Discussion Papers In Economics And Business

A Political Economy Model of Earnings Mobility and Redistribution Policy

Ryo Arawatari and Tetsuo Ono

Discussion Paper 08-18-Rev.4

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

A Political Economy Model of Earnings Mobility and Redistribution Policy

Ryo Arawatari and Tetsuo Ono

Discussion Paper 08-18-Rev.4

October 2012

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

A Political Economy Model of Earnings Mobility and Redistribution Policy

Ryo Arawatari*
Nagoya University

Tetsuo Ono†
Osaka University

October 9, 2012

Abstract

This paper presents a politico-economic model that includes a mutual link between life cycle earnings mobility and redistributive politics. The model demonstrates that when an economy features a high opportunity of upward mobility and high risk of downward mobility, it attains a unique equilibrium where unskilled, low-income agents support a low redistribution because of the hope of upward mobility in future. In contrast, the economy attains multiple equilibria when mobility opportunity and risk are low: one is an unskilled-majority equilibrium defined by low mobility and the other is a skilled-majority equilibrium defined by high mobility. The paper gives a comparison between the political equilibrium and the social planner's allocation in terms of mobility, and shows that the skilled-majority equilibrium realizes mobility close to the optimal one.

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 during their youth retain their skills over their life cycle without any additional effort, and thus face no risk of downward mobility such as job loss or demotion. In addition, agents who fail in education during their youth have no second opportunity of becoming skilled at a later stage of their life and thus must accept their low-income status throughout their life. 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).[1] 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 both upward mobility chance and 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

[1] 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.

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, 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, 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 the 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 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, representing the US. 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 the future (Benabou and Ok, 2001; Alesina and La Ferrara, 2005).

In contrast, when γ is below the threshold value, the economy attains multiple equilibria, representing some European countries. One equilibrium is an unskilled-majority equilibrium with taxation on the old, representing those Continental European countries which feature high inequality and low mobility; and the other is a skilled-majority equilibrium with no taxation on the old, representing Scandinavian countries defined by low inequality and high mobility. Which equilibrium is realized as an outcome depends on the expectations of the 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 during youth because, for skilled agents, one's status in youth is less likely to persist into old age. Given this feature, a natural prediction is that in an economy with a high prospect of downward mobility the majority will be unskilled agents who support taxation on the skilled old. The former prediction is true, but the latter is not. A higher θ implies a lower 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 an equilibrium supporting the POUM hypothesis. Upward mobility opportunity and downward mobility risk produce qualitatively similar properties with respect to political equilibrium characterization: this is the result which was not shown in Hassler, Storesletten and Zilibotti (2007) and Arawatari and Ono (2009).

Another noteworthy feature of the political equilibrium is that the relative number of upwardly and downwardly mobile agents are completely different between the two types of unskilled-majority equilibrium. The number of upwardly mobile agents is larger than the number of downwardly mobile agents in the unskilled-majority equilibrium with no taxation on the old, representing the US. However, the opposite result holds in the unskilled-majority equilibrium with taxation on the old, representing some Continental European countries. Whether the old are taxed or not critically affects the relative number of upwardly and downwardly mobile agents. The model prediction would provide one possible explanation for the difference in earnings mobility between the US and some Continental European countries.

We consider normative aspects of the political equilibrium with earnings mobility, which was not fully investigated in the previous studies. We characterize a Ramsey allocation defined as a feasible plan chosen by a benevolent planner who can commit to a policy sequence. 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 skilled-majorty equilibrium, representing Scandinavian countries, attain mobility levels close to the optimal ones. We also find that the two types of unskilled-majority equilibrium share similar implications of optimality with respect to downward mobility: they attain lower numbers of downwardly mobile agents compared to those in the Ramsey allocation. However, they show different implications of optimality with respect to upward mobility. The unskilled-majority equilibrium with no taxation on the old, representing the US, shows a higher level of upward mobility whereas the unskilled-majority equilibrium with 100% taxation on the old, representing some Continental European countries, show a lower level of upward mobility than that in the Ramsey allocation.

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, 2012; 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 debt accumulation (Song, Storesletten and Zilibotti, 2012), intergenerational risk sharing (D'Amato and Galasso, 2010) and intergenerational mobility (Arawatari and Ono, 2012). 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 properties. Section 4 characterizes the Ramsey allocation and considers normative aspects of the political equilibrium. Section 5 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.[2] Each generation has a unit mass.

Consider the 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 during 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: skilled and unskilled. Our model differs 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, skilled agents 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 skilled agents is given by $(e_{t+1}^{os})^2$,

[2]In the current framework, the first and second periods of life corresponds to the young and the middle life, respectively. For example, the first-period of life includes ages of 25-44 years and the second-period of life includes ages of 45-64 years. However, we follow the conventional terminology in a two-period overlapping-generation model, and use the terms, youth and old age, throughout the paper.

where e^{os}_{t+1} is the probability of being skilled and $1 - e^{os}_{t+1}$ is the probability of being unskilled. Examples of educational investment in the later stage of life include part-time study at a university, job-training programs and studying for a promotional examination.

With unskilled agents, they may have an opportunity for upward mobility as in Arawatari and Ono (2009). They remain 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 unskilled agents is given by $(e^{ou}_{t+1})^2$, where e^{ou}_{t+1} is the probability of being skilled and $1 - e^{ou}_{t+1}$ 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): unskilled agents may have an opportunity for upward mobility, but skilled agents 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.

Based on the description so far, we can summarize the expected utility functions of agents alive at time t as follows:

$$V^{os}_t = (1 - \theta\gamma)(1 - \tau^o_t) + \theta\gamma\{e^{os}_t(1 - \tau^o_t) - (e^{os}_t)^2\} + s_t, \tag{1}$$

$$V^{ou}_t = \gamma\{e^{ou}_t \cdot (1 - \tau^o_t) - (e^{ou}_t)^2\} + s_t, \tag{2}$$

$$\begin{aligned} V^y_t = {} & e^y_t \cdot (1 - \tau^y_t) - (e^y_t)^2 + s_t \\ & + \beta \cdot \left[e^y_t \left\{(1 - \theta\gamma)(1 - \tau^o_{t+1}) + \theta\gamma \left(e^{os}_{t+1}(1 - \tau^o_{t+1}) - (e^{os}_{t+1})^2\right)\right\}\right. \\ & \left. + (1 - e^y_t)\gamma \left\{e^{ou}_{t+1} \cdot (1 - \tau^o_{t+1}) - (e^{ou}_{t+1})^2\right\} + s_{t+1}\right], \end{aligned} \tag{3}$$

where V^{os}_t, V^{ou}_t and V^y_t 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^{os}_t, V^{ou}_t and V^y_t 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^{os}_t to maximize V^{os}_t; an unskilled old agent chooses e^{ou}_t to maximize V^{ou}_t; and a young agent in period t chooses e^y_t to maximize V^y_t by taking account of the optimal investments in his/her old age, e^{os}_{t+1} and

e^{ou}_{t+1}. Therefore, optimal investments by the old and the young are given by, respectively:

$$e^{os*}(\tau^o_t) = e^{ou*}(\tau^o_t) = \frac{(1-\tau^o_t)}{2}, \tag{4}$$

$$e^{y*}(\tau^y_t, \tau^o_{t+1}) = \frac{1}{2} \cdot \left[(1-\tau^y_t) + \beta \left\{ (1-\theta\gamma)(1-\tau^o_{t+1}) - \frac{\gamma(1-\theta)}{4}(1-\tau^o_{t+1})^2 \right\} \right], \tag{5}$$

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

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 equal to the probability of being failed in education in youth:

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

The proportion, u_{t+1}, depends on the tax levied on the skilled young agents in period t, τ^y_t, and the tax levied on the skilled old agents in period $t+1$, τ^o_{t+1}.

In this economy, there is earnings mobility over the life cycle. Let M^{up}_{t+1} 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^{down}_{t+1} 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^{up}_{t+1} and M^{down}_{t+1} are calculated as, respectively:

$$M^{up}_{t+1} = (1-e^y_t)\gamma e^{ou}_{t+1}, \; M^{down}_{t+1} = e^y_t \theta\gamma(1-e^{os}_{t+1}).$$

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:

$$2s_t = W(\tau^o_t, u_t) + Z(\tau^y_t, \tau^o_{t+1}),$$

where:

$$W(\tau^o_t, u_t) \equiv [(1-u_t)\{(1-\theta\gamma) + \theta\gamma e^{os*}(\tau^o_t)\} + u_t\gamma e^{ou*}(\tau^o_t)] \cdot \tau^o_t;$$
$$Z(\tau^y_t, \tau^o_{t+1}) \equiv e^{y*}(\tau^y_t, \tau^o_{t+1}) \cdot \tau^y_t.$$

The left-hand side, denoted by $2s_t$, represents aggregate transfers to the young and the old. On the right-hand side, the first term, denoted by $W(\tau^o_t, u_t)$, is the tax revenue financed by the skilled old agents; and the second term, denoted by $Z(\tau^y_t, \tau^o_{t+1})$, 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. Section 3.4 demonstrates the mobility in the political equilibria.

3.1 Definition of Political Equilibrium

Following Hassler et al. (2003) and Hassler, Storesletten and Zilibotti (2007), we assume that agents vote over current taxes at the beginning of each period but that only the old vote. Under this assumption, we exclusively focus on the political conflict between the rich and the poor within a generation; the conflict between the young and the old is abstracted away from the analysis.

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:

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

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 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. $T^o(u_t) = \begin{cases} \arg\max_{\tau_t^o\in[0,1]} V_t^{os} & \text{if } u_t \le 1/2, \\ \arg\max_{\tau_t^o\in[0,1]} V_t^{ou} & \text{if } u_t > 1/2, \end{cases}$

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 V_t^{os}. This function has the following properties: $\partial V_t^{os}(\tau_t^o, u_t)/\partial \tau_t^o|_{\tau_t^o=0} \leq 0$ and $\partial^2 V_t^{os}(\tau_t^o, u_t)/\partial \tau_t^{o2} < 0$.[3] These properties imply that V_t^os is maximized at $\tau_t^o = 0$: the skilled old pay more than they receive because unskilled agents pay no tax even though the revenue is distributed equally between skilled and unskilled agents. Therefore, the skilled old prefer $\tau_t^o = 0$, implying that $T^o(u_t) = 0$ if the majority are skilled agents:

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

Alternatively, suppose that the unskilled old are in the majority: $u_t > 1/2$. The objective function of the majority is given by $V_t^{ou}(\tau_t^o, u_t)$. The second derivative of this function with respect to τ_t^o is:

$$\frac{\partial^2 V_t^{ou}(\tau_t^o, u_t)}{\partial \tau_t^{o2}} = \frac{\gamma}{2}(1-\theta)(1-u_t) > 0,$$

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

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

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

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

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

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

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

Condition (7) means that if the number of unskilled agents is larger/smaller than the threshold level, $1-\gamma/\{2(1-\theta\gamma)\}$, 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 $u_t > (<) 1-\gamma/\{2(1-\theta\gamma)\}$. If $u_t = 1-\gamma/\{2(1-\theta\gamma)\}$, 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.

Given (6) and (7), we can summarize the mapping satisfying equilibrium condition 1 as follows.

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

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

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

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

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

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

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

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:

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

where:

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

The 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 $u_{t+1}=1-e^{y*}(\tau_t^y,0)\leq 1/2$ or $1-\gamma/\left\{2(1-\theta\gamma)\right\}\leq u_{t+1}=1-e^{y*}(\tau_t^y,0)\leq 1$, that is, if $\tau_t^y\leq\tau^{s0}$ or $\tau^{u0}\leq\tau_t^y$. The former condition implies that a low tax burden on the young does not dampen their motivation of educational investment, thereby resulting in a skilled majority who support no taxation on the old. The latter condition implies that a high tax burden on the young gives them a disincentive for educational investment, which results in an unskilled majority. While their income status is low, the unskilled support no taxation on the old because of the hope of upward mobility in their old age.

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 $1/2<u_{t+1}=1-e^{y*}(\tau_t^y,1)\leq 1-\gamma/\left\{2(1-\theta\gamma)\right\}$, that is, if $\tau_t^y\leq\tau^{u1}$. The condition implies that the tax burden is low for the young. However, the low tax burden does not stimulate an incentive in unskilled agents towards educational investment due to the expectation of high tax burden in their old age. Therefore, the majority become the unskilled who support 100% taxation on the skilled old.

As illustrated in Figure 2, there are multiple, self-fulfilling expectations of U for the set of $\tau_t^y\leq\min\left\{\tau^{s0},\tau^{u1}\right\}$. 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 $\tau_t^y:\tau^e\leq\min\left\{\tau^{s0},\tau^{u1}\right\}$. 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 are the unskilled old. However, for $\tau_t^y\leq\tau^e$, the majority are either skilled or unskilled depending on agents' expectations. Panel (a) in Figure 2 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.

[Figure 2 about here.]

Given the definition of τ^e, the solution is given by:

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

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 are unskilled, there are two types 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 are 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 skilled majority at time $t+1$ ($u_{t+1} \le 1/2$).

Before proceeding to the analysis, we note the following properties of the tax revenue function $Z(\cdot,\cdot)$ 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

$$\tau_t^y = \frac{1}{2} \cdot [1 + \beta\{(1-\theta\gamma) - \gamma(1-\theta)/4\}] \text{ and } \tau_t^y = \frac{1}{2},$$

respectively. Given these properties with equilibrium conditions 1 and 2, revenue from the young can be illustrated as in Figure 3.

Panel (a) in Figure 3 illustrates a case that produces the unskilled-majority equilibrium with no taxation on the old: that is, the tax revenue from the young is maximized under the expectation of $\tau_{t+1}^o = 0$. There are two possible cases: an interior solution where the revenue from the young is maximized by setting the tax rate, $\tau_t^y = \arg\max Z(\tau_t^y, 0)$, that attains the top of the Laffer curve; and a corner solution case where the revenue from the young might be maximized at $\tau_t^y = \tau^{u0}$. Panel (a) illustrates the latter case.

[Figure 3 about here.]

Panel (b) in Figure 3 also illustrates a case that produces the unskilled-majority equilibrium. However, this case differs from that illustrated in Panel (a) 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 (b) 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 $Z(\tau^y, 1)$: $\tau_t^y = \arg\max Z(\tau_t^y, 1) = 1/2$.

Finally, Panel (c) in Figure 3 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 (c) illustrates the case where the revenue from the young is maximized at $\tau = \tau^e$. However, when τ^e is set to be low, the unskilled-majority equilibrium is realized as demonstrated in Panel (b). Therefore, there are multiple equilibria depending on the expectations of agents.

Based on the abovementioned argument, we can derive the condition for the existence of each type of equilibrium. However, a complete characterization of the political equilibria requires a lot of space for presentation. In order to save space and to simplify the presentation, we demonstrate the numerical results for the equilibria, leaving a full characterization of the equilibria in Appendix B. Figure 4 illustrates the political equilibria where θ is fixed at 0.8 and γ and β vary between the ranges of $(0, 0.5555)$ and $(0, 1)$, respectively.[4] Figure 5 illustrates the political equilibria where γ is fixed at 0.4 and both θ and β vary between the range of $(0, 1)$.

[Figures 4 and 5 about here.]

First, let us consider the effect of the parameter γ by utilizing Figure 4. For example, let $\beta = 0.5438$.[5] There is a threshold value of γ, $\hat{\gamma} = 0.3043$, such that the equilibrium is characterized by the unskilled majority who support no taxation on the old when γ is above the threshold value. A higher γ 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

[4] The upper bound of γ in Figure 4 is given by 0.5555 under Assumption 1: $\gamma < 1/(1+\theta) = 1/(1+0.8) \approx 0.5555$.

[5] The selection of β is as follows. We assume a generation to be 20 years in length. The first and the second period correspond to, for example, ages 25-44 and 45-64 years, respectively. Our selection of one-period discount factor is 0.97. Because the agents under the current assumption plan over generations that span 20 years, we discount the future by $(0.97)^{20} \approx 0.5438$.

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 chances 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 become unskilled when agents attach a low value to old-age utility, whereas they become 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 5 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 5, given $\beta = 0.5438$, 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. 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.

3.4 Mobility in the Political Equilibria

By using the numerical result demonstrated in Section 3.3, we compute the numbers of upwardly and downwardly mobile agents. The solid curves and shaded areas in Figures 6 and 7 depict how the numbers of upwardly and downwardly mobile agents are affected by the parameters γ and θ, respectively.[6] We should note that the numbers of mobile agents in the skilled-majority equilibria are illustrated by the shaded area in Figures 6 and 7 because the size of redistribution depends on the agents' expectations which take continuum values.

[Figures 6 and 7 about here.]

From the figures, we can make the following observation which holds in general for each type of equilibria (for formal proof, see Appendix B). First, consider the unskilled-majority equilibrium with no taxation on the old. The number of mobile agents are plotted by the solid curve for the range of $\gamma \in [0.3043, 1)$ in Figure 6 and by that for the range of $\theta \in [0.2613, 1)$ in Figure 7. In this equilibrium, the number of upwardly mobile agents is larger than the number of downwardly mobile agents.

[6] Dotted curves in the figures depict the corresponding values in the Ramsey allocation investigated in the next section.

The mechanism behind the abovementioned 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 $(1-u)\theta\gamma/2$.

Next, consider the unskilled-majority equilibrium with taxation on the old. The mobility in this equilibrium is depicted for the range of $\gamma \in (0, 0.3043)$ in Figure 6 and for the range of $\theta \in (0, 0.2613)$ in Figure 7. In this equilibrium, there is no upwardly mobile agent; however, there are some downwardly mobile agents as illustrated by the inferior of the shaded area in Panel (a). This result is qualitatively opposite to that in the unskilled-majority equilibrium with no taxation on the old. The difference in mobility between the two types of unskilled-majorty equilibrium comes from whether the old are taxed or not.

Finally, consider the skilled-majority equilibrium where the number of mobile agents are illustrated by the shaded area in Figures 6 and 7. In this 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, $\theta\gamma e^y$; that is, it depends on the expectations of agents represented by the parameter τ^e.

4 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 4.1). The Ramsey allocation derived here will be compared with the political equilibria in order to consider the normative aspect of the politics (Subsection 4.2).

4.1 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 Appendix A 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 1: *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:

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

Proof. See Appendix A.

Proposition 1 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.

Given the result in Proposition 1, we can calculate mobility in the Ramsey allocation. In particular, M^{up} and M^{down} in period $t \geq 1$ are given by:

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

Dotted curves in Figures 6 and 7 depict how M^{up} and M^{down} in the Ramsey allocation are affected by the parameters γ and θ, respectively.

4.2 Normative Aspect of Political Equilibria

Let us now discuss the normative aspect of the political equilibria in terms of mobility and redistribution, based on the observation in Figures 6 and 7. First, consider the unskilled-majority equilibrium with no taxation on the old, where the POUM hypothesis is supported. From the viewpoint of redistribution, this equilibrium bears some resemblance to the Ramsey allocation in the case of $\lambda < \beta$, where the tax burden falls on the young. However, there is a remarkable difference in mobility between the political equilibrium and the Ramsey allocation: the Ramsey allocation requires a larger number of downwardly mobile agents and a smaller number of upwardly mobile agents compared to those in the corresponding political equilibrium. In other words, the political equilibrium realizes excessively high upward mobility and low downward mobility from the viewpoint of optimality.

The abovementioned result comes from the fact that the tax burden on the young is lower in the Ramsey allocation than in the political equilibria. A lower tax burden in the Ramsey allocation results in a larger number of agents being skilled in youth. This implies a larger number of potential agents who experience downward mobility in old age; and it also implies a smaller number of potential agents who can get chances for upward mobility in old age. Given these factors, the Rasemy allocation requires higher downward mobility and lower upward mobility compared to those in the corresponding political equilibrium.

Next, consider the unskilled-majority equilibrium with taxation on the old. >From the viewpoint of redistribution, 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.

There is also a difference in mobility. The downward mobility is lower in the political equilibrium than in the Ramsey allocation. In addition, there are some agents who can move up the income ladder in the Rasemy allocation, while there are no such agents in the political economy because of 100% taxation on the old. Therefore, the unskilled-majority equilibrium with 100% taxation on the old attains excessively low downward and upward mobility from the viewpoint of optimality. In summary, the two types of unskilled-majority equilibrium analyzed so far share similar implications of optimality with respect to downward mobility, but have different implications with respect to upward mobility.

Finally, consider the skilled-majority equilibrium. This equilibrium bears some resemblance to the Ramsey tax in terms of redistribution 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. In the situation of multiple equilibria, the political economy may or may not attain an allocation similar to the Ramsey allocation depending on expectations of agents and the planner's weight to the young. In addition, the quantitative 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.

5 Conclusion

This paper presents a simple theoretical model that includes life cycle 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 equilibrium. When the opportunity and risk are high, the economy is more likely to attain a unique, unskilled-majority equilibrium representing the US, where the POUM hypothesis is supported. However, when the opportunity and risk are low, the economy attains multiple equilibria: an unskilled-majority equilibrium that features low earnings mobility, representing some Continental European countries, and a skilled-majority equilibrium that features high earnings mobility, representing Scandinavian countries. Which equilibrium is realized as an outcome depends on the expectations of the agents.

Second, we characterize the Ramsey allocation that defines an optimal allocation, and compare the political equilibrium with the Ramsey allocation to evaluate the optimality of the political equilibrium in terms of earnings mobility. We find that the skilled-majority equilibrium, representing Scandinavian countries, attains close-to-optimal mobility levels.

We also find that two types of unskilled-majority equilibrium attain lower downward mobility from the viewpoint of optimality, but show different implications of optimality with respect to upward mobility. In particular, the equilibrium representing the US shows higher upward mobility while the equilibrium representing some Continental European countries shows lower upward mobility compared to the optimal one.

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.

6 Appendix A

6.1 Proof of Proposition 1

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:

$$
\begin{aligned}
\Omega = \beta(1-u_0) \cdot & \left[\underbrace{(1-\theta\gamma)(1-\tau_0^o) + \frac{\theta\gamma}{4}(1-\tau_0^o)^2}_{\text{(a.1)}} + \underbrace{\frac{1}{2}W(\tau_0^o, u_0)}_{\text{(a.2)}} + \underbrace{\frac{1}{2}Z(\tau_0^y, \tau_1^o)}_{\text{(b.1)}} \right] \\
+ \beta u_0 \cdot & \left[\underbrace{\frac{\gamma}{4}(1-\tau_0^o)^2}_{\text{(a.3)}} + \underbrace{\frac{1}{2}W(\tau_0^o, u_0)}_{\text{(a.4)}} + \underbrace{\frac{1}{2}Z(\tau_0^y, \tau_1^o)}_{\text{(b.2)}} \right] \\
+ \lambda \cdot & \left[\underbrace{e^{y*}(\tau_0^y, \tau_1^o) \cdot (1-\tau_0^y) - (e^{y*}(\tau_0^y, \tau_1^o))^2}_{\text{(b.3)}} + \underbrace{\frac{1}{2}W(\tau_0^o, u_0)}_{\text{(a.5)}} + \underbrace{\frac{1}{2}Z(\tau_0^y, \tau_1^o)}_{\text{(b.4)}} \right. \\
+\beta \cdot & \left\{ \underbrace{e^{y*}(\tau_0^y, \tau_1^o) \left\{ (1-\theta\gamma)(1-\tau_1^o) + \frac{\theta\gamma}{4}(1-\tau_1^o)^2 \right\} + (1 - e^{y*}(\tau_0^y, \tau_1^o)) \frac{\gamma}{4}(1-\tau_1^o)^2}_{\text{(b.5)}} \right. \\
& \left. \left. + \underbrace{\frac{1}{2}W(\tau_1^o, u_1)}_{\text{(b.6)}} + \underbrace{\frac{1}{2}Z(\tau_1^y, \tau_2^o)}_{\text{(c.1)}} \right\} \right]
\end{aligned}
$$

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

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

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

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

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

$$\left. \left. + \frac{1}{2}W(\tau_3^o, u_3) + \frac{1}{2}Z(\tau_3^y, \tau_4^o) \right\} \right] + \sum_{t=3}^{\infty} \lambda^{t+1} V_t^y$$

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 1

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

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

Therefore, the solution is given by:[7]

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

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 obtain:

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

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 1

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

$$\tau^y : \partial L/\partial \tau^y - \xi_1^y + \xi_0^y = 0, \tag{16}$$
$$\tau^o : \partial L/\partial \tau^o - \xi_1^o + \xi_0^o = 0, \tag{17}$$

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

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

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

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

$$
\begin{aligned}
\tau^y :& \frac{1}{2}(\beta+\lambda)e^{y*}(\tau^y,\tau^o) - \frac{1}{4}(\beta+\lambda)\tau^y + \frac{1}{4}(\beta+\lambda)\lambda\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}\tau^o \\
&= \xi_1^y - \xi_0^y, \qquad (18)\\
\tau^o :& \frac{1}{2}(\beta+\lambda)\left[\frac{1}{2}\beta\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\tau^o)(1-\theta)\right\}\tau^y\right.\\
&+\lambda\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-2\tau^o)\right\}e^{y*}(\tau^y,\tau^o)\\
&\left.+\frac{\lambda\gamma}{2}(1-2\tau^o) - \frac{\lambda\beta}{2}\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}^2\tau^o\right]\\
&+\lambda\left[e^{y*}(\tau^y,\tau^o)\beta\left\{(-1)(1-\theta\gamma) + \frac{\gamma}{2}(1-\tau^o)(1-\theta)\right\} - \frac{\beta\gamma}{2}(1-\tau^o)\right]\\
&= \xi_1^o - \xi_0^o. \qquad (19)
\end{aligned}
$$

First, assume that $\tau^o = \xi_1^y = \xi_1^o = \xi_0^y = 0$. Then, from (18) and (19), 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 (18) and (19), we obtain:

$$
\begin{aligned}
\tau^o &= \frac{(\lambda-\beta)\left[\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}e^{y*}(0,\tau^o) + \frac{\gamma}{2}(1-\tau^o)\right]}{\frac{1}{2}(\beta+\lambda)\left[\gamma - \gamma(1-\theta)e^{y*}(0,\tau^o) + \beta\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}^2\right]}, \qquad (20)\\
\xi_0^y &= \frac{1}{4}(\lambda-\beta)\left[1+\beta\left\{(1-\theta\gamma)(1-\tau^o) - \frac{\gamma}{4}(1-\theta)(1-\tau^o)^2\right\}\right]\\
&\quad + \frac{1}{4}(\beta+\lambda)\lambda\left\{(1-\theta\gamma) - \frac{\gamma}{2}(1-\theta)(1-\tau^o)\right\}\tau^o,
\end{aligned}
$$

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

The left-hand and right-hand sides of (20), 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 (20), 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:

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

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 Appendix B (Not for Publication)

In this appendix, we provide a formal characterization of the political equilibria demonstrated in Section 3. For notational convenience, we define:

$$\phi \equiv \frac{1}{(1-\theta\gamma) - \frac{\gamma}{4}(1-\theta)};$$
$$\arg\max Z(\tau_t^y, 0) = \tau_{\max}^{Z0} \equiv \frac{1}{2} \cdot \left[1 + \beta \left\{(1-\theta\gamma) - \gamma(1-\theta)/4\right\}\right].$$

7.1 Unskilled-majority Equilibrium with No Taxation on the Old

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

Proposition B1:

(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:*

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

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 $M^{up} > M^{down}$.

Proof.

(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 $\tau^o_{t+1} = 0$, 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 $\tau^y_t = \tau^{Z0}_{\max}$.

To establish that setting $\tau^y_t = \tau^{Z0}_{\max}$ is optimal for the decisive voter, we note the following properties of the function Z : (a) $Z(\tau^y_t, 0)$ is concave in τ^y_t and is maximized at $\tau^y_t = \arg\max_{\tau^y_t \in [0,1]} Z(\tau^y_t, 0) = \tau^{Z0}_{\max}$; and (b) $Z(\tau^y_t, 0) > Z(\tau^y_t, 1) \forall \tau^y_t \in (0,1]$. These properties imply that setting $\tau^y_t = \tau^{Z0}_{\max}$ is optimal if this setting is feasible under the expectation of $\tau^o_{t+1} = 0$, i.e., if $\tau^{u0} \leq \tau^{Z0}_{\max}$. 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 B1.

Given $\tau^y = \tau^{Z0}_{\max}$ 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 $\tau^y_t = \tau^{u0}$ and expect $\tau^o_{t+1} = 0$. 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 $\tau^o_{t+1} = 0$, 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 $\tau^y_t = \tau^{u0}$.

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

$$\beta \leq \phi \cdot \frac{\gamma}{1 - \theta\gamma}. \tag{21}$$

This is the second assumption given in Proposition B1(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:

$$\phi \cdot \left[1 - 2\frac{1-\gamma(1+\theta)}{1-\theta\gamma}\right] < \beta. \tag{22}$$

This is the first assumption given in Proposition B1(ii). If (22) 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 (21) and (22), the revenue from the young is illustrated in panel (a) of Figure 3. 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 B1(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 B1(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 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 B2:

Suppose that the following conditions hold:

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

and:

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

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 $M^{up} = 0 < M^{down}$.

Proof.

The first assumption in the statement of Proposition B2, $\gamma \leq 1/(2+\theta)$, 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 B2 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 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 B3:

(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 $\tau^e \gtreqless (1-\theta)/(1+\theta) + \tau^{s0}$.

Proof.

(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 B3,

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

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 B3,

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

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 B1, 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 B3 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 B1, 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 B3 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 $\tau^e \gtreqless (1-\theta)/(1+\theta) + \tau^{s0}$. ■

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] Arawatari, R., and Ono, T., 2012. Inequality, mobility and redistributive politics. *Journal of Economic Theory*, forthcoming.

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

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

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

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

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

[10] 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

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

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

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

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

[15] 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

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

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

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

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

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

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

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

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

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

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

[26] 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

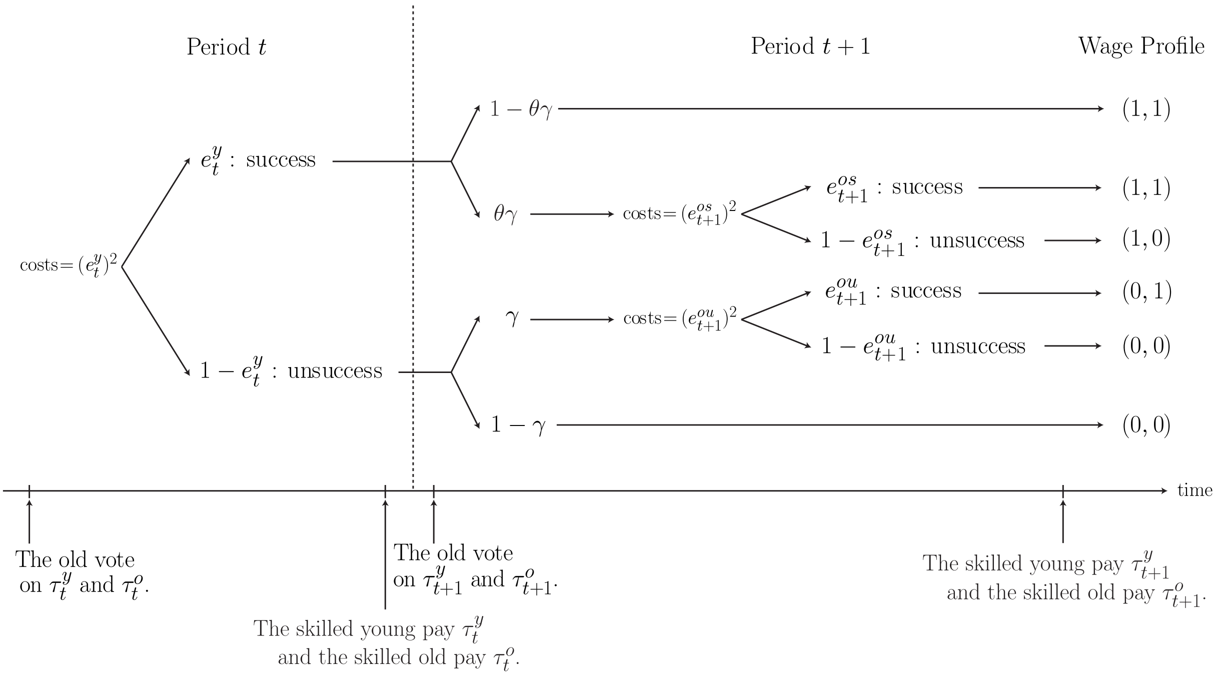

Figure 1: The figure illustrates the timing of events and the distribution of the skilled and the unskilled for generation t.

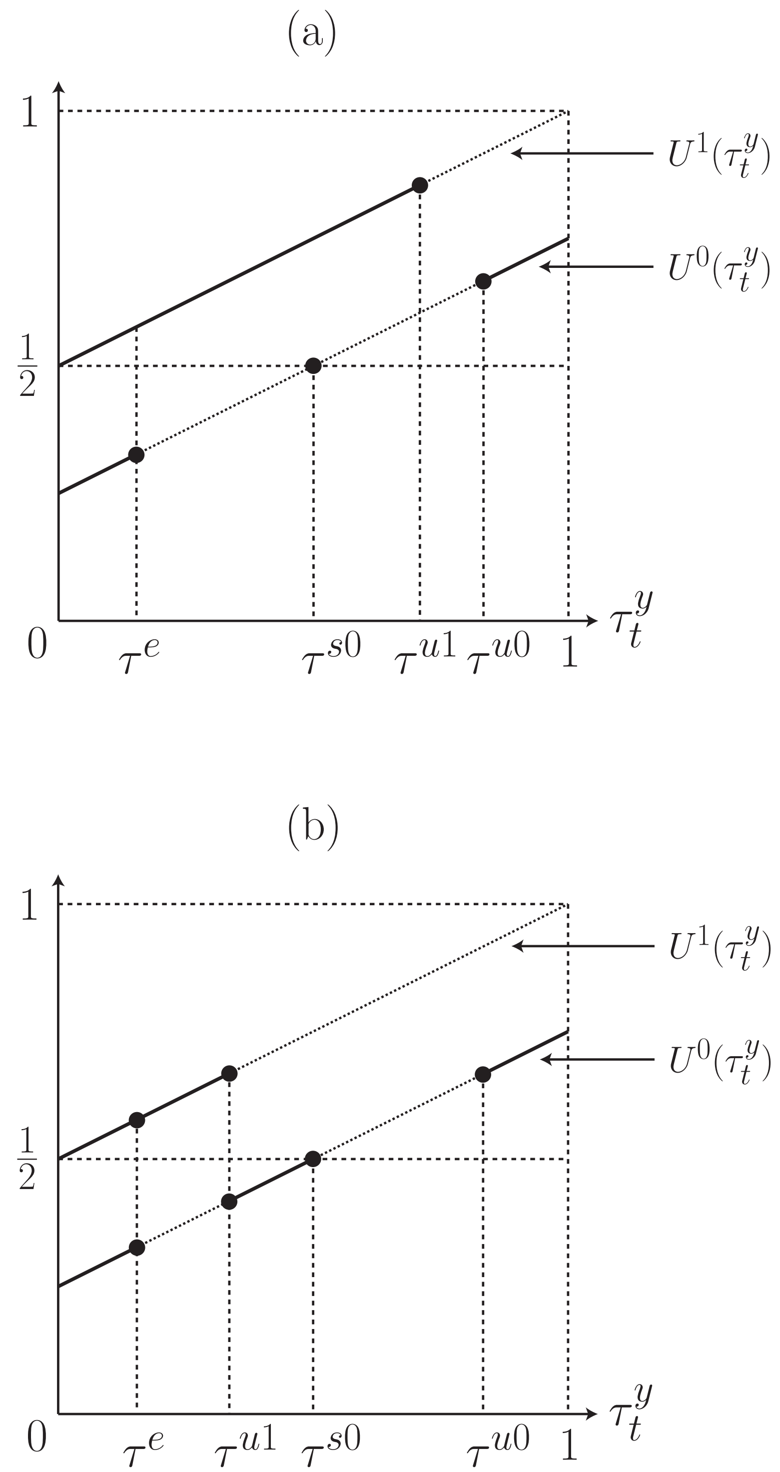

Figure 2: Panel (a) illustrates the case of $\tau^{s0} < \tau^{u1}$: τ^e is the highest tax rate that produces the skilled majority. Panel (b) illustrates the case of $\tau^{s0} \geq \tau^{u1}$: τ^{s0} is the highest tax rate that produces the skilled majority.

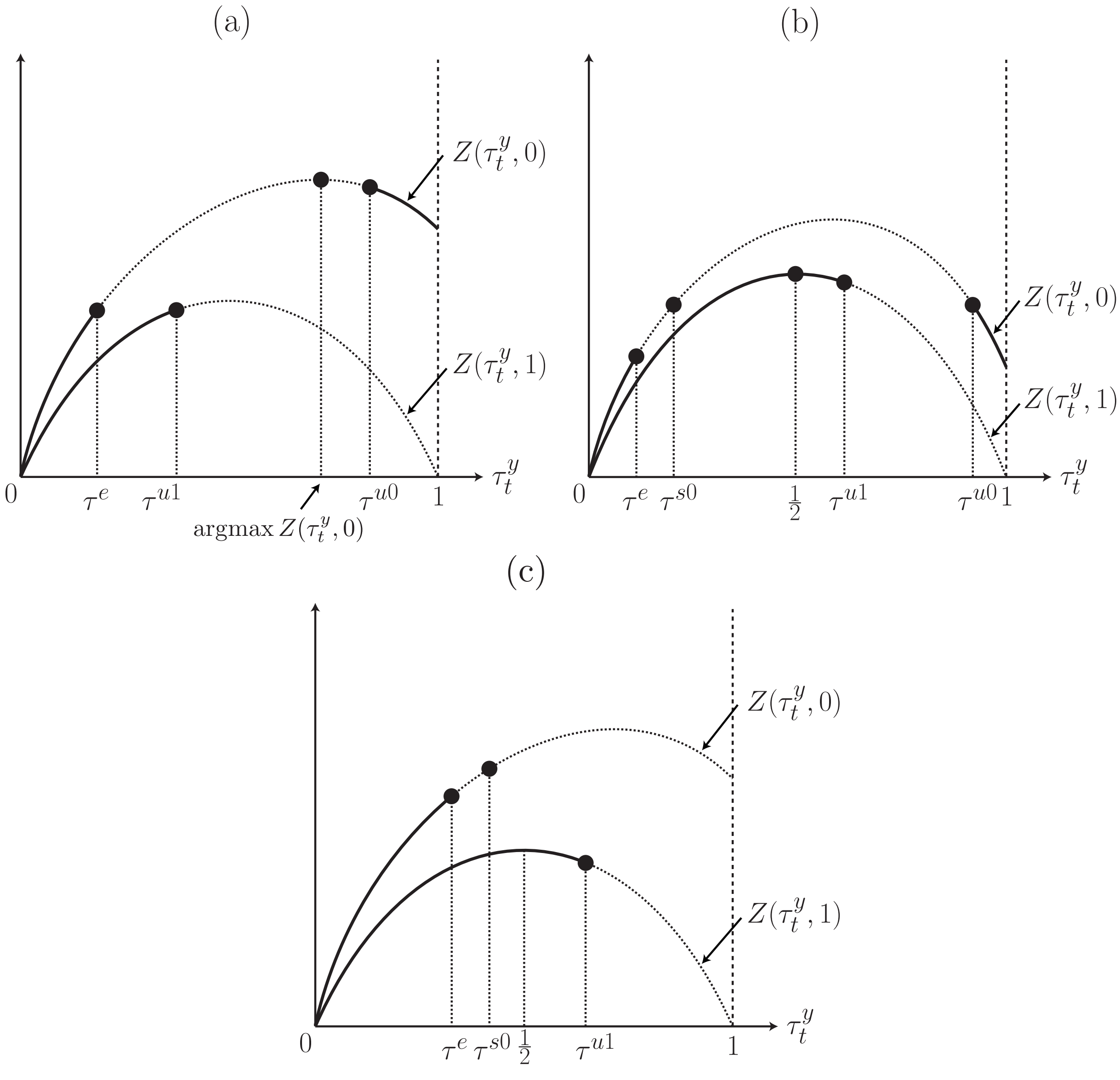

Figure 3: The unskilled-majority equilibrium with no taxation on the old (panel (a)); the unskilled-majority equilibrium with taxation on the old (panel (b)); and the skilled-majority equilibrium with no taxation on the old (panel (c)).

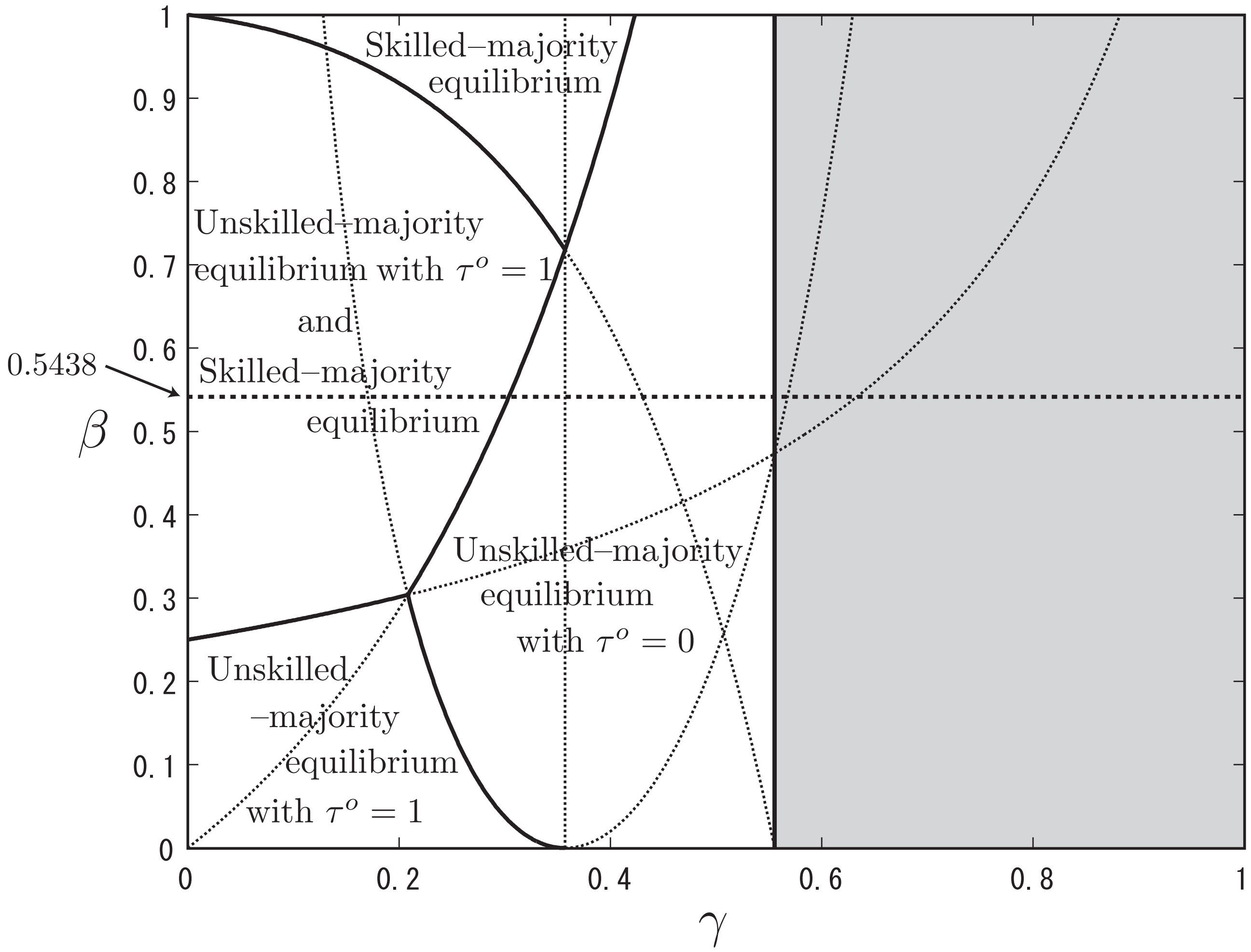

Figure 4: The figure displays the set of parameters (β, γ), where θ is fixed at 0.8, classified according to the characterization of political equilibria. The parameter ϕ is defined by $\phi \equiv \frac{1}{(1-\theta\gamma)-\frac{\gamma}{4}(1-\theta)}$.

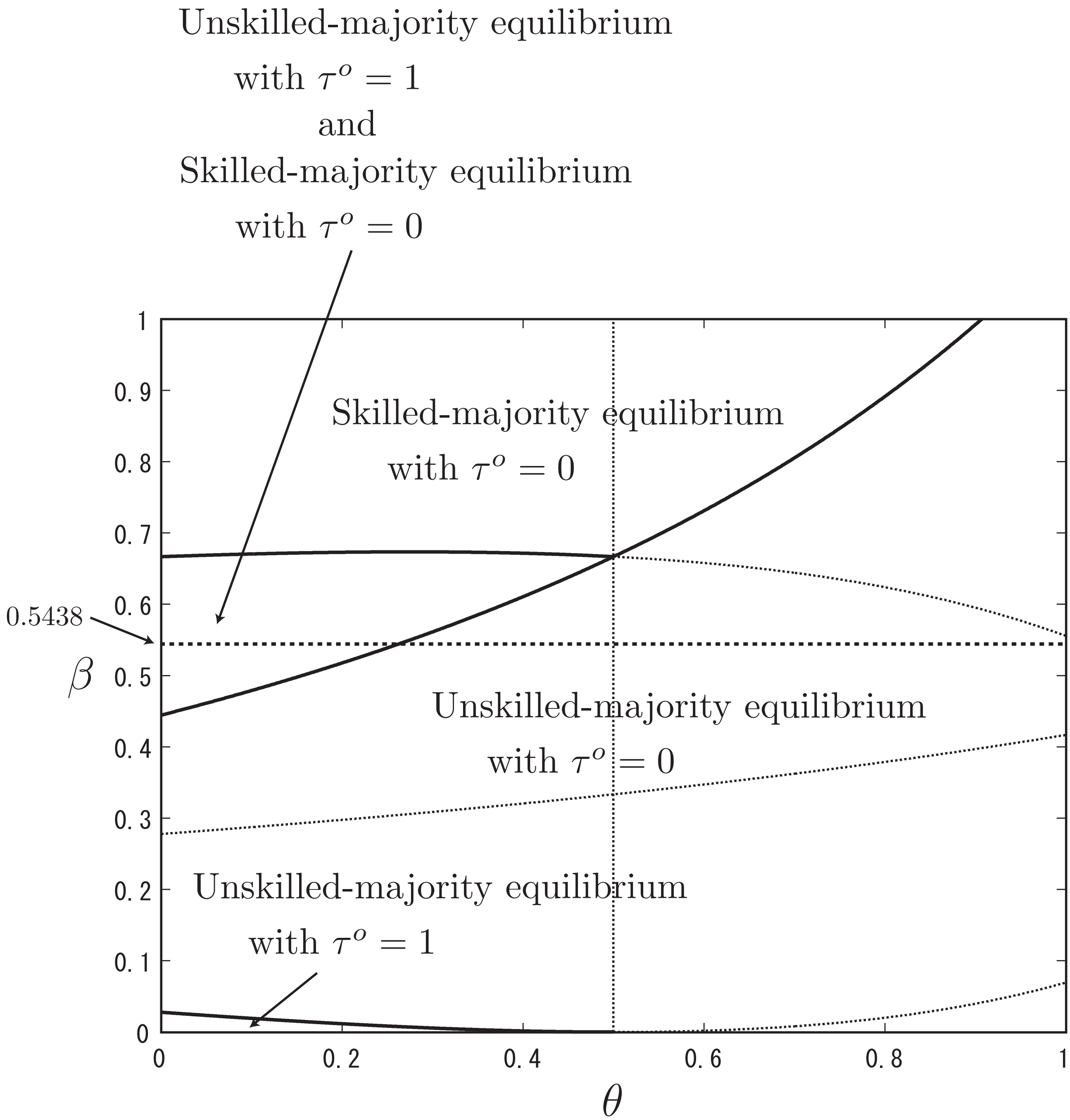

Figure 5: The figure displays the set of parameters (β, θ), where γ is fixed at 0.4, classified according to the characterization of political equilibria.

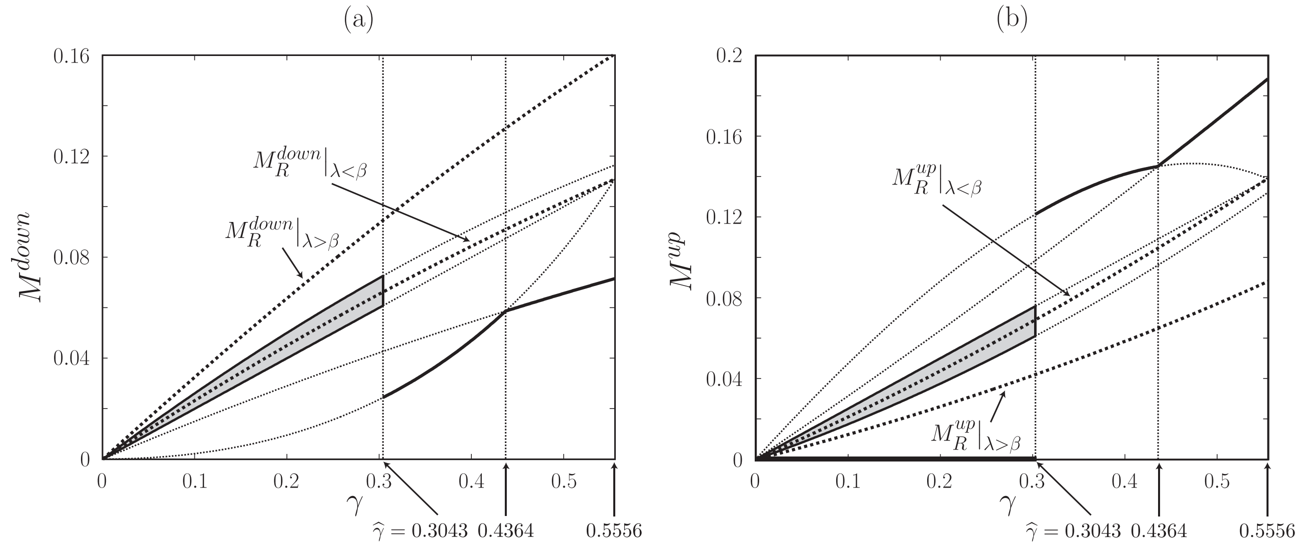

Figure 6: Solid curves and shaded area depict how the parameter γ affects the downward mobility (panel (a)) and the upward mobility (panel (b)) in the political equilibria. Dotted curves depict the corresponding values in the Ramsey allocation investigated in Section 4.

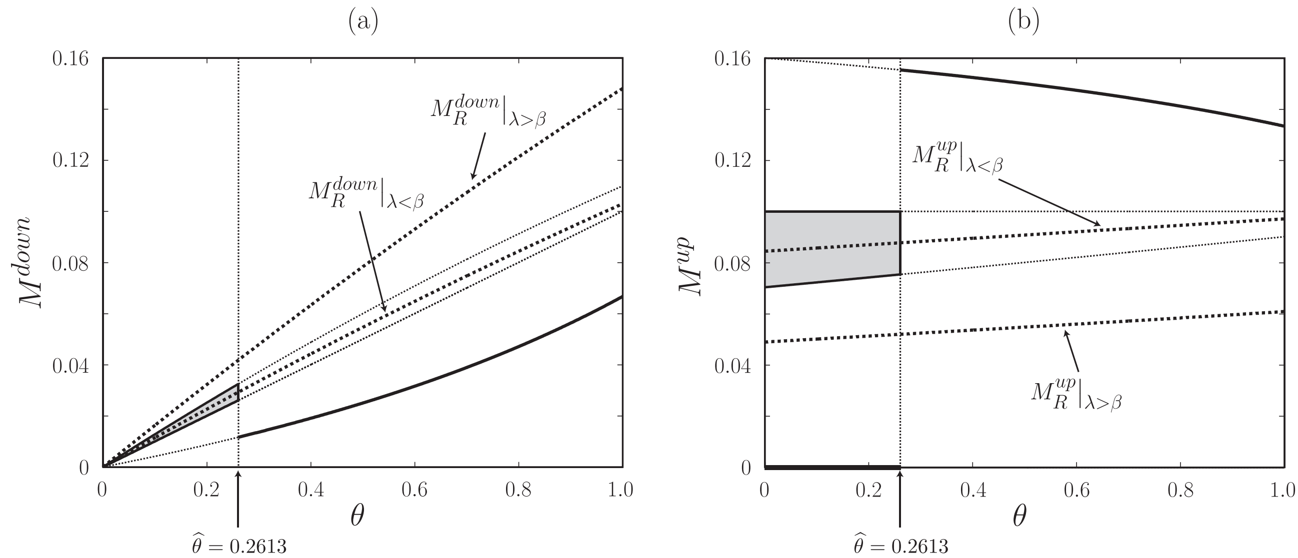

Figure 7: Solid curves and shaded area depict how the parameter θ affects the downward mobility (panel (a)) and the upward mobility (panel (b)) in the political equilibria. Dotted curves M_R^j ($j = down, up$) depict the corresponding values in the Ramsey allocation investigated in Section 4. In particular, $M_R^j\big|_{\lambda<\beta}$ and $M_R^j\big|_{\lambda>\beta}$ denote the numbers of mobility when $\lambda < \beta$ and $\lambda > \beta$, respectively.