the assumptions in Statement (iii) hold, young agents have little incentive for educational investment. There no longer remains a skilled-majority equilibrium; that is, there is a unique unskilled-majority equilibrium, as in Statement (iii).

The model prediction of multiple equilibria in Statement (ii) is consistent with the empirical evidence for OECD countries. As demonstrated in Daval (2003), we can classify countries in the OECD into two groups in terms of their implicit taxes on continued work and the labor force participation of the elderly. The first, featuring high tax rates and low labor force participation rates, includes Austria, France, Germany and Italy. The second, featuring low tax rates and high labor force participation rates, includes Canada, Norway, Sweden, the United Kingdom and the United States. The former group of countries corresponds to the unskilled-majority equilibrium with $\tau_t^o = \tau^o$, $l^{s*}(\tau^o) = 1 - \tau^o$ and $l^{u*}(\tau^o) = w(1 - \tau^o)$, whereas the latter group of countries corresponds to the skilled-majority equilibrium with $\tau_t^o = 0$, $l^{s*}(0) = 1$ and $l^{u*}(0) = w$. The current model then provides at least one possible explanation for the differences in implicit tax on continued work and labor force participation rates of the elderly across the OECD. The key to this explanation is the multiple, self-fulfilling expectations of agents.

We should note that Norway and Sweden are much less unequal than Canada, the United Kingdom and the United States, even though all present similar levels of taxes and labor force participation of the elderly. This evidence of similarity at first appears counterintuitive. However, the similarity can be explained, using our framework, by the expectation-based tax rate on the young in the skilled-majority equilibrium. That is, people in Norway and Sweden have similar expectations θ as in those in Canada, the United Kingdom, and the United States. This similarity of expectations leads to the same pattern of taxation and comparable labor force participation of the elderly for the two different groups of countries.

4 Numerical Analysis

Until now, we have characterized the political equilibrium and qualitatively assessed the impact of wage inequality and the expectations of agents on the determination of the taxation and labor force participation rate of the elderly. To facilitate understanding, this section numerically investigates how wage inequality and the expectations of agents affect the pension–GDP ratio as well as the expected utility of the young.

For the purpose of the analysis, we assume a generation to be 20 years in duration. The first and second periods correspond to, for example, ages 26–45 and 46–65 years, respectively. The parameters μ and τ^o are assumed to be 0.8 and 0.25, respectively. Our selection of β is based on the single-period discount rate of 0.95. Because the agents under the current assumption plan over generations that span 20 years, we discount the future by $(0.95)^{20}$.

In the current environment, there are three threshold levels of w, as illustrated in Figure 6: $w = 0.0090, 0.2227$, and 0.3642. For $w \in (0, 0.0090)$, there exists a unique skilled-majority equilibrium as in Proposition 1(i); for $w \in (0.0090, 0.2227)$, there exist multiple political equilibria: one is the skilled-majority equilibrium as in Proposition 1(ii), and the other is the unskilled-majority equilibrium as in Proposition 2(i); for $w \in (0.2227, 0.3624)$, there exists a unique unskilled majority equilibrium as in Proposition 2(i); and for $w \in (0.3642, 1)$, there exists a unique unskilled-majority equilibrium as in Proposition 2(ii). We hereafter focus on the range $(0.0090, 0.2227)$ that creates multiple political equilibria because these provide an explanation for the cross-country differences in the tax and labor force participation rate of the elderly as presented in the previous section.

4.1 Cross-country Differences in the Pension–GDP Ratio

Panel (a) in Figure 8 demonstrates how the pension–GDP ratio is affected by wage inequality in the current sample economy. The upper curve shows the pension–GDP ratio in the unskilled-majority equilibrium, whereas the lower set of downward-sloping curves indicates the pension–GDP ratio in the skilled-majority equilibria. As illustrated, the unskilled-majority equilibrium attains a higher pension–GDP ratio than the skilled-majority equilibrium. This is mainly because the tax on the old in the former equilibrium is higher than that in the latter equilibrium.

[Figure 8 about here.]

The numerical result in Panel (a) fits the empirical evidence of Panel (b) in the following ways. First, the evidence indicates that the ratios in Austria, France, Germany and Italy are about twice those in Canada, Norway, Sweden, the United Kingdom and the United States. Multiple equilibria in the current framework can provide a possible explanation for this difference in the pension–GDP ratio between the two groups of countries sharing similar levels of wage inequality.[1]

Second, the evidence indicates a negative correlation between wage equality and the pension–GDP ratio in most countries, except Austria, France, Germany and Italy. A more equal society generally realizes a lower pension–GDP ratio. This correlation can be explained in the current framework by focusing on the set of downward-sloping curves in the skilled-majority equilibria, where more equality in wages is associated with a lower tax on the young[2], which results in a lower per capita tax revenue and thus a smaller pension.

4.2 Expected Utility of the Young

Panel (a) in Figure 9 illustrates the expected utility of the young. As shown, the young obtain a higher expected utility in the skilled-majority equilibrium than in the unskilled-majority equilibrium. This is mainly because the skilled-majority equilibrium requires a lower tax burden on the old and thus realizes a higher probability of being successful in education than the unskilled-majority equilibrium.

[Figure 9 about here.]

One noteworthy feature is that effects of wage equality on expected utility are different between the two equilibria. In the skilled-majority equilibrium, more equality (i.e., a higher w) leads to the higher expected utility of the young. However, in the unskilled-majority equilibrium, more equality results in lower expected utility. These differences in the consequences of wage equality arise from the tax on the young (see Panel (b) in Figure 9).

[1] We should note that the range of w in Panel (a) is qualitatively different from that in Panel (b). This difference comes from the assumption of the fixed tax rate on the old, τ^o, and the model specification of the disutility functions of education and the labor supply. Changes in τ^o and the utility functions would derive the range of w that fits the empirical evidence.

[2] The mechanism underlying this effect is explained in Subsection 4.2.

In the unskilled-majority equilibrium, more equality results in a higher tax rate on the young. A higher w gives agents a disincentive to invest in education because they can obtain a higher wage, even when they fail in education and become poor. This results in fewer skilled agents and thus a smaller tax base. Given the shrinkage of the tax base, the decisive voter chooses to impose a higher tax on the young to maintain tax revenue and thus the level of pension benefit. Therefore, more equality results in a higher tax rate on the young and thus the lower expected utility of the young in the unskilled-majority equilibrium.

In the skilled-majority equilibrium, more equality results in a lower tax rate on the young and thus the higher expected utility of the young. This opposing result to the unskilled-majority equilibrium comes about in the following way. As indicated above, a higher wage for the unskilled gives agents a disincentive to invest in education, thereby resulting in fewer skilled agents. A lower tax on the young offsets this negative size effect and thus keeps the majority skilled. Therefore, more equality results in a lower tax rate on the young and thus a higher expected utility of the young in the skilled-majority equilibrium.

5 Extensions and Further Analysis

The analysis thus far has assumed that: (i) opportunities to continue to work are common to both the skilled and the unskilled; and (ii) the same pension benefits are paid to the skilled and the unskilled. This section briefly considers how the analysis and the results are changed when each of these assumptions is relaxed.

5.1 Different Opportunities for Working in Old Age

The analysis in the previous sections assumed that opportunities to continue to work are common to both the skilled and the unskilled. However, the skilled (i.e., the rich) may have more opportunities to work than the unskilled (i.e., the poor) because the rich are healthier than the poor, as observed in some empirical studies (see Smith, 1999, and the references therein). This subsection introduces differences in opportunities into the model and briefly considers the consequences.

Assume that the probability of obtaining opportunities to work in old age is μ for the unskilled (as in the original model), whereas it is $\gamma\mu(> \mu)$ for the skilled where $\gamma > 1$ holds. Under this assumption, the utility function of the unskilled old is still given by (2), whereas the utility function of the skilled old is given by:

$$V_t^{os} = \gamma\mu \left\{ l_t^s(1-\tau_t^o) - \frac{(l_t^s)^2}{2} \right\} + b_t.$$

Educational investment by the young is now given as follows:

$$e^{y*}(\tau_t^y, \tau_{t+1}^o) = (1-w)\left[(1-\tau_t^y) + \frac{\beta\mu}{2}\left\{ \frac{\gamma-1}{1-w} + (1+w) \right\}(1-\tau_{t+1}^o)^2 \right].$$

This equation indicates that a larger γ results in a higher probability of becoming skilled in youth. The economy would then be more likely to attain the skilled-majority equilibrium if the skilled could obtain a higher probability of opportunities to work in old age.

5.2 Contribution-related Pension

Pension schemes in many OECD countries feature a mixture of Beveridgean and Bismarckian characteristics. The purely Beveridgean system provides a lump-sum benefit to every agent, irrespective of his/her contribution (as in the model in previous sections). In contrast, the purely Bismarckian system gives the benefit to each agent depending on his/her contribution in their youth.

Under a mixture of the Beveridgean and Bismarckian systems, the government budget constraint is given by:

$$[(1-u_t)+u_t\{\alpha w+(1-\alpha)\}]\,b_t=\tilde{W}(\tau_t^o,u_t)+\tilde{Z}(\tau_t^y,\tau_{t+1}^o),$$

where the parameter $\alpha\in[0,1]$ represents the Bismarckian factor, and $\tilde{W}(\tau_t^o,u_t)$ and $\tilde{Z}(\tau_t^y,\tau_{t+1}^o)$ are the tax revenues from the old and young, respectively:

$$\tilde{W}(\tau_t^o,u_t)\equiv\frac{W(\tau_t^o,u_t)}{(1-u_t)+u_t\{\alpha w+(1-\alpha)\}};$$
$$\tilde{Z}(\tau_t^y,\tau_{t+1}^o)\equiv\frac{Z(\tau_t^y,\tau_{t+1}^o)}{(1-u_t)+u_t\{\alpha w+(1-\alpha)\}}.$$

The pension system becomes purely Beveridgean if $\alpha=0$; the system is purely Bismarckian if $\alpha=1$.

The objective functions of the old then change to:

$$V_t^{os}=\frac{\mu}{2}(1-\tau_t^o)^2+\tilde{W}(\tau_t^o,u_t)+\tilde{Z}(\tau_t^y,\tau_{t+1}^o),$$
$$V_t^{ou}=\frac{\mu}{2}(w)^2(1-\tau_t^o)^2+\tilde{W}(\tau_t^o,u_t)+\tilde{Z}(\tau_t^y,\tau_{t+1}^o).$$

The rich old prefer $\tau_t^o=0$, whereas the poor old prefer $\tau_t^o=\tau^o$ under Assumption 1. The qualitative properties of the analysis and the results of previous sections would then not change if a contribution-related pension system were introduced into the model.

6 Conclusion

This paper focuses on the negative correlation between the labor force participation rate of the elderly and the implicit tax on continued work in older ages. Some previous studies demonstrate the mechanism of how a higher tax gives the elderly an incentive to retire early and thus results in their lower participation rate. However, few studies show why some countries choose a lower tax rate while others choose a higher tax rate. We develop a politico-economic model with endogenous retirement decisions to answer this question.

A key feature of our analysis that is of importance to the result is the multiple, self-fulfilling expectations of agents. When young agents hold expectations of a higher social security benefit, it provides a disincentive to engage in educational investment, thereby resulting in a lower proportion of the skilled. This implies a majority of unskilled individuals choosing a higher social security benefit. This induces retirement and results in the lower labor force participation rate of the elderly. The opposite applies when young agents have expectations of lower social security in their old age. This finding of multiple political equilibria well fits the empirical evidence for the OECD.

We undertake numerical analysis to consider how wage inequality between the skilled and the unskilled affects the size of the pension as well as the expected utility of the young. The results show that the predictions of our model well fit observations of the pension–GDP ratio in the OECD. The results also show that a reduction in wage inequality has opposite effects on the expected utility of the young depending on the equilibria.

To obtain these results, we imposed a restriction on the set of tax rates available for the old. Because of this restriction, we were able to obtain an analytically tractable solution for the model. While we would expect that expanding the set of tax rates available would provide additional insights, we defer this to future work.

7 Appendix

7.1 Proof of Proposition 1

(i) Suppose that at time t, agents know that $\tau_t^y = \tau^{ys}$ and expect $\tau_{t+1}^o = 0$. Then:

$$u_{t+1} = 1 - (1-w)\left\{(1-\tau^{ys}) + \frac{\beta\mu(1+w)}{2}\right\} \leq \frac{1}{2}.$$

By (7), this implies that $\tau_{t+1}^o = 0$, fulfilling initial expectations. Therefore, there exists a skilled-majority equilibrium with $\tau_{t+1}^o = 0$ if the decisive voter finds it optimal to set $\tau_t^y = \tau^{ys}$.

To establish that setting $\tau_t^y = \tau^{ys}$ is optimal under the expectation of $\tau_{t+1}^o = 0$ for the decisive voter, we note the following properties of the function $Z : Z(\tau_t^y, 0)$ attain the top of the Laffer curve at $\tau_t^y = \tau^{ys}$; and $Z(\tau_t^y, 0) > Z(\tau_t^y, \tau^o) \forall \tau_t^y \in [0,1]$. These properties imply that setting $\tau_t^y = \tau^{ys}$ is optimal for any θ if $\tau^{ys} \leq \underline{\tau}^y (\leq \theta)$ (see Panel (a) in Figure 5). The inequality $\tau^{ys} \leq \underline{\tau}^y$ is rewritten as:

$$w(2-w) \leq \left\{(1-\tau^o)^2 - \frac{1}{2}\right\}(1+w)(1-w)^2\beta\mu,$$

where the left-hand side is positive; and the right-hand side is positive if and only if $\tau^o \in \left(0, (2-\sqrt{2})/2\right)$. Therefore, $\tau^{ys} \leq \underline{\tau}^y$ is rewritten as $\tau^o \in \left(0, (2-\sqrt{2})/2\right)$ and $\beta\mu \geq f^1(w) \equiv w(2-w)/[(1+w)(1-w)^2\left\{(1-\tau^o)^2 - 1/2\right\}]$.

(ii) Suppose that at time t, agents know that $\tau_t^y = \theta$ and expect $\tau_{t+1}^o = 0$. Then:

$$u_{t+1} = 1 - (1-w)\left\{(1-\theta) + \frac{\beta\mu(1+w)}{2}\right\} < \frac{1}{2}.$$

By (7), this implies that $\tau_{t+1}^o = 0$, fulfilling initial expectations. Therefore, there exists a skilled-majority equilibrium with $\tau_{t+1}^o = 0$ if the decisive voter finds it optimal to set $\tau_t^y = \theta$.

To establish that setting $\tau_t^y = \theta$ is optimal for the decisive voter, we first note that the first assumption in Statement (ii) lies opposite to the assumption in Statement (i); that is, $\tau^{ys} > \underline{\tau}^y$ holds under the first assumption. Under this condition, the decisive voter finds it optimal to choose $\tau_t^y = \max(\theta, \tau^{ys})$ from the viewpoint of revenue maximization when he/she expects $\tau_{t+1}^o = 0$.

Given the result above, either of the following options can oppose the choice of $(\tau_t^y, \tau_{t+1}^o) = (\theta, 0)$: setting $\tau_t^y = 1$ under the expectation of $\tau_{t+1}^o = \tau^o$ if $\tau^{yu} > 1$; or setting $\tau_t^y = \tau^{yu}$ under the expectation of $\tau_{t+1}^o = \tau^o$ if $\tau^{yu} \leq 1$. The concerned choice is sustained against the first option if $Z(\theta, 0) \geq Z(1, \tau^o)$; that is:

$$(\theta)^2 - 2\tau^{ys}\theta - (1 - 2\tau^{yu}) \leq 0 \Leftrightarrow \theta \geq \tilde{\theta},$$

where:

$$\tilde{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys}-1)^2 + 2(\tau^{ys} - \tau^{yu})}.$$

As θ is bounded above $\bar{\tau}^y$, $Z(\theta, 0) \geq Z(1, \tau^o)$ holds if and only if $\theta \in [\tilde{\theta}, \bar{\tau}^y]$. The set $[\tilde{\theta}, \bar{\tau}^y]$ is nonempty if and only if $\tilde{\theta} \leq \bar{\tau}^y$, which is rewritten as the second assumption in Statement (ii).

Setting $\tau_t^y = \theta$ under the expectation of $\tau_{t+1}^o = 0$ is sustained against the second option if $Z(\theta, 0) \geq Z(\tau^{yu}, \tau^o)$; that is:

$$(\theta)^2 - 2\tau^{ys}\theta + (\tau^{yu})^2 \leq 0 \Leftrightarrow \theta \geq \hat{\theta},$$

where:

$$\hat{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys})^2 - (\tau^{yu})^2}.$$

As θ is bounded above $\bar{\tau}^y$, $Z(\theta, 0) \geq Z(\tau^{yu}, \tau^o)$ holds if and only if $\theta \in [\hat{\theta}, \bar{\tau}^y]$. The set $[\hat{\theta}, \bar{\tau}^y]$ is nonempty if and only if $\hat{\theta} \leq \bar{\tau}^y$, which is rewritten as the third assumption in Statement (ii). ■

7.2 Proof of Proposition 2

(i) The second assumption $\beta\mu \leq f^4(w)$ is equivalent to $\tau^{yu} \leq 1$. Given the expectation of $\tau_{t+1}^o = \tau^o$, agents can choose τ_t^y to attain the top of the Laffer curve $Z(\tau_t^y, \tau^o)$.

Suppose that at time t, agents know that $\tau_t^y = \tau^{yu}$ and expect $\tau_{t+1}^o = \tau^o$. Then, $u_{t+1} = U(\tau^{yu}, \tau^o) > 1/2$. By (7), this implies that $\tau_{t+1}^o = \tau^o$, fulfilling initial expectations. Therefore, there exists an unskilled-majority equilibrium with $\tau_{t+1}^o = \tau^o$ if the decisive voter finds it optimal to set $\tau_t^y = \tau^{yu}$.

Given the properties of the payoff functions, we have the following two options: setting $\tau_t^y = \tau^{ys}$ under the expectation of $\tau_{t+1}^o = 0$; or setting $\tau_t^y = \theta$ under the expectation of $\tau_{t+1}^o = 0$. When the first option is available, setting $\tau_t^y = \tau^{yu}$ under the expectation of $\tau_{t+1}^o = \tau^o$ is not sustained against the first option because $Z(\tau^{ys}, 0) > Z(\tau^{yu}, \tau^o)$ holds. Therefore, we impose the condition that the first option is unfeasible, $\tau^{ys} \geq \underline{\tau}^y$, which is rewritten as:

$$1 - (1-w)^2 \geq \left\{(1-\tau^o)^2 - \frac{1}{2}\right\}(1+w)(1-w)^2\beta\mu,$$

where the left-hand side is always positive as $w \in [0, 1)$; and the right-hand side is positive if and only if $\tau^o \in \big(0, (2-\sqrt{2})/2\big)$. Therefore, the above inequality condition holds if the right-hand side is either nonpositive, i.e., $\tau^o \in \big[(2-\sqrt{2})/2, 1/2\big)$, or positive, i.e., $\tau^o \in \big(0, (2-\sqrt{2})/2\big)$ and $\beta\mu < f^1(w)$. These conditions are summarized as the first assumption in Statement (i).

Setting $\tau_t^y = \tau^{yu}$ under the expectation of $\tau_{t+1}^o = \tau^o$ is sustained against the second option if $Z(\tau^{yu}, \tau^o) > Z(\theta, 0)$; that is, if

$$\left[w + (1-w)^2\left\{(1-\tau^{yu}) + \frac{\beta\mu}{2}(1+w)(1-\tau^o)^2\right\}\right]\tau^{yu}$$
$$> \left[w + (1-w)^2\left\{(1-\theta) + \frac{\beta\mu}{2}(1+w)\right\}\right]\theta.$$

This is rewritten as:

$$(\theta)^2 - 2\tau^{ys}\theta + (\tau^{yu})^2 > 0 \Leftrightarrow \theta < \hat{\theta},$$

where:

$$\hat{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys})^2 - (\tau^{yu})^2}.$$

Therefore, $Z(\tau^{yu}, \tau^o) > Z(\theta, 0)$ holds if and only if $\theta < \hat{\theta}$ as illustrated by Panel (b) in Figure 7.

(ii) The second assumption $\beta\mu > f^4(w)$ is equivalent to $\tau^{yu} > 1$. Given the expectation of $\tau_{t+1}^o = \tau^o$, agents cannot choose $\tau_t^y = \tau^{yu}$ to attain the top of the Laffer curve. The tax revenue from the young is maximized at $\tau_t^y = 1$ as long as the expectation is $\tau_{t+1}^o = \tau^o$.

Suppose that at time t, agents know that $\tau_t^y = 1$ and expect that $\tau_{t+1}^o = \tau^o$. Then, $u_{t+1} = U(1, \tau^o) > 1/2$. By (7), this implies that $\tau_{t+1}^o = \tau^o$, fulfilling initial expectations. Therefore, there exists an unskilled-majority equilibrium with $\tau_{t+1}^o = \tau^o$ if the decisive voter finds it optimal to set $\tau_t^y = 1$.

Given the properties of the payoff functions, we have the following two options: setting $\tau_t^y = \tau^{ys}$ under the expectation of $\tau_{t+1}^o = 0$; or setting $\tau_t^y = \theta$ under the expectation of $\tau_{t+1}^o = 0$. When the first option is available, setting $\tau_t^y = 1$ under the expectation of $\tau_{t+1}^o = \tau^o$ is not sustained against the first option, because $Z(\tau^{ys}, 0) > Z(1, \tau^o)$ holds. Therefore, we impose the condition that the first option is unfeasible, $\tau^{ys} \geq \underline{\tau}^y$, which is rewritten as the first assumption in Statement (ii).

Setting $\tau_t^y = 1$ under the expectation of $\tau_{t+1}^o = \tau^o$ is sustained against the second option if $Z(1, \tau^o) > Z(\theta, 0)$; that is, if

$$\left[w + (1-w)^2 \frac{\beta\mu}{2}(1+w)(1-\tau^o)^2\right] \cdot 1$$
$$> \left[w + (1-w)^2 \left\{(1-\theta) + \frac{\beta\mu}{2}(1+w)\right\}\right] \cdot \theta.$$

This is rewritten as:

$$(\theta)^2 - 2\tau^{ys}\theta - (1 - 2\tau^{yu}) > 0 \Leftrightarrow \theta < \tilde{\theta},$$

where:

$$\tilde{\theta} \equiv \tau^{ys} - \sqrt{(\tau^{ys} - 1)^2 + 2(\tau^{ys} - \tau^{yu})}.$$

Therefore, $Z(1, \tau^o) > Z(\theta, 0)$ holds if and only if $\theta < \tilde{\theta}$, as illustrated by Panel (a) in Figure 7. ■

References

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

[2] Casamatta, G., Cremer, H., and Pestieau, P., 2005. Voting on pensions with endogenous retirement age. *International Tax and Public Finance* 12, 7–28.

[3] Casamatta, G., Cremer, H., and Pestieau, P., 2006. Is there a political support for the double burden on prolonged activity? *Economics of Governance* 7, 143–154.

[4] Conde-Ruiz, J.I., and Galasso, V., 2003. Early retirement. *Review of Economic Dynamics* 6, 12–36.

[5] Conde-Ruiz, J.I., and Galasso, V., 2004. The macroeconomics of early retirement. *Journal of Public Economics* 88, 1849–1869.

[6] Conde-Ruiz, J.I., Galasso, V., and Profeta, P., 2005. The evolution of early retirement. CEPR w.p. no. 4863.

[7] Crawford, V.P., and Lilien, D.M., 1981. Social security and the retirement decisions. *Quarterly Journal of Economics* 95, 505–529.

[8] Daval, R., 2003. Retirement behaviour in OECD countries: Impact of old-age pension schemes and other social transfer programmes. *OECD Economic Studies* 37, 7–50.

[9] Feldstein, M., and Liebman, J.B., 2002. Social security. In: Auerbach, A.J., and Feldstein, M. (Eds), Handbook of Public Economics Volume 4, Elsevier Science, Amsterdam.

[10] Fenge, R., and Pestieau, P., 2005. *Social Security and Early Retirement.* CESifo Book Series, MIT Press.

[11] Galasso, V., 2008. Postponing retirement: the political effect of aging. *Journal of Public Economics* 92, 2157–2169.

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

[13] Gruber, J., and Wise, D. (Eds.), 1999. Social Security and Retirement around the World. University of Chicago Press, Chicago.

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

[15] Hassler, J., Storesletten, K., and Zilibotti, F., 2007. Democratic public good provision. *Journal of Economic Theory* 133, 127–151.

[16] Lacomba, J.A., and Lagos, F.M., 2006. Population aging and legal retirement age. *Journal of Population Economics* 19, 507–519.

[17] Lacomba, J.A., and Lagos, F.M., 2007. Political election on legal retirement. *Social Choice and Welfare* 20, 1–17.

[18] OECD, 2009a. *Education at a Glance 2009*. OECD, Paris.

[19] OECD, 2009b. *Pensions at a Glance 2009: Retirement -Income Systems in OECD Countries*. OECD, Paris.

[20] Profeta, P., 2002. Retirement and social security in a probabilistic voting model. *International Tax and Public Finance* 9, 331–348.

[21] Sheshinski, E., 1978. A model of social security and retirement decisions. *Journal of Public Economics* 10, 337–360.

[22] Smith, J.P., 1999. Healthy bodies and thick wallets: The dual relation between health and economic status. *Journal of Economic Perspectives* 13, 145–166.

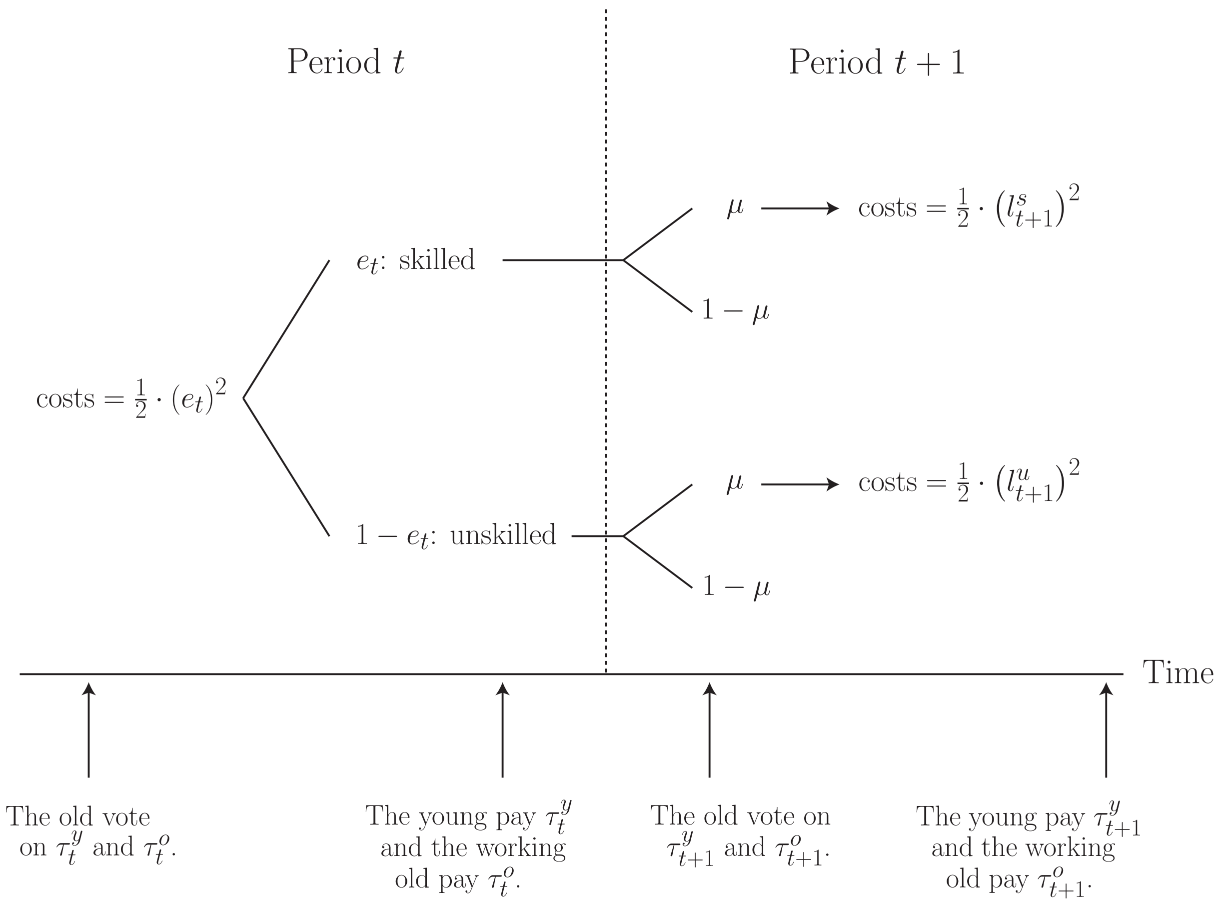

Figure 1: The timing of events and the distribution of the skilled and unskilled.

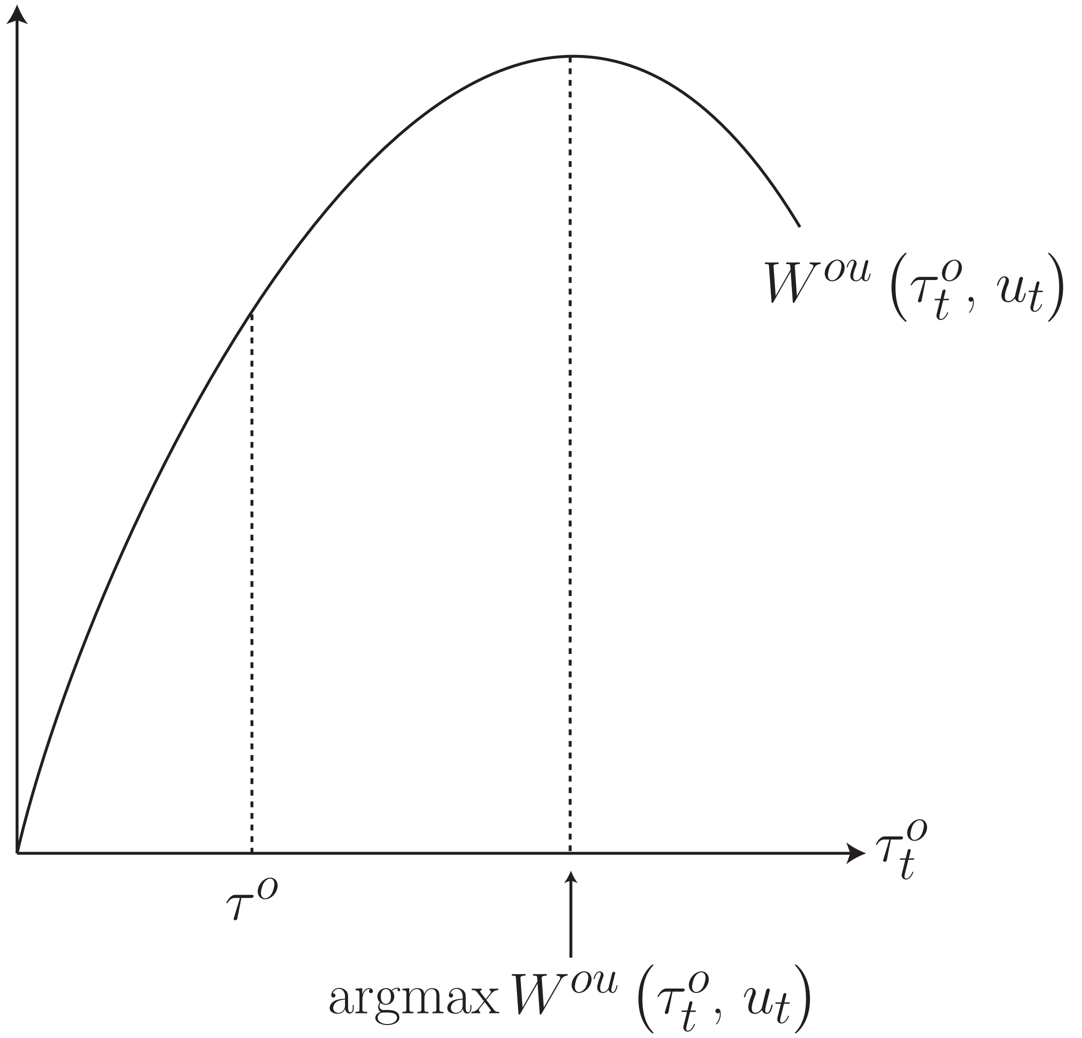

Figure 2: The range of politically available tax rates for the old under Assumption 1.

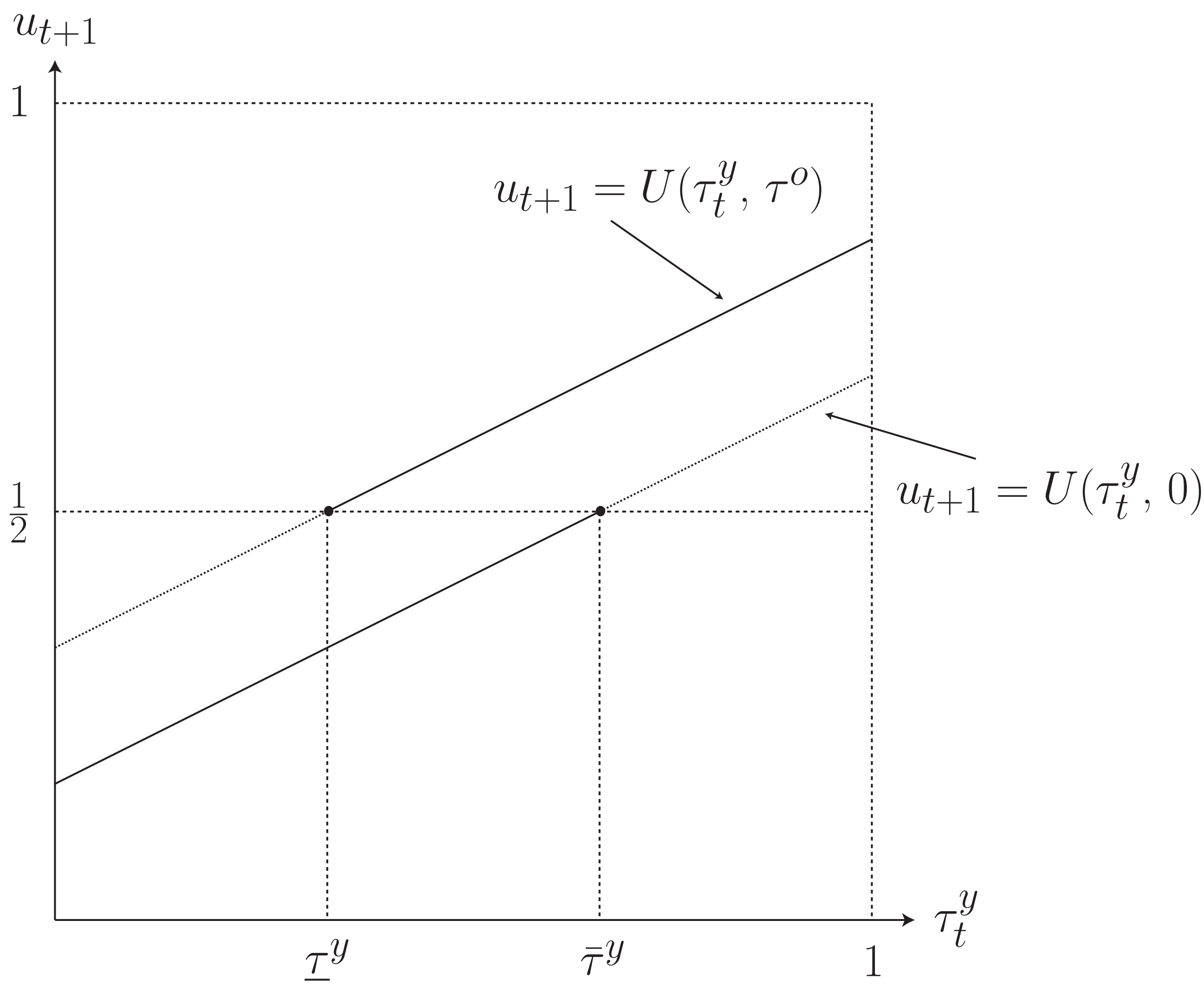

Figure 3: An example of solutions to the functional equation (9).

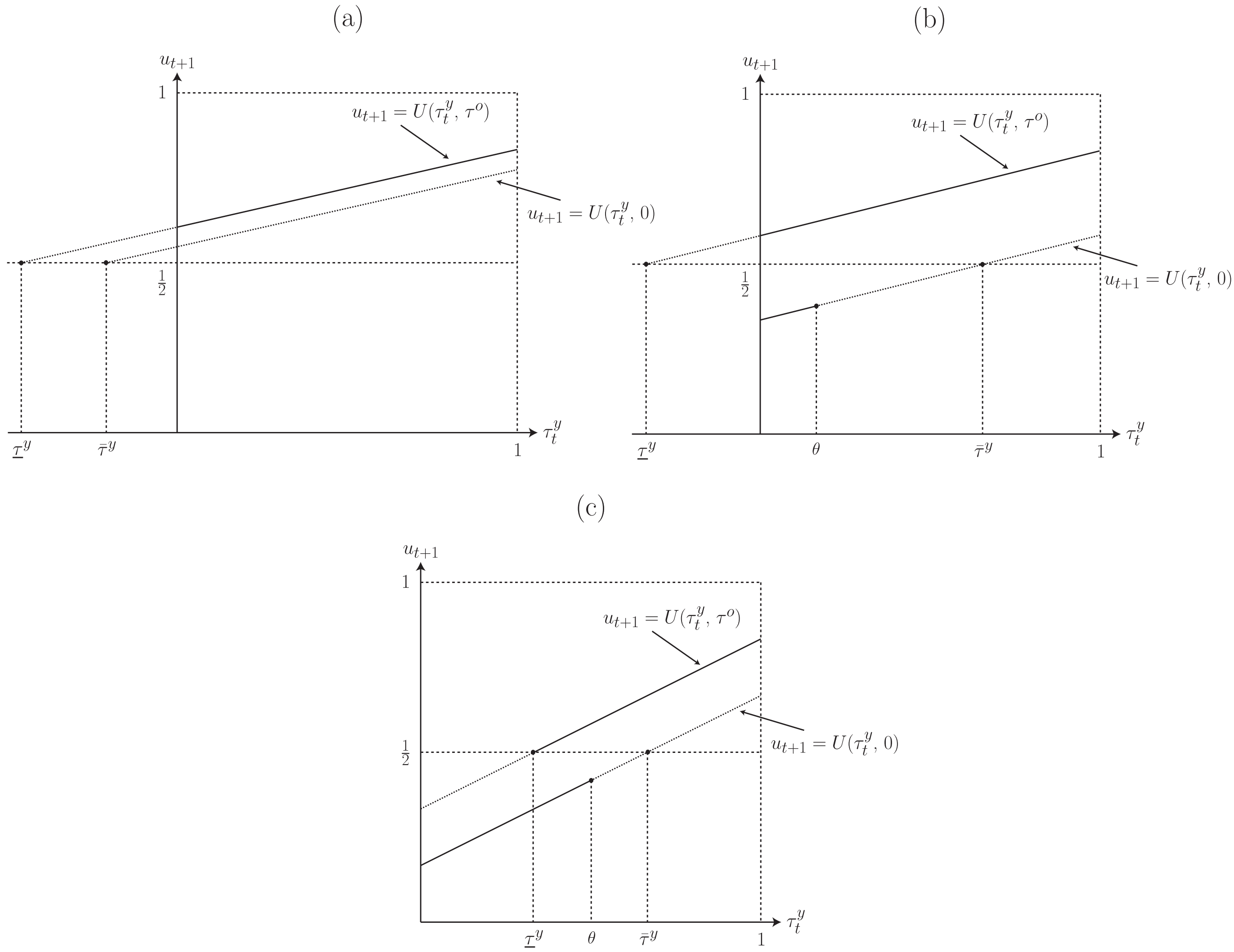

Figure 4: Three possible cases of the solution (10).

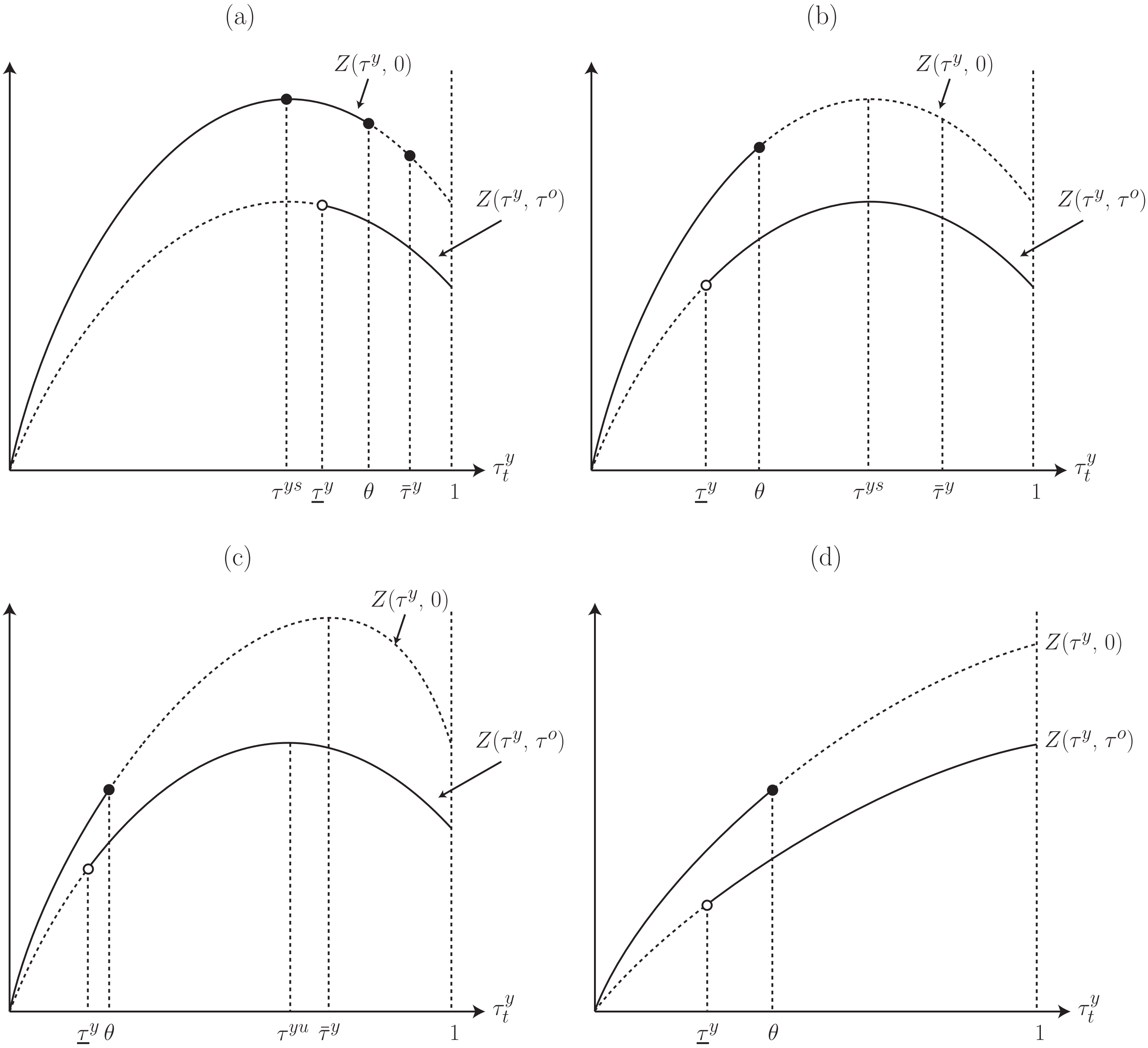

Figure 5: Revenue functions of the young. The skilled-majority equilibrium is illustrated in Panels (a) and (b). Panel (a) is the case of $\tau^{ys} \leq \underline{\tau}^y$; Panel (b) is the case of $\tau^{ys} > \underline{\tau}^y$. The unskilled-majority equilibrium is illustrated in Panels (c) and (d). Panel (c) is the case of $\tau^{yu} \leq 1$; Panel (d) is the case of $\tau^{yu} > 1$.

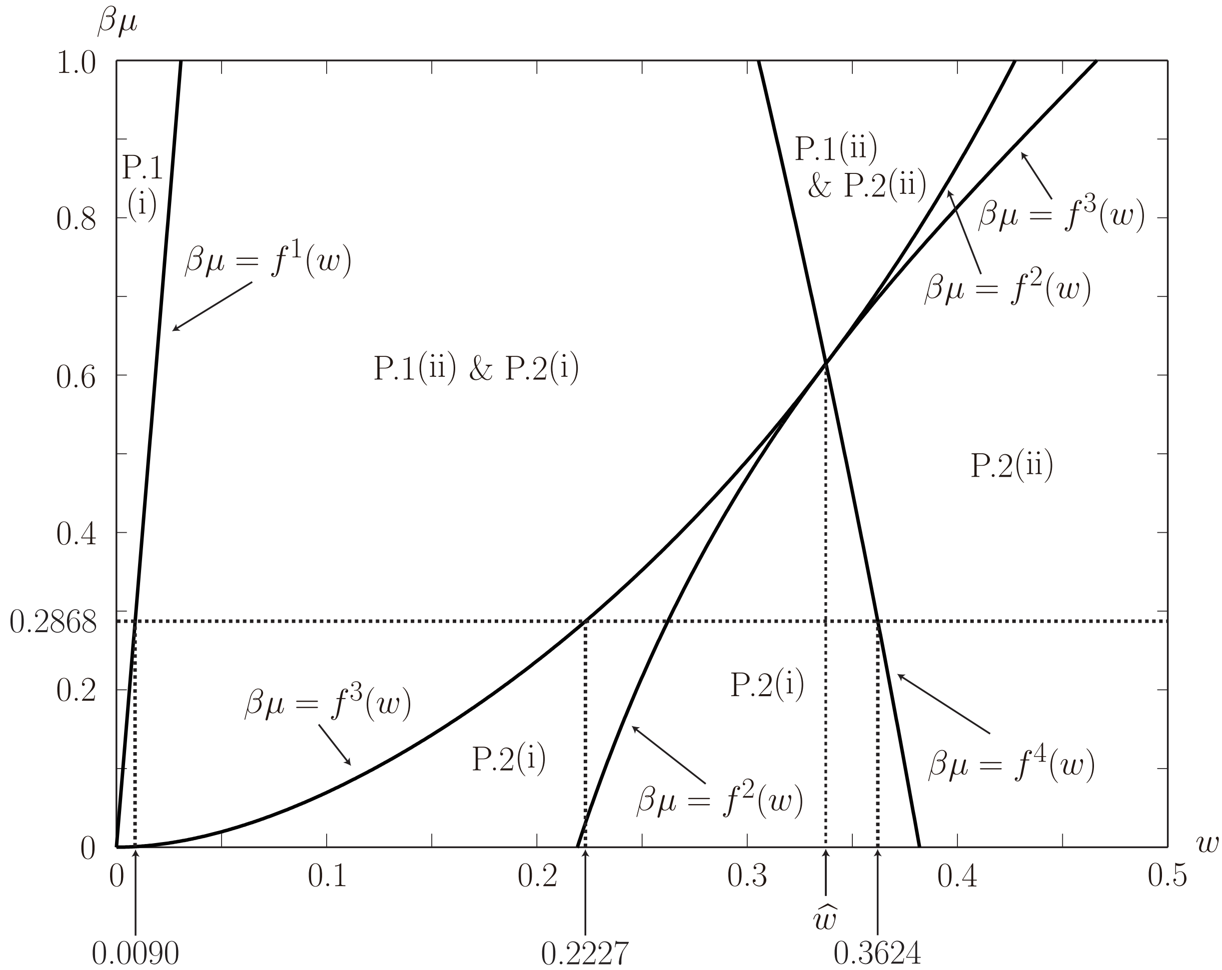

Figure 6: Illustration of function $f^i (i = 1, 2, 3, 4)$ under the assumption of $\tau^o = 0.25$.

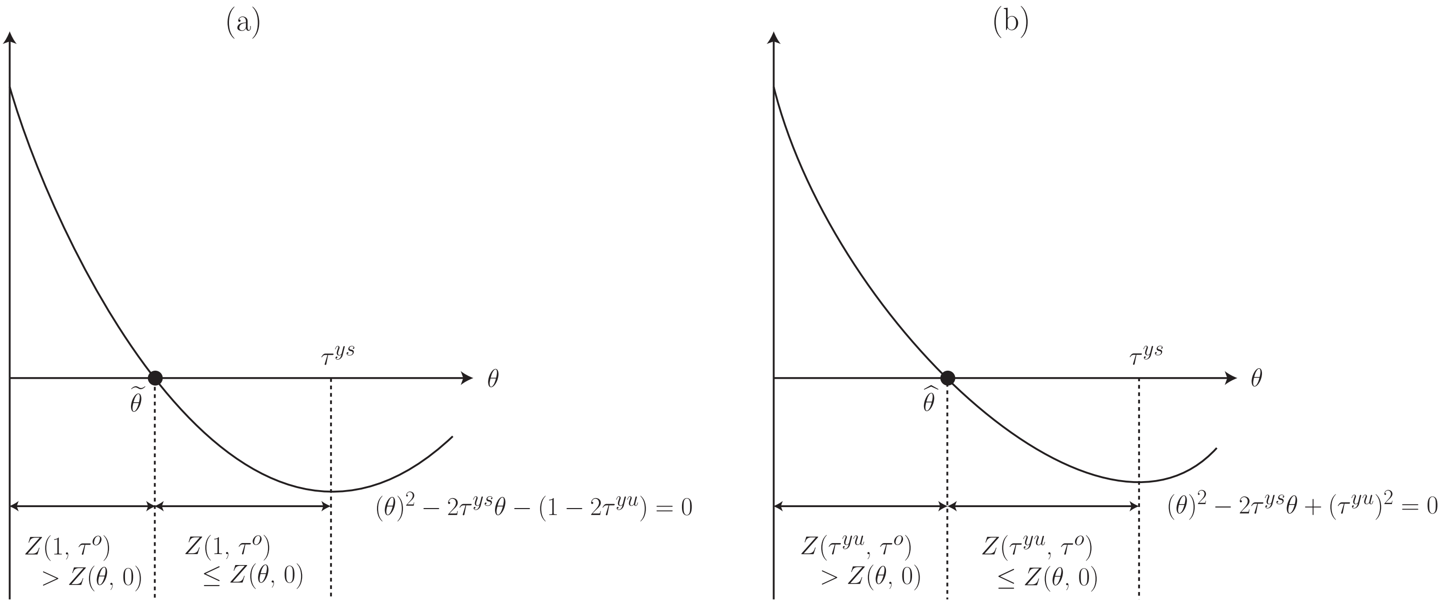

Figure 7: Determination of $\tilde{\theta}$ (Panel (a)) and $\hat{\theta}$ (Panel (b)).

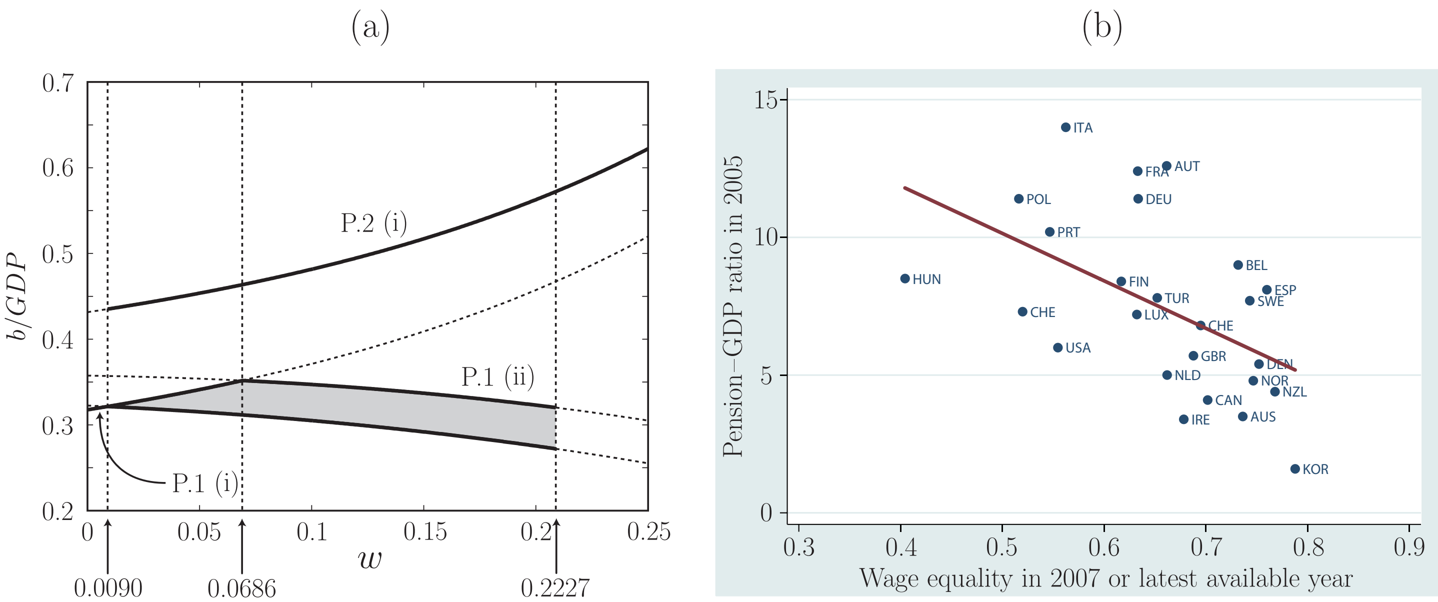

Figure 8: Panel (a) demonstrates the numerical result of the pension–GDP ratio in the current example; Panel (b) plots the cross-country data on wage equality and the pension–GDP ratio. **Source:** OECD (2009a) for wage equality; OECD (2009b) for the pension–GDP ratio.

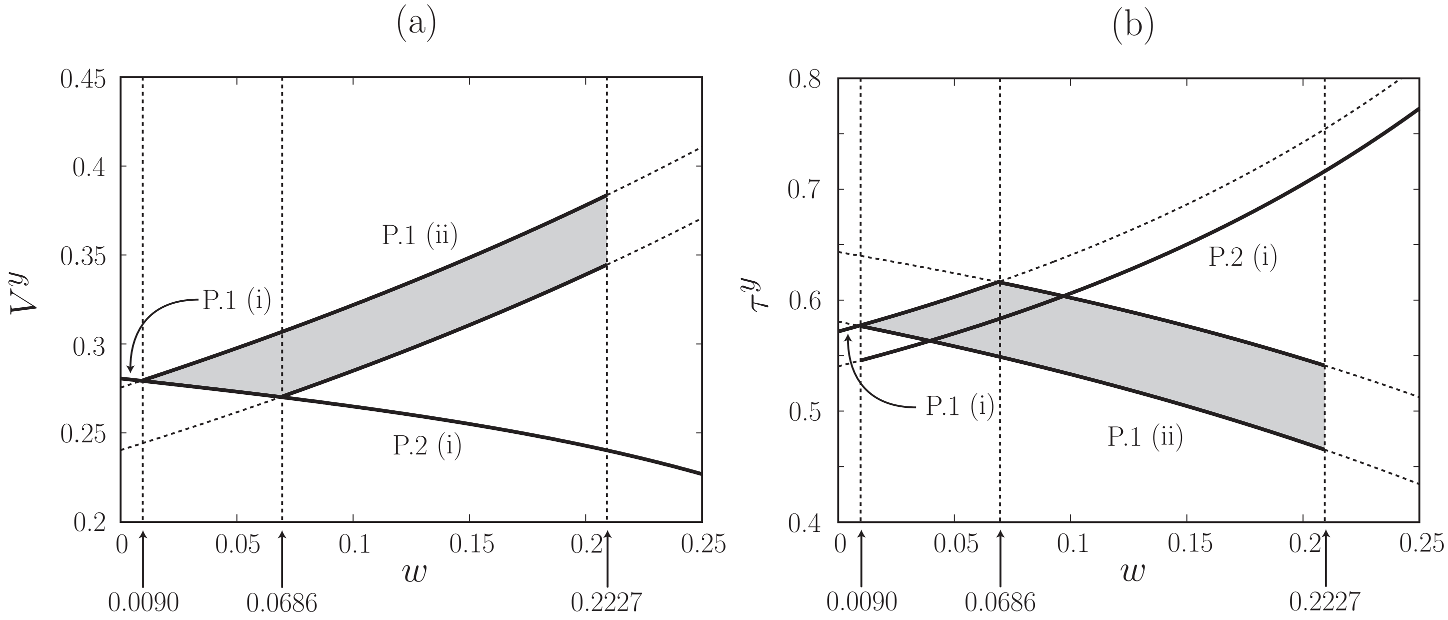

Figure 9: Panel (a) depicts the expected utility of the young; Panel (b) depicts the tax rate on the young.