



Discussion Papers In Economics And Business

Dynamic Political Economy of Redistribution Policy:
The Role of Education Costs

Ryo Arawatari and Tetsuo Ono

Discussion Paper 07-31-Rev.

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

Dynamic Political Economy of Redistribution Policy:
The Role of Education Costs

Ryo Arawatari and Tetsuo Ono

Discussion Paper 07-31-Rev.

November 2007

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

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

Dynamic Political Economy of Redistribution Policy: The Role of Education Costs*

RYO ARAWATARI[†] AND TETSUO ONO[‡]

November 2007

Abstract

This paper focuses on how education costs affect the political determination of redistribution policy via individual decision-making on education. For cases of high costs, there are multiple equilibria: the high-tax equilibrium featured by the minority of highly educated individuals and a large size of the government, and the low-tax equilibrium featured by the majority of highly educated individuals and a small size of the government. For cases of low costs, there is a unique equilibrium featured by the majority of highly educated individuals and a large size of the government.

Keywords: Markov perfect equilibrium; Dynamic political economy; Redistribution policy; Education costs

JEL Classification: D72; D78; E62

*This paper was formally circulated under the title “Markov Perfect Political Equilibrium with Social Security: The Role of Education Cost.” We would like to thank Hideki Konishi, Ryuichi Tanaka and participants at the Seminar on Economic Theory and Econometrics, Osaka Prefecture University, February 2007, the Macroeconomics Conference for Young Professionals, Osaka, March 2007, Japanese Economic Association Annual Spring Meeting at Osaka Gakuin University, June 2007, Osaka University Forum at Groningen, June 2007, Far Eastern Meeting of the Econometric Society at Taipei, July 2007, and European Meeting of the Econometric Society at Budapest, August 2007 for their useful comments and suggestions. Arawatari acknowledges financial support from the Research Fellowships for Young Scientists of the Japan Society for the Promotion of Science (JSPS). Ono acknowledges financial support from the JSPS through a Grant-in-Aid for Young Scientists (B) (No.17730131) and the 21st Century COE Program (Osaka University). All remaining errors are ours.

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

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

1 Introduction

For the past few decades, there has been growing concern about the political process of the determination of redistribution policy in macroeconomics (see Galasso and Profeta, 2002, and the references therein). Several earlier papers analyzed this issue by assuming myopic voting behavior (for example, Verbon and Verhoeven, 1992; Alesina and Rodrik, 1994; Meijdam and Verbon, 1996) or once-and-for-all voting (for example, Boadway and Wildasin, 1989; Bertola, 1993; Casamatta et al., 2000). In other words, they investigated the issue under the environment of the absence of dynamic interaction between redistribution policy and individual decision-making. However, as represented by the pay-as-you-go public pension, redistribution policy has a dynamic or an intergenerational aspect in nature, implying that there is a feedback mechanism between redistribution policy and individual decision-making. Expectations about future redistribution policy affect the current individuals' decision-making on saving or educational investment, which, in turn, has an effect on the future distribution of income and, thus, on the future voting behavior over redistribution policy. Therefore, there is a need to incorporate this mechanism into the analysis of political economy of redistribution.

Recently, Hassler et al. (2003) and Hassler et al. (2007) have provided politico-economic frameworks that incorporate this feedback mechanism, and have shown that the mechanism results in multiple political equilibria: the pro-welfare state and the anti-welfare state. In the pro-welfare state, expectations of higher future redistribution lead to lower educational investments and, thus, to a lower proportion of highly educated individuals. This implies a larger size of low-income individuals, which, in turn, increases future demand for redistribution, resulting in a higher tax rate and a larger size of the government. In the antiwelfare state, expectations of lower future redistribution lead to higher educational investments and, thus, to a higher proportion of highly educated individuals. This implies a larger size of high-income individuals, which, in turn, decreases future demand for redistribution, resulting in a lower tax rate and a smaller size of the

government.¹

Their two predictions provide explanations for the empirical observations in many OECD countries. Figure 1 presents a scatter plot of graduation rates of tertiary education and the ratios of public social expenditure to GDP in 2003 for OECD countries.² We take the graduation rate as a proxy variable of the proportion of highly educated individuals and the ratio of public social expenditure to GDP as a proxy variable of the size of the government. The prediction of the pro-welfare state is consistent with empirical observations in Continental European countries like Austria, France, Germany and Italy, and the prediction of the antiwelfare state is consistent with observations in Australia, Iceland, Ireland, Japan, the United Kingdom and the United States. Therefore, the theory of Hassler et al. (2003) and Hassler et al. (2007) contributes to the understanding of the difference with respect to welfare programs among some OECD countries that share similar political backgrounds. However, their model does not fully explain the empirical observations in Nordic countries (Denmark, Finland, Norway, and Sweden) that experience higher graduation rates and larger sizes of governments among OECD countries. The aim of this paper is to provide a framework that provides the *third prediction*, which was not shown by Hassler et al. (2003) and Hassler et al. (2007).

[Figure 1 around here.]

For the purpose of analysis, we focus on education costs in the model of Hassler et al.

¹The difference between Hassler et al. (2003) and Hassler et al. (2007) is that Hassler et al. (2003) assume age-independent taxation whereas Hassler et al. (2007) assume age-dependent taxation. The age-independent taxation creates a dynamic connection between the taxation of the old and the educational investment of the young, which produces the persistence effect of the state variable, i.e., the proportion of the highly educated individuals. In contrast, the age-dependent taxation breaks this link and, thus, produces a stationary property of equilibrium that makes the analysis more tractable. However, the assumption of age-dependent taxation still keeps the property of the feedback mechanism on which we focus here.

²The data on graduation rates of tertiary education (type A) are based on OECD (2005). The tertiary type A course is largely theory-based and designed to provide sufficient qualifications for entry to advanced research programs and professions with high skill requirements. There is another course (type B) that is shorter than the type A and it focuses on practical, technical or occupational skills for direct entry into the labor market. We here focus on tertiary type A education from the viewpoint of suitability for our analysis. The data on public social expenditure to GDP are based on OECD (2007). Public social expenditure comprises cash benefits, direct in-kind provision of goods and services, and tax breaks with social purposes.

(2007) and extend their model by generalizing the education cost function. Hassler et al. (2007) assumed that the cost of investment in education e , which can be interpreted as the disutility of educational effort, is given by $(e)^2$, while we assume that it is given by $\eta \cdot (e)^2$ where $\eta > 0$. For cases of high values of η , with Hassler et al. (2007) ($\eta = 1$) as a special case, individuals have a weak incentive to invest in education, which results in a similar equilibrium characterization to Hassler et al. (2007): the economy attains the pro-welfare state or the antiwelfare state depending on expectations by individuals. However, for cases of low values of η , individuals have a strong incentive to invest in education even if they expect a high future tax rate and, thus, a low return from the investment. Therefore, for cases of low education costs, the economy attains a state featured by a large proportion of highly educated individuals and a large size of the government. The generalization of the education cost function creates the third prediction, which was not presented by Hassler et al. (2003) and Hassler et al. (2007).

The organization of this paper is as follows. Section 2 presents the model. Section 3 characterizes the political equilibrium and explores empirical implications of our result. Section 4 characterizes the Ramsey allocation to evaluate the efficiency of the political equilibrium. Section 5 provides concluding remarks.

2 The Model

The model is based on that developed by Hassler et al. (2007). Time is discrete and denoted by $t = 1, 2, \dots$. The economy consists of a continuum of agents living for two periods. Each generation has a unit mass. Agents are, at birth, of two types: high ability and low ability, in proportions $\mu (> 1/2)$ and $1 - \mu (\leq 1/2)$, respectively. The role of the assumption $\mu > 1/2$ is discussed in Section 3.

High-ability agents can affect their prospects in life by an educational investment. In particular, they either become rich or poor, and by undertaking a costly investment can increase the probability e of becoming rich. The cost of investment, which can be interpreted as the disutility of educational effort, is given by $\eta \cdot (e)^2$, $\eta > 0$: Hassler et

al. (2007) considered the special case of $\eta = 1$. Low-ability agents make no investment choice: irrespective of their private actions, they are deemed to be in poverty. Rich agents earn a high wage, normalized to unity, in both periods, whereas poor agents earn a low wage, normalized to zero.

The government provides 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.³ The tax rates are determined before the young agents decide on their investment. The government budget balances in every period.⁴

There is no storage technology in this economy. Each individual uses up his/her endowments within a period. Therefore, the expected utility of agents alive at time t is given as follows:

$$\begin{aligned} V_t^{os} &= 1 - \tau_t^o + s_t, \\ V_t^{ou} &= V_t^{ol} = s_t, \\ V_t^y &= e_t \cdot \{(1 - \tau_t^y) + \beta(1 - \tau_{t+1}^o)\} + (s_t + \beta s_{t+1}) - \eta(e_t)^2, \\ V_t^{yl} &= (s_t + \beta s_{t+1}), \end{aligned}$$

where V_t^{os} , V_t^{ou} , and V_t^{ol} denote the utility of the old who were successful in youth, the old who were unsuccessful in youth, and old low-ability types in period t , and V_t^y and V_t^{yl} denote expected utility of young high-ability and young low-ability agents born in period t . Note that V_t^y is computed prior to individual success or failure. The parameter $\beta \in (0, 1)$ is the discount factor.

³We can alternatively assume age-independent taxation (Hassler et al., 2003) that creates a dynamic connection between the taxation of the old and the investment of the young (see Hassler et al. (2007) Section 6 for a discussion). Because our focus is on the role of education costs, we follow the assumption of age-dependent taxes in Hassler et al. (2007).

⁴Hassler et al. (2007) assumed that the government provides transfers (s) as well as public goods (g). We abstract from public goods provision in our analysis for simplicity of exposition. However, this simplification does not qualitatively affect the result shown below.

Given these preferences, the optimal investment of young high-ability agents is:

$$e^*(\tau_t^y, \tau_{t+1}^o) = \frac{1}{2\eta}(1 + \beta - \tau_t^y - \beta\tau_{t+1}^o). \quad (1)$$

Because high-ability agents are identical ex ante, agents of the same cohort choose the same investment, implying that the proportion of old poor among high-ability agents in period $t + 1$ is given by:

$$u_{t+1} \equiv 1 - e^*(\tau_t^y, \tau_{t+1}^o) = \frac{1}{2\eta}(2\eta - (1 + \beta) + \tau_t^y + \beta\tau_{t+1}^o).$$

Thus, the proportion of old poor among high-ability agents in period $t + 1$, u_{t+1} , depends on the income tax rate levied on the young rich in period t , τ_t^y , and the discounted tax rate levied on the old rich in period $t + 1$, $\beta\tau_{t+1}^o$.⁵

The government runs a balanced budget in every period, implying that the budget can be expressed as:

$$\begin{aligned} 2s_t &= \mu(1 - u_t)\tau_t^o + \mu e^*(\tau_t^y, \tau_{t+1}^o)\tau_t^y \\ &= W(u_t, \tau_t^o) + Z(\tau_t^y, \tau_{t+1}^o), \end{aligned} \quad (2)$$

where $W(u_t, \tau_t^o) \equiv \mu(1 - u_t)\tau_t^o$ is the tax revenue financed by the old, and $Z(\tau_t^y, \tau_{t+1}^o) \equiv \mu e^*(\tau_t^y, \tau_{t+1}^o)\tau_t^y$ is the tax revenue financed by the young.

3 Political Equilibrium

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

⁵When we characterize a political equilibrium (defined below) in the next section, we check whether or not the levels of e^* and u are given within the range $(0, 1)$.

the old, and discuss the role of education costs in determining the properties of political equilibrium. Section 3.4 discusses empirical implications of our model.

3.1 Definition of Political Equilibrium

Following Hassler et al. (2007), it is assumed that agents vote on current taxes at the beginning of each period but that only the old vote.⁶ The respective utility of the old rich and the old poor can be written as:

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

Whatever the level of u_t , the old rich would never tax themselves to finance transfers because the marginal cost of taxation is greater than the marginal benefit of taxation: $\partial V_t^{os} / \partial \tau_t^o = -1 + \mu(1 - u_t)/2 < 0$. However, the old poor prefer taxation on the old because they pay no tax but benefit from transfer: $\partial V_t^{ou} / \partial \tau_t^o = \mu(1 - u_t)/2 > 0$. Thus, the rich prefer a zero tax rate on the old, while the poor prefer 100% taxation on the old. As for the choice of τ_t^y , the old agents have the same political aim irrespective of their type: they wish to maximize the revenue from the young, $Z(\tau_t^y, \tau_{t+1}^o)$.

We now provide the definition of political equilibrium. The current paper focuses on stationary Markov-perfect equilibria, where the state of the economy is summarized by the proportion of the current old poor among high ability agents (u_t).

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

⁶For discussion about the role of this assumption, see Section IIB in Hassler et al. (2003) and Section 6 in Hassler et al. (2007).

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

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

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

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

The first equilibrium condition requires that τ_t^o maximizes the utility of the decisive old voter who is either the old rich (if $u_t \leq 1 - 1/2\mu$) or the old poor (if $u_t > 1 - 1/2\mu$). We assume that the rich decide τ_t^o in the case of an equal number of rich and poor voters ($u_t = 1 - 1/2\mu$).⁷ The second equilibrium condition implies that all young agents choose their investments optimally, given τ_t^y and τ_{t+1}^o , and that agents hold rational expectations about future taxes and distributions of types. The third equilibrium condition requires that τ_t^y maximizes the objective function of the decisive voter.

3.2 The Determination of T^o and U

We now solve the three equilibrium conditions recursively. The mapping $T^o(\cdot)$ that satisfies equilibrium condition 1 is given by:

$$T^o(u_t) = \begin{cases} 0 & \text{if } u_t \leq 1 - 1/2\mu, \\ 1 & \text{if } u_t > 1 - 1/2\mu. \end{cases} \quad (3)$$

The decisive voter will set $\tau_t^o = 0$ if the rich are the majority: $\mu(1 - u_t) \geq 1/2$, i.e., $u_t \leq 1 - 1/2\mu$, whereas it will set $\tau_t^o = 1$ if the poor are the majority: $\mu(1 - u_t) < 1/2$, i.e., $u_t > 1 - 1/2\mu$.

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

⁷Recall that we have assumed $\mu > 1/2$ in Section 2. If $\mu \leq 1/2$, then $\mu(1 - u_t) < 1/2$ always holds: there is no opportunity for the rich to become the majority of voters. In order to present a case where the rich are the majority, we have adopted the assumption of $\mu > 1/2$.

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

$$U(\tau_t^y) = \frac{1}{2\eta} (2\eta - (1 + \beta) + \tau_t^y + \beta T^o(U(\tau_t^y))). \quad (4)$$

Because $T^o(\cdot) \in \{0, 1\}$ is given by (3), any solution of (4) must be a combination of the two linear functions: $U(\tau_t^y) = (2\eta - (1 + \beta) + \tau_t^y) / 2\eta$ if $T^o = 0$ and $U(\tau_t^y) = (2\eta - 1 + \tau_t^y) / 2\eta$ if $T^o = 1$.

Under the assumption of rational expectations about future taxes and distributions of types, any solution to the functional equation (4) is given by:

$$U(\tau_t^y) = \begin{cases} (2\eta - (1 + \beta) + \tau_t^y) / 2\eta & \text{if } \tau_t^y \leq 1 - \eta/\mu \\ \{(2\eta - (1 + \beta) + \tau_t^y) / 2\eta, (2\eta - 1 + \tau_t^y) / 2\eta\} & \text{if } 1 - \eta/\mu < \tau_t^y \leq 1 - \eta/\mu + \beta \\ (2\eta - 1 + \tau_t^y) / 2\eta & \text{if } 1 - \eta/\mu + \beta < \tau_t^y. \end{cases} \quad (5)$$

Suppose that young agents in period t expect $\tau_{t+1}^o = 0$ ($\tau_{t+1}^o = 1$): the rich (poor) old are the majority of voters in period $t + 1$. Under this expectation, young agents choose their investments as $e^*(\tau_t^y, 0)$ ($e^*(\tau_t^y, 1)$). This expectation is rational if $\mu e^*(\tau_t^y, 0) \geq 1/2$ (if $\mu e^*(\tau_t^y, 1) < 1/2$), that is, if $\tau_t^y \leq 1 - \eta/\mu + \beta$ (if $1 - \eta/\mu < \tau_t^y$). Figure 2 illustrates possible cases of U that satisfy the second equilibrium condition (5). The solid lines show the graphs of U satisfying equilibrium condition 2. Since our focus is on the educational costs, these cases are classified according to the size of η/μ .

[Figure 2 around here.]

Suppose that the costs of education are high such that $\eta/\mu > 1$ (see Panels (a) and (b)). A high η implies that educational investment is costly, thereby giving the high-skilled individuals a disincentive to invest in education. Thus, a high η results in a low probability of being successful, and the majority in the next period can be poor $\forall \tau_t^y \in [0, 1]$. Alternatively, suppose that the costs of education are low such that $\eta/\mu \in (0, 1]$ (see Panels (c) and (d)). The high-skilled individuals have a strong incentive to invest in education, which results in a high probability of being successful. Thus, the

majority in the next period can be rich for $\tau_t^y \in [0, \min(1 - \eta/\mu + \beta, 1)]$. In particular, if $\tau_t^y \leq 1 - \eta/\mu$, the majority in the next period are always the rich.

Hassler et al. (2007) ignored the cases displayed in Panels (c) and (d) by assuming a specific form of education cost function with $\eta = 1$. If $\eta = 1$, there is no $\tau_t^y \in [0, 1]$ that satisfies $\tau_t^y \leq 1 - \eta/\mu$: there is no case in which the majority, in the next period, are always the rich. Their analysis was therefore limited to the cases illustrated in Panels (a) and (b). The current paper extends the model of Hassler et al. (2007) by generalizing the cost function of education, and shows that in the cases presented in Panels (c) and (d), there is a political equilibrium whose outcome is qualitatively different from Hassler et al. (2007).

As depicted in Figure 2, there are multiple, self-fulfilling expectations of U for a certain set of τ_t^y . Which U arises in equilibrium depends on the expectations of agents. To illustrate U in equilibrium, let us define the critical rate of τ_t^y : $\theta \in (\max(0, 1 - \eta/\mu), \min(1 - \eta/\mu + \beta, 1)]$. The rate θ , which depends on the expectations of agents, is the highest tax rate that yields a majority of old rich in the next period. For $\tau_t^y > \theta$, the majority is the poor. However, for $\tau_t^y \leq \theta$, the majority in the next period is either the rich or the poor depending on expectations by agents. Thus, the equilibrium features a monotonic or a nonmonotonic function U . Without loss of generality, we hereafter focus our attention on equilibria featuring a monotonic function U (Figure 3).⁸ The function U is thus given by:

$$U(\tau_t^y) = \begin{cases} (2\eta - (1 + \beta) + \tau_t^y) / 2\eta & \text{if } \tau_t^y \leq \theta, \\ (2\eta - 1 + \tau_t^y) / 2\eta & \text{if } \tau_t^y > \theta. \end{cases} \quad (6)$$

[Figure 3 around here.]

⁸As shown in Hassler et al. (2007), this restriction does not qualitatively affect the equilibrium outcome.

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

Given the characterization of T^o and U satisfying equilibrium conditions 1 and 2, respectively, we now consider the political determination of τ_t^y that satisfies equilibrium condition 3. Because there are two possible solutions of U , we introduce the corresponding definitions of political equilibria: a *high-tax equilibrium* featured by $\tau^o = 1$ and a *low-tax equilibrium* featured by $\tau^o = 0$. In the high-tax equilibrium, agents choose τ_t^y to induce a majority of poor at $t + 1$; they maximize the tax revenue financed by the young $Z(\tau_t^y, 1)$. In the low-tax equilibrium, agents choose τ_t^y to induce a majority of rich at $t + 1$; they maximize the tax revenue financed by the young $Z(\tau_t^y, 0)$. We first characterize a high-tax equilibrium.

Proposition 1. *Suppose that $\beta < (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$ holds. There exists a set of high-tax equilibria such that $\forall t$, $T^o(u_t)$ is given by (3), $T^y = 1/2$, and U is given by (6). The equilibrium outcome is unique and such that $\forall t$, $\tau_t^y = 1/2$, $\tau_t^o = 1$, $u_t = 1 - 1/4\eta$, and $2s_t = 3\mu/8\eta$.*

Proof. See Appendix.

Suppose that agents choose τ_t^y to induce a future majority of poor. In cases illustrated in Panels (a) and (b) of Figure 2, agents can choose $\tau_t^y \in [0, 1]$ that maximizes $Z(\tau_t^y, 1)$:

$$\tau_t^y = \arg \max_{\tau_t^y \in [0, 1]} Z(\tau_t^y, 1) = \frac{1}{2}.$$

The young agents are taxed on the top of the Laffer curve, conditional on their expecting 100% taxation when old. In the case illustrated in Panel (a), voting that induces a future majority of poor is the only option; a high-tax equilibrium is sustained by setting $\tau_t^y = 1/2$. However, in the case illustrated in Panel (b), voting that induces a future majority of rich can be an alternative option for $\tau_t^y \in [0, 1 + \beta - \eta/\mu]$. In this case, a high-tax equilibrium is sustained if $Z(1/2, 1) > Z(\tau_t^y, 0)$ for all τ_t^y inducing a future majority of rich. $Z(\tau_t^y, 0)$

is maximized by setting $\tau_t^y = \theta$ because (i) setting $\arg \max_{\tau_t^y \in [0,1]} Z(\tau_t^y, 0) = (1 + \beta)/2$ is impossible in the current case, and (ii) $\theta (\leq 1 + \beta - \eta/\mu)$ is the highest tax rate that induces a future majority of rich. Therefore, a high-tax equilibrium is sustained if $Z(1/2, 1) > Z(\theta, 0)$, i.e., if the expectation parameter θ is set below the critical level: $\theta < \tilde{\theta}(\beta) \equiv \left(1 + \beta - \sqrt{\beta(\beta + 2)}\right) / 2$. This implies that, for the case illustrated in Panel (b), the existence of the high-tax equilibrium depends on the expectations by agents.

In cases illustrated in Panels (c) and (d) of Figure 2, we still have an option that induces a future majority of poor by setting $\tau_t^y = 1/2$. However, there is an alternative option that induces a future majority of rich by choosing $\tau_t^y \in [0, \min(1 + \beta - \eta/\mu, 1)]$. In particular, if $(1 + \beta)/2 \leq 1 - \eta/\mu$, agents can definitely induce a future majority of rich by setting $\tau_t^y = \arg \max_{\tau_t^y \in [0,1]} Z(\tau_t^y, 0) = (1 + \beta)/2$. Because $Z((1 + \beta)/2, 0) > Z(1/2, 1)$ holds, a high-tax equilibrium is not sustained as an equilibrium. Based on this consideration, we can conclude that a high-tax equilibrium is sustained if (i) setting $\tau_t^y = (1 + \beta)/2$ is not possible: $(1 + \beta)/2 > 1 - \eta/\mu$, (ii) the expectation θ is low such that a future majority of poor will be induced: $\theta \in (1 - \eta/\mu, (1 + \beta)/2)$, and (iii) $Z(1/2, 1) > Z(\theta, 0)$, i.e., $\theta < \tilde{\theta}(\beta)$. The second and the third conditions require that θ is set within the range $(1 - \eta/\mu, \tilde{\theta}(\beta))$. The condition $\beta < (\eta/\mu - 1/2)^2 / (1 - \eta/\mu)$, which includes the first condition, guarantees that the range is not empty. Figure 4 illustrates an example of the graph of the tax revenue from the young, satisfying these three conditions.

[Figure 4 around here.]

Next, we characterize a low-tax equilibrium.

Proposition 2

- (i) *Suppose that $\beta \leq 1 - 2\eta/\mu$ holds. There exists a set of low-tax equilibria such that $\forall t$, $T^o(u_t)$ is given by (3), $T^y = (1 + \beta)/2$, and U is given by (6). The equilibrium outcome is unique and such that $\forall t$, $\tau_t^y = (1 + \beta)/2$, $\tau_t^o = 0$, $u_t = (2\eta - (1 + \beta)/2) / 2\eta$, and $2s_t = \mu(1 + \beta)^2 / 8\eta$.*

(ii) Suppose that $\beta > 1 - 2\eta/\mu$ and $\beta \geq (\mu/\eta)(\eta/\mu - 1/2)^2$ hold. There exists a set of low-tax equilibria such that $\forall t$, $T^o(u_t)$ is given by (3), $T^y = \theta \in \left(\max(1 - \eta/\mu, \tilde{\theta}(\beta)), \min((1 + \beta)/2, 1 + \beta - \eta/\mu) \right]$, and U is given by (6). The equilibrium outcome is indeterminate and such that $\forall t$, $\tau_t^y = \theta$, $\tau_t^o = 0$, $u_t = (2\eta - (1 + \beta) + \theta)/2\eta$, and $2s_t = \mu(1 - \theta + \beta)^2/2\eta$.

Proof. See Appendix.

Suppose that agents choose τ_t^y to induce a future majority of rich. In the case illustrated in Panel (a) of Figure 2, such a choice is unavailable for agents. In the case illustrated in Panel (b), the decisive voter prefers the highest tax rate on the young that induces a future majority of rich, $\tau_t^y = \theta (\leq 1 + \beta - \eta/\mu)$, because the revenue from the young $Z(\tau_t^y, 0)$ is increasing in τ_t^y in the range $(0, \theta)$. Therefore, the choice of $\tau_t^y = \theta$ under the expectation of $\tau_{t+1}^o = 0$ is sustained as an equilibrium if $Z(\theta, 0) \geq Z(1/2, 1)$, i.e., if $\theta \geq \tilde{\theta}(\beta)$. The condition $\beta \geq (\mu/\eta)(\eta/\mu - 1/2)^2$ implies that the range $[\tilde{\theta}(\beta), 1 + \beta - \eta/\mu]$ is not empty.

Next, consider the cases illustrated in Panels (c) and (d) of Figure 2. If $1 - \eta/\mu < (1 + \beta)/2$, i.e., if $\beta > 1 - 2\eta/\mu$, agents have an option that induces a future majority of rich by setting $\tau_t^y = \theta$. The tax rate on the young depends on the expectations by agents (see Panel (a) of Figure 5). In particular, when the expectation is given by $\theta = (1 + \beta)/2$, the young are taxed on the top of the Laffer curve. However, if $(1 + \beta)/2 \leq 1 - \eta/\mu$, i.e., if $\beta \leq 1 - 2\eta/\mu$, expectations no longer affect the determination of τ_t^y . Voters can choose $\tau_t^y = (1 + \beta)/2$ and maximize the revenue from the young by taxing on the top of the Laffer curve (see Panel (b) of Figure 5).

[Figure 5 around here.]

Figure 6 illustrates the parameter conditions derived in Propositions 1 and 2. The next proposition summarizes the results established so far. The proof of the proposition is immediate from Figure 6.

[Figure 6 around here.]

Proposition 3

- (i) *If $\beta < (\mu/\eta)(\eta/\mu - 1/2)^2$, there exists a unique high-tax equilibrium as in Proposition 1.*
- (ii) *If $(\mu/\eta)(\eta/\mu - 1/2)^2 \leq \beta < (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$, the equilibrium is indeterminate. There exist both a high-tax equilibrium as in Proposition 1 and a set of low-tax equilibria as in Proposition 2(ii).*
- (iii) *If $\beta > 1 - 2\eta/\mu$ and $\beta \geq (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$, there exists a set of low-tax equilibria as in Proposition 2(ii).*
- (iv) *If $\beta \leq 1 - 2\eta/\mu$, there exists a unique low-tax equilibrium as in Proposition 2(i).*

As illustrated in Figure 5, given $\mu \in (1/2, 1]$, the range of η/μ is limited to $[1, 2]$ if $\eta = 1$. By assuming the cost function of education with $\eta = 1$, Hassler et al. (2007) focused on this range and showed two patterns of political equilibria, as in Proposition 3(i) and 3(ii). Contrary to their analysis, however, we focus on a wider range of η/μ by assuming a generalized cost function. Under this extended framework, we can derive two patterns of political equilibria, as in Proposition 3(iii) and 3(iv), which were not shown by Hassler et al. (2007).

The mechanism behind our main finding is intuitive and can be understood as follows. First, suppose that the investment cost is high such that $\eta/\mu > 1$. In this case, it is less attractive for agents to invest in education. When agents attach a low value to the future such that $\beta < (\mu/\eta)(\eta/\mu - 1/2)^2$, there is a future majority of poor irrespective of policy. However, when agents attach a high value to the future, such that $\beta \geq (\mu/\eta)(\eta/\mu - 1/2)^2$, agents can undertake strategic political behavior to induce a future majority of rich. The equilibrium is indeterminate in this case.

Second, suppose that the investment cost is low such that $\eta/\mu \leq 1$. In this case, it is more attractive for agents to invest in education. In particular, when the cost is sufficiently low such that $\beta \geq (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$, agents have an incentive to invest

in education even if they expect a high future tax rate and, thus, a low return from the investment. Then, it becomes possible to tax heavily on the young without compromising the future political equilibrium. Therefore, in this case, there is a future majority of rich irrespective of policy.

3.4 Empirical Implications

The results established in Propositions 1–3 show that the level of government spending and the distribution of rich and poor are affected by the parameter representing educational cost, η . To investigate the empirical implications of η , consider two levels of η , η_H and $\eta_L (< \eta_H)$. Suppose that given η_H , there exist both a low-tax equilibrium and a high-tax equilibrium, as in Proposition 3(ii). Suppose also that, given η_L , there exists a set of low-tax equilibria as in Proposition 3(iii) or a unique low-tax equilibrium as in Proposition 3(iv).

The equilibrium level of government spending differs across political equilibria. Consider first the case of $\eta = \eta_H$ that leads to multiple equilibria. In the high-tax equilibrium, the level of government spending is given by $3\mu/8\eta_H$. In the low-tax equilibrium, the level, given by $\mu(1 - \theta + \beta)\theta/2\eta_H$, depends on expectations by agents. The lowest level is $\mu/8\eta_H$ at $\theta = \tilde{\theta}(\beta)$, which is lower than $3\mu/8\eta_H$. The highest level is $\mu(1 + \beta)^2/8\eta_H$ at $\theta = (1 + \beta)/2$, which exceeds $3\mu/8\eta_H$ if β is high. Therefore, the high-tax equilibrium is featured by less or more spending depending on expectations by agents, θ , and the discount factor, β .

As already shown by Hassler et al. (2007), the case of η_H provides an explanation for the cross-country difference in the size of the government spending and the distribution of rich and poor. To provide an empirical viewpoint, we take the entry and graduation rates of tertiary education as proxy variables of the size of the rich μe^* . Then, the high-tax equilibrium corresponds to some Continental European countries like Austria, France, Germany and Italy, featured with large sizes of governments and low graduation rates; the low-tax equilibrium corresponds to Australia, Iceland, Ireland, Japan, the United King-

dom and the United States, featured with small sizes of government and high graduation rates (see Figure 1).

Although the case of η_H can describe the cross-country difference among some OECD countries, it does not fit the empirical fact of Nordic countries (Denmark, Finland, Norway and Sweden); they have large sizes of government and high graduation rates as presented in Figure 1. However, the case of η_L can provide an explanation for the Nordic countries. In this case, the majority is rich, and the level of government spending is given by $2s_t = \mu(1 - \theta + \beta)^2/2\eta_L$ or $2s_t = \mu(1 + \beta)^2/8\eta_L$, which is larger than that in the high-tax equilibrium, $2s_t = 3\mu/8\eta_H$, for a small η_L and a large β . Therefore by focusing on educational costs, we can show the equilibrium that fits the empirical fact of the Nordic countries, which was not shown by Hassler et al. (2003) and Hassler et al. (2007).

We have shown the existence of equilibrium featured by the majority of rich and a large size of government by considering the cases of low education costs. We can show the existence of a qualitatively similar equilibrium by allowing the wage of the successful agents, w , to differ from unity. A higher wage inequality results in a unique equilibrium where the majority of rich prefer a large size of government and imposes a high tax rate on the young (Hassler et al., 2003). However, this case does not fit the empirical fact of the Nordic countries that are featured by low inequality (Björklund et al., 2002). Therefore, focusing on education costs rather than wage inequality is a key to explaining the empirical fact of the Nordic countries in the framework of Hassler et al. (2003) and Hassler et al. (2007).

4 Ramsey Allocation

In this section, we first characterize the Ramsey allocation, and then compare it with the political equilibria presented in Section 3. Following Hassler et al. (2007), we define a Ramsey allocation as a feasible plan chosen by a benevolent social planner who can

commit to a policy sequence at time 0. The allocation solves the following problem:

$$\begin{aligned} & \max_{\{s_t, \tau_t^y, \tau_t^o\}_{t=0}^{\infty}} [\beta\mu(1-u_0)V^{os}(s_0, \tau_0^o) + \beta(\mu u_0 + (1-\mu)H)V^{ou}(s_0, \tau_0^o) \\ & + \sum_{t=0}^{\infty} \lambda^{t+1} \{ \mu V^y(e_t, s_t, s_{t+1}, \tau_t^y, \tau_{t+1}^o) + (1-\mu)H V^y(s_t, s_{t+1}) \}], \end{aligned}$$

where $\lambda \in [0, 1)$ is a discount factor, $\beta (> 0)$ is a weight on the initial generation of old agents, and $H (\geq 1)$ is a bias in the planner's preference towards low-ability agents. We employ the same form of a social welfare function as Hassler et al. (2007) in order to show how our extension affects the characterization of the Ramsey allocation.

Given the educational investment (1) and the government budget constraint (2), the problem can be rewritten as a simple static problem:

$$\max_{\tau_0^o \in [0, 1]} \beta\mu(1-u_0)(1-\tau_0^o) + (\beta + \lambda)A\mu(1-u_0)\tau_0^o + \frac{L}{1-\lambda}, \quad (7)$$

where

$$\begin{aligned} L \equiv & \max_{\{\tau^y \in [0, 1], \tau^o \in [0, 1]\}} [\lambda\mu \{ (1-\tau^y + \beta(1-\tau^o))e^*(\tau^y, \tau^o) - \eta \cdot e^*(\tau_t^y, \tau_{t+1}^o)^2 \} \\ & + (\beta + \lambda)A\mu\tau^ye^*(\tau^y, \tau^o) + \lambda(\beta + \lambda)A\mu\tau^oe^*(\tau^y, \tau^o)], \end{aligned}$$

where $A \equiv (\mu + (1-\mu)H)/2$ is the planner's marginal value of expenditures. 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*

$$\tau_0^o = \begin{cases} 1 & \text{if } A > \frac{\beta}{\beta + \lambda}, \\ 0 & \text{otherwise,} \end{cases}$$

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

1. Suppose that $\lambda < \beta$. Then,

$$\tau^y = \max \left\{ 0, (1 + \beta) \frac{A - \lambda/(\beta + \lambda)}{2A - \lambda/(\beta + \lambda)} \right\} \text{ and } \tau^o = 0.$$

2. Suppose that $\lambda > \beta$. Then,

$$\tau^y = \max \left\{ 0, \frac{A(1 - \lambda) - \lambda/(\beta + \lambda)}{2A - \lambda/(\beta + \lambda)} \right\} \text{ and } \tau^o = \min \left\{ \frac{1 + \beta}{\beta} \cdot \frac{A - \beta/(\beta + \lambda)}{2A - \beta/(\beta + \lambda)}, 1 \right\}.$$

3. Suppose that $\lambda = \beta$. Then,

$$\begin{cases} \tau^y = \tau^o = 0 & \text{if } A \leq \frac{1}{2} \\ \tau^y + \beta\tau^o = \frac{(2A-1)(1+\beta)}{4A-1} & \text{if } A > \frac{1}{2}. \end{cases}$$

Proof. See Appendix.

Proposition 4 shows that the parameter η does not affect the determination of tax rates in the Ramsey allocation. This result depends on the cost function specified by a quadratic form. With this feature in mind, we now summarize the solution to the Ramsey problem as follows. First, as regards the choice of τ_0^o , there is no interaction with future variables. The planner chooses τ_0^o to maximize the weighted sum of the utilities of the initial old agents. Thus, the planner sets $\tau_0^o = 1$ if $A > \beta/(\beta + \lambda)$ and $\tau_0^o = 0$ otherwise. Second, the choice of τ^y and τ^o depends on the magnitude correlation between β and λ . If $\lambda < \beta$ (part 1 of Proposition 4), the tax burden falls on the young; if $\lambda = \beta$ (part 3 of Proposition 4), the distribution of the tax burden is indeterminate; and if $\lambda > \beta$ (part 2 of Proposition 4), most of the tax burden falls on the old. In the case of $\lambda > \beta$, for a range of a small A , the nonnegativity constraint on the young binds; for range of a large A , the constraint of $\tau^o \leq 1$ is binding. In words, the tax burden is increased by the planner's higher value to the low-ability agents.

Given the characterization of the Ramsey allocation above, we now evaluate the efficiency of the political equilibria. First, consider the high-tax equilibrium. The high-tax

equilibrium shows a qualitative similarity to the Ramsey allocation in the case of $\lambda > \beta$, in that the main tax burden falls on the old. However, the tax rates in the high-tax equilibrium are quantitatively different from those in the Ramsey allocation in the following two points. First, when A is low, the Ramsey planner does not tax the old at the maximum rate. Second, in the Ramsey allocation with $\lambda > \beta$, the young are subject to a tax $\tau^y \leq (1 - \lambda)/2$, while they are taxed at a 50% rate in the high-tax equilibrium.

Next, consider the low-tax equilibrium featured by no taxation on the old (part 2). This equilibrium shows a qualitative similarity to the Ramsey allocation in the case of $\lambda < \beta$, in that the old bear no tax burden. When the political tax rate on the young depends on the expectations by agents as in Proposition 2(ii) the low-tax equilibrium with $\tau^y = \theta$ may realize the same tax rates as those in the Ramsey allocation, depending on the expectation θ . However, when the political tax rate on the young is determinate as in Proposition 2(i), the low-tax equilibrium with $\tau^y = (1 + \beta)/2$ realizes a higher tax burden as compared with the Ramsey allocation.⁹ Therefore, for low levels of η , there is no possibility that the political equilibrium resembles the Ramsey allocation.

5 Conclusion

This paper has extended the model of Hassler et al. (2007) by generalizing the cost function of education, which is summarized as follows. First, for high costs, with Hassler et al. (2007) as a special case, individuals have a weak incentive to invest in education, which results in a similar equilibrium characterization to that found in Hassler et al. (2007). The model presents multiple equilibria: the high-tax equilibrium (the pro-welfare state) featured by the minority of highly educated individuals and a large size of the government; and the low-tax equilibrium (the antiwelfare state) featured by the majority of highly educated individuals and a small size of the government. Which state is realized

⁹Assume that $\lambda < \beta$ and $A > \lambda/(\beta + \lambda)$: the Ramsey allocation is featured by $\tau^y = (1 + \beta)\{A - \lambda/(\beta + \lambda)\}/\{2A - \lambda/(\beta + \lambda)\}$ and $\tau^o = 0$. Assume also that $\beta < 1 - 2\eta/\mu$: the political equilibrium is featured by $\tau^y = (1 + \beta)/2$ and $\tau^o = 0$ as in Proposition 2(i). By direct calculation, we find that $(1 + \beta)\{A - \lambda/(\beta + \lambda)\}/\{2A - \lambda/(\beta + \lambda)\} < (1 + \beta)/2$ is equivalent to $A - \lambda/(\beta + \lambda) < A - \lambda/2(\beta + \lambda)$, which holds for any A, β and λ .

depends on expectations by individuals.

Second, for low costs, individuals have a strong incentive to invest in education irrespective of future redistribution policy. In this case, there is a unique equilibrium featured by the majority of highly educated individuals and the large size of the government, which fits the empirical facts of the Nordic countries (Denmark, Finland, Norway and Sweden). Furthermore, this unique equilibrium may require an excess tax burden on the young compared with the Ramsey allocation.

The analysis in this paper focuses on the characterization of political equilibria under a system of lump-sum general transfers paid to both the rich and the poor; we have abstracted from targeted transfers to the elderly or to the poor. However, the paper has provided a framework for understanding the role of education costs that affect the political determination of redistribution policy via individual decisions on education.

6 Appendix

6.1 Proof of Proposition 1

Suppose that at time t , agents know that $\tau_t^y = 1/2$ and expect that $\tau_{t+1}^o = 1$. Then, $1 - e^*(1/2, 1) = u_{t+1} = 1 - 1/4\eta = 1 - (1/4\mu)(\mu/\eta)$. Because the condition $\beta < (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$ ensures that $\eta/\mu > 1/2$ (see Figure 6), $u_{t+1} = 1 - (1/4\mu)(\mu/\eta) > 1 - 1/2\mu$ holds for all t . By (3), this implies that $\tau_{t+1}^o = 1$, fulfilling initial expectations. Therefore, there exists a high-tax equilibrium 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 note the following properties of the function Z : (i) $Z(\tau_t^y, 0)$ is concave in τ_t^y and is maximized at $\tau_t^y = (1 + \beta)/2$; (ii) $Z(\tau_t^y, 0) > Z(\tau_t^y, 1)$. Given these properties, we derive the conditions that (i) setting $\tau_t^y = (1 + \beta)/2$ is not available under the expectation $\tau_{t+1}^o = 0$; (ii) $Z(\tau_t^y, 0)$ is maximized at $\tau_t^y = \theta$; and (iii) $Z(\theta, 0) < Z(1/2, 1)$ hold. These conditions are given by (i) $(1 + \beta)/2 \leq 1 - \eta/\mu$, (ii) $\theta \in (\max(0, 1 - \eta/\mu), (1 + \beta)/2)$, and (iii) $\theta < \tilde{\theta}(\beta)$, respectively. The second and third conditions require that θ is set within the range $(\max(0, 1 - \eta/\mu), \tilde{\theta}(\beta))$. The condition $\beta < (\eta/\mu - 1/2)^2/(1 - \eta/\mu)$, which guarantees that the set $(\max(0, 1 - \eta/\mu), \tilde{\theta}(\beta))$ is not empty, includes the first condition, as illustrated in Figure 6. ■

6.2 Proof of Proposition 2

(i) Suppose that in period t , agents know that $\tau_t^y = (1 + \beta)/2$ and expect that $\tau_{t+1}^o = 0$. Then $u_{t+1} = 1 - e^*((1 + \beta)/2, 0) = (2\eta - (1 + \beta)/2) \leq 1 - 1/2\mu$ if and only if $\beta \geq -1 + 2\eta/\mu$. The condition $\beta \geq -1 + 2\eta/\mu$ holds under the assumption of $\beta \leq 1 - 2\eta/\mu$ (see Figure 6). By (3), the condition $u_{t+1} \leq 1 - 1/2\mu$ implies that $\tau_{t+1}^o = 0$, fulfilling initial expectations. Given that $Z(\tau_t^y, 0) > Z(\tau_t^y, 1) \forall \tau_t^y \in [0, 1]$ and that $\arg \max_{\tau_t^y \in [0, 1]} Z(\tau_t^y, 0) = (1 + \beta)/2$, setting $\tau_t^y = (1 + \beta)/2$ is optimal for the decisive voter.

(ii) Under the assumption that U is monotonic, $T^o(U(\theta)) = 0$ implies that $\forall \tau_t^y \leq \theta$, $T^o(U(\tau_t^y)) = 0$ and $U(\tau_t^y) = (2\eta - (1 + \beta) + \tau_t^y)/2\eta$. Consequently, the relevant payoff

function in the range $\tau_t^y \leq \theta$ is $Z(\tau_t^y, 0)$. The function $Z(\tau_t^y, 0)$ is a hump-shaped function of τ_t^y , with a maximum at $(1 + \beta)/2$.

Given the first assumption $\beta > 1 - 2\eta/\mu$, i.e., $1 - \eta/\mu < (1 + \beta)/2$, we can consider the following two cases: (a) the case of $1 - \eta/\mu < (1 + \beta)/2 < 1 + \beta - \eta/\mu$ and (b) the case of $1 + \beta - \eta/\mu \leq (1 + \beta)/2$. In case (a), $Z(\tau_t^y, 0)$ is increasing in the range $\tau_t^y < (1 + \beta)/2$ and decreasing in the range $\tau_t^y \geq (1 + \beta)/2$. In case (b), $Z(\tau_t^y, 0)$ is increasing in the range $\tau_t^y < 1 + \beta - \eta/\mu$. Therefore, the decisive voters prefer $\tau_t^y = \theta \in (1 - \eta/\mu, \min(1 + \beta - \eta/\mu, (1 + \beta)/2)]$ to any $\tau_t^y < \theta$ under the assumption of $\beta > 1 - 2\eta/\mu$.

An alternative option for the decisive voter is to set $\tau_t^y = 1/2$ under the expectation of $\tau_{t+1}^o = 1$. The choice of $\tau_t^y = \theta$ under the expectation of $\tau_{t+1}^o = 0$ is sustained as an equilibrium if $Z(\theta, 0) \geq Z(1/2, 1)$, i.e., if $\theta \geq \tilde{\theta}(\beta)$, where

$$\tilde{\theta}(\beta) \equiv \frac{1 + \beta - \sqrt{\beta(\beta + 2)}}{2} < \frac{1 + \beta}{2}.$$

Therefore, there exists a set of low-tax equilibria featured by $T^o(u_t)$ is given by (3), $T^y = \theta$ and U is given by (6) if θ is set within the range:

$$\left(\max(1 - \eta/\mu, \tilde{\theta}(\beta)), \min((1 + \beta)/2, 1 + \beta - \eta/\mu) \right].$$

The condition $\beta \geq (\mu/\eta)(\eta/\mu - 1/2)^2$, which is rewritten as $\tilde{\theta}(\beta) \leq 1 + \beta - \eta/\mu$, implying that the above range is not empty. ■

6.3 Proof of Proposition 4

It is immediate from (7) that $\tau_0^o = 1$ if $A > \beta/(\beta + \lambda)$ and $\tau_0^o = 0$ otherwise. The solution of the pair (τ^y, τ^o) is derived by solving (8). The solution must satisfy the following first-order conditions:

$$\tau^y : \frac{\partial L}{\partial \tau^y} - \xi^y + \theta^y = 0, \quad (8)$$

$$\tau^o : \frac{\partial L}{\partial \tau^o} - \xi^o + \theta^o = 0, \quad (9)$$

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

Assume, first, that $\tau^o = \xi^y = \xi^o = \theta^y = 0$. Then, we obtain

$$\begin{cases} \tau^y = (1 + \beta) \frac{A - \lambda / (\beta + \lambda)}{2A - \lambda / (\beta + \lambda)}, \\ \theta^o = \frac{\mu A^2}{2\eta} \cdot \frac{(1 + \beta)(\beta - \lambda)(\beta + \lambda)^2}{2A(\beta + \lambda) - \lambda}, \end{cases}$$

where $\tau^y < 1$ and $\theta^o > 0$ as long as $\beta > \lambda$. This establishes part 1.

Assume, next, that $\tau^y = \xi^y = \xi^o = \theta^o = 0$. Then, we obtain

$$\begin{cases} \tau^o = \frac{1 + \beta}{\beta} \cdot \frac{A - \beta / (\beta + \lambda)}{2A - \beta / (\beta + \lambda)}, \\ \theta^y = \frac{\mu A^2}{2\eta} \cdot \frac{(1 + \beta)(\lambda - \beta)(\beta + \lambda)^2}{\beta(2A(\beta + \lambda) - \beta)}, \end{cases}$$

where $\theta^y > 0$ as long as $\beta < \lambda$. Furthermore, $\tau^o < 1$ holds as long as $A < \beta / (1 - \beta)(\beta + \lambda)$. If $A \geq \beta / (1 - \beta)(\beta + \lambda)$, then $\tau^o = 1$ and $\xi^o > 0$. However, it remains to be checked whether $\tau^y = 0$ continues to be a solution. Given $\tau^o = 1$, we obtain

$$\frac{\partial L}{\partial \tau^y} - \xi^y + \theta^y = \frac{1}{2\eta} (-\lambda\mu(1 - \tau^y) + (\beta + \lambda)A\mu(1 - \lambda + 2\tau^y) + \theta^y) = 0.$$

The solution features $\tau^y = 0$ and $\theta^y > 0$ if and only if $A \leq \lambda / (1 - \lambda)(\beta + \lambda)$. For larger values of A , the solution instead features $\theta^y = 0$ and

$$\tau^y = \frac{A(1 - \lambda) - \lambda / (\beta + \lambda)}{2A - \lambda / (\beta + \lambda)}.$$

This establishes part 2.

Finally, consider the case where $\beta = \lambda$. Then, the objective function can be written as:

$$L(x) = \beta\mu\{(1 + \beta - x) \cdot e^*(x) - \eta \cdot e^*(x)^2\} + 2\beta A\mu x \cdot e^*(x),$$

where $x \equiv \tau^y + \beta\tau^o$. The first-order condition yields the result in part 3. \blacksquare

References

- [1] Alesina, A., and Rodrik, D., 1994. Distributive Politics and Economic Growth. *Quarterly Journal of Economics* 109, 465–490.
- [2] Bertola, G., 1993. Factor Shares and Savings in Endogenous Growth. *American Economic Review* 83, 1184–1198.
- [3] Björklund, A., Eriksson, T., Jantti, M., Osterbacka, E., and Raaum, O. 2002. Brother Correlations in Earnings in Denmark, Finland, Norway and Sweden Compared to the United States. *Journal of Population Economics* 15, 757–772.
- [4] Boadway, R.W., and Wildasin, D.E., 1989. A Median Voter Model of Social Security. *International Economic Review* 30, 307–328.
- [5] Casamatta, G., Cremer, H., and Pestieau, P., 2000. The Political Economy of Social Security. *Scandinavian Journal of Economics* 102, 503–522.
- [6] Galasso, V., and Profeta, P., 2002. The Political Economy of Social Security: A Survey. *European Journal of Political Economy* 18, 1–29.
- [7] Hassler, J., Rodriguez Mora, J.V., Storesletten, K., and Zilibotti, F., 2003. The Survival of the Welfare State. *American Economic Review* 93, 87–112.
- [8] Hassler, J., Storesletten, K., and Zilibotti, F., 2007. Democratic Public Good Provision. *Journal of Economic Theory* 133, 127–151.
- [9] Meijdam, L., and Verbon, H.A.A., 1996. Aging and Political Decision Making on Public Pensions. *Journal of Population Economics* 9, 141–158.
- [10] OECD, 2005. *Education at a Glance 2005*. OECD, Paris.
- [11] OECD, 2007. *OECD Factbook 2007 Economic, Environmental and Social Statistics*. OECD, Paris.

- [12] Verbon, H.A.A., and Verhoeven, M.J.M., 1992. Decision Making on Pension Schemes under Rational Expectations. *Journal of Economics (Zeitschrift fuer Nationaloekonomie)* 56, 71–97.



Figure 1: Figure 1 is a scatter plot of graduation rates of tertiary education and the ratios of public social expenditure to GDP in 2003 for OECD countries. The solid vertical line in the plot indicates the average graduation rate among OECD countries; the solid horizontal line indicates the average public social expenditure to GDP among OECD countries.

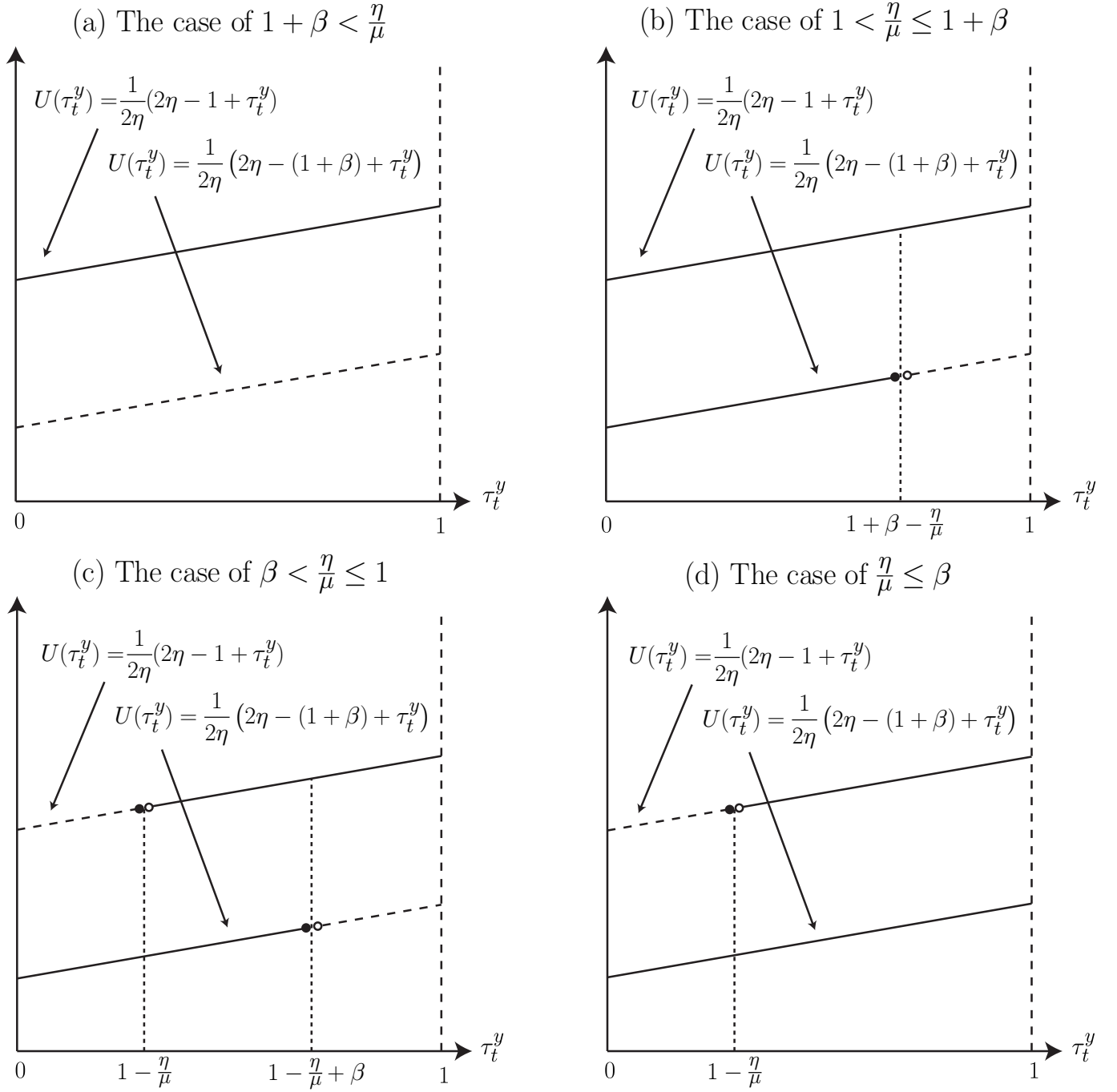


Figure 2: The figures represent the equilibrium decision rule $u_{t+1} = U(\tau_t^y)$. The solid lines show the graphs of U satisfying equilibrium condition 2. The panel (a) illustrates the case of $1 + \beta < \frac{\eta}{\mu}$; the panel (b) illustrates the case of $1 < \frac{\eta}{\mu} \leq 1 + \beta$; the panel (c) illustrates the case of $\beta < \frac{\eta}{\mu} \leq 1$; the panel (d) illustrates the case of $\frac{\eta}{\mu} \leq \beta$.

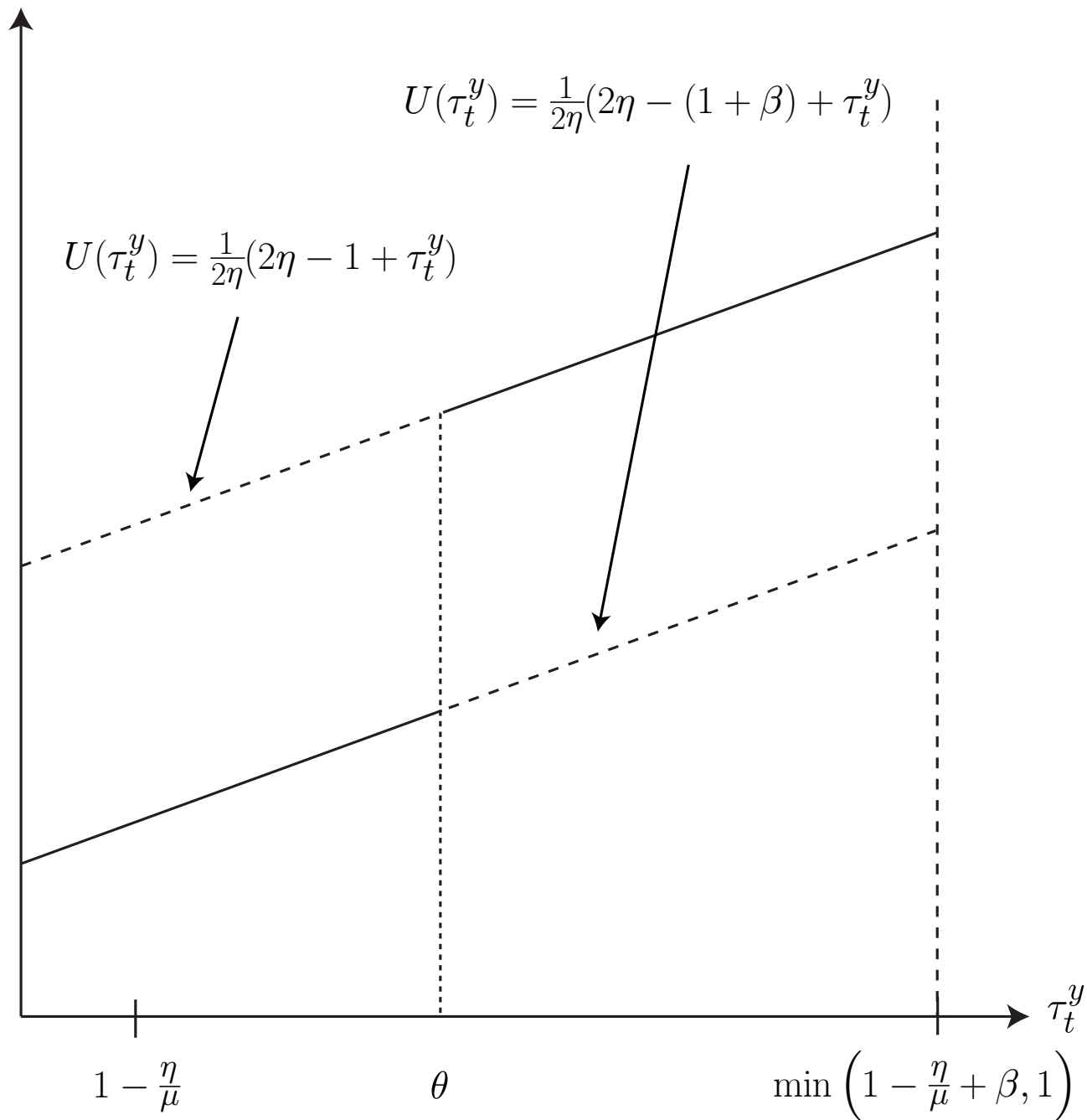


Figure 3: The figure illustrates an example of a monotonic function U for the case of $\frac{\eta}{\mu} \leq 1$.

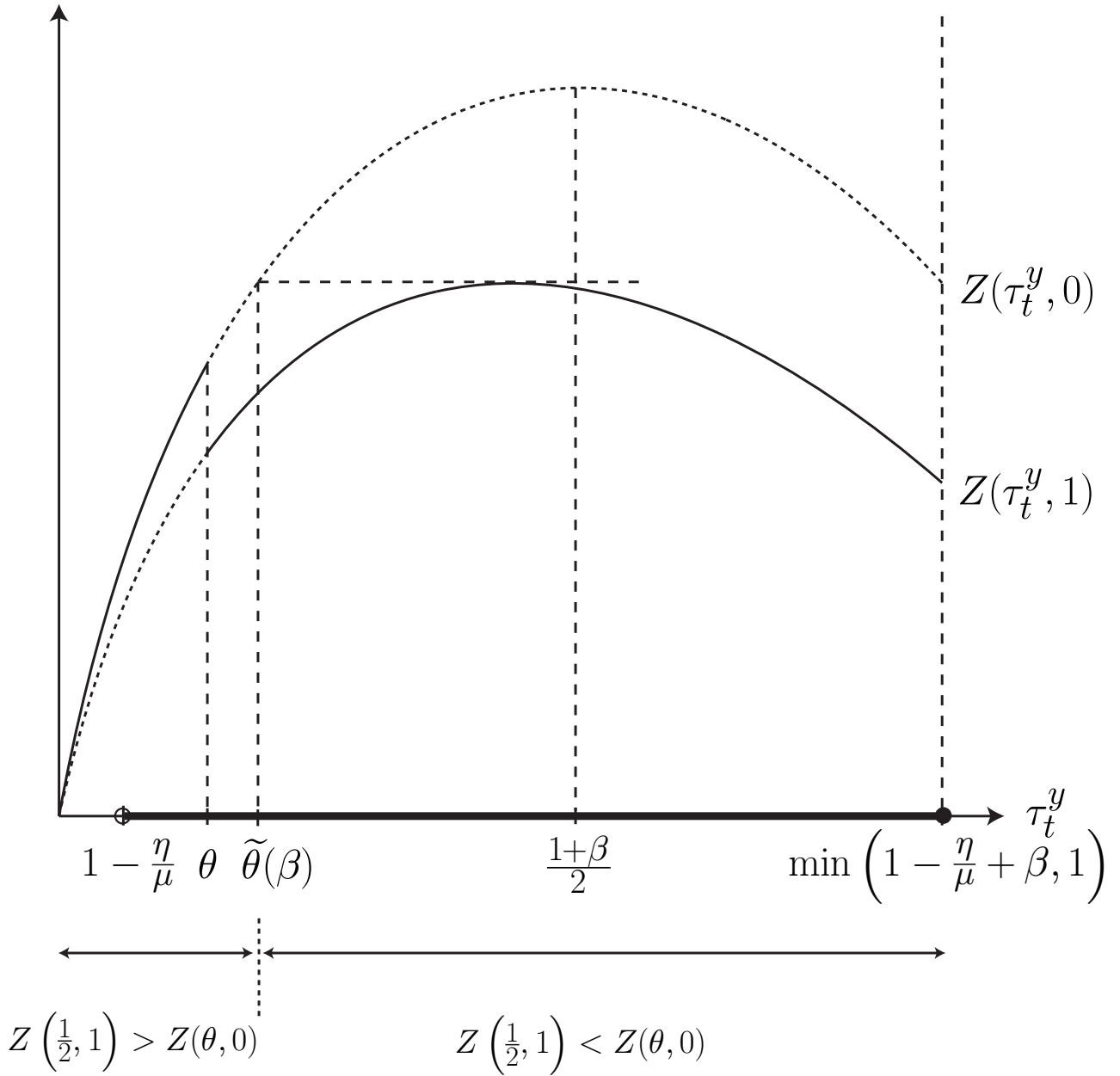


Figure 4: The solid curves illustrate the available tax revenue from the young, Z , under the assumption of $\beta < \frac{(\eta/\mu - 1/2)^2}{1 - \eta/\mu}$, i.e., $1 - \frac{\eta}{\mu} < \tilde{\theta}(\beta)$.

Fig.5 (a)

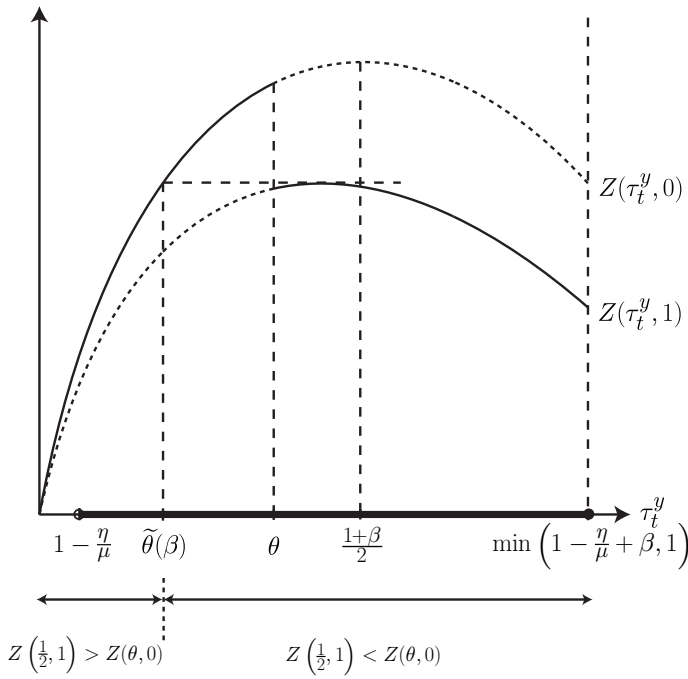


Fig.5 (b)

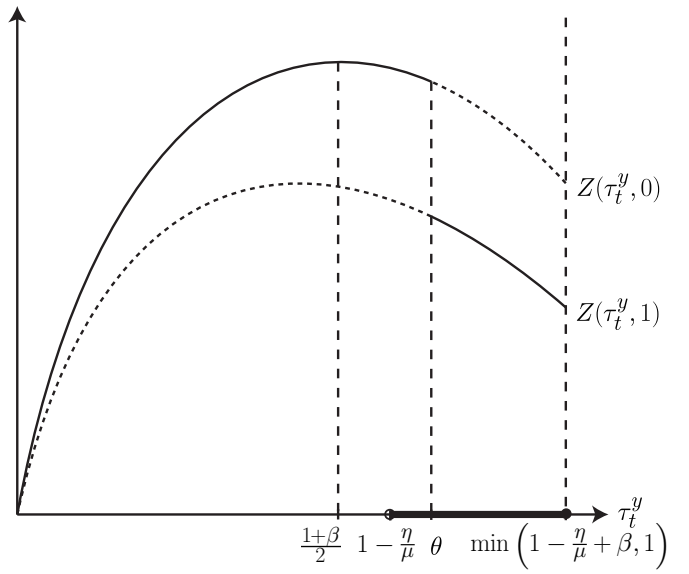


Figure 5: The solid curves illustrate the available tax revenue from the young, Z . The panel (a) represents the case of $\beta > 1 - 2\eta/\mu$, i.e., $1 - \eta/\mu < (1 + \beta)/2$. The panel (b) represents the case of $\beta \leq 1 - 2\eta/\mu$, i.e., $1 - \eta/\mu \geq (1 + \beta)/2$.

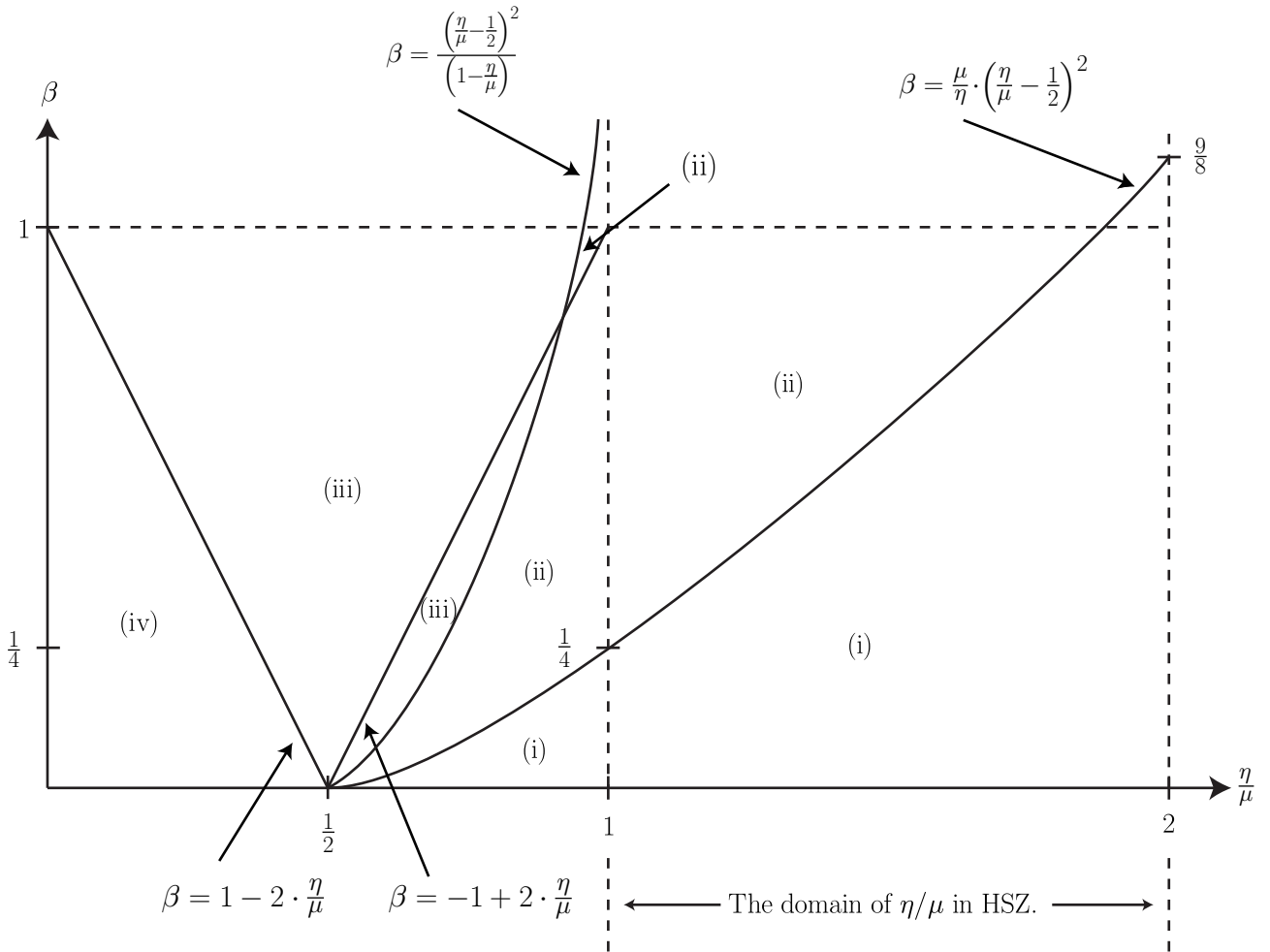


Figure 6: The figure displays the set of parameters (β, η, μ) classified according to the characterization of political equilibrium in Propositions 1 and 2. The areas (i), (ii), (iii) and (iv) satisfy the parameter conditions given in Proposition 3 (i), (ii), (iii) and (iv), respectively.