GCOE Discussion Paper Series

Global COE Program
Human Behavior and Socioeconomic Dynamics

Discussion Paper No.211

A Political Economy Model of Earnings Mobility and Redistribution Policy

Ryo Arawatari and Tetsuo Ono

August 2011

GCOE Secretariat
Graduate School of Economics
OSAKA UNIVERSITY
1-7 Machikaneyama, Toyonaka, Osaka, 560-0043, Japan

A Political Economy Model of Earnings Mobility and Redistribution Policy

Ryo Arawatari* Tetsuo Ono†

Nagoya University Osaka University

August 23, 2011

Abstract

This paper presents a politico-economic model that includes a mutual link between earnings mobility and redistributive politics. The model demonstrates that an economy attains a unique, unskilled-majority equilibrium where unskilled, low-income agents support a low, rather than a high, redistribution when the economy is featured by a high opportunity of upward mobility and downward mobility risk. In contrast, the economy attains multiple equilibria when mobility opportunity and risk are low: one is an unskilled-majority equilibrium supporting high redistribution and the other is a skilled-majority equilibrium supporting low redistribution. Which equilibrium arises depends on the expectations of agents. The paper gives a comparison between the political equilibrium outcome and the social planner's allocation in terms of mobility and redistribution policy.

Key words: earnings mobility; political economy; redistribution;

JEL Classification: D30; D72; H20

*Graduate School of Economics, Nagoya University, Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan. E-mail: arawatari@soec.nagoya-u.ac.jp

†Graduate School of Economics, Osaka University, 1-7, Machikaneyama, Toyonaka, Osaka 560-0043, Japan. Email: tono@econ.osaka-u.ac.jp

1 Introduction

Expectations of redistribution affect individuals' decisions on educational investment. Their decisions determine the distribution of skilled and unskilled agents and thus inequality among agents, which in turn has an impact upon individuals' votes over redistribution. This feedback mechanism between individual decisions and redistributive politics could produce multiple equilibria (Glomm and Ravikumar, 1995; Saint-Paul and Verdier, 1997; Benabou, 2000). Based on the concept of a stationary Markov-perfect equilibrium, Hassler et al. (2003) and Hassler, Storesletten and Zilibotti (2007) capture the feedback mechanism and demonstrate multiple equilibria that explain the cross-country variations in welfare programs among democratic countries sharing similar economic backgrounds.

While the analysis by Hassler et al. (2003) and Hassler, Storesletten and Zilibotti (2007) provided a key insight into redistributive politics, it leaves the earning mobility issue untouched. In their framework, agents who succeed in education in youth retain their skills over the life cycle without any additional effort, and thus face no risk of downward mobility. In addition, agents who fail in education in youth have no second opportunity of becoming skilled in a later stage of their life and thus must accept their low-income status throughout their life. Thus, the following questions arise: how do upward mobility opportunity and downward mobility risk expected in the future affect current decisions on education and thus on voting over redistribution policy, and what is the consequence of mobility opportunity and risk for the efficiency of allocation?

Several studies have attempted to answer these questions. Early work by Piketty (1995) and Quadrini (1999) considered the effect of earnings mobility on agents' preferences for redistribution. However, a mutual link between mobility and redistributive politics is omitted from their analysis because of the assumption of exogenous mobility or idiosyncratic shocks to mobility. In an earlier study (Arawatari and Ono, 2009), we considered a mutual link between upward mobility and redistributive politics by introducing an upward mobility opportunity into the framework of Hassler, Storesletten and Zilibotti (2007). However, the downward mobility is omitted from the analysis, and the efficiency of mobility in the political equilibrium allocation in the presence of earnings mobility is left untouched.

While previous studies contribute to our knowledge and understanding of mobility and redistributive politics, the following issues still remain unresolved: (i) how do redistributive politics interact with mobility and distribution of income in the presence of downward mobility risk, and (ii) how does the political equilibrium outcome depart from the commitment solution (called the Ramsey allocation) with respect to earnings mobility and redistribution? Answers to these questions will provide more general insights into mobility and redistributive politics.

For the purpose of analysis, we adopt a framework based on that developed by Hassler, Storesletten and Zilibotti (2007) and that extended by Arawatari and Ono (2009). We further extend this framework by introducing downward mobility risk of agents. In particular, we consider agents living in two periods, youth and old age. In youth, agents undertake educational investments that determine whether their status in youth is skilled (i.e., rich) or unskilled (i.e., poor). At the beginning of old age, the unskilled agents have an opportunity of upward mobility with a probability γ, and can increase the probability of becoming skilled via reinvestment in education. By contrast, the skilled agents are at risk of downward mobility with a probability $\gamma \times \theta$, but they can reduce the probability of becoming poor by reinvestment in education. The expectations of redistribution affect agents' decisions on education, which in turn determines voting behavior over redistribution policy and thus mobility in the economy.

Focusing on the two key parameters, γ and θ, we first present the political equilibrium allocation via majority voting, and investigate qualitatively (in Section 3) and quantitatively (in Section 4) how the two parameters affect the political equilibrium outcome. When the upward mobility opportunity is high such that γ is above the threshold value, the economy attains a unique, unskilled-majority equilibrium with no taxation on the old. A high prospect of upward mobility in the future gives agents a disincentive to invest in education in youth, but they support no taxation on the old because of the prospect of upward mobility (POUM) in old age (Benabou and Ok, 2001; Alesina and La Ferrara, 2005). In contrast, when γ is below the threshold value, the economy attains multiple equilibria as demonstrated in Hassler, Storelsetten and Zilibotti (2007): an unskilled-majority equilibrium with taxation on the old and a skilled-majority equilibrium with no taxation on the old. Which outcome is realized in equilibrium depends on the expectations of agents.

The parameter θ, representing the risk of downward mobility, also affects agents' decisions on education. In particular, a higher θ gives agents a disincentive to invest in education in youth because the status in youth is less likely to persist into old age for the skilled agents. Given this feature, a natural prediction is that in an economy with a high prospect of downward mobility, the majority is the unskilled, and they support taxation on the skilled old. The former prediction is true, but the latter is not. A higher θ implies a smaller number of skilled old and thus less redistributive benefit from taxation on the skilled old. The redistributive benefit is outweighed by the expected tax burden of the unskilled who may become the skilled via reinvestment in education. Therefore, a higher downward mobility risk is more likely to realize the equilibrium supporting the POUM hypothesis. Upward mobility opportunity and downward mobility risk produce qualitatively similar properties with respect to political equilibrium outcomes.

In order to consider normative aspects of the political equilibrium, we characterize a Ramsey allocation defined as a feasible plan chosen by a benevolent planner who can commit to a policy sequence (in Section 5). The planner is assumed to choose an allocation to maximize the discounted sum of the utility functions of the successive generations. By comparing the political equilibrium with the Ramsey allocation, we find that the political economy featured by a unique, unskilled-majority equilibrium with no taxation on the old, which supports the POUM hypothesis, bears some resemblance to the Ramsey allocation when the planner attaches a low weight to the young. In contrast, the political economy featured by multiple equilibria may or may not attain an allocation similar to the Ramsey allocation depending on the expectations of agents as well as the planner's weight to the young.

Besides the literature mentioned above, the current paper is also related to the literature on the dynamic political economy of redistribution in overlapping-generations models with the concept of a stationary Markov-perfect equilibrium. The literature includes studies demonstrating a unique equilibrium pinned down by the initial expectation (Grossman and Helpman, 1998; Azariadis and Galasso, 2002) and multiple, self-fulfilling expectations of agents (Hassler et al., 2003). These studies are extended by introducing capital accumulation (Forni, 2005; Gonzalez-Eiras and Niepelt, 2008, 2011; Song, 2009a), retirement decisions of the elderly (Arawatari and Ono, 2011; Conde-Ruiz, Galasso and Profeta, 2011), ideology shifts (Song, 2009b), risk-averse agents (Hassler et al., 2005), wage inequality (Chen and Song, 2009), public goods (Hassler, Storesletten and Zilibotti, 2007), public debt accumulation (Song, Storesletten and Zilibotti, 2007) and intergenerational risk sharing (D'Amato and Galasso, 2010). These studies assumed that the economic status of each agent persists into the future, thereby removing the effects of earnings mobility over the life cycle. In contrast to these studies, the current paper includes earnings mobility over the life cycle, which plays a key role in redistributive politics and its efficiency.

The organization of this paper is as follows. Section 2 develops the model. Section 3 characterizes the political equilibria and investigates their qualitative properties. Section 4 undertakes a numerical analysis. Section 5 characterizes the Ramsey allocation and considers normative aspects of the political equilibrium. Section 6 provides concluding remarks.

2 The Model

The model is a two-period-lived overlapping-generations model based on that developed by Hassler, Storesletten and Zilibotti (2007) and extended by Arawatari and Ono (2009). Time is discrete and denoted by $t = 0, 1, 2, \cdots$. The economy consists of a continuum of agents living for two periods, youth and old age. Each generation has a unit mass.

Consider young agents born in period t. They are, at birth, identical. However, they can affect their prospects in life with educational investment. In particular, they become either skilled or unskilled, and by undertaking costly investment, can increase the probability e_t^y of becoming skilled in youth. Skilled agents earn a high wage, normalized to unity, whereas unskilled agents earn a low wage, normalized to zero. Because of this assumption regarding wages, the probability e_t^y is set within a range $[0, 1]$ without any additional assumptions. The cost of investment in youth is given by $(e_t^y)^2$. This cost is measured in terms of disutility; the financial constraint of investment is omitted from the analysis. Figure 1 illustrates the timing of events and the distribution of the skilled and the unskilled for generation t.

[Figure 1 about here.]

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) and Arawatari and Ono (2009) in that in old age, skilled agents may have a risk of downward mobility. In particular, the skilled can retain their status without any additional effort with probability $1-\theta\gamma \in [0, 1]$; however, with probability $\theta\gamma \in [0, 1]$, they need to reinvest in education to keep their status in old age. The cost for the skilled is given by $(e_{t+1}^{os})^2$, where e_{t+1}^{os} is the probability of being skilled and $1 - e_{t+1}^{os}$ is the probability of being unskilled.

As for the unskilled, they may have an opportunity for upward mobility as in Arawatari and Ono (2009). They are unskilled in old age with probability $1 - \gamma \in [0, 1]$; however, with probability $\gamma \in [0, 1]$, they have a second opportunity to reinvest in education. The cost for the unskilled is given by $(e_{t+1}^{ou})^2$, where e_{t+1}^{ou} is the probability of being skilled and $1 - e_{t+1}^{ou}$ is the probability of being unskilled.

The parameter $\theta \in [0, 1]$, solely capturing the effect of downward mobility risk, plays a key role in our analysis. We introduce this parameter to distinguish between the upward and downward mobility effects, and to focus on the role of downward mobility, which has not yet been analyzed in previous studies. We should note that if $\theta = 0$, the current model is similar to that of Arawatari and Ono (2009): the unskilled may have an opportunity for upward mobility, but the skilled are faced with no downward risk. We should also note that if $\gamma = 0$, the current model is similar to that of Hassler, Storesletten and Zilibotti (2007). Our model includes the cases of Hassler, Storesletten and Zilibotti (2007) and Arawatari and Ono (2009) as special ones.

There is no storage technology in this economy. Each agent uses his/her endowment within a period. The government provides lump-sum transfers, s, financed by taxes levied on the rich. The tax rates are age dependent: τ^o for the old and τ^y for the young. The tax rates are determined before the young agents decide on their investments.

The expected utility functions of agents alive at time t are given as follows:

$$
\begin{aligned}
V_t^{os} &= (1-\theta\gamma)(1-\tau_t^o) + \theta\gamma\{e_t^{os}(1-\tau_t^o) - (e_t^{os})^2\} + s_t, && (1)\\
V_t^{ou} &= \gamma\{e_t^{ou}\cdot(1-\tau_t^o) - (e_t^{ou})^2\} + s_t, && (2)\\
V_t^{y} &= e_t^y\cdot(1-\tau_t^y) - (e_t^y)^2 + s_t && (3)\\
&\quad + \beta\left[e_t^y\left\{(1-\theta\gamma)(1-\tau_{t+1}^o) + \theta\gamma\left(e_{t+1}^{os}(1-\tau_{t+1}^o) - (e_{t+1}^{os})^2\right)\right\}\right.\\
&\quad \left. + (1-e_t^y)\gamma\left\{e_{t+1}^{ou}\cdot(1-\tau_{t+1}^o) - (e_{t+1}^{ou})^2\right\} + s_{t+1}\right],
\end{aligned}
$$

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 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. The parameter $\beta \in (0,1)$ is a discount factor.

Given these preferences, a skilled old agent chooses e_t^{os} to maximize V_t^{os}; an unskilled old agent chooses e_t^{ou} to maximize V_t^{ou}; and a young agent in period t chooses e_t^y to maximize V_t^y by taking account of the optimal investments in his/her old age, e_{t+1}^{os} and e_{t+1}^{ou}. Therefore, optimal investments by the old and the young are given by, respectively:

$$
\begin{aligned}
e^{os*}(\tau_t^o) &= e^{ou*}(\tau_t^o) = \frac{(1-\tau_t^o)}{2}, && (4)\\
e^{y*}(\tau_t^y, \tau_{t+1}^o) &= \frac{1}{2}\cdot\left[(1-\tau_t^y) + \beta\left\{(1-\theta\gamma)(1-\tau_{t+1}^o) - \frac{\gamma(1-\theta)}{4}(1-\tau_{t+1}^o)^2\right\}\right], && (5)
\end{aligned}
$$

where $e^{oj*} \in [0, 1/2]$, $j = s, u$ and $e^{y*} \in [0,1)$ hold for given τ_t^o, τ_t^y and τ_{t+1}^o.

Because young agents are ex ante identical, agents of the same cohort choose the same investment. This implies that at the beginning of period $t+1$, the proportion of the unskilled old is given by:

$$
u_{t+1} \equiv 1 - e^{y*}(\tau_t^y, \tau_{t+1}^o) = 1 - \frac{1}{2}\cdot\left[(1-\tau_t^y) + \beta\left\{(1-\theta\gamma)(1-\tau_{t+1}^o) - \frac{\gamma(1-\theta)}{4}(1-\tau_{t+1}^o)^2\right\}\right].
$$

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

In this economy, there is earnings mobility over the life cycle. Let M_{t+1}^{up} denote the number of upwardly mobile agents from period t to period $t+1$, i.e., the number of agents who are unskilled in youth (in period t) but become skilled in old age (in period $t+1$). Let M_{t+1}^{down} denote the number of downwardly mobile agents from period t to period $t+1$, i.e., the number of agents who are skilled in youth (in period t) but become unskilled in old age (in period $t+1$). Then, M_{t+1}^{up} and M_{t+1}^{down} are calculated as, respectively:

$$M_{t+1}^{up} = (1-e_t^y)\gamma e_{t+1}^{ou},\ M_{t+1}^{down} = e_t^y \theta\gamma(1-e_{t+1}^{os}).$$

If $M_{t+1}^{up} > (<) M_{t+1}^{down}$, generation t experiences an increase (decrease) in the number of the skilled (unskilled) within a generation.

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

$$\begin{aligned} 2s_t &= [(1-u_t)\{(1-\theta\gamma)+\theta\gamma e^{os*}(\tau_t^o)\} + u_t\gamma e^{ou*}(\tau_t^o)]\cdot\tau_t^o + e^{y*}(\tau_t^y,\tau_{t+1}^o)\cdot\tau_t^y \\ &= W(\tau_t^o,u_t) + Z(\tau_t^y,\tau_{t+1}^o), \end{aligned}$$

where:

$$\begin{aligned} W(\tau_t^o,u_t) &\equiv [(1-\theta\gamma)(1-u_t)+(\gamma/2)(1-\tau_t^o)\{1-(1-\theta)(1-u_t)\}]\cdot\tau_t^o; \\ Z(\tau_t^y,\tau_{t+1}^o) &\equiv (1/2)\cdot\left[(1-\tau_t^y)+\beta\left\{(1-\theta\gamma)(1-\tau_{t+1}^o)-(\gamma/4)(1-\theta)(1-\tau_{t+1}^o)^2\right\}\right]\cdot\tau_t^y. \end{aligned}$$

$W(\tau_t^o,u_t)$ is the tax revenue financed by the skilled old agents, and $Z(\tau_t^y,\tau_{t+1}^o)$ is the tax revenue financed by the skilled young agents.

3 Political Equilibria

This section characterizes 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 beginning of each period and that only the old agents vote on current taxes. Our assumption about voting implies that we focus on intragenerational conflict over redistribution rather than intergenerational conflict.

With the optimal investments $e^{os*}(\tau_t^o), e^{ou*}(\tau_t^o)$ and $e^{y*}(\tau_t^y,\tau_{t+1}^o)$ and the government budget constraint, the indirect utility functions of the skilled and the unskilled old are given by, respectively:

$$V_t^{os} = (1-\theta\gamma)(1-\tau_t^o) + \frac{\theta\gamma}{4}(1-\tau_t^o)^2 + \frac{1}{2}\cdot\left\{W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o)\right\},$$
$$V_t^{ou} = \frac{\gamma}{4}(1-\tau_t^o)^2 + \frac{1}{2}\cdot\left\{W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o)\right\}.$$

The term in the first line, $(1-\theta\gamma)(1-\tau_t^o)$, is the expected after-tax income of the skilled old, the term $\theta\gamma(1-\tau_t^o)^2/4$ in the first line is the expected net benefit from reinvestment in education for the skilled old, the term $\gamma(1-\tau_t^o)^2/4$ in the second line is the expected net benefit from the second challenge for the unskilled old and the term $(W+Z)/2$ observed in both lines is the lump-sum transfer.

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:

$$W^{dec}(\tau_t^o, u_t) = \begin{cases} W^{os}(\tau_t^o, u_t) \equiv (1-\theta\gamma)(1-\tau_t^o) + \frac{\theta\gamma}{4}(1-\tau_t^o)^2 + \frac{1}{2}\cdot W(\tau_t^o, u_t) & \text{if } u_t \le 1/2, \\ W^{ou}(\tau_t^o, u_t) \equiv \frac{\gamma}{4}(1-\tau_t^o)^2 + \frac{1}{2}\cdot W(\tau_t^o, u_t) & \text{if } u_t > 1/2 \end{cases}$$

2. $U(\tau_t^y) = 1 - e^{y*}(\tau_t^y, \tau_{t+1}^o)$, 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 an equal number of skilled and unskilled 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 distributions of types. The third equilibrium condition requires that the decisive 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 $W^{os}(\tau_t^o, u_t) = (1-\theta\gamma)(1-\tau_t^o) + (\theta\gamma/4)\cdot(1-\tau_t^o)^2 + (1/2)\cdot W(\tau_t^o, u_t)$. This function has the following properties: $\partial W^{os}(\tau_t^o, u_t)/\partial \tau_t^o|_{\tau_t^o=0} \leq 0$ and $\partial^2 W^{os}(\tau_t^o, u_t)/\partial \tau_t^{o2} < 0$.[1] These properties imply that $W^o s(\tau_t^o, u_t)$ is maximized at $\tau_t^o = 0$: the skilled old pay more than they receive because the unskilled agents pay no tax but the revenue is distributed equally between the skilled and unskilled agents. Therefore, the skilled old prefer $\tau_t^o = 0$, implying that $T^o(u_t) = 0$ if the majority is the skilled:

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

Alternatively, suppose that the unskilled old are in the majority: $u_t > 1/2$. The objective function of the majority is given by $W^{ou}(\tau_t^o, u_t) = (\gamma/4)\cdot(1-\tau_t^o)^2 + (1/2)\cdot W(\tau_t^o, u_t)$. The first and second derivatives of $W^{ou}(\tau_t^o, u_t)$ with respect to τ_t^o are given by, respectively:

$$\begin{aligned}
\frac{\partial W^{ou}(\tau_t^o, u_t)}{\partial \tau_t^o} &= -\frac{\gamma}{2}(1-\tau_t^o) + \frac{1}{2}\cdot\left[(1-\theta\gamma)(1-u_t) + \frac{\gamma}{2}(1-2\tau_t^o)\{1-(1-\theta)(1-u_t)\}\right]; \\
\frac{\partial^2 W^{ou}(\tau_t^o, u_t)}{\partial \tau_t^{o2}} &= \frac{\gamma}{2}(1-\theta)(1-u_t) > 0,
\end{aligned}$$

implying that the unskilled old prefer $\tau_t^o = 0$ or 1 depending on the relative size of $W^{ou}(0, u_t)$ and $W^{ou}(1, u_t)$. Given that $W^{ou}(0, u_t) \gtreqless W^{ou}(1, u_t) \Leftrightarrow u_t \gtreqless 1-\gamma/\left\{2(1-\theta\gamma)\right\}$, the tax on the old when the majority is the unskilled is given by:

$$T^o(u_t)\begin{cases} = 0 & \text{if} \quad u_t > 1-\frac{\gamma}{2(1-\theta\gamma)} \\ \in \{0,1\} & \text{if} \quad u_t = 1-\frac{\gamma}{2(1-\theta\gamma)} \\ = 1 & \text{if} \quad u_t < 1-\frac{\gamma}{2(1-\theta\gamma)} \end{cases} \tag{7}$$

[1]The first and the second derivatives of $W^{os}(\tau_t^o, u_t)$ with respect to τ_t^o are given by, respectively:

$$\begin{aligned}
\partial W^{os}(\tau_t^o, u_t)/\partial \tau_t^o &= (-1)(1-\theta\gamma) - (\theta\gamma/2)\cdot(1-\tau_t^o) \\
&\quad + (1/2)\left[(1-\theta\gamma)(1-u_t) + (\gamma/2)\cdot(1-2\tau_t^o)\{1-(1-\theta)(1-u_t)\}\right], \\
\partial^2 W^{os}(\tau_t^o, u_t)/\partial \tau_t^{o2} &= -(\gamma/2)\cdot(1-\theta)u_t < 0.
\end{aligned}$$

We evaluate the first derivative at $\tau_t^o = 0$ and obtain:

$$\partial W^{os}(\tau_t^o, u_t)/\partial \tau_t^o|_{\tau_t^o=0} = -(1/2)\left[1-(\theta\gamma/2)\right] - (1/2)\left[1-(\gamma/2)(1+\theta)\right]u_t \leq 0,$$

where the inequality holds under the assumption of $\gamma \in [0,1]$ and $\theta \in [0,1]$.

The condition (7) means that if the number of unskilled agents is larger (smaller) than the threshold level, $1-\gamma/\left\{2(1-\theta\gamma)\right\}$, then the expected marginal benefit from taxation is smaller (larger) than the expected marginal cost of taxation. The unskilled agents know that the size of the tax base, $1-u_t$, is smaller (larger) such that they can get less (more) than they pay. Therefore, they prefer $\tau_t^o = 0 (= 1)$ if $u_t > (<) 1-\gamma/\left\{2(1-\theta\gamma)\right\}$. If $u_t = 1-\gamma/\left\{2(1-\theta\gamma)\right\}$, then the expected marginal benefit is equal to the expected marginal cost; they are indifferent as to whether there is 100% taxation or no taxation.

Based on the preferences over taxation of the old given by (6) and (7), the mapping that satisfies equilibrium condition 1 is summarized as follows.

(a) The case of $\gamma(1+\theta) \geq 1$ (i.e., $1/2 \geq 1-\gamma/\left\{2(1-\theta\gamma)\right\}$):

$$T^o(u_t) = 0 \ \forall \ u_t \in [0,1]. \tag{8}$$

(b) The case of $\gamma(1+\theta) < 1$ (i.e., $1/2 < 1-\gamma/\left\{2(1-\theta\gamma)\right\}$):

$$T^o(u_t) \begin{cases} = 0 & \text{if} \quad u_t \leq \frac{1}{2} \text{ or } 1-\frac{\gamma}{2(1-\theta\gamma)} < u_t \leq 1 \\ \in \{0,1\} & \text{if} \quad u_t = 1-\frac{\gamma}{2(1-\theta\gamma)} \\ = 1 & \text{if} \quad \frac{1}{2} < u_t < 1-\frac{\gamma}{2(1-\theta\gamma)} \end{cases} \tag{9}$$

Case (a) is trivial because there is no taxation on the old regardless of the status of a decisive voter. In what follows, we exclusively focus on case (b) by making the following assumption:

Assumption 1: $\gamma(1+\theta) < 1$.

[Figure 2 about here.]

Figure 2 illustrates (9), which satisfies equilibrium condition 1 under Assumption 1. Suppose that the decisive voter is a skilled old agent: $u_t \leq 1/2$. He/she sets $\tau_t^o = 0$ because his/her marginal cost of taxation is greater than his/her marginal benefit of taxation. Alternatively, suppose that the decisive voter is an unskilled old agent: $u_t > 1/2$. He/she sets either $\tau_t^o = 0$ or 1 depending on the magnitude of the correlation between the marginal benefit and cost of taxation. Under Assumption 1, the marginal benefit is smaller (larger) than the marginal cost for $u_t \in (1-\gamma/\left\{2(1-\theta\gamma)\right\}, 1]$ ($u_t \in (1/2, 1-\gamma/\left\{2(1-\theta\gamma)\right\})$).[2]

Next, we rewrite equilibrium condition 2 by substituting in the optimal investment

[2]If $u_t = 1-\gamma/2(1-\theta\gamma)$, the marginal benefit is equal to the marginal cost. The decisive voter is indifferent between $\tau_t^o = 0$ and 1.

$e^{y*}(\tau_t^y, \tau_{t+1}^o)$. This yields the following functional equation:

$$U(\tau_t^y) = 1 - \frac{1}{2} \cdot \left[(1-\tau_t^y) + \beta \left\{ (1-\theta\gamma)(1 - T^o(U(\tau_t^y))) - \frac{\gamma(1-\theta)}{4}(1 - T^o(U(\tau_t^y)))^2 \right\} \right], \tag{10}$$

where $T^o(\cdot)$ is given by (9) under Assumption 1. We derive the solution to the functional equation (10) by assuming rational expectations. Any solution to the functional equation (10) is given by:

$$U(\tau_t^y) = \begin{cases} 1 - \frac{1}{2} \cdot \left[(1-\tau_t^y) + \beta \left\{ (1-\theta\gamma) - \frac{\gamma(1-\theta)}{4} \right\} \right] & \text{if} \quad \tau_t^y \le \tau^{s0} \text{ or } \tau^{u0} \le \tau_t^y \\ 1 - \frac{1}{2} \cdot (1-\tau_t^y) & \text{if} \quad \tau_t^y \le \tau^{u1} \end{cases} \tag{11}$$

where:

$$\tau^{s0} \equiv \beta \left\{ (1-\theta\gamma) - \frac{\gamma(1-\theta)}{4} \right\},$$
$$\tau^{u0} \equiv \frac{1-\gamma(1+\theta)}{1-\theta\gamma} + \beta \left\{ (1-\theta\gamma) - \frac{\gamma(1-\theta)}{4} \right\},$$
$$\tau^{u1} \equiv \frac{1-\gamma(1+\theta)}{1-\theta\gamma}.$$

τ^{s0} denotes the upper bound of τ_t^y that realizes the skilled majority with the expectation of $\tau_{t+1}^o = 0$; τ^{u0} denotes the lower bound of τ_t^y that realizes the unskilled majority with the expectation of $\tau_{t+1}^o = 0$; and τ^{u1} denotes the upper bound of τ_t^y that realizes the unskilled majority with the expectation of $\tau_{t+1}^o = 1$.

Interpretation of (11) is as follows. Suppose that agents in period t expect $\tau_{t+1}^o = 0$. Under this expectation, young agents choose their investment as $e^{y*}(\tau_t^y, 0)$. By (9), this expectation is rational if $u_{t+1} = 1 - e^{y*}(\tau_t^y, 0) \le 1/2$ or $1 - \gamma/\{2(1-\theta\gamma)\} \le u_{t+1} = 1 - e^{y*}(\tau_t^y, 0) \le 1$, that is, if $\tau_t^y \le \tau^{s0}$ or $\tau^{u0} \le \tau_t^y$. Next, suppose that the young agents in period t expect $\tau_{t+1}^o = 1$. Under this expectation, young agents choose their investment as $e^{y*}(\tau_t^y, 1)$. By (9), their expectation is rational if $1/2 < u_{t+1} = 1 - e^{y*}(\tau_t^y, 1) \le 1 - \gamma/\{2(1-\theta\gamma)\}$, that is, if $\tau_t^y \le \tau^{u1}$. Figure 3 illustrates the solution (11).

[Figure 3 about here.]

As illustrated in Figure 3, there are multiple, self-fulfilling expectations of U for the set of $\tau_t^y \le \min\{\tau^{s0}, \tau^{u1}\}$. Which U arises in equilibrium depends on the expectations of agents. To illustrate U in equilibrium, we follow the method of Hassler, Storesletten and Zilibotti (2007) and introduce the critical rate of $\tau_t^y : \tau^e \le \min\{\tau^{s0}, \tau^{u1}\}$. The rate τ^e, which depends on the expectations of agents, is the highest tax rate that can yield an unskilled old majority. For $\tau_t^y > \tau^e$, the majority is the unskilled old. However, for $\tau_t^y \leq \tau^e$, the majority is either the skilled or the unskilled depending on agents' expectations. Panel (a) in Figure 3 illustrates the case where the tax rate τ^e is the highest rate that produces the skilled majority; panel (b) illustrates the case where τ^{s0} is the highest tax rate that produces the skilled majority.

Given the definition of τ^e, the function is now reduced as follows. The solution is given by:

$$U(\tau_t^y) = \begin{cases} \{U^0(\tau_t^y), U^1(\tau_t^y)\} & \text{if} \quad \tau_t^y \leq \tau^e \\ U^1(\tau_t^y) & \text{if} \quad \tau^e < \tau_t^y \leq \tau^{u1} \\ U^0(\tau_t^y) & \text{if} \quad \begin{array}{c} \tau^{u1} < \tau_t^y \leq \tau^{s0}, \\ \text{or } \tau^{u0} \leq \tau_t^y \leq 1 \end{array} \end{cases} \tag{12}$$

where $U^0(\tau_t^y)$ and $U^1(\tau_t^y)$ are defined by:

$$U^0(\tau_t^y) \equiv 1 - \frac{1}{2} \cdot \left[(1-\tau_t^y) + \beta \left\{ (1-\theta\gamma) - \frac{\gamma(1-\theta)}{4} \right\} \right];$$
$$U^1(\tau_t^y) \equiv 1 - \frac{1}{2}(1-\tau_t^y).$$

The superscripts "0" and "1" imply 0% and 100% taxation on the old, respectively.

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

Given the characterization of T^o and U satisfying equilibrium conditions 1 and 2, respectively, we now consider the tax rate on the young, τ_t^y, that satisfies equilibrium condition 3. Because there are two possible cases of majority, we introduce corresponding definitions of the political equilibria: an *unskilled-majority equilibrium* and a *skilled-majority equilibrium*.

The first equilibrium condition given by (9) implies that when the majority is the unskilled, there are two sorts of unskilled-majority equilibria: one is the equilibrium where agents expect no taxation on the old ($\tau_{t+1}^o = 0$) and choose τ_t^y to induce an unskilled majority at time $t+1$ ($u_{t+1} > 1/2$); the other is the equilibrium where agents expect taxation on the old ($\tau_{t+1}^o = 1$) and choose τ_t^y to induce an unskilled majority at time $t+1$ ($u_{t+1} > 1/2$). In contrast, when the majority is the skilled, there is a skilled-majority equilibrium where agents expect no taxation on the old ($\tau_{t+1}^o = 0$) and choose τ_t^y to induce a majority of the skilled at time $t+1$ ($u_{t+1} \leq 1/2$).

Before proceeding to the analysis, we note the following properties in order to find τ_t^y that satisfies equilibrium condition 3: (i) $Z(\tau_t^y, 0) > Z(\tau_t^y, 1)$ for any $\tau_t^y \in [0,1]$; and (ii) $Z(\tau_t^y, 0)$ and $Z(\tau_t^y, 1)$ attain the tops of the Laffer curves at $\tau_t^y = \tau_{\max}^{Z0}$ and $\tau_t^y = 1/2$, respectively, where:

$$\tau_{\max}^{Z0} \equiv \frac{1}{2} \cdot \left[1 + \beta \left\{(1-\theta\gamma) - \gamma(1-\theta)/4\right\}\right].$$

Given these properties with equilibrium conditions 1 and 2, revenue from the young is illustrated in Figures 4–6.

Figure 4 illustrates the case that produces the unskilled-majority equilibrium with no taxation on the old: that is, revenue from the young is maximized under the expectation of $\tau_{t+1}^{o} = 0$. In particular, panel (a) shows an interior solution where the revenue from the young is maximized by setting the tax rate, $\tau_{\max}^{Z0} = \arg\max Z(\tau_t^y, 0)$, that attains the top of the Laffer curve. Such a choice is impossible in the case illustrated in panel (b): the revenue from the young might be maximized at $\tau_t^y = \tau^{u0}$.

[Figure 4 about here.]

Figure 5 also illustrates the case that produces the unskilled-majority equilibrium. However, the current case differs from that illustrated in Figure 4 in that it is rational to expect taxation on the old. That is, the revenue from the young is represented by $Z(\tau^y, 1)$. Panel (a) illustrates an interior solution where the revenue from the young is maximized by setting the tax rate that attains the top of the Laffer curve $Z(\tau^y, 1)$: $\tau_t^y = \arg\max Z(\tau_t^y, 1) = 1/2$. Panel (b) illustrates the case where such a choice is impossible.

[Figure 5 about here.]

Finally, Figure 6 illustrates the case that produces the skilled-majority equilibrium. As shown in the previous section, the skilled agents might form the majority when the tax rate on the young is below the critical rate τ^e. Panel (a) illustrates the case where τ^e is the highest tax rate that realizes the skilled majority. The revenue from the young is maximized at $\tau = \tau^e$ when the majority is the skilled. In contrast, panel (b) illustrates the case where the maximizing tax rate is τ^{s0}, which is higher than τ^e.

[Figure 6 about here.]

Based on the above-mentioned argument, we will derive the condition for the existence of each type of equilibrium. For notational convenience, we define:

$$\phi \equiv \frac{1}{(1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)}.$$

3.3.1 Unskilled-majority Equilibrium with No Taxation on the Old

We first characterize an unskilled-majority equilibrium with no taxation on the old.

Proposition 1:

(i) *Suppose that the following condition holds:*

$$\beta \leq \phi \cdot \left[1 - 2 \cdot \frac{1-\gamma\cdot(1+\theta)}{1-\theta\gamma}\right].$$

There exists a set of unskilled-majority equilibria with no taxation on the old such that $\forall t$, T^o *is given by* (9), $U(\tau_t^y)$ *is given by* (12) *and* $T^y = \tau_{\max}^{Z0}$. *The equilibrium outcome is unique, such that* $\forall t$, $\tau_t^y = \tau_{\max}^{Z0}$, $\tau_t^o = 0$, $u_t = 1 - (1/4)\cdot [1 + \beta\{(1-\theta\gamma) - (\gamma/4)(1-\theta)\}]$ *and* $M^{up} > M^{down}$.

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

$$\text{(a) } \phi \cdot \left[1 - 2 \cdot \frac{1-\gamma\cdot(1+\theta)}{1-\theta\gamma}\right] < \beta;\ \text{(b) } \beta \leq \phi \cdot \frac{\gamma}{1-\theta\gamma};\ \textit{and}$$
$$\text{(c) } \phi \cdot \left[\frac{1-\theta\gamma}{4\gamma} - \frac{1-\gamma(1+\theta)}{1-\theta\gamma}\right] \leq \beta.$$

There exists a set of unskilled-majority equilibria with no taxation on the old such that $\forall t$, T^o *is given by* (9), $U(\tau_t^y)$ *is given by* (12) *and* $T^y = \tau^{u0}$. *The equilibrium outcome is unique, such that* $\forall t$, $\tau_t^y = \tau^{u0}$, $\tau_t^o = 0$, $u_t = 1 - \gamma/2(1-\theta\gamma)$ *and* $M^{up} > M^{down}$.

[Figures 7 and 8 about here.]

Areas P.1(i) and P.1(ii) in Figures 7 and 8 indicate the sets of parameters satisfying the equilibrium conditions in statements (i) and (ii) of Proposition 1, respectively. In particular, Figure 7 illustrates the set of parameters (β, γ) for a given θ; Figure 8 illustrates the set of parameters (β, θ) for a given γ.

The assumption in statement (i) ensures that the decisive voter can choose an interior solution $\tau_t^y = \tau_{\max}^{Z0}$. The young are taxed at the top of the Laffer curve, conditional on their expecting no taxation when old as illustrated in panel (a) in Figure 4.

The first assumption in statement (ii) is the exact opposite of the assumption in statement (i). This assumption implies that the decisive voter is unable to set an interior solution $\tau_t^y = \tau_{\max}^{Z0}$. The second assumption implies that it is feasible to set a corner solution given by $\tau_t^y = \tau^{u0}$. Therefore, the revenue from the young is maximized by setting $\tau_t^y = \tau^{u0}$ as long as the majority is the unskilled who expect no taxation on the old.

In this situation, there are, at most, two alternatives that may dominate the above-mentioned corner solution: voting that induces a majority of the unskilled by setting $\tau_t^y = \arg\max_{\tau_t^y \in [0,1]} Z(\tau_t^y, 1) = 1/2$; and voting that induces a skilled majority by setting $\tau_t^y = \tau^e (\leq \tau^{s0})$. The former alternative cannot dominate the corner solution if it provides less tax revenue from the young; that is, if $Z(1/2, 1) \leq Z(\tau^{u0}, 0)$. The third assumption in statement (ii) ensures that this condition holds. The latter alternative cannot dominate the corner solution because it holds that $Z(\tau^{u0}, 0) > Z(\tau^{s0}, 0) \geq Z(\tau^e, 0)$.

As illustrated in Figure 7, given the discount factor β, a higher γ is required for the existence of the unskilled-majority equilibrium with no taxation on the old. A high γ implies that status in youth is less likely to persist into old age. Because efforts of young individuals will be neutralized with a higher probability in the future, they have a disincentive to invest in education; the unskilled agents form a majority. However, the unskilled do not support a high tax rate on the old because of their hope for upward mobility through educational investment in the later stage of life. Hence, given more opportunity of upward mobility, there is a unique equilibrium that supports the POUM hypothesis.

The parameter θ, solely capturing the effect of downward mobility risk, also affects the realization of the unskilled-majority equilibrium with no taxation on the old. A higher θ implies that the status in youth is less likely to persist into old age for the skilled agents. This gives the young a disincentive to invest in education, thereby resulting in an unskilled-majority equilibrium.

As a decisive voter, the unskilled choose no taxation on the skilled old rather than taxation on them when θ is high. The reason for this choice is as follows. A higher θ implies a smaller number of skilled old and thus a smaller redistribution benefit from taxation on the skilled old. Given a high θ, the redistributive benefit is outweighed by the expected tax burden of the unskilled who become the skilled via reinvestment in education in old age. Therefore, the equilibrium supporting the POUM hypothesis is more likely to exist when θ is high.

Proposition 1 shows that the number of upwardly mobile agents is larger than the number of downwardly mobile agents. The mechanism behind this result is as follows. First, γu is the number of unskilled who have an opportunity of becoming skilled in old age. This is larger than the number of skilled who have a possibility of becoming unskilled in old age, $\theta\gamma(1-u)$, because $u > 1/2$ holds in an unskilled-majority equilibrium. Second, for the unskilled who have opportunities of upward mobility, the probability of becoming skilled is equal to the probability of becoming unskilled for the skilled who face the risk of downward mobility: $1 - e^{os} = e^{ou} = 1/2$. Therefore, the number of upwardly mobile agents is given by $u\gamma/2$, which is greater than the number of downwardly mobile agents given by $(1-u)\theta\gamma/2$.

3.3.2 Unskilled-majority Equilibrium with Taxation on the Old

The next proposition provides a characterization of an unskilled-majority equilibrium with taxation on the old.

Proposition 2:

Suppose that the following conditions hold:

$$\gamma \leq \frac{1}{2+\theta};\ \beta < \phi \cdot \frac{1-\gamma(1+\theta)}{1-\theta\gamma},$$

and:

$$\text{(a) } \beta > \phi \cdot \frac{\gamma}{1-\theta\gamma},\ \textit{or } \text{(b) } \beta \leq \phi \cdot \frac{\gamma}{1-\theta\gamma} \textit{ and } \beta < \phi \cdot \left[\frac{1-\theta\gamma}{4\gamma} - \frac{1-\gamma(1+\theta)}{1-\theta\gamma}\right].$$

There exists a set of unskilled-majority equilibria with taxation on the old such that $\forall t$, T^o *is given by* (9), $U(\tau_t^y)$ *is given by* (12) *and* $T^y = 1/2$. *The equilibrium outcome is unique, such that* $\forall t$, $\tau^y = 1/2, \tau^o = 1, u = 3/4$ *and* $M^{up} = 0 < M^{down}$.

Area P.2 in Figures 7 and 8 indicate the set of parameters satisfying the equilibrium conditions in Proposition 2. The first condition is derived from $1/2 \leq \tau^{u1}$, implying that given the expectation of $\tau_{t+1}^o = 1$, agents can choose $\tau_t^y = \arg\max Z(\tau_t^y, 1) = 1/2$ to attain the top of the Laffer curve $Z(\tau_t^y, 1)$. In addition, it is infeasible to set $\tau_t^y = \arg\max Z(\tau_t^y, 0)$, which attains the top of the Laffer curve $Z(\tau_t^y, 0)$.

In this situation, there are, at most, two alternatives that may dominate the current option: voting that induces a skilled majority by setting $\tau_t^y = \tau^e (\leq \tau^{s0})$ with the expectation of $\tau_{t+1}^o = 0$, and voting that induces an unskilled majority by setting $\tau_t^y = \tau^{u0}$ with the expectation of $\tau_{t+1}^o = 0$. The second condition, derived from $\tau^{u1} > \tau^{s0}$, implies that for the range of $\tau_t^y \in [0, \tau^{s0}]$, there are multiple, self-fulfilling expectations of agents. The current option $(\tau_t^y, \tau_{t+1}^o) = (1/2, 1)$ dominates the first alternative option $(\tau_t^y, \tau_{t+1}^o) = (\tau^e, 0)$ with $\tau^e \leq \tau^{s0}$ as long as the expectation τ^e is low. Whether τ^e is set to be low depends on the expectations of agents.

The first part of the third assumption, denoted by (a), implies that setting $\tau_t^y = \tau^{u0}$ with the expectation of $\tau_{t+1}^o = 0$ is unavailable for the decisive voter; the second part, denoted by (b), implies that such a setting is available but produces a lower revenue from the young: $Z(\tau^{u0}, 0) < Z(1/2, 1)$. Therefore, under the third assumption, the current option $(\tau_t^y, \tau_{t+1}^o) = (1/2, 1)$ dominates the second alternative option $(\tau_t^y, \tau_{t+1}^o) = (\tau^{u0}, 0)$.

Proposition 2 shows that in the unskilled-majority equilibrium with taxation on the old, the number of upwardly mobile agents is smaller than the number of downwardly mobile agents. This result is opposite to that in the unskilled-majority equilibrium with no taxation on the old as in Proposition 1. The opposite result comes from the fact that there are no upwardly mobile agents because of 100% taxation on the old. Whether the old are taxed or not critically affects the relative size of upwardly and downwardly mobile agents.

3.3.3 Skilled-majority Equilibrium with No Taxation on the Old

Having established the unskilled-majority equilibrium, we next provide the existence of a skilled-majority equilibrium.

Proposition 3:

(i) *Suppose that the following condition holds:*

$$\beta > \max\left\{\phi \cdot \frac{1-\gamma(1+\theta)}{1-\theta\gamma}, \phi \cdot \frac{\gamma}{1-\theta\gamma}\right\}.$$

There exists a set of skilled-majority equilibria with no taxation on the old such that $\forall t$, T^o *is given by (9),* $U(\tau_t^y)$ *is given by (12) and* $T^y = \tau^{s0}$. *The equilibrium outcome is unique, such that* $\forall t$, $\tau^y = \tau^{s0}, \tau^o = 0, u = 1/2$ *and* $M^{up} \geq M^{down}$. *The equality of* $M^{up} = M^{down}$ *holds if and only if* $\theta = 1$.

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

$$\beta > \phi \cdot \frac{\gamma}{1-\theta\gamma}, \beta \leq \phi \cdot \frac{1-\gamma(1+\theta)}{1-\theta\gamma} \text{ and } \beta \geq \phi \cdot \left(\frac{1}{4}\right).$$

There exists a set of skilled-majority equilibria with no taxation on the old such that $\forall t$, T^o *is given by (9),* $U(\tau_t^y)$ *is given by (12) and* $T^y = \tau^e (\leq \tau^{s0})$. *The equilibrium outcome is indeterminate, such that* $\forall t$, $\tau^y = \tau^e, \tau^o = 0, u = 1 - [(1-\tau^e) + \beta\{(1-\theta\gamma) - \gamma(1-\theta)/4\}]/2$ *and* $M^{up} \gtreqless M^{down}$ *if and only if* $\tau^e \gtreqless (1-\theta)/(1+\theta) + \tau^{s0}$.

The area marked P.3(i) in Figures 7 and 8 indicates the set of parameters satisfying the equilibrium conditions in statement (i) of Proposition 3. The first assumption in statement (i) is derived from the condition $\tau^{s0} > \tau^{u1}$. This condition implies that under the expectation of $\tau_{t+1}^o = 0$, the decisive voter can choose $\tau_t^y = \tau^{s0}$ irrespective of the expectation of τ^e. If the condition fails to hold, there are multiple expectations within the range of $\tau_t^y \in [0, \tau^{s0}]$; the decisive voter cannot choose $\tau_t^y = \tau^{s0}$ definitely under the expectation of $\tau_{t+1}^o = 0$. That is, the choice of τ^y depends on the expectations of agents.

The second assumption in statement (i) implies that it is infeasible for the decisive voter to set the two tax rates on the young, demonstrated in Proposition 1, under the expectation of $\tau^o_{t+1} = 0$. If possible, either of them dominates the current choice $\tau^y_t = \tau^{s0}$. Therefore, under the second assumption, the available choice for the decisive voter is limited to the range of $[0, \tau^{s0}]$ as long as he/she expects $\tau^o_{t+1} = 0$. Given that $Z(\tau^y_t, 0)$ is increasing in τ^y_t for that range, setting $\tau^y_t = \tau^{s0}$ is the revenue-maximizing behavior under the expectation of $\tau^o_{t+1} = 0$.

The area marked P.3(ii) in Figures 7 and 8 indicates the sets of parameters satisfying the equilibrium conditions in statement (ii) of Proposition 3. The first assumption in statement (ii) implies that it is impossible for the decisive voter to set the two tax rates on the young, demonstrated in Proposition 1, under the expectation of $\tau^o_{t+1} = 0$. If possible, either of them dominates the current choice $\tau^y_t = \tau^e (\leq \tau^{s0}\})$. Given that $Z(\tau^y_t, 0)$ is increasing in τ^y_t for $\tau^y_t \in [0, \tau^e)$, setting $\tau^y_t = \tau^e$ is the revenue-maximizing behavior under the expectation of $\tau^o_{t+1} = 0$ as long as $[0, \tau^e)$ is a feasible range of τ^y_t.

Given the above argument, the remaining alternative options for the decisive voter are to set $\tau^y_t = \tau^{s0}$ irrespective of expectations on τ^y_t under the expectation of $\tau^o_{t+1} = 0$, or to set $\tau^y_t = \arg\max Z(\tau^y_t, 1) = 1/2$ under the expectation of $\tau^o_{t+1} = 1$. The second assumption in statement (ii) implies that the option of $(\tau^y_t, \tau^o_{t+1}) = (\tau^{s0}, 0)$ is unavailable for the decisive voter; otherwise, this option dominates the current choice: $Z\left(\tau^{s0}, 0\right) \geq Z(\tau^e, 0)$.

The third assumption in statement (ii) ensures that there exists a $\tau^y_t = \tau^e$ that sustains the choice of $(\tau^y_t, \tau^o_{t+1}) = (\tau^e, 0)$ against the choice of $(\tau^y_t, \tau^o_{t+1}) = (1/2, 1)$ as an equilibrium: $Z(\tau^e, 0) \geq Z(1/2, 1)$. This inequality condition is rewritten as $\tau^e \geq \tilde{\tau}$ where:

$$\tilde{\tau} \equiv \frac{1 + \tau^{s0} - \sqrt{[1 + \tau^{s0}]^2 - 1}}{2}.$$

As τ^e is bounded above τ^{s0}, τ^e must be set within the range $[\tilde{\tau}, \tau^{s0}]$ in order that $Z(\tau^e, 0) \geq Z(1/2, 1)$ holds. The third assumption ensures that the set $[\tilde{\tau}, \tau^{s0}]$ is nonempty.

In the skilled-majority equilibrium, there is no taxation on the old. This implies that the probability of becoming skilled for the unskilled who have opportunity for upward mobility is equal to the probability of becoming unskilled for the skilled who face the risk of downward mobility: $e^{ou} = 1 - e^{os} = 1/2$. Thus, the relative number of upwardly and downwardly mobile agents depends on the number of unskilled with a mobility opportunity, $\gamma(1 - e^y)$, and the number of skilled with mobility risk, $\theta\gamma e^y$.

When the economy attains the equilibrium in statement (i) of Proposition 3, the tax rate on the young is $\tau^y = \tau^{s0}$, which results in $e^y = 1/2$. Thus, the number of unskilled young, $\gamma(1 - e^y)$, is larger than or equal to the number of skilled young, $\theta\gamma e^y$; an equality holds if and only if $\theta = 1$. When the economy attains the equilibrium in statement (ii) of Proposition 3, the tax rate on the young is $\tau^y = \tau^e$, which depends on the expectations of agents. A higher τ^e leads to a lower probability of becoming skilled in youth. Thus,

$\gamma(1 - e^y) > \theta\gamma e^y$ (that is, $M^{up} > M^{down}$) is more likely to hold if τ^e is set to be higher by agents.

3.4 Effects of Mobility Opportunity and Risk on the Characterization of Political Equilibria

In what follows, we provide an interpretation of the results established in Propositions 1–3 by focusing on γ and θ, respectively. First, let us consider the effect of the parameter γ by utilizing Figure 7 where θ is fixed at $\theta = 0.8$. Given β, there is a threshold value of β such that the equilibrium is characterized by the unskilled majority who support no taxation on the old as demonstrated in Proposition 1. When γ is above the threshold value, agents have a high probability of second opportunities for upward mobility. This implies that agents have excellent prospects of upward mobility in old age; this prospect gives agents a disincentive to invest in education in youth. Therefore, a high γ leads to a majority of unskilled agents who prefer no taxation on the old. The POUM hypothesis, supported by the US data (Benabou and Ok, 2001; Alesina and La Ferrara, 2005) holds when γ is above the threshold value.

When γ is below the threshold value, there is no equilibrium that supports the POUM hypothesis. This is because given few opportunities for second opportunities, the status in youth is highly persistent in old age. The unskilled and skilled agents are expected to remain unskilled and skilled in old age with a high probability; therefore, the unskilled agents prefer taxation on the skilled old whereas the skilled prefer no taxation on the skilled old. The majority becomes the unskilled when agents attach a low value to old-age utility, whereas it becomes the skilled when they attach a high value to old-age utility. The equilibrium outcome depends on the expectations of agents.

Next, consider the effect of θ by utilizing Figure 8 where γ is fixed at 0.4. By fixing the value of γ, we can focus exclusively on the effect of downward mobility risk, and we are insulated from the prospect of upward mobility captured by the parameter γ. As demonstrated in Figure 8, given β, the economy is more likely to be in the unskilled-majority equilibrium with no taxation on the old when the risk of downward mobility is higher. Therefore, a higher probability of downward mobility also produces the equilibrium that supports the POUM hypothesis. The mechanism behind this result is given in the paragraph following Proposition 1. Although the parameters γ and θ have different implications for mobility, they lead to similar results with regard to the emergence of the equilibrium supporting the POUM hypothesis.

4 Numerical Analysis

So far, we have characterized the political equilibrium and qualitatively assessed the impacts of mobility opportunity (represented by γ) as well as downward mobility risk (represented by θ) on the determination of tax rates and earnings mobility. To facilitate an understanding of the political equilibria, this section investigates numerically how opportunity and risk affect the expected utility of the young via tax rates (τ^y and τ^o) and the probability of success in youth (e^y).

For the purpose of analysis, we assume a generation to be 20 years in length. The first and the second periods correspond to, for example, ages 25–44 and 45–64 years, respectively.[3] Our selection of β is 0.98. Because the agents under the current assumption plan over generations that span 20 years, we discount the future by $(0.98)^{20} \approx 0.6676$.

4.1 Effects of Upward Mobility Opportunity

The solid curves and shaded area in Figure 9 illustrate how the parameter γ, representing the mobility opportunity in the economy, affects the tax rate on the young (panel (a)), the probability of becoming skilled in youth (panel (b)) and the expected utility of the young (panel (c)).

[Figure 9 about here.]

In the current environment where θ is fixed at 0.8, there is a threshold value of γ, denoted by $\hat{\gamma} = 0.3433$, that divides the range of γ into two intervals. One interval, which is less than the threshold value, is called the "low-mobility economy" the other interval, which is more than the threshold value, is called the "high-mobility economy." The low-mobility economy is featured by multiple equilibria: there is the skilled-majority equilibrium as in Proposition 3(ii) and the unskilled-majority equilibrium as in Proposition 2. The better-opportunity economy is featured by a unique, unskilled-majority equilibrium supporting the POUM hypothesis as in Proposition 1. In particular, the equilibrium has a corner solution as in Proposition 1(ii) for a smaller subinterval of γ whereas it has an interior solution as in Proposition 1(i) for a larger subinterval of γ.

4.1.1 Tax Rate on the Young

Panel (a) in Figure 9 depicts how the tax rate on the young is affected by the parameter γ. Let us first consider the low-mobility economy featured by multiple equilibria. The

[3] The length of this period is shorter than the usual assumption, for example, 30 years. We adopt the shorter period because the current model assumes that agents work in both periods of life. Taking a longer period does not qualitatively affect the result shown below.

tax rate on the young in the skilled-majority equilibrium depends on the expectations of agents. The young in the skilled-majority equilibrium may or may not bear a heavier tax burden than in the unskilled-majority equilibrium depending on the expectations of agents.

At the threshold value of γ, an increase in γ leads to a jump in the tax rate. This sharp increase comes from a change in the choice of the old who want to maximize the tax revenue from the young. When γ is below the threshold value, young individuals have a strong incentive to invest in education. The old utilize this incentive by imposing a lower tax rate on the young and thus expanding the tax base (i.e., the number of skilled young). In contrast, when γ is above the threshold value, young individuals have a weak incentive for educational investment even if their tax burden is small. In this case, the old impose a higher tax rate on the young to increase tax revenue from each skilled young individual.

Next, consider the high-mobility economy featured by a unique unskilled-majority equilibrium supporting the POUM hypothesis. For this range of γ, an increase in γ leads to a decrease in the tax rate on the young. In particular, a sharp decrease is observed at first; then a gradual decrease is observed. The difference between the two cases comes from the property of the equilibrium tax rates. The tax rate on the young is featured by a corner solution in the former case, whereas it is featured by an interior solution (that is, the top of the Laffer curve) in the latter case. This difference in solution properties results in a quantitative change in the value of γ that divides the range supporting the POUM hypothesis into two subintervals.

4.1.2 Probability of Becoming Skilled in Youth

Given the determination of the tax rate on the young described above, we now consider how the probability of becoming skilled in youth, e^y, is affected by the parameter γ. The probability is given by $e^{y*} = \left[(1-\tau^y) + \beta\left\{(1-\theta\gamma)(1-\tau^o) - \gamma(1-\theta)(1-\tau^o)^2/4\right\}\right]/2$. This probability indicates that the parameter γ has two effects on e^y : economic and political effects. The economic effect implies that an increase in γ leads to less persistence of agents' status in youth. This gives young individuals a disincentive to invest in education, thereby resulting in a lower probability of becoming skilled in youth. On the other hand, the political effect implies that an increase in γ affects the determination of the tax rates τ^y and τ^o, which in turn affects the decision on educational investment.

Panel (b) of Figure 9 illustrates how the probability of becoming skilled in youth, e^y, is affected by the parameter γ. For the low-mobility economy featured by multiple equilibria, the tax burden over the life cycle is heavier in the unskilled-majority equilibrium than in the skilled-majority equilibrium because the tax rate on the old is 100% in the former equilibrium whereas it is zero in the latter equilibrium. Because of this tax-burden effect, the probability of becoming successful in education in the unskilled-majority equilibrium is lower than that in the skilled-majority equilibrium.

Next, consider the high-mobility economy featured by a unique unskilled-majority equilibrium supporting the POUM hypothesis. For the range of γ that realizes a corner solution, the political effect is positive and large because an increase in γ leads to a sharp decrease in the tax rate on the young. The positive political effect dominates the negative economic effect, which results in an increase in e^y. However, for the range of γ that realizes an interior solution, the political effect is positive but small because an increase in γ leads to an imperceptible decrease in the tax rate on the young. The negative economic effect dominates the positive political effect, which results in a decrease in e^y. Although patterns of taxation are qualitatively similar between the two cases of solutions, the quantitative effects of the mobility opportunity on the probability of becoming skilled in youth differ between them.

4.1.3 Expected Utility of the Young

Panel (c) of Figure 9 illustrates how the parameter γ affects the expected utility of the young. First, consider the low-mobility economy featured by multiple equilibria. The young obtain a higher level of expected utility in the skilled-majority equilibrium than in the unskilled-majority equilibrium because the skilled-majority equilibrium requires a lower tax burden for the old and thus realizes a higher probability of becoming skilled in youth.

Next, consider the high-mobility economy featured by a unique unskilled-majority equilibrium supporting the POUM hypothesis. In this economy, the effect on the utility parallels the effect on the probability of becoming skilled in youth. In particular, for the range of γ that attains a corner solution, an increase in γ improves the expected utility of the young through an increase in the probability of becoming skilled in youth. However, for the range of γ that attains an interior solution, an increase in γ decreases the expected utility of the young through a decrease in the probability of becoming skilled.

In the unskilled-majority equilibrium, a greater opportunity of mobility improves the expected utility of the young only if the economy attains the corner-solution equilibrium supporting the POUM hypothesis. This implication heavily depends on the endogenous determination of the tax rates via voting. When the tax rates are exogenously given, the expected utility is definitely decreased by increased opportunity of mobility because the political effect is the only source of an increase in the probability of becoming skilled in youth and an improvement in the expected utility of the young. Our analysis and results suggest that redistributive politics could play an important role when we evaluate the effect of mobility opportunity from the viewpoint of utility.

4.2 Effects of Downward Mobility Risk

So far, we have conducted a numerical analysis by taking θ, representing the downward mobility risk for the skilled individuals, as given. In order to demonstrate the effect of downward mobility risk clearly, we set γ equal to 0.4, and quantitatively assess the impact of a change in θ. The solid curves and shaded area in Figure 10 illustrate how the parameter θ affects the tax rate on the young (panel (a)), the probability of becoming skilled in youth (panel (b)) and the expected utility of the young (panel (c)).

[Figure 10 about here.]

As in the analysis of γ, there is a threshold value of θ, given by $\hat{\theta} = 0.5015$, that divides the range of θ into two intervals. One interval, which is less than 0.5015, is called the "low-risk economy"; the other interval, which is more than 0.5015, is called the "high-risk economy". The low-risk economy is featured by multiple equilibria: these are the skilled-majority equilibrium as in Proposition 3(ii) and the unskilled-majority equilibrium as in Proposition 2. The high-risk economy is featured by a unique unskilled-majority equilibrium with a corner solution, supporting the POUM hypothesis as in Proposition 1(ii).

Figure 10 demonstrates that the effects of θ on the tax rate on the young, the probability of becoming skilled in youth and the expected utility of the young are qualitatively similar to the effects of γ. For example, in the low-risk economy, the expected utility is higher in the skilled-majority equilibrium than in the unskilled-majority equilibrium; in the high-risk economy featuring a corner solution supporting the POUM hypothesis, a higher risk improves the expected utility of the young. In other words, mobility opportunity represented by γ and downward mobility risk represented by θ produce qualitatively similar economic and political consequences.

5 Ramsey Allocation

In this section, we characterize a *Ramsey allocation* as a feasible plan chosen by a benevolent social planner who can commit to a policy sequence at time zero (Subsection 5.1). The Ramsey allocation derived here will be compared with the political equilibria in order to consider the normative aspect of the politics (Subsection 5.2).

5.1 Characterization of the Ramsey Allocation

The Ramsey allocation solves the following problem:

$$\max \beta \left\{(1-u_0)V^{os}(s_0,\tau_0^o) + u_0 V^{ou}(s_0,\tau_0^o)\right\} + \sum_{t=0}^{\infty} \lambda^{t+1} V_t^y(e_t, s_t, s_{t+1}, \tau_t^y, \tau_{t+1}^o),$$

where $\lambda \in (0,1)$ is a discount factor and $u_0 \in [0,1]$ is the initial distribution. The term $(1-u_0)V^{os}$ is the utility of the initial skilled old V^{os} multiplied by their proportion $1-u_0$, the term $u_0 V^{ou}$ is the utility of the initial unskilled old V^{ou} multiplied by their proportion u_0 and the term $\sum_{t=0}^{\infty} \lambda^{t+1} V_t^y$ is the discounted sum of the utility functions of the successive young generations.

Given the educational investments (4) and (5) and the government budget constraint, the problem can be rewritten as a simple static problem (see the appendix for the derivation of the following expression):

$$\max_{\tau_0^o \in [0,1]} L_0 + \frac{L}{1-\lambda}, \tag{13}$$

where:

$$L_0 \equiv \beta(1-\tau_0^o)\left[(1-u_0)(1-\theta\gamma) + \frac{\gamma}{4}(1-\tau_0^o)\left\{(1-u_0)\theta + u_0\right\}\right] + \frac{1}{2}(\beta+\lambda)W(\tau_0^o, u_0) \tag{14}$$

and:

$$L \equiv \frac{1}{2}(\beta+\lambda)\left[Z(\tau^y,\tau^o) + \lambda W(\tau^o, 1-e^y(\tau^y,\tau^o))\right] + \lambda\left[(e^y(\tau^y,\tau^o))^2 + \frac{\beta\gamma}{4}(1-\tau^o)^2\right]. \tag{15}$$

The problem implies that after the initial choice of τ_0^o, the problem reduces to a sequence of identical static optimization problems over τ^y and τ^o. The next proposition characterizes the solution of the Ramsey problem.

Proposition 4: *The allocation solving the Ramsey problem has:*

$$\tau_0^o = \begin{cases} 0 & \text{if } \lambda \leq \beta \\ \min\left\{1, \frac{\lambda-\beta}{\lambda\gamma}\left(\frac{(1-u_0)(1-\theta\gamma)}{(1-u_0)\theta+u_0} + \frac{\gamma}{2}\right)\right\} & \text{if } \lambda > \beta \end{cases}$$

and a constant sequence of taxes, τ^y and τ^o, given by:

$$(\tau^y,\tau^o) = \begin{cases} \left(\frac{\beta-\lambda}{2\beta}\left[1+\beta\left((1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right)\right], 0\right) & \text{if } \lambda < \beta \\ (0,0) & \text{if } \lambda = \beta \\ (0, \min\{1, \hat{\tau}^o\}) & \text{if } \lambda > \beta \end{cases}$$

where $\hat{\tau}^o$ *is a* τ^o *that satisfies:*

$$\tau^o = \frac{(\lambda-\beta)\left[\{(1-\theta\gamma)-(\gamma/2)(1-\theta)(1-\tau^o)\}e^y(0,\tau^o)+(\gamma/2)(1-\tau^o)\right]}{(1/2)(\beta+\lambda)\left[\gamma-\gamma(1-\theta)e^y(0,\tau^o)+\beta\{(1-\theta\gamma)-(\gamma/2)(1-\theta)(1-\tau^o)\}^2\right]}.$$

Proposition 4 states that the tax rates in the Ramsey allocation depend on the relative magnitude between β and λ. For $\lambda > \beta$, the planner attaches a larger weight to the young and a smaller weight to the old: the tax burden falls on the old. For $\lambda = \beta$, the planner attaches the same weights to the young and the old, but finds it optimal to set no tax on both of them. For $\lambda < \beta$, the planner attaches a larger weight to the old and a smaller weight to the young: the tax burden falls on the young.

Based on the solution of the Ramsey problem in Proposition 1, we can show under what condition the number of upwardly mobile agents is more or less than the number of downwardly mobile agents in the Ramsey allocation. In particular, we focus on mobile agents after period 1; the period-0 mobile agents are abstracted from the analysis because their number is determined by the initial distribution u_0.

In period $t \geq 1$, the numbers of upwardly and downwardly mobile agents in the Ramsey allocation are compared as follows. The derivation of the following result is given in the Appendix.

(i) When $\lambda \leq \beta$:

$$M^{up} \gtreqless M^{down} \text{ if and only if}$$
$$\lambda \lesseqgtr \hat{\lambda} \equiv \frac{4-(1+\theta)\left[1+\beta\{(1-\theta\gamma)-(1/4)\gamma(1-\theta)\}\right]}{(1+\theta)\left[1+\beta\{(1-\theta\gamma)-(1/4)\gamma(1-\theta)\}\right]}; \tag{16}$$

(ii) When $\lambda > \beta$:

$$M^{up} \gtreqless M^{down} \text{ if and only if}$$
$$\frac{1-\beta\{(1-\theta\gamma)(1-\tau^o)-(1/4)\gamma(1-\theta)(1-\tau^o)^2\}}{1+\beta\{(1-\theta\gamma)(1-\tau^o)-(1/4)\gamma(1-\theta)(1-\tau^o)^2\}} \gtreqless \theta\frac{1+\tau^o}{1-\tau^o},$$

where $\tau^o = \min\{1, \hat{\tau}^o\}$.

For $\lambda \leq \beta$, there is no tax burden on the old and the tax burden falls entirely on the young (Proposition 4). The former property implies that the probability of being skilled in old age is high so that the number of upwardly mobile agents is large. The latter property implies that the probability of being skilled in youth is low so that the number of potentially upwardly mobile agents is large. Given these two features, a natural prediction is that the number of upwardly mobile agents is larger than the number of downwardly mobile agents.

This prediction is not necessarily true. When λ is sufficiently low such that λ is below the critical value, $\hat{\lambda}$, given in (16), the prediction is actually true. However, for λ above the critical value, the prediction fails to hold: the number of upwardly mobile agents is smaller than the number of downwardly mobile agents. The mechanism behind this counterintuitive result is as follows. A larger weight on the young leads to a lower tax rate on the young. This results in a higher probability of being successful in youth and thus a larger number of agents who potentially experience downward mobility in old age. Because of this size effect, the number of downwardly mobile agents becomes larger than the number of upwardly mobile agents when λ is above the critical value.

For $\lambda > \beta$, there is no tax burden on the young and all the tax burden falls on the old (Proposition 4). The former property implies that the probability of being skilled in youth is high so that the number of potential downwardly mobile agents is large. The latter property implies that the probability of being skilled in old age is low because of the heavy tax burden. Given these two features, the number of downwardly mobile agents is larger than the number of upwardly mobile agents as long as the effect of the latter property is large such that a 100% tax rate is levied on the old. However, the opposite result holds if the tax rate on the old is less than 100% and the downward mobility risk, represented by θ, is low. Therefore, the relative size of upwardly and downwardly mobile agents depends on the Ramsey tax rate on the old as well as the downward mobility risk when $\lambda > \beta$.

5.2 Normative Aspect of Political Equilibria

Based on the characterization of the Ramsey allocation above, let us discuss the normative aspects of the political equilibria. Figures 9 and 10 illustrate numerical examples that compare the political equilibria with the Ramsey allocation in terms of the tax rate on the young, the probability of becoming skilled in youth, the expected utility of the young and upward and downward mobility.

First, consider the unskilled-majority equilibrium with no taxation on the old (Proposition 1). This equilibrium bears some resemblance to the Ramsey allocation in the case of $\lambda < \min(\beta, \hat{\lambda})$, where the tax burden falls on the young and the number of upwardly mobile agents is larger than the number of downwardly mobile agents (see Proposition 4). This implies that when the economy is featured by higher prospects of upward and downward mobility such that the POUM hypothesis is supported, the political economy attains an allocation qualitatively similar to the Ramsey allocation as long as the Ramsey planner attaches a low weight to the young such that $\lambda < \min(\beta, \hat{\lambda})$.

Next, consider the unskilled-majority equilibrium with taxation on the old (Proposition 2). The equilibrium has a resemblance to the Ramsey allocation in the case of $\lambda > \beta$, where the tax burden falls on the old. However, there is a difference in that the Ramsey planner does not impose the tax on the young, while they are taxed in the political equilibrium presented in Proposition 2. There is also a difference in mobility: there are some upwardly mobile agents in the Ramsey allocation, while there are no upwardly mobile agents in the political equilibrium in Proposition 2.

Finally, consider the skilled-majority equilibrium (Proposition 3). This equilibrium bears some resemblance to the Ramsey tax in the case of $\lambda < \beta$, where the tax burden falls on the young. However, this resemblance is not firm because under the set of parameters that attain the skilled-majority equilibrium, there may also be the unskilled-majority equilibrium presented in Proposition 2. In the situation of multiple equilibria, the political economy may or may not attain an allocation qualitatively similar to the Ramsey allocation depending on expectations of agents and the planner's weight to the young. In addition, the qualitative property of mobility in the political equilibrium may or may not be similar to that in the Ramsey allocation depending on the expectations of agents.

6 Conclusion

This paper presented a simple theoretical model that includes earnings mobility and redistributive politics. The model demonstrates a mutual link between mobility and redistributive politics, and gives a comparison between the political equilibrium and the Ramsey allocation in terms of mobility and redistributive policy.

The contribution of this paper is twofold. First, the model draws a distinction between upward mobility opportunity and downward mobility risk, but it shows that two factors play similar roles in the characterization of political equilibria. That is, the economy is more likely to attain a unique, unskilled-majority equilibrium supporting the POUM hypothesis when the opportunity and risk are higher. Second, efficiency of the political equilibria in terms of mobility and redistributive policy depends on the expectations of agents.

The results established in the paper have the following empirical and policy implications. Empirical evidence shows that among Western European countries, there is an appreciable difference in redistribution and earnings mobility. In particular, the US is characterized by higher earnings mobility and lower redistribution compared with Western European countries such as Denmark, France, Germany, Italy, Sweden and the UK (OECD, 1996). In the US, the current poor do not support high redistribution because of the hope of upward mobility in the future. By contrast, in Europe, people who fail in education or acquisition of high skills are supported by high redistribution. Furthermore, even in Western European countries, there is also a cross-country difference in earnings mobility and redistribution (Atkinson, Bourguignon and Morrisson, 1992; OECD, 1996). For example, the UK is featured by higher mobility and lower redistribution compared with the other countries. The difference is also observed between Scandinavian and continental European countries.

The analysis and the results in this paper provide one possible explanation for the above-mentioned cross-country differences. Keys to the explanation are (i) upward mobility opportunity and downward mobility risk; and (ii) multiple, self-fulfilling expectations of agents. A high-mobility economy featured by a unique, unskilled-majority equilibrium with no taxation on the old, which supports the POUM hypothesis, gives an explanation for the US economy with low redistribution supported by the majority of low-income agents. In contrast, a low-mobility economy featured by multiple equilibria gives one possible explanation for the cross-country differences among European countries.

Our numerical analysis provided an interesting policy implication for mobility; that is, an increase in the mobility opportunity may improve the expected utility of the young under a certain condition. This result depends heavily on the political determination of redistribution taxes. An increase in mobility opportunity might induce voters to set a lower tax rate on the young. This gives the young an incentive to invest more in educational investment, which results in an improvement in the expected utility of the young. Such a positive mechanism never holds in an economy without an endogenous response of tax rates via voting. Our result therefore indicates that the mobility opportunity might be beneficial from the viewpoint of utility when the mobility is associated with redistributive politics.

To obtain these results, we simplified the analysis by adopting a simple lump-sum transfer scheme. We did not consider alternative policy methods, for example, transfers that target the elderly or the poor. In addition, we did not consider differences in ability by assuming homogeneous agents. However, we believe that this paper provides a tractable framework for explaining the cross-country differences in earnings mobility and redistribution policy and for examining the efficiency implications of mobility opportunity and risk.

7 Appendix

7.1 Proof of Proposition 1

(i) Suppose that at time t, agents know that $\tau_t^y = \arg\max Z(\tau_t^y, 0) \equiv \tau_{\max}^{Z0}$ and expect $\tau_{t+1}^o = 0$. Then:

$$\begin{aligned} u_{t+1} &= 1 - e^{y*}\left(\tau_{\max}^{Z0}, 0\right) \\ &= 1 - \frac{1}{4}\left[1 + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right] \\ &> 1/2, \end{aligned}$$

where the last inequality comes from $\beta \in (0,1]$ and $\gamma, \theta \in [0,1]$. By (9), this implies that $\tau_{t+1}^o = 0$, fulfilling initial expectations. Therefore, there exists an unskilled-majority equilibrium with no taxation on the old if the decisive voter finds it optimal to set $\tau_t^y = \tau_{\max}^{Z0}$.

To establish that setting $\tau_t^y = \tau_{\max}^{Z0}$ is optimal for the decisive voter, we note the following properties of the function Z : (a) $Z(\tau_t^y, 0)$ is concave in τ_t^y and is maximized at $\tau_t^y = \arg\max_{\tau_t^y \in [0,1]} Z(\tau_t^y, 0) = \tau_{\max}^{Z0}$; and (b) $Z(\tau_t^y, 0) > Z(\tau_t^y, 1) \forall \tau_t^y \in (0,1]$. These properties imply that setting $\tau_t^y = \tau_{\max}^{Z0}$ is optimal if this setting is feasible under the expectation of $\tau_{t+1}^o = 0$, i.e., if $\tau^{u0} \leq \tau_{\max}^{Z0}$. The inequality is rewritten as:

$$\beta \leq \phi \cdot \left[1 - 2 \cdot \frac{1 - \gamma \cdot (1+\theta)}{1 - \theta\gamma}\right],$$

which is equivalent to the assumption in statement (i) of Proposition 1. Panel (a) of Figure 4 illustrates the revenue from the young under Proposition 1(i).

Given $\tau^y = \tau_{\max}^{Z0}$ and $\tau^o = 0$, e^y, e^{os} and e^{ou} are calculated as:

$$e^y = \frac{1}{4}\left[1 + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right], e^{os} = e^{ou} = \frac{1}{2},$$

which leads to:

$$\begin{aligned} M^{up} &= \frac{\gamma}{2}\left[1 - \frac{1}{4}\left[1 + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right]\right], \\ M^{down} &= \frac{\theta\gamma}{8}\left[1 + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right]. \end{aligned}$$

Direct calculation yields $M^{up} - M^{down} > 0$.

(ii) Suppose that at time t, agents know that $\tau_t^y = \tau^{u0}$ and expect $\tau_{t+1}^o = 0$. Then:

$$\begin{aligned} u_{t+1} &= 1 - e^{y*}\left(\tau^{u0}, 0\right) \\ &= 1 - \frac{\gamma}{2(1-\theta\gamma)} \\ &> 1/2, \end{aligned}$$

where the last inequality holds under Assumption 1. By (9), this implies that $\tau_{t+1}^o = 0$, fulfilling initial expectations. Therefore, there exists an unskilled-majority equilibrium with no taxation on the old if the decisive voter finds it optimal to set $\tau_t^y = \tau^{u0}$.

To establish that setting $\tau_t^y = \tau^{u0}$ is optimal for the decisive voter, we first note that setting $\tau_t^y = \tau^{u0}$ is feasible if and only if $\tau_t^y = \tau^{u0} \leq 1$, i.e.:

$$\beta \leq \phi \cdot \frac{\gamma}{1-\theta\gamma}. \tag{17}$$

This is the second assumption given in Proposition 1(ii).

Second, setting $\tau_t^y = \arg\max_{\tau_t^y \in [0,1]} Z(\tau_t^y, 0)$ is infeasible if and only if $\tau_{\max}^{Z0} < \tau^{u0}$, that is:

$$\phi \cdot \left[1 - 2\frac{1-\gamma(1+\theta)}{1-\theta\gamma}\right] < \beta. \tag{18}$$

This is the first assumption given in Proposition 1(ii). If (18) fails to hold, $\tau_t^y = \tau^{u0}$ is dominated by $\tau_t^y = \tau_{\max}^{Z0}$ because $Z(\tau_t^y, 0)$ is maximized at $\tau_t^y = \tau_{\max}^{Z0}$.

Given conditions (17) and (18), the revenue from the young is illustrated in panel (b) of Figure 4. The relevant payoff function is $Z(\tau_t^y, 0)$ for $\tau_t^y \leq \tau^e$ and $Z(\tau_t^y, 1)$ for $\tau_t^y \leq \tau^{u1}$. Given the properties such that $Z(\tau_t^y, 0)$ is increasing in τ_t^y for $\tau_t^y \in (0, \tau^e)$ with $\tau^e \leq \tau^{s0}$ and decreasing in τ_t^y for $\tau_t^y \in (\tau^{u0}, 1)$, and $\arg\max Z(\tau_t^y, 1) = 1/2$, an alternative option is to choose (i) $\tau_t^y = \tau^e$ under the expectation of $\tau_{t+1}^o = 0$, or (ii) $\tau_t^y = 1/2$ under the expectation of $\tau_{t+1}^o = 1$. The original option dominates the first alternative if $Z(\tau^{u0}, 0) \geq Z(\tau^{s0}, 0)$ holds, i.e., if the second assumption in Proposition 1(ii) holds. The original option dominates the second alternative if $Z(\tau^{u0}, 0) \geq Z(1/2, 1)$ holds, i.e., if the third assumption in Proposition 1(ii) holds.

Given $\tau^y = \tau^{u0}$ and $\tau^o = 0$, e^y, e^{os} and e^{ou} are calculated as:

$$e^y = \frac{\gamma}{2(1-\theta\gamma)}, e^{os} = e^{ou} = \frac{1}{2},$$

which leads to:

$$M^{up} = \frac{\gamma}{2}\left[1 - \frac{\gamma}{2(1-\theta\gamma)}\right], M^{down} = \frac{\theta(\gamma)^2}{4(1-\theta\gamma)}.$$

Direct calculation yields $M^{up} - M^{down} > 0$. ■

7.2 Proof of Proposition 2

The first assumption in the statement of Proposition 2, $\gamma \leq 1/(2+\theta)$, is rewritten as $1/2 \leq \tau^{u1}$. The assumption therefore implies that it is feasible to set $\tau_t^y = 1/2$ under the expectation of $\tau_{t+1}^o = 1$.

Suppose that at time t, agents know that $\tau_t^y = 1/2$ and expect that $\tau_{t+1}^o = 1$. Then, $1 - e^{y*}(1/2, 1) = u_{t+1} = 3/4$. By (9), this implies that $\tau_{t+1}^o = 1$, fulfilling initial expectations. Therefore, there exists an unskilled-majority equilibrium with $\tau_{t+1}^o = 1$ if the decisive voter finds it optimal to set $\tau_t^y = 1/2$.

To establish that setting $\tau_t^y = 1/2$ is optimal for the decisive voter, we first note that under the first assumption, it always holds that $\tau_{\max}^{Z0} < \tau^{u0}$, because this inequality is rewritten as:

$$\frac{-1+\gamma(2+\theta)}{2(1-\theta\gamma)} < \frac{1}{2}\beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\},$$

where the left-hand side is nonpositive as long as the first assumption, $\gamma \leq 1/(2+\theta)$, holds, while the right-hand side is positive. Therefore, it is infeasible to set $\tau_t^y = \tau_{\max}^{Z0}$ with the expectation of $\tau_{t+1}^o = 0$ under the assumption of $\gamma \leq 1/(2+\theta)$.

Given this result, there are two alternative options: setting $\tau_t^y = \tau^e$ under the expectation of $\tau_{t+1}^o = 0$, and setting $\tau_t^y = \tau^{u0}$ under the expectation of $\tau_{t+1}^o = 0$ provided that this alternative option is feasible. The second assumption in the statement of Proposition 2 is rewritten as $\tau^{s0} < \tau^{u1}$. Under this condition, there are multiple, self-fulfilling expectations of agents as long as $\tau_t^y \leq \tau^{s0}$. The current option, $(\tau_t^y, \tau_{t+1}^o) = (1/2, 1)$, dominates the first alternative option, $(\tau_t^y, \tau_{t+1}^o) = (\tau^e, 0)$ as long as τ^e is set to be low such that $Z(\tau^e, 0) < Z(1/2, 1)$.

The first part of the third assumption, denoted by (a), is rewritten as $\tau^{u0} > 1$. Under this condition, it is infeasible to choose the second alternative option $(\tau_t^y, \tau_{t+1}^o) = (\tau^{u0}, 0)$. In contrast, under the second part of the third assumption, denoted by (b), the second alternative option is feasible but does not dominate the current option because the assumption (b) is rewritten as:

$$Z\left(\tau^{u0}, 0\right) < Z\left(\frac{1}{2}, 1\right).$$

With $(\tau_t^y, \tau_{t+1}^o) = (1/2, 1)$, M^{up} and M^{down} are calculated as $M^{up} = 0$ and $M^{down} = \theta\gamma/4$. ■

7.3 Proof of Proposition 3

(i) Suppose that at time t, agents know that $\tau_t^y = \tau^{s0}$ and expect $\tau_{t+1}^o = 0$. Then, $1 - e^{y*}(\tau^{s0}, 0) = u_{t+1} = 1/2$. By (9), 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 $\tau_t^y = \tau^{s0}$.

To establish that setting $\tau_t^y = \tau^{s0}$ is optimal for the decisive voter, we first note that the first assumption in statement (i) of Proposition 3,

$$\beta > \phi \cdot \frac{1-\gamma(1+\theta)}{1-\theta\gamma},$$

is derived from the condition $\tau^{s0} > \tau^{u1}$. This condition implies that under the expectation of $\tau_{t+1}^{o} = 0$, the decisive voter can choose $\tau_t^y = \tau^{s0}$ irrespective of the expectation of τ^e.

The second assumption in statement (i) of Proposition 3,

$$\beta > \phi \cdot \frac{\gamma}{1-\theta\gamma},$$

is rewritten as $1 < \tau^{u0}$, implying that it is infeasible for the decisive voter to set the two tax rates on the young, $\tau_t^y = \tau_{\max}^{Z0}$ and τ^{u0}, demonstrated in Proposition 1, under the expectation of $\tau_{t+1}^{o} = 0$. Therefore, the available choice for the decisive voter is limited to the range of $[0, \tau^{s0}]$ as long as he/she expects $\tau_{t+1}^{o} = 0$. Given that $Z(\tau_t^y, 0)$ is increasing in τ_t^y for that range, setting $\tau_t^y = \tau^{s0}$ is the revenue-maximizing behavior under the expectation of $\tau_{t+1}^{o} = 0$.

Given $(\tau_t^y, \tau_{t+1}^{o}) = (\tau^{s0}, 0)$, we obtain $e^y = 1/2$ and $e^{ou} = e^{os} = 1/2$. M^{up} and M^{down} are calculated as $M^{up} = (1-e^y)\gamma e^{ou} = \gamma/4$ and $M^{down} = e^y\theta\gamma(1-e^{os}) = \theta\gamma/4$. $M^{up} \geq M^{down}$ holds where an equality holds if and only if $\theta = 1$.

(ii) Suppose that at time t, agents know that $\tau_t^y = \tau^e (\leq \tau^{s0})$ and expect $\tau_{t+1}^{o} = 0$. Then, $1 - e^{y*}(\tau^e, 0) = u_{t+1} < 1/2$. By (9), 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 $\tau_t^y = \tau^e$.

To establish that setting $\tau_t^y = \tau^e$ is optimal for the decisive voter, we first note that the first assumption in statement (ii) of Proposition 3 is rewritten as $1 < \tau^{u0}$, implying that it is infeasible for the decisive voter to set the two tax rates on the young, demonstrated in Proposition 1, under the expectation of $\tau_{t+1}^{o} = 0$. Given that $Z(\tau_t^y, 0)$ is increasing in τ_t^y for $\tau_t^y \in [0, \tau^e)$, setting $\tau_t^y = \tau^e$ is the revenue-maximizing behavior under the expectation of $\tau_{t+1}^{o} = 0$ as long as $[0, \tau^e)$ is a feasible range of τ_t^y.

Given the above argument, the remaining alternative options for the decisive voter are to set $\tau_t^y = \tau^{s0}$ irrespective of expectations on τ_t^y under the expectation of $\tau_{t+1}^{o} = 0$, or to set $\tau_t^y = \arg\max Z(\tau_t^y, 1) = 1/2$ under the expectation of $\tau_{t+1}^{o} = 1$. The second assumption in statement (ii) of Proposition 3 is rewritten as $\tau^{s0} \leq \tau^{u1}$, implying that the option of $(\tau_t^y, \tau_{t+1}^{o}) = (\tau^{s0}, 0)$ is unavailable for the decisive voter; otherwise, this option dominates the current choice: $Z(\tau^{s0}, 0) \geq Z(\tau^e, 0)$.

The third assumption in statement (ii) ensures that there exists a $\tau_t^y = \tau^e$ that sustains the choice of $(\tau_t^y, \tau_{t+1}^o) = (\tau^e, 0)$ against the choice of $(\tau_t^y, \tau_{t+1}^o) = (1/2, 1)$ as an equilibrium: $Z(\tau^e, 0) \geq Z(1/2, 1)$. This inequality condition is rewritten as $\tau^e \geq \tilde{\tau}$ where:

$$\begin{aligned}\tilde{\tau} &\equiv \frac{1 + \beta\left\{(1-\theta\gamma) - (\gamma/4)\cdot(1-\theta)\right\} - \sqrt{\left[1 + \beta\left\{(1-\theta\gamma) - (\gamma/4)\cdot(1-\theta)\right\}\right]^2 - 1}}{2} \\ &= \frac{1 + \tau^{s0} - \sqrt{\left[1+\tau^{s0}\right]^2 - 1}}{2}\end{aligned}$$

As τ^e is bounded above τ^{s0}, τ^e must be set within the range $[\tilde{\tau}, \tau^{s0}]$ in order that $Z(\tau^e, 0) \geq Z(1/2, 1)$ holds. The third assumption ensures that the set $[\tilde{\tau}, \tau^{s0}]$ is nonempty.

Given that $\tau^y = \tau^e$ and $\tau^o = 0$, e^y, e^{os} and e^{ou} are calculated as:

$$\begin{aligned}M^{up} &= \frac{\gamma}{2}\left[1 - \frac{1}{2}\left\{(1-\tau^e) + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right\}\right], \\ M^{down} &= \frac{\theta\gamma}{4}\left[(1-\tau^e) + \beta\left\{(1-\theta\gamma) - \frac{\gamma(1-\theta)}{4}\right\}\right].\end{aligned}$$

Direct calculation leads to $M^{up} \gtreqless M^{down}$ if and only if $\tau^e \gtreqless (1-\theta)/(1+\theta) + \tau^{s0}$. ■

7.4 Proof of Proposition 4

We first show that the Ramsey problem is written as a static optimization problem given by (13)–(15). To show this, we calculate the indirect utility functions of the initial old and the young in generation t :

$$\begin{aligned}V_0^{os} &= (1-\theta\gamma)(1-\tau_0^o) + \frac{\theta\gamma}{4}(1-\tau_0^o)^2 + \frac{1}{2}\{W(\tau_0^o, u_0) + Z(\tau_0^y, \tau_1^o)\}, \\ V_0^{ou} &= \frac{\gamma}{4}(1-\tau_0^o)^2 + \frac{1}{2}\{W(\tau_0^o, u_0) + Z(\tau_0^y, \tau_1^o)\}, \\ V_t^y &= e^{y*}(\tau_t^y, \tau_{t+1}^o)\cdot(1-\tau_t^y) - (e^{y*}(\tau_t^y, \tau_{t+1}^o))^2 + \frac{1}{2}\{W(\tau_t^o, u_t) + Z(\tau_t^y, \tau_{t+1}^o)\} \\ &\quad + \beta\left[e^{y*}(\tau_t^y, \tau_{t+1}^o)\left\{(1-\theta\gamma)(1-\tau_{t+1}^o) + \frac{\theta\gamma}{4}(1-\tau_{t+1}^o)^2\right\}\right. \\ &\quad \left. + \left(1 - e^{y*}(\tau_t^y, \tau_{t+1}^o)\right)\frac{\gamma}{4}(1-\tau_{t+1}^o)^2 + \frac{1}{2}\{W(\tau_{t+1}^o, u_{t+1}) + Z(\tau_{t+1}^y, \tau_{t+2}^o)\}\right].\end{aligned}$$

We substitute these functions into the social welfare function, denoted by Ω, to obtain:

$$\Omega = \beta(1-u_0)\cdot\left[\underbrace{(1-\theta\gamma)(1-\tau_0^o)+\frac{\theta\gamma}{4}(1-\tau_0^o)^2}_{\text{(a.1)}}+\underbrace{\frac{1}{2}W(\tau_0^o,u_0)}_{\text{(a.2)}}+\underbrace{\frac{1}{2}Z(\tau_0^y,\tau_1^o)}_{\text{(b.1)}}\right]$$

$$+\beta u_0\cdot\left[\underbrace{\frac{\gamma}{4}(1-\tau_0^o)^2}_{\text{(a.3)}}+\underbrace{\frac{1}{2}W(\tau_0^o,u_0)}_{\text{(a.4)}}+\underbrace{\frac{1}{2}Z(\tau_0^y,\tau_1^o)}_{\text{(b.2)}}\right]$$

$$+\lambda\cdot\left[\underbrace{e^{y*}(\tau_0^y,\tau_1^o)\cdot(1-\tau_0^y)-(e^{y*}(\tau_0^y,\tau_1^o))^2}_{\text{(b.3)}}+\underbrace{\frac{1}{2}W(\tau_0^o,u_0)}_{\text{(a.5)}}+\underbrace{\frac{1}{2}Z(\tau_0^y,\tau_1^o)}_{\text{(b.4)}}\right.$$

$$+\beta\cdot\left\{\underbrace{e^{y*}(\tau_0^y,\tau_1^o)\left\{(1-\theta\gamma)(1-\tau_1^o)+\frac{\theta\gamma}{4}(1-\tau_1^o)^2\right\}+(1-e^{y*}(\tau_0^y,\tau_1^o))\,\frac{\gamma}{4}(1-\tau_1^o)^2}_{\text{(b.5)}}\right.$$

$$\left.\left.+\underbrace{\frac{1}{2}W(\tau_1^o,u_1)}_{\text{(b.6)}}+\underbrace{\frac{1}{2}Z(\tau_1^y,\tau_2^o)}_{\text{(c.1)}}\right\}\right]$$

$$+\lambda^2\cdot\left[\underbrace{e^{y*}(\tau_1^y,\tau_2^o)\cdot(1-\tau_1^y)-(e^{y*}(\tau_1^y,\tau_2^o))^2}_{\text{(c.2)}}+\underbrace{\frac{1}{2}W(\tau_1^o,u_1)}_{\text{(b.7)}}+\underbrace{\frac{1}{2}Z(\tau_1^y,\tau_2^o)}_{\text{(c.3)}}\right.$$

$$+\beta\cdot\left\{\underbrace{e^{y*}(\tau_1^y,\tau_2^o)\left\{(1-\theta\gamma)(1-\tau_2^o)+\frac{\theta\gamma}{4}(1-\tau_2^o)^2\right\}+(1-e^{y*}(\tau_1^y,\tau_2^o))\,\frac{\gamma}{4}(1-\tau_2^o)^2}_{\text{(c.4)}}\right.$$

$$\left.\left.+\underbrace{\frac{1}{2}W(\tau_2^o,u_2)}_{\text{(c.5)}}+\frac{1}{2}Z(\tau_2^y,\tau_3^o)\right\}\right]$$

$$+\lambda^3\cdot\left[e^{y*}(\tau_2^y,\tau_3^o)\cdot(1-\tau_2^y)-(e^{y*}(\tau_2^y,\tau_3^o))^2+\underbrace{\frac{1}{2}W(\tau_2^o,u_2)}_{\text{(c.6)}}+\frac{1}{2}Z(\tau_2^y,\tau_3^o)\right.$$

$$+\beta\cdot\left\{e^{y*}(\tau_2^y,\tau_3^o)\left\{(1-\theta\gamma)(1-\tau_3^o)+\frac{\theta\gamma}{4}(1-\tau_3^o)^2\right\}+(1-e^{y*}(\tau_2^y,\tau_3^o))\,\frac{\gamma}{4}(1-\tau_3^o)^2\right.$$

$$\left.\left.+\frac{1}{2}W(\tau_3^o,u_3)+\frac{1}{2}Z(\tau_3^y,\tau_4^o)\right\}\right]$$

$$+\sum_{t=3}^{\infty}\lambda^{t+1}V_t^y$$

where the terms (a1)–(a5) include τ_0^o, the terms (b.1)–(b.7) include τ_0^y and/or τ_1^o, the terms (c.1)–(c.6) include τ_1^y and/or τ_2^o, and so on. Given this feature, the equation above is rewritten as:

$$\Omega = L_0 + \sum_{t=1}^{\infty} \lambda^t L_t,$$

where:

$$L_0 \equiv \beta(1-\tau_0^o)\left[(1-u_0)(1-\theta\gamma) + \frac{\gamma}{4}(1-\tau_0^o)\{(1-u_0)\theta + u_0\}\right] + \frac{1}{2}(\beta+\lambda)W(\tau_0^o, u_0),$$
$$L_t \equiv \frac{1}{2}(\beta+\lambda)\left[Z(\tau_{t-1}^y, \tau_t^o) + \lambda W(\tau_t^o, 1-e^{y*}(\tau_{t-1}^y, \tau_t^o))\right] + \lambda\left[\left(e^{y*}(\tau_{t-1}^y, \tau_t^o)\right)^2 + \frac{\beta\gamma}{4}(1-\tau_t^o)^2\right].$$

The function L_0 is the sum of the terms (a1)–(a5), the function L_1 is the sum of the terms (b1)–(b7), the function L_2 is the sum of the terms (c1)–(c6), and so on. The function L_t ($t \geq 1$) indicates that the solution of (τ_{t-1}^y, τ_t^o) that maximizes L_t is stationary over time. Therefore, we can write the Ramsey problem as (13)–(15).

Part 1 of Proposition 4

The solution of τ_0^o is derived by solving (14). The first derivative of L_0 with respect to τ_0^o is:

$$\begin{aligned}\frac{\partial L_0}{\partial \tau_0^o} &= (-\beta)\left[(1-u_0)(1-\theta\gamma) + \frac{\gamma}{4}(1-\tau_0^o)\left\{(1-u_0)\theta + u_0\right\}\right] + (-\beta)(1-\tau_0^o)\frac{\gamma}{4}\left\{(1-u_0)\theta + u_0\right\} \\ &+ \frac{1}{2}(\beta+\lambda)\left[(1-u_0)(1-\theta\gamma) + \frac{\gamma}{2}(1-\tau_0^o)\left\{1-(1-\theta)(1-u_0)\right\}\right] \\ &+ (-1)\frac{1}{2}(\beta+\lambda)\cdot\frac{\gamma}{2}\cdot\left\{1-(1-\theta)(1-u_0)\right\}\tau_0^o.\end{aligned}$$

Therefore, the solution is given by:[4]

$$\begin{cases} \tau_0^o = 0 & \text{if } \partial L_0/\partial \tau_0^o|_{\tau_0^o=0} \leq 0 \\ \tau_0^o = 1 & \text{if } \partial L_0/\partial \tau_0^o|_{\tau_0^o=1} \geq 0 \\ \tau_0^o = \frac{\lambda-\beta}{\lambda\gamma}\left[\frac{(1-u_0)(1-\theta\gamma)}{(1-u_0)\theta+u_0} + \frac{\gamma}{2}\right] & \text{otherwise.} \end{cases}$$

The condition $\partial L_0/\partial \tau_0^o|_{\tau_0^o=0} \leq 0$ is rewritten as $\lambda \leq \beta$. The condition $\partial L_0/\partial \tau_0^o|_{\tau_0^o=1} \geq 0$ is reduced to $(1-\theta\gamma)(1-u_0)(\lambda-\beta) \geq (\gamma/2)(\beta+\lambda)\left\{(1-u_0)\theta + u_0\right\}$. Therefore, we

[4]The second-order condition, $\partial^2 L_0/\partial \tau_0^{o2} < 0$, is satisfied:

$$\begin{aligned}\frac{\partial^2 L_0}{\partial \tau_0^{o2}} &= \frac{\beta\gamma}{2}\left\{(1-u_0)\theta + u_0\right\} - \frac{1}{2}(\beta+\lambda)\gamma\left\{1-(1-\theta)(1-u_0)\right\} \\ &\leq \frac{\beta\gamma}{2}\left\{(1-u_0)\theta + u_0\right\} - \frac{\beta\gamma}{2}\left\{1-(1-\theta)(1-u_0)\right\} \\ &= 0.\end{aligned}$$

obtain:

$$\begin{cases} \tau_0^o = 0 & \text{if} \quad \lambda \le \beta \\ \tau_0^o = 1 & \text{if} \quad \lambda > \beta \text{ and } u_0 \le \frac{(1-\theta\gamma)(\lambda-\beta)-(\gamma\theta/2)(\beta+\lambda)}{(1-\theta\gamma)(\lambda-\beta)+(1-\theta)(\gamma/2)(\beta+\lambda)} \\ \tau_0^o = \frac{\lambda-\beta}{\lambda\gamma}\left[\frac{(1-u_0)(1-\theta\gamma)}{(1-u_0)\theta+u_0} + \frac{\gamma}{2}\right] & \text{if} \quad \text{otherwise.} \end{cases}$$

The last two solutions are summarized as:

$$\tau_0^o = \min\left\{1, \frac{\lambda-\beta}{\lambda\gamma}\left[\frac{(1-u_0)(1-\theta\gamma)}{(1-u_0)\theta+u_0} + \frac{\gamma}{2}\right]\right\} \text{ if } \lambda > \beta.$$

Part 2 of Proposition 4

Next, we derive the solution of the pair (τ^y, τ^o) by solving (15). The solution must satisfy the following first-order conditions:

$$\tau^y : \partial L/\partial \tau^y - \xi_1^y + \xi_0^y = 0, \tag{19}$$

$$\tau^o : \partial L/\partial \tau^o - \xi_1^o + \xi_0^o = 0, \tag{20}$$

where ξ_0^y and ξ_0^o are Kuhn–Tucker multipliers associated with the constraints $\tau^y \ge 0$ and $\tau^o \ge 0$, respectively, whereas ξ_1^y and ξ_1^o are the Kuhn–Tucker multipliers associated with the constraints $\tau^y \le 1$ and $\tau^o \le 1$, respectively.

The first-order conditions with respect to τ^y and τ^o are given by, respectively:

$$\begin{aligned} \tau^y :& \frac{1}{2}(\beta+\lambda)e^{y*}(\tau^y,\tau^o) - \frac{1}{4}(\beta+\lambda)\tau^y + \frac{1}{4}(\beta+\lambda)\lambda\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}\tau^o \\ &= \xi_1^y - \xi_0^y, \end{aligned} \tag{21}$$

$$\begin{aligned} \tau^o :& \frac{1}{2}(\beta+\lambda)\left[\frac{1}{2}\beta\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\tau^o)(1-\theta)\right\}\tau^y \right. \\ &+\lambda\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-2\tau^o)\right\}e^{y*}(\tau^y,\tau^o) \\ &\left. +\frac{\lambda\gamma}{2}(1-2\tau^o) - \frac{\lambda\beta}{2}\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}^2\tau^o\right] \\ &+\lambda\left[e^{y*}(\tau^y,\tau^o)\beta\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\tau^o)(1-\theta)\right\} - \frac{\beta\gamma}{2}(1-\tau^o)\right] \\ &= \xi_1^o - \xi_0^o. \end{aligned} \tag{22}$$

First, assume that $\tau^o = \xi_1^y = \xi_1^o = \xi_0^y = 0$. Then, from (21) and (22), we obtain:

$$\begin{aligned} \tau^y &= \frac{1}{2\beta}(\beta-\lambda)\left\{1 + \beta((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta))\right\} \in (0,1), \\ \xi_0^o &= (\beta-\lambda)\left[\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)\right\}e^{y*}(\tau^y,\tau^o)\frac{1}{2}(\beta+\lambda) + \frac{\lambda\gamma}{4}\right], \end{aligned}$$

where $\xi_0^o > 0$ as long as $\beta > \lambda$. If $\beta = \lambda$, then $(\tau^y, \tau^o) = (0, 0)$.

Next, assume that $\tau^y = \xi_1^y = \xi_1^o = \xi_0^o = 0$. Then, from (21) and (22), we obtain:

$$\tau^o = \frac{(\lambda - \beta)\left[\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\} e^{y*}(0, \tau^o) + \frac{\gamma}{2}(1-\tau^o)\right]}{\frac{1}{2}(\beta+\lambda)\left[\gamma - \gamma(1-\theta)e^{y*}(0,\tau^o) + \beta\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}^2\right]}, \tag{23}$$

$$\begin{aligned}\xi_0^y &= \frac{1}{4}(\lambda-\beta)\left[1 + \beta\left\{(1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2\right\}\right] \\ &\quad + \frac{1}{4}(\beta+\lambda)\lambda\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}\tau^o,\end{aligned}$$

where $\tau^o > 0$ and $\xi_0^y > 0$ as long as $\lambda > \beta$.

The left-hand and right-hand sides of (23), denoted by LHS and RHS, respectively, have the following properties:

$$\begin{aligned}&\frac{\partial LHS}{\partial \tau^o} > 0,\ LHS|_{\tau^o=0} = 0,\ LHS|_{\tau^o=1} = 1; \\ &\frac{\partial RHS}{\partial \tau^o} < 0,\ RHS|_{\tau^o=0} > 0,\ RHS|_{\tau^o=1} = \frac{(\lambda-\beta)(1-\theta\gamma)}{(\beta+\lambda)\left[\gamma\left\{1 - (1/2)(1-\theta)\right\} + \beta(1-\theta\gamma)^2\right]}.\end{aligned}$$

Given these properties, there exists a τ^o, denoted by $\hat{\tau}^o$, that satisfies (23), where $\hat{\tau}^o$ is defined in Proposition 1. If $LHS|_{\tau^o=1} > RHS|_{\tau^o=1}$, then $\tau^o = \hat{\tau}^o$; otherwise, $\tau^o = 1$:

$$\tau^o = \min(\hat{\tau}^o, 1).$$

It remains to be checked whether $\tau^y = 0$ continues to be a solution. Given $\tau^o = \min(\hat{\tau}^o, 1)$, we obtain:

$$\begin{aligned}\xi_0^y &= -\left.\partial L/\partial \tau^y\right|_{\tau^y=0} \\ &= \frac{1}{4}(\lambda-\beta)\left[1 + \beta\left\{(1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2\right\}\right] \\ &\quad + \frac{1}{4}(\beta+\lambda)\lambda\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}\tau^o \\ &> 0,\end{aligned}$$

where the last inequality holds under the assumption of $\lambda > \beta$. Therefore, we obtain $(\tau^y, \tau^o) = (0, \min(\hat{\tau}^o, 1))$ if $\lambda > \beta$. ■

7.5 Mobility in the Ramsey Allocation

By definition, M^{up} and M^{down} in period $t \geq 1$ are given by:

$$\begin{aligned} M^{up} &= (1-e^y)\gamma e^{ou} \\ &= \left[1 - \frac{1}{2}\left[(1-\tau^y) + \beta\left\{(1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2\right\}\right]\right]\frac{\gamma}{2}(1-\tau^o), \\ M^{down} &= e^y\gamma(1-e^{os}) \\ &= \frac{1}{2}\left[(1-\tau^y) + \beta\left\{(1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2\right\}\right]\theta\gamma\left\{1 - \frac{1}{2}(1-\tau^o)\right\}. \end{aligned}$$

(i) Suppose that $\lambda \leq \beta$.

If $\lambda < \beta$, the pair of Ramsey tax rates is:

$$(\tau^y, \tau^o) = \left(\frac{\beta - \lambda}{2\beta}\left[1 + \beta\left\{(1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right\}\right], 0\right).$$

Then, direct calculation leads to:

$$\begin{aligned} (M^{up}, M^{down}) = \Bigg(& \frac{\gamma}{2}\left[1 - \frac{\beta+\lambda}{4\beta}\left\{1 + \beta\left((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right)\right\}\right], \\ & \frac{\theta\gamma(\beta+\lambda)}{8\beta}\left\{1 + \beta\left((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right)\right\}\Bigg), \end{aligned}$$

which results in:

$$M^{up} \gtreqless M^{down} \Leftrightarrow \lambda \lesseqgtr \hat{\lambda} \equiv \frac{4 - (1+\theta)\left[1 + \beta\left\{(1-\theta\gamma) - (1/4)\gamma(1-\theta)\right\}\right]}{(1+\theta)\left[1 + \beta\left\{(1-\theta\gamma) - (1/4)\gamma(1-\theta)\right\}\right]}. \quad (24)$$

If $\lambda = \beta$, the pair of Ramsey tax rates is $(\tau^y, \tau^o) = (0, 0)$. Then, direct calculation leads to:

$$(M^{up}, M^{down}) = \left(\frac{\gamma}{2}\left[1 - \frac{1}{2}\left\{1 + \beta\left((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right)\right\}\right], \frac{\theta\gamma}{4}\left\{1 + \beta\left((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right)\right\}\right),$$

which results in:

$$M^{up} \gtreqless M^{down} \Leftrightarrow 1 \gtreqless \frac{1+\theta}{2}\left\{1 + \beta\left((1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)\right)\right\}. \quad (25)$$

Setting $\lambda = \beta$ in (24) leads to (25). The results when $\lambda \leq \beta$ are summarized as those in (24).

(ii) Suppose that $\lambda > \beta$. The pair of Ramsey tax rates is $(\tau^y, \tau^o) = (0, \min(1, \hat{\tau}^o))$.

Then, direct calculation leads to:

$$(M^{up}, M^{down}) = \left(\frac{\gamma}{4}(1-\tau^o) \left\{ 1 - \beta \left((1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2 \right) \right\}, \right.$$
$$\left. \frac{\theta\gamma}{4}(1+\tau^o) \left\{ 1 + \beta \left((1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2 \right) \right\} \right),$$

which results in:

$$\begin{aligned} & M^{up} \gtreqless M^{down} \\ & \Leftrightarrow (1-\tau^o) \left[1 - \beta \left\{ (1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2 \right\} \right] \\ & \gtreqless \theta(1+\tau^o) \left[1 + \beta \left\{ (1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2 \right\} \right]. \end{aligned} \tag{26}$$

If $\tau^o = 1$, then the left-hand side is zero whereas the right-hand side is positive: $M^{up} < M^{down}$ holds. If $\tau^o < 1$ and $\theta = 0$, then $M^{up} > M^{down}$ holds; if $\tau^o < 1$ and $\theta = 1$, then $M^{up} < M^{down}$ holds. ■

References

[1] Alesina, A., and La Ferrara, E., 2005. Preferences for redistribution in the land of opportunities. *Journal of Public Economics* 89, 897–931.

[2] Arawatari, R., and Ono, T., 2009. A second chance at success: a political economy perspective. *Journal of Economic Theory* 144, 1249–1277. doi:10.1016/j.jet.2008.11.002.

[3] Arawatari, R., and Ono, T., 2011. Retirement and social security: a political economy perspective. *Economics of Governance*, forthcoming. DOI: 10.1007/s10101-011-0099-x

[4] Atkinson, A.B., Bourguignon, F., and Morrisson, C., 1992. *Empirical Studies of Earnings Mobility*. Chur, Switzerland, Harwood Academic Publishers.

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

[6] Benabou, R., 2000. Unequal societies: income distribution and the social contract. *American Economic Review* 90, 96–129.

[7] Benabou, R., and Ok, E.A., 2001. Social mobility and the demand for redistribution; the POUM hypothesis. *Quarterly Journal of Economics* 116, 447–487.

[8] Chen, K., and Song, Z.M., 2009. Markovian social security in unequal societies. Available at SSRN: http://ssrn.com/abstract=1367476

[9] Conde-Ruiz, J.I., Galasso, V., and Profeta, P., 2011. The role of income effects in early retirement. *Journal of Public Economic Theory*, forthcoming. Available at http://www.accessecon.com/pubs/JPET/JPET-10-00184S.pdf

[10] D'Amato, M., and Galasso, V., 2010. Political intergenerational risk sharing. *Journal of Public Economics* 94, 628–637.

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

[12] Glomm, G., and Ravikumar, B., 1995. Endogenous public policy and multiple equilibria. *European Journal of Political Economy* 11, 653–662.

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

[14] Gonzalez-Eiras, M., and Niepelt, D., 2011. Ageing, government budgets, retirement, and growth. *European Economic Review*, forthcoming. doi:10.1016/j.euroecorev.2011.05.007

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

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

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

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

[19] OECD, 1996. *Employment Outlook 1996*. Paris, OECD.

[20] Piketty, T., 1995. Social mobility and redistributive politics. *Quarterly Journal of Economics* 110, 551–584.

[21] Quadrini, V., 1999. Growth, learning and redistributive policies. *Journal of Public Economics* 74, 263–297.

[22] Saint-Paul, G., and Verdier, T., 1997. Power, distributive conflicts, and multiple growth paths. *Journal of Economic Growth* 2, 155–168.

[23] Song, Z.M., 2009a. The dynamics of inequality and social security in general equilibrium. Available at SSRN: http://ssrn.com/abstract=1367491

[24] Song, Z.M., 2009b. Persistent ideology and the determination of public policy over time. Available at SSRN: http://ssrn.com/abstract=1367474

[25] Song, Z.M., Storesletten, K., and Zilibotti, F., 2007. Rotten parents and disciplined children: a politico-economic theory of public expenditure and debt. Available at SSRN: http://ssrn.com/abstract=996388

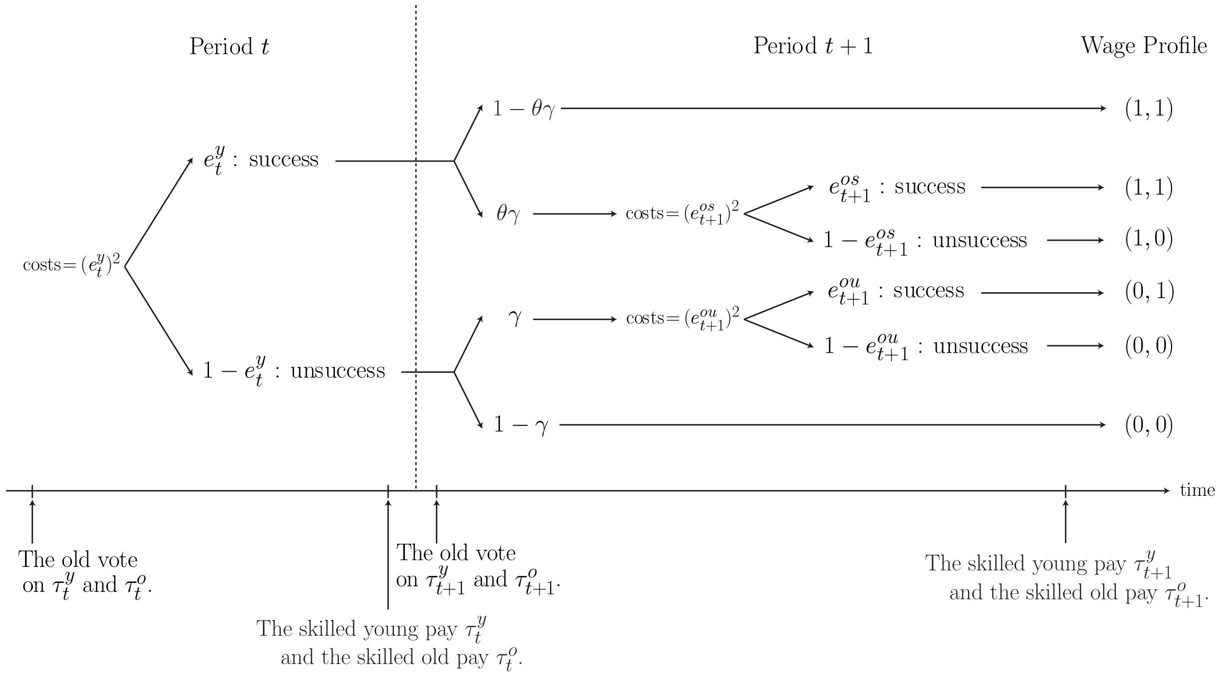


Figure 1: The figure illustrates the timing of events and the distribution of the skilled and the unskilled for generation t.

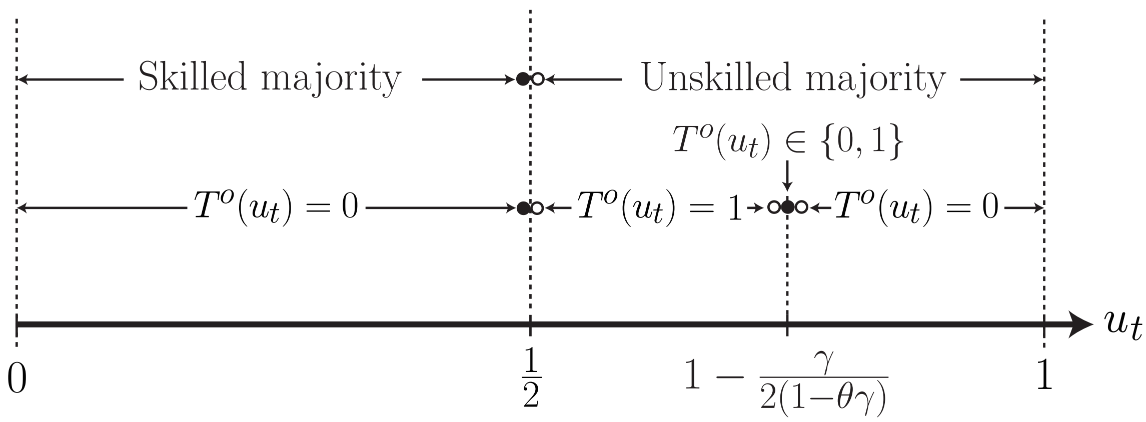


Figure 2: The figure illustrates the determination of τ^o under Assumption 1.

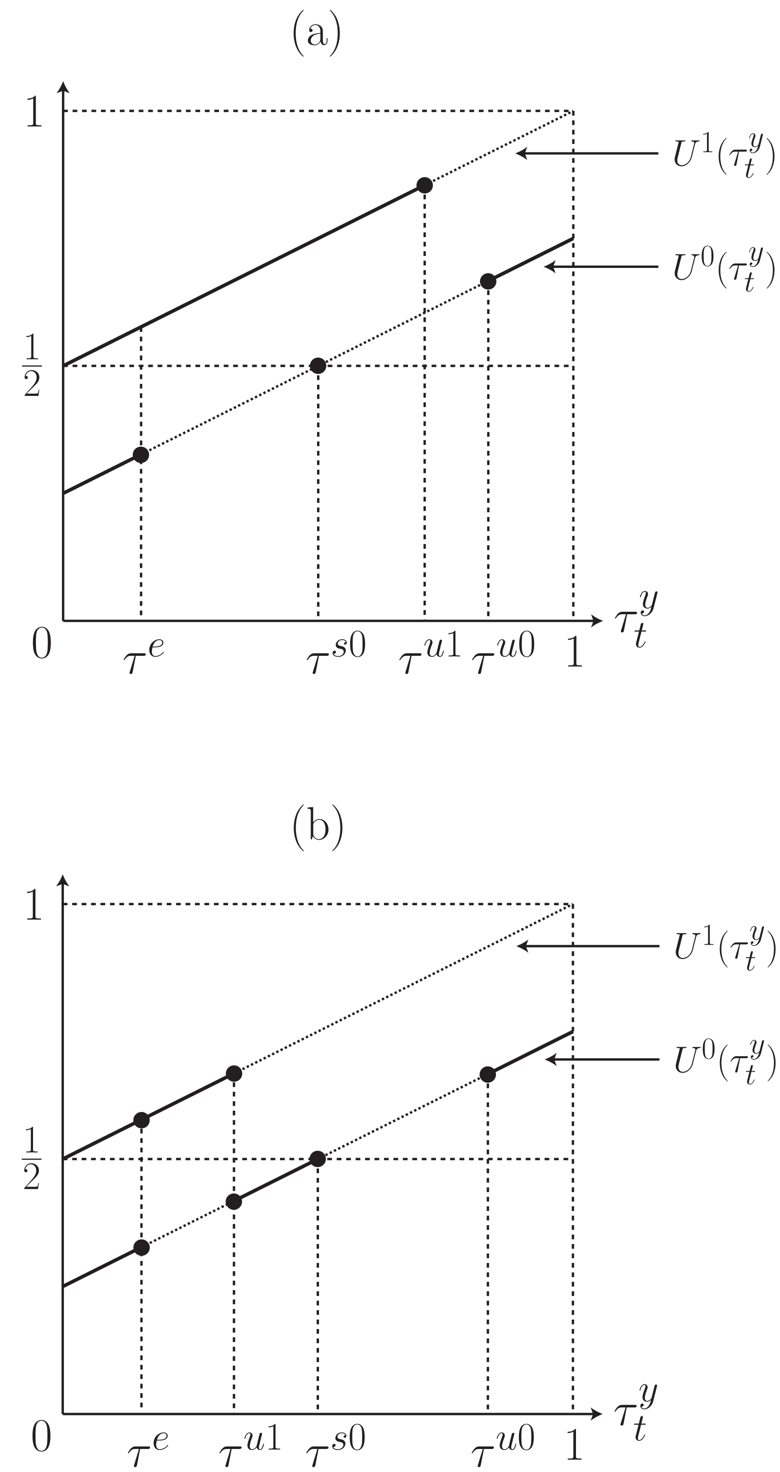
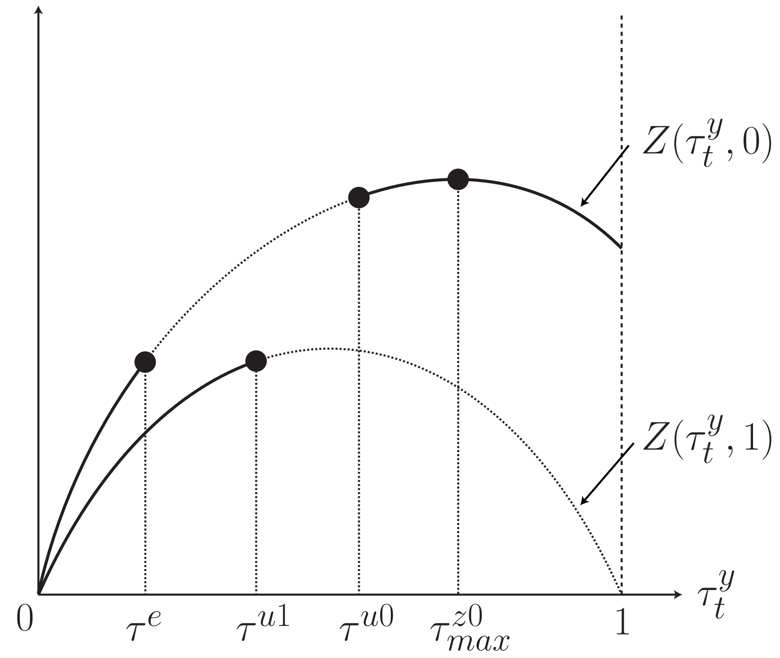


Figure 3: Panel (a) illustrates the case of $\tau^{s0} < \tau^{u1}$: τ^e is the highest tax rate that produces the skilled majority. Panel (b) illustrates the case of $\tau^{s0} \geq \tau^{u1}$: τ^{s0} is the highest tax rate that produces the skilled majority.

(a)



(b)

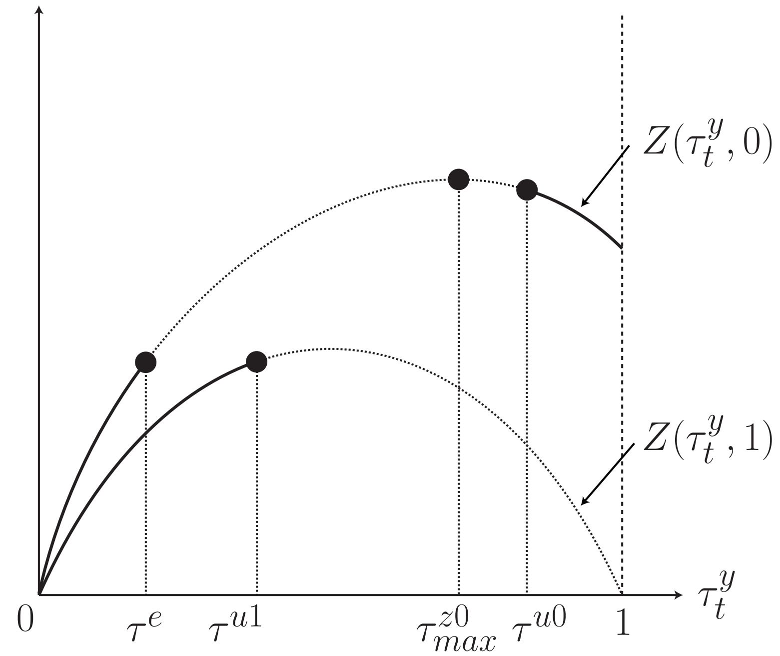


Figure 4: The unskilled-majority equilibrium with no taxation on the old. Panel (a) depicts an interior solution of $\tau = \tau^{z0}_{\max}$; panel (b) depicts a corner solution of $\tau = \tau^{u0}$.

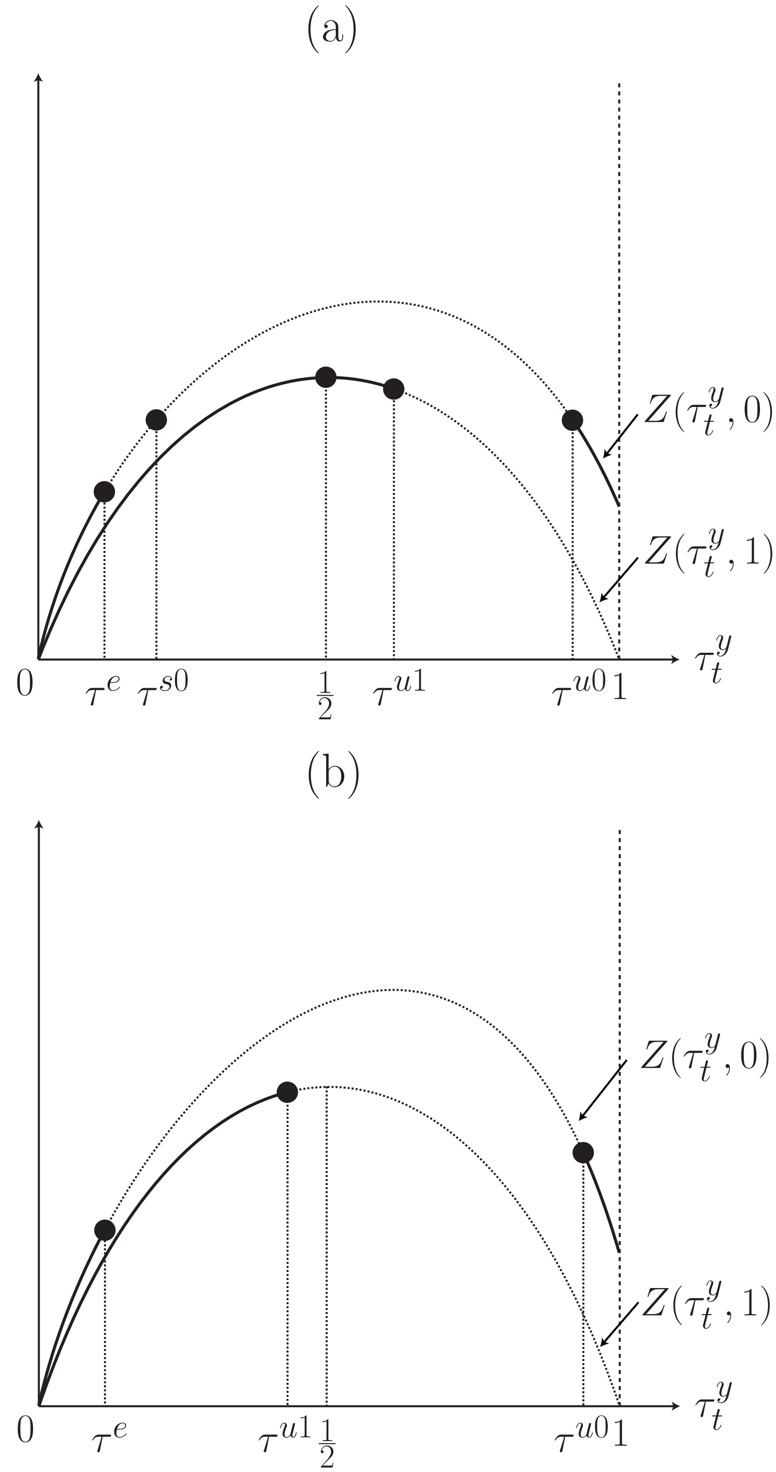


Figure 5: The unskilled-majority equilibrium with taxation on the old. Panel (a) depicts an interior solution of $\tau^y = \arg\max Z(\tau^y, 1) = 1/2$; panel (b) depicts a corner solution of $\tau^y = \tau^{u1}$.

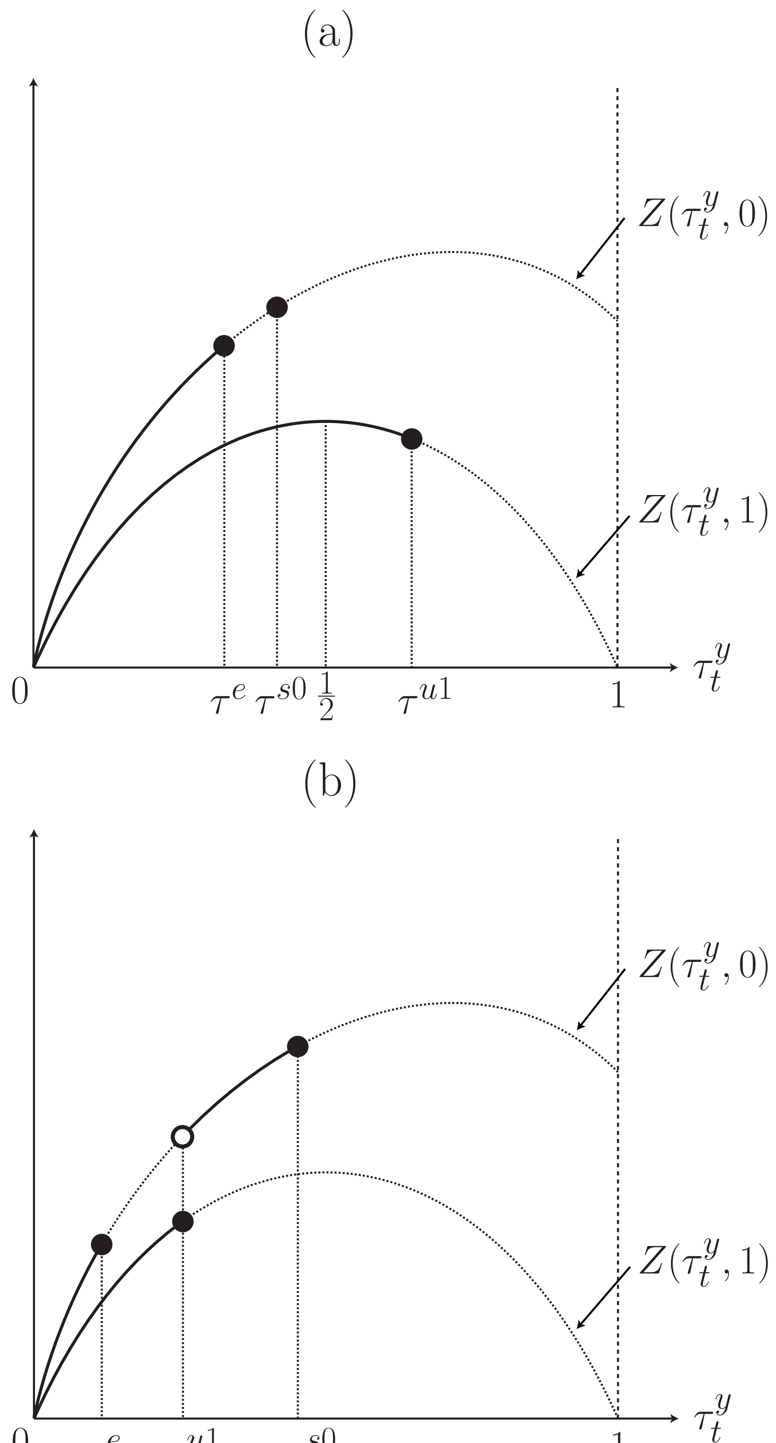


Figure 6: The skilled-majority equilibrium with no taxation on the old. Panel (a) depicts the case where the solution is $\tau^y = \tau^e$; panel (b) depicts the case where the solution is $\tau^y = \tau^{s0}$.

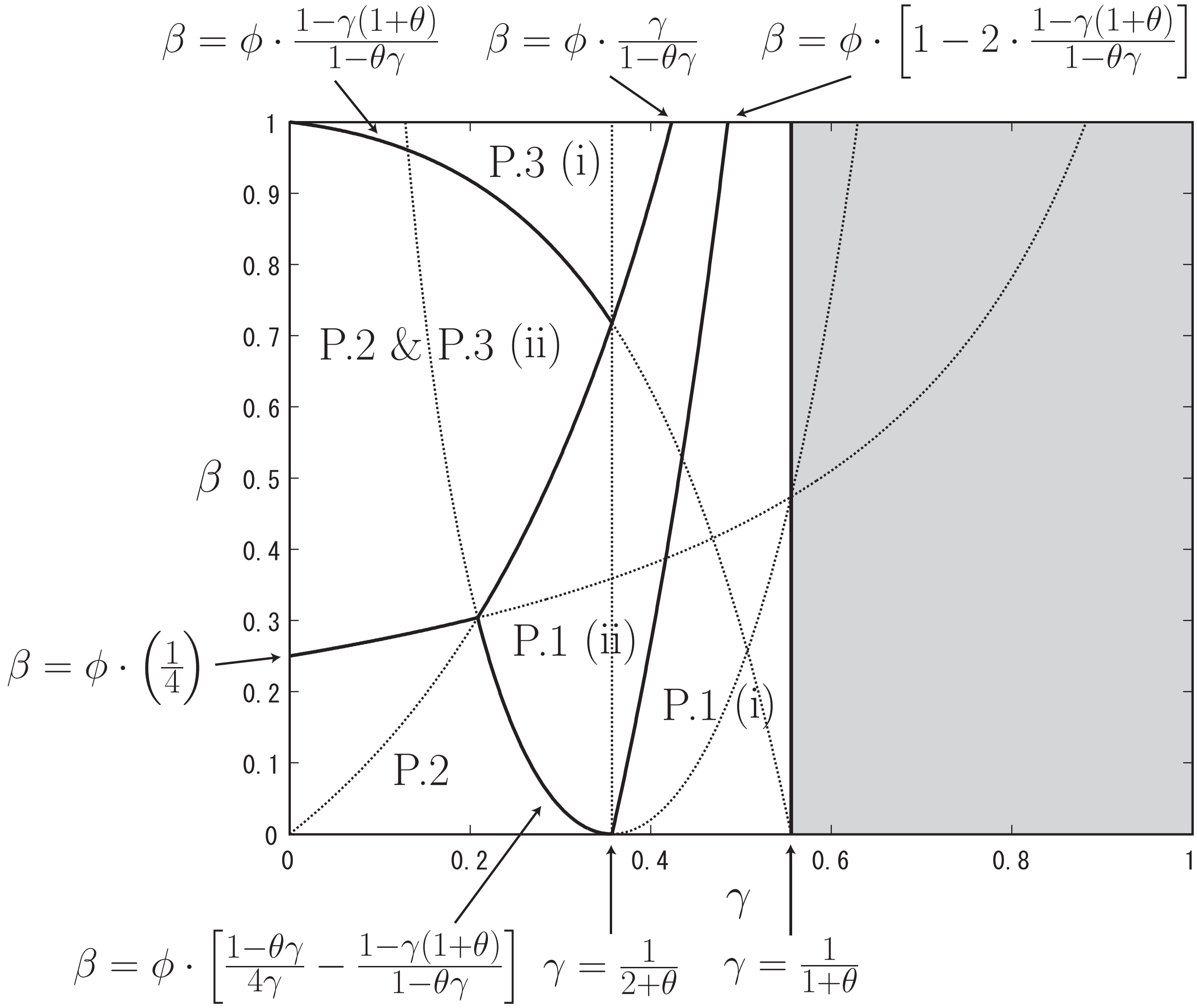


Figure 7: The figure displays the set of parameters (β, γ), where θ is fixed at 0.8, classified according to the characterization of political equilibria.

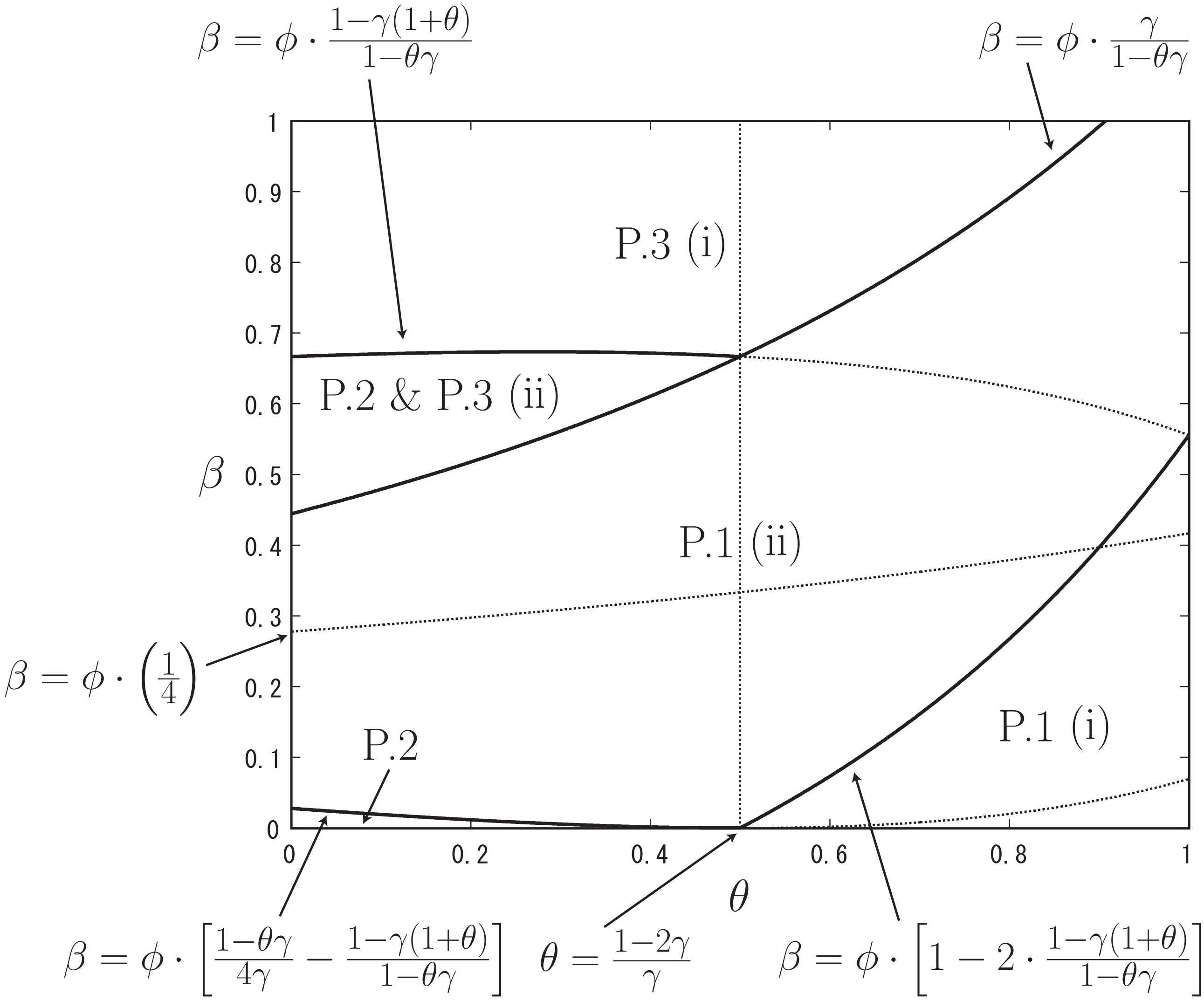


Figure 8: The figure displays the set of parameters (β, θ), where γ is fixed at 0.4, classified according to the characterization of political equilibria.

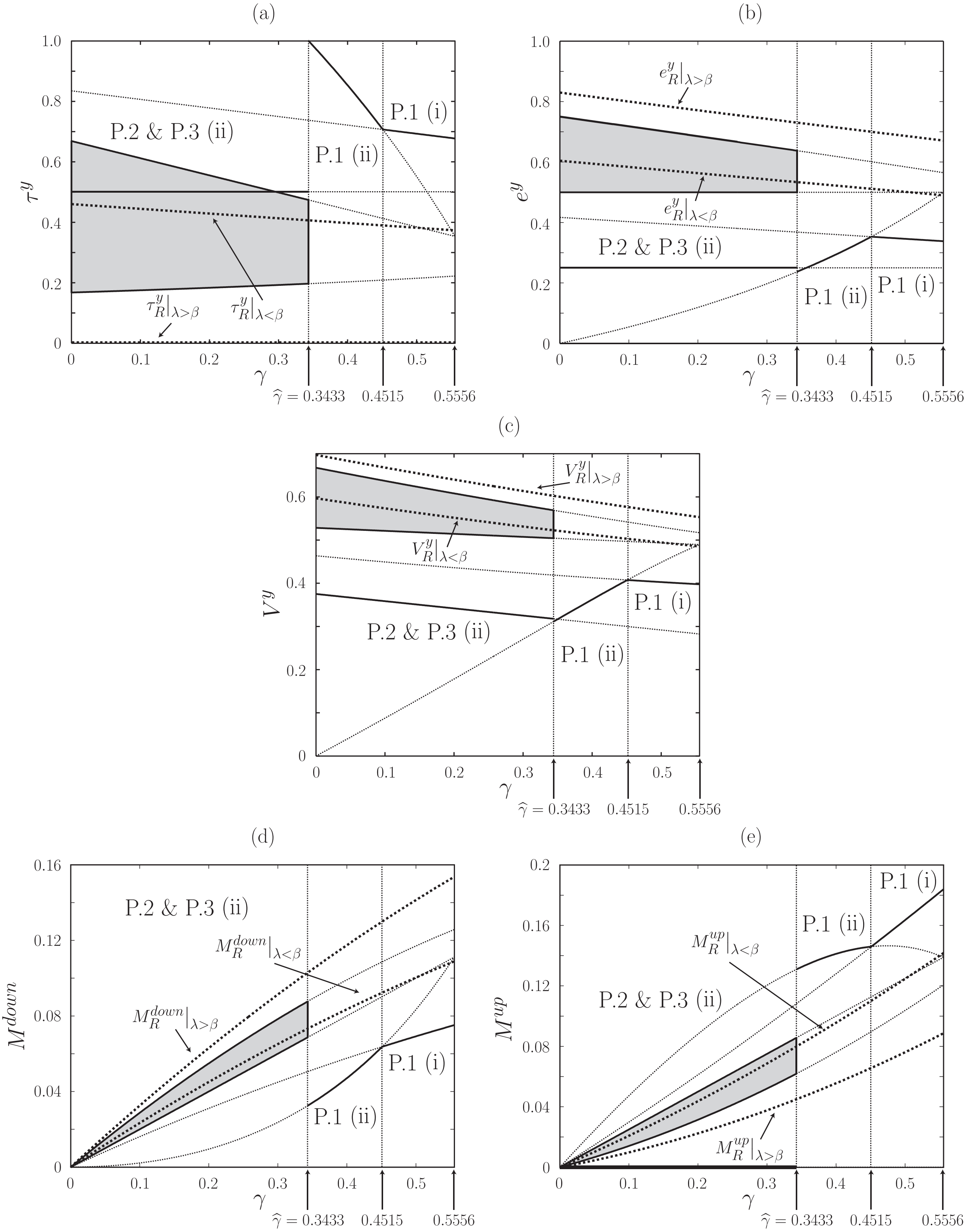


Figure 9: Solid curves and shaded area depict how the parameter γ affects the tax rate on the young (panel (a)), the probability of being skilled in youth (panel (b)), the expected utility of the young (panel (c)), the downward mobility (panel (d)) and the upward mobility (panel (e)) in the political equilibria. Dotted curves depict the corresponding values in the Ramsey allocation investigated in Section 5.

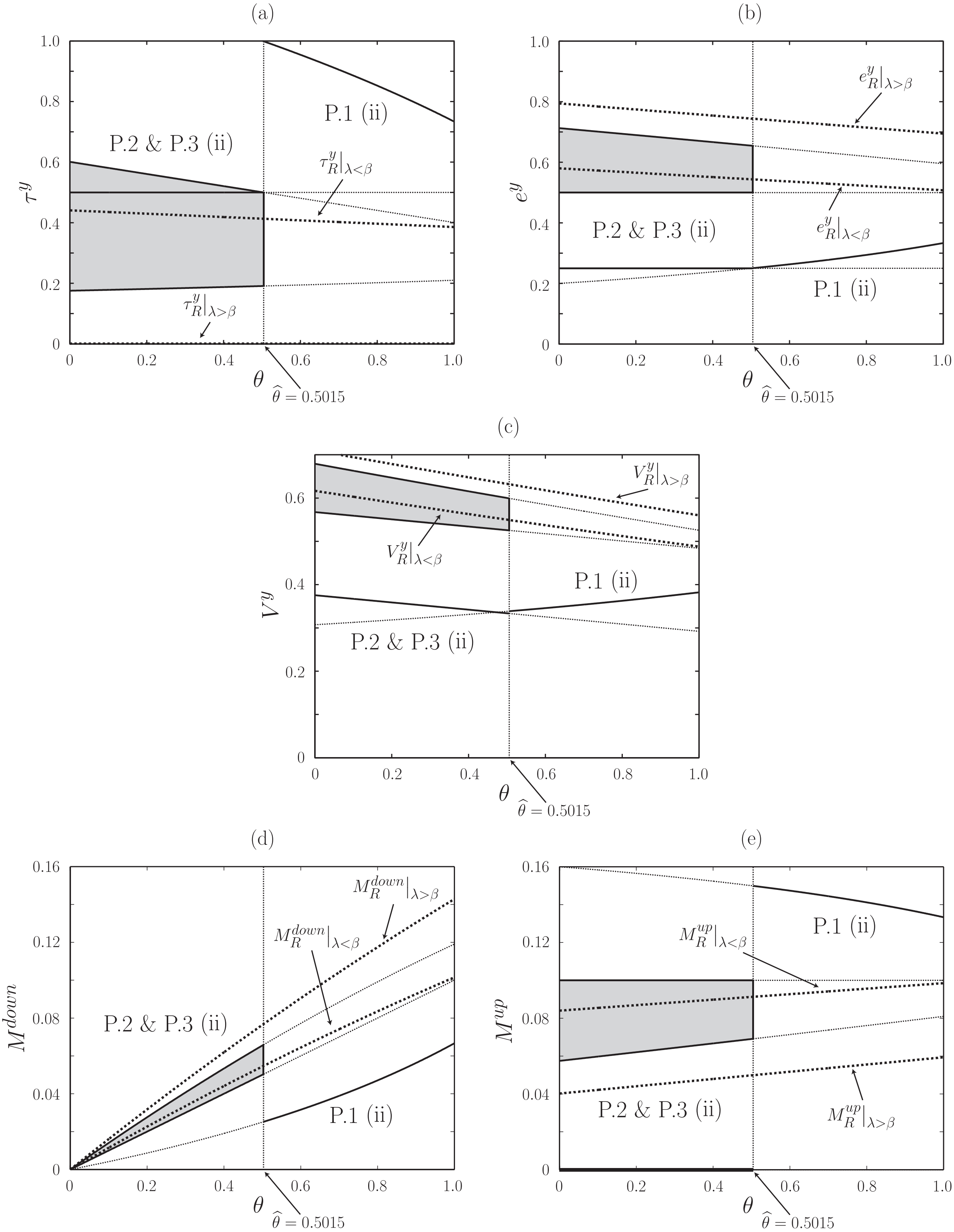


Figure 10: Solid curves and shaded area depict how the parameter θ affects the tax rate on the young (panel (a)), the probability of being skilled in youth (panel (b)), the expected utility of the young (panel (c)), the downward mobility (panel (d)) and the upward mobility (panel (e)) in the political equilibria. Dotted curves depict the corresponding values in the Ramsey allocation investigated in Section 5.