CCES Discussion Paper Series
Center for Research on Contemporary Economic Systems

Graduate School of Economics
Hitotsubashi University

CCES Discussion Paper Series, No.64
December 2016

Differential Income Taxation and Tiebout Sorting
(revised)

Takuya Obara
(Hitotsubashi University)

Naka 2-1, Kunitachi, Tokyo 186-8601, Japan
Phone: +81-42-580-9076 Fax: +81-42-580-9102
URL: http://www.econ.hit-u.ac.jp/~cces/index.htm
E-mail: cces@econ.hit-u.ac.jp

Differential Income Taxation and Tiebout Sorting*

Takuya Obara†

December 1, 2016

Abstract

This study examines optimal nonlinear income taxes when individuals differ in their preference for a public good and labor productivity. We consider two regions, of which one provides a higher quality public service than the other, thus inducing individuals to "vote with their feet." In addition, the government implements a region-specific income tax schedule to reflect the difference in benefits from the public service between the regions in the tax system. We show that if two characteristics are independently distributed and the first derivative of the social welfare function is strictly convex, the marginal tax rate in the region providing the higher quality public service is lower since the participation effect is greater than the mechanical effect. Further, we numerically find that the correlation can be substantial in differentiating income tax schedules, although labor mobility weakens the differentiation of marginal tax rates on the basis of a positive correlation between the two characteristics.

JEL Classification: H20, H41
Keywords: Extensive margin, Optimal nonlinear income taxation, Tagging

*I thank my supervisor Motohiro Sato for the invaluable comments and discussions. I also thank Nobuo Akai, Hiroki Arato, Shun-ichiro Bessho, Takeo Hori, Kaname Miyagishima, Hikaru Ogawa, and Yoshihiro Takamatsu as well as the seminar participants at Aoyama Gakuin University, the Summer Workshop on Economic Theory 2016 at Otaru University of Commerce, and the Economic Theory and Policy Workshop at Tokyo Metropolitan University. This study is financially supported by Grants-in-Aid for Scientific Research [Grant no. 26285065] and Hitotsubashi University.

†Graduate School of Economics, Hitotsubashi University, Naka 2-1, Kunitachi, Tokyo 186-8601, Japan. Email: ed142003@g.hit-u.ac.jp

1. Introduction

Since Mirrlees (1971) seminal work, the optimal income taxation model has considered a situation in which the government designs a redistributive tax system when individuals have private information in terms of their labor productivities. While their labor productivities are unobservable to the government, there are several individual characteristics that the government can observe, such as age, gender, and disability status, which are correlated with their labor productivities. Akerlof (1978) shows that the use of categorical information (also called "tagging") is welfare improving from the viewpoint of utilitarianism, since it allows redistribution not only within each tagged group but also between groups.[1] Therefore, if the government reflects observable characteristics that are correlated with abilities in the tax system, it can reinforce the redistributive tax system.

The objective of this study is to examine how income tax schedules should be differentiated between two regions with different amenities resulting from the quality of local public goods.[2] These regions can potentially be a tag. Bayer and McMillan (2012) show that heterogeneity in housing characteristics, including local public goods such as school quality and crime, lead to increases in income stratification. Verdugo (2016) investigates how a policy allowing immigrants with children to live in public housing in France affects their location choices and shows that cities with higher public housing stocks attract more low-skilled immigrants, implying that spatial differences in public housing may cause income stratification. Using a panel dataset of European regions from 17 countries, Kessing and Strozzi (2016) empirically find that the level of public employment is significantly higher in low productivity regions. Therefore, the difference in the quality of public goods across regions is useful as a tag that is correlated with income levels.

We consider an economy that comprises individuals who differ in their preference for a public good and labor productivity. Individuals make a labor supply decision on the basis of nonlinear income taxes and a binary one regarding which region to live in

[1] It is well known that tagging violates the principle of horizontal equity and therefore, is limited in practice. However, Weinzierl (2014) uses the equal sacrifice principle as a comprehensive criterion in that tagging, not the horizontal equity principle, is limited and shows that tagging is justified because the deviation from the equal sacrifice principle is small when observable characteristics are strongly correlated with abilities.

[2] Our study does not consider tax competition among governments since we suppose that the responsibility for redistributive taxation is devolved to a supranational government, such as the European Union. Given that the idea of deeper fiscal integration is suggested in the European policy agenda, Bargain et al. (2013) estimate the effect of replacing with an integrated tax and transfer system on redistribution and fiscal stabilization. Kessing et al. (2015) theoretically examine the optimal nonlinear income tax schedules in each member state designed by an integrated government. In contrast, Morelli et al. (2012), Bierbrauer et al. (2013), and Lehmann et al. (2014) analyze the nonlinear income tax competition between two governments.

without incurring mobility costs.[3] The difference in tax burdens and amenities from local public goods between two regions separates the population into two categories: one that obtains more benefits from a public good and has higher tax burdens and the other that gains lower benefits and has lower tax burdens. The government designs differential income tax schemes for the two regions to maximize social welfare while taking account of the labor supply decision on the intensive margin and participation decision on the extensive margin.

First, we examine the case in which the government is allowed to implement a lump-sum transfer between two regions, although it cannot differentiate marginal income tax rates. We characterize optimal marginal income tax rates and optimal level of lump-sum transfer. The former result is similar to that derived in Mirrlees (1971). The latter result is characterized as the Ramsey inverse elasticity rule and the direction of transfer is determined by the government's redistributive tastes and the correlation between two characteristics. In particular, if there is no correlation between the characteristics, the inter-regional transfer from the region with higher quality public goods to that with lower quality public goods is desirable. Second, we allow the government to introduce the differentiation of marginal tax rates into the tax system and find that the shape of optimal income tax schedules crucially depends on the government's redistributive tastes and the correlation between two characteristics. Using a tax perturbation method, we analytically demonstrate that, if public goods preferences and labor productivities are independently distributed and the first derivative of the social welfare function is strictly convex, the marginal income tax rate for individuals who receive higher amenities from local public goods is lower. This is because the decrease in tax burdens on the region with higher quality public goods leads to a greater welfare gain generated by inducing individuals to access the region (participation effect) than the welfare loss done by the decrease in tax receipts (mechanical effect). Therefore, introducing the differentiation of marginal income tax rates reinforces redistribution. Further, we numerically assess whether our tax perturbation method is reasonable to understand the shape of optimal income schedules and present the implication of introducing the correlation between two characteristics.

This study draws from the growing body of literature examining separated income tax schedules for groups divided by observable characters, or the so-called "tagging" (e.g., Akerlof (1978), Immonen et al. (1998), Viard (2001), Boadway and Pestieau (2006), Cremer et al. (2010), Mankiw and Weinzierl (2010)). In particular, Boadway and Pestieau (2006) and Cremer et al. (2010) analytically examine optimal income taxation with tagging in an economy comprising two groups, of which one has a higher

[3] Our study is part of large body of papers dealing with optimal nonlinear income taxes in random participation models with multidimensional heterogeneity (e.g., Jacquet et al. (2013), Rothchild and Scheuer (2013, 2016), Lehmann et al. (2014), Blumkin et al. (2015)). These studies assume individuals differ in ability and other characteristics, such as migration cost or work cost, and do not allow the government to separate income tax schedules. By contrast, this study investigates tagging assuming individuals differ in ability and public goods preferences.

proportion of high-ability individuals, and conclude that, in this case, the tax system with inter-group transfers will be more redistributive compared to standard optimal taxation model of Mirrlees (1971) and Saez (2001). The crucial difference is that we consider a variable category as the tag, which means that individuals engage in decision making on the extensive and intensive margin. In other words, the government must pay attention to two types of distortion in individual labor supply when implementing income taxation. By contrast, much of the previous literature supposes that a tagged group is immutable, that is, individuals respond along the intensive margin only.[4] Therefore, we aim to explore how responses along the extensive margin affect the differentiation of income taxes.

This paper is not the first to examine differential income taxation in a variable category (e.g., Kleven et al. (2006, 2009), Gomes et al. (2014), Kessing et al. (2015)). Our study is closely related to Kleven et al. (2006, 2009), who examine how the government should differentiate income tax schedules, considering whether the spouse works as a tag, and numerically investigate the impact of introducing the correlation between ability and work cost on the tax system.[5] They show that the household in which the spouse (does not) works faces lower (higher) marginal tax rates under no correlation between two characteristics and that introducing the correlation is not significant, that is, the theoretical result does not overturn. Our study differs in two ways from their framework. First, the applications of our findings pertain to the design of the optimal inter-regional transfer program related to the difference in the quality of public goods between regions. Second, it attempts to clarify the difference in tagging between immutable and variable categories and shows that the government must take account of the participation effect in addition to the mechanical effect. We numerically demonstrate that the participation effect caused by labor mobility decreases (increases) the marginal tax rate in region A (region B), which implies that it weakens the differentiation of marginal tax rates on the basis of a positive correlation. However, compared to Kleven et al. (2006, 2009), we find that differentiation due to the mechanical effect slightly remains if the correlation between characteristics is strong, that is, the correlation can be an important parameter as in previous studies examining tagging on immutable categories. To the best of our knowledge, there is no theoretical model that

[4]Indeed, previous works have considered demographic characteristics such as age and gender or health conditions including illnesses or disabilities as observable characters. In this case, individuals do not make decisions along the extensive margin since they cannot change groups.

[5]Gomes et al. (2014) examine the effect of sector-specific income taxes on production efficiency when individuals with sector-specific abilities choose a sector to work in. They characterize a sufficient condition in which production inefficiency is optimal, although they do not investigate how income tax schemes are differentiated across two sectors. On the other hand, Kessing et al. (2015) investigates differential income taxation on two regions from the viewpoint of the central government, as in the present study. They assume that one is the more productive region, that is, if individuals live in this region, their productivities are enhanced, and numerically find that the shape of optimal differential income taxation dramatically changes in response to migration elasticity. However, they ignore the correlation between two characteristics in numerically analyzing optimal marginal income tax rates.

elucidates the gap in policy implications for immutable and variable categories in the presence of the correlation.

This remainder of this paper is organized as follows. Section 2 describes the framework of the basic model. Section 3 characterizes differential income tax schedules with non-differentiated marginal tax rates, which is the benchmark result in our study. Section 4 shows that it is desirable to introduce differentiated marginal tax rates and section 5 presents the numerical results. Section 6 offers concluding remarks.

2. Model

We consider an economy consisting of individuals who are characterized by the preference for a public good and labor productivity denoted by θ and w. The two types of characteristics (θ, w) are distributed according to the cumulative distribution function $F(\theta, w)$ with the strictly positive and continuously differentiable density function $f(\theta, w)$ over $[\underline{\theta}, \overline{\theta}] \times [\underline{w}, \overline{w}]$. We assume that $0 = \underline{\theta} < \overline{\theta} < \infty$ and $0 < \underline{w} < \overline{w} < \infty$. The size of population is normalized to 1. We consider two regions indexed by $i = A, B$ and there is an exogenous and a same type of public good in each region.[6] The quality of public goods is denoted by G_i. Without loss of generality, indicator A represents a region in which there is a higher quality public good and indicator B a region in which there is a lower quality public good, that is, $G_A > G_B$. In the model, we suppose local public goods that satisfy non-excludability and rivalness such as city parks, roadways, education, or health services. As used in Diamond (1998), the utility function of individuals in region i is described by

$$U_i = \theta G_i + x_i - v(\ell_i) \tag{1}$$

where x_i denotes the private consumption of individuals in region i, and ℓ_i is the labor supply of individuals in region i. On the other hand, $v(\cdot)$ denotes the disutility of labor supply and is strictly increasing, strictly convex, and continuously differentiable.

The government can observe the labor income of individuals in region i, denoted by $z_i \equiv w\ell_i$, and thus, can levy nonlinear income taxes depending on each region, denoted by $T_i(z_i)$. The budget constraint which individuals in region i face is given by $x_i = z_i - T_i(z_i)$.

[6] For simplicity, we assume that public goods are exogenous. Despite doing so, the change in the assumption does not affect our main conclusion and the provision rule of public goods is expressed by the modified Samuelson condition. Therefore, whether public goods are endogenously determined is not of significance.

2.1 Intensive margin

Individuals with type vector (θ, w) in region i choose the amount of labor supply by solving the following optimization problem:

$$max_{\ell_i}\ U_i = \theta G_i + w\ell_i - T_i(w\ell_i) - v(\ell_i)$$

The first-order condition yields

$$\frac{v'(\ell_i)}{w} = 1 - T_i'(w\ell_i) \qquad \forall w \tag{2}$$

where $v'(\cdot) \equiv \frac{\partial v}{\partial \ell}$ denotes the marginal disutility of labor.

Let us denote the indirect utility function of individuals in region A by $\theta G_A + V_A(w) \equiv \theta G_A + x_A(w) - v(\ell_A(w))$ and those in region B as $\theta G_B + V_B(w) \equiv \theta G_B + x_B(w) - v(\ell_B(w))$, where $x_i(w)$ and $\ell_i(w)$ are the private consumption and labor supply of individuals in region i with labor productivity w.

We define the elasticity of labor supply with respect to the net-of-tax wage rate $1 - T_i'$ as

$$\epsilon_i \equiv \frac{1 - T_i'}{\ell_i} \frac{\partial \ell_i}{\partial 1 - T_i'} = \frac{v'(\ell_i)}{\ell_i v''(\ell_i)} \tag{3}$$

From optimized individual behavior (equation (2)), we have $\epsilon \equiv \epsilon_A = \epsilon_B$ if $T_A' = T_B'$.

2.2 Extensive margin

Individuals choose a region to live in without migration cost, which means that labor is perfectly mobile. Individuals with type vector (θ, w) obtain utility $\theta G_A + V_A(w)$ if they have access to region A and utility $\theta G_B + V_B(w)$ if they access region B. Therefore, they access region A if and only if

$$\theta \geq \frac{V_B(w) - V_A(w)}{\Delta G} \equiv \hat{\theta}(w) \tag{4}$$

where $\Delta G \equiv G_A - G_B$ denotes the difference in the quality of public goods between regions. We interpret $\hat{\theta}(w)$ as the net gain from living in region B. This means if the preference for the public good by individuals with labor productivity w is greater (lower) than the threshold $\hat{\theta}(w)$, they (do not) access region A.

Here, we denote the entire labor productivity and preference for a public good density by $f(w)$ and $f(\theta)$. If θ and w are independently distributed, the density of joint distribution $f(\theta, w)$ is expressed by $f(w)f(\theta)$. For each labor productivity, the conditional density of the preference for a public good in region A is $f_A^c(w) \equiv \int_{\hat{\theta}(w)}^{\overline{\theta}} f(\theta|w)d\theta$ and that in region B is $f_B^c(w) \equiv \int_{\underline{\theta}}^{\hat{\theta}(w)} f(\theta|w)d\theta$. Therefore, the skill density in region

i is $f_i^c(w)f(w)$ denoted by $f_i(w)$ and the corresponding cumulative distribution function is $\int_{\underline{w}}^{w} f_i(x)dx$ denoted by $F_i(w)$. The entire population in region i is $\int_{\underline{w}}^{\overline{w}} f_i(w)dw$ denoted by N_i.

2.3 Government

The budget constraint of the government takes the following form:

$$\int_{\underline{w}}^{\overline{w}} T_A(z_A(w))f_A(w)dw + \int_{\underline{w}}^{\overline{w}} T_B(z_B(w))f_B(w)dw = \phi(G_A, N_A) + \phi(G_B, N_B) \quad (5)$$

Each term on the left-hand side represents the aggregate revenue from income taxes imposed on individuals in region i. On the other hand, $\phi(\cdot)$ is a strictly increasing, strictly convex, and continuously differentiable cost function of a public good that captures not only provision cost but also congestion cost.[7] Here, we define $\frac{\partial \phi}{\partial N_i} \equiv \phi_{N_i}$ as the marginal congestion cost.

We focus on the Bergson-Samuelson criterion, which is represented as follows:

$$\mathcal{W} \equiv \int_{\underline{w}}^{\overline{w}} \left[\int_{\hat{\theta}(w)}^{\bar{\theta}} W(\theta G_A + V_A(w))f(\theta, w)d\theta + \int_{\underline{\theta}}^{\hat{\theta}(w)} W(\theta G_B + V_B(w))f(\theta, w)d\theta \right] dw \quad (6)$$

where W is a strictly increasing and concave function, that is, $W' > 0$ and $W'' < 0$.

In the second best environment, the government cannot observe labor productivity, which is individuals' private information. As per the revelation principle, it suffices to induce individuals to reveal their true types of labor productivity to maximize the objectives of the government. As shown in Mirrlees (1971), the first-order incentive compatibility constraint in region i is given by[8]

$$V_i'(w) = \frac{\ell_i(w)}{w} v'(\ell_i(w)) \qquad \forall w \quad (7)$$

This is the necessary condition to meet the incentive constraint. Hereafter, we assume that the sufficient condition is satisfied, that is, the Spence-Mirrlees condition and monotonicity conditions hold.

Before investigating the property of the optimal tax system, it is useful to define the marginal social welfare weight for individuals with labor productivity w in region i

[7] To express the congestion effect, the functional form is followed by McGuire (1974) model. The assumption means that the number of residents causes the production effect. On the other hand, Buchanan (1965) model assumes that the number of residents directly affects the perceived amount of public good, that is, the congestion effect can be observed in the utility function.

[8] Since the government can observe the region in which individuals live, they cannot mimic those in the other region. Therefore, we consider only the incentive constraint within a region.

denoted by $g_i(w)$.

$$g_A(w) \equiv \frac{\int_{\hat{\theta}(w)}^{\overline{\theta}} W'(\theta G_A + V_A(w))f(\theta|w)d\theta}{\gamma f_A^c(w)}, \ g_B(w) \equiv \frac{\int_{\underline{\theta}}^{\hat{\theta}(w)} W'(\theta G_B + V_B(w))f(\theta|w)d\theta}{\gamma f_B^c(w)}$$

g_i measures the relative value of the government that gives an additional 1\$ to individuals with labor productivity w in region i. Thus, if the government has redistributive tastes, g_i is decreasing in w, which allows income tax schedules to be progressive in region i. Moreover, as shown later, these parameters are crucially related to the optimal redistribution between two regions as well as within each region.

3. Non-differentiated marginal tax rates

First, we illustrate the benchmark case in which the government designs differential income tax schedules with the same marginal tax rates, that is, $T' \equiv T'_A = T'_B$. In this case, we allow the government to make the lump-sum transfer E within two regions, where $E \equiv T_A - T_B$ and E is constant in w. Before characterizing the optimal tax policy, we show that it suffices to satisfy the following constraints to solve the optimization problem.

It is sufficient to meet either the incentive constraint in region A or B. Under the same marginal tax rates, the labor supply of individuals in region A is the same as that in region B from equation (2), that is, $\ell \equiv \ell_A = \ell_B$ and $z \equiv z_A = z_B$. Therefore, each incentive constraint coincides. Without loss of generality, we take account of the incentive constraint in region B, that is,

$$V'_B(w) = \frac{\ell(w)}{w} v'(\ell(w)) \qquad \forall w \tag{8}$$

Second, the threshold $\hat{\theta}(w)$ becomes constant in w. From the definition of $\hat{\theta}(w)$, the first derivative of $\hat{\theta}(w)$ is $\frac{V'_B(w) - V'_A(w)}{\Delta G}$. Since $V'_A(w) = V'_B(w)$ holds from the incentive constraint given that $\ell_A = \ell_B$, $\hat{\theta}(w)$ takes a constant value defined as $\hat{\theta}$. This result allows for a further interpretation of equation (4). In this case, equation (4) can be rewritten as

$$\theta \Delta G \geq E = \hat{\theta} \Delta G \tag{9}$$

That is, individuals prefer to access region A if benefit $\theta \Delta G$ they draw from the additional enjoyment of a public good exceeds additional taxes E.

Finally, using $E = \hat{\theta}\Delta G$, budget constraint (5) can be rewritten as follows:

$$\int_{\hat{\theta}}^{\overline{\theta}} \hat{\theta} \Delta G f(\theta) d\theta + \int_{\underline{w}}^{\overline{w}} T_B(z(w)) f(w) dw = \phi(G_A, N_A) + \phi(G_B, N_B) \tag{10}$$

In addition, substituting $V_A = -\hat{\theta}\Delta G + V_B$ into social welfare function (6), the following denoted by $\hat{\mathcal{W}}$ is obtained:

$$\hat{\mathcal{W}} \equiv \int_{\underline{w}}^{\overline{w}} \left[\int_{\hat{\theta}}^{\overline{\theta}} W([\theta - \hat{\theta}]\Delta G + \theta G_B + V_B(w)) f(\theta, w) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W(\theta G_B + V_B(w)) f(\theta, w) d\theta \right] dw \tag{11}$$

In sum, the government faces with the problem of choosing $V_B(w)$, $\ell(w)$, and $\hat{\theta}$ to maximize social welfare function (11) subject to budget constraint (10) and incentive constraint (8):

$$\begin{aligned} \max_{V_B(w), \ell_B(w), \hat{\theta}} \quad \hat{\mathcal{W}} \quad s.t. \quad & V_B'(w) = \frac{\ell(w)}{w} v'(\ell(w)) \quad and \\ & \int_{\hat{\theta}}^{\overline{\theta}} \hat{\theta}\Delta G f(\theta) d\theta + \int_{\underline{w}}^{\overline{w}} T_B(z(w)) f(w) dw = \phi(G_A, N_A) + \phi(G_B, N_B) \end{aligned} \tag{12}$$

The corresponding Lagrangian is

$$\begin{aligned} \mathcal{L} = & \hat{\mathcal{W}} + \int_{\underline{w}}^{\overline{w}} \lambda(w) \left[\frac{\ell(w)}{w} v'(\ell(w)) - V_B'(w) \right] dw \\ & + \gamma \left[\int_{\hat{\theta}}^{\overline{\theta}} \hat{\theta}\Delta G f(\theta) d\theta + \int_{\underline{w}}^{\overline{w}} T_B(z(w)) f(w) dw - \phi(G_A, N_A) - \phi(G_B, N_B) \right] \end{aligned} \tag{13}$$

where γ is the Lagrangian multiplier in the resource constraint and $\lambda(w)$ is the co-state variable in the incentive constraint.

The first-order conditions are given in Appendix A. Rearranging the first-order conditions, we can obtain the following.

Proposition 1. *Under non-differentiated marginal tax rates, the optimal marginal income tax rate and optimal level of lump-sum transfer are characterized by*

$$\frac{T'(z(w))}{1 - T'(z(w))} = \left[1 + \frac{1}{\epsilon} \right] \frac{1}{w f(w)} \int_{w}^{\overline{w}} [1 - \overline{g}(x)] f(x) dx \tag{14}$$

$$\frac{E - (\phi_{N_A} - \phi_{N_B})}{E} = \frac{1}{\eta N_A} \left[N_A \int_{\underline{w}}^{\overline{w}} g_B(w) f_B^c(w) f(w) dw - N_B \int_{\underline{w}}^{\overline{w}} g_A(w) f_A^c(w) f(w) dw \right] \tag{15}$$

where $\overline{g}(x) \equiv f_A^c(x) g_A(x) + f_B^c(x) g_B(x)$ is the average social marginal welfare weight for individuals with labor productivity w and $\eta \equiv -\frac{\partial 1 - F(\frac{E}{G})}{\partial E} \frac{E}{1 - F(\frac{E}{G})}$ is the migration elasticity with respect to E in region A.

These derivations are also included in Appendix A. Equation (14) is the traditional formula for optimal marginal income tax rate obtained by Mirrlees (1971) under no income effect. The heuristic derivation is followed by Saez (2001).

Equation (15) is the Ramsey inverse elasticity rule in terms of lump-sum transfers. The amount of lump-sum transfers charged is determined by two main terms. First, the elasticity of demands for additional taxes η in the denominator represents distortions, that is, a decrease in individuals accessing region A, created by imposing additional taxes. If η is highly inelastic, the level of lump-sum transfers tends to increase. Second, the numerator expresses the net welfare gains from the redistribution between regions and the first and second terms in the numerator describe the government's redistributive tastes for each region. If the government prefers to redistribute from region A to B, that is, the first term in the bracket on the right-hand side is greater than the second term, additional taxes are charged above the marginal congestion cost to increase revenues from lump-sum transfers and raise consumption levels. Therefore, whether the level of lump-sum transfer deviates from net marginal congestion cost $\phi_{N_A} - \phi_{N_B}$ crucially depends on the sign of the numerator. Since the Ramsey formula above is very general, making the sign of the numerator ambiguous, we present a special case in which the sign is determined by placing assumptions on the correlation between two characteristics.

3.1 Heuristic derivation and interpretation of the Ramsey inverse elasticity formula

Here, we provide the heuristic derivation for equation (15) to help with intuition. We suppose a situation in which the government marginally increases additional taxes E. Let dE be a small tax reform for the lump-sum transfer. First, a small reform, such that E increases, distorts the decision making on the extensive margin. That is, individuals with lower preferences for a public good tend to access region B, which amounts to the size of $f(\hat{\theta})d\hat{\theta}$. Therefore, revenues from additional taxes E decrease. In addition, the decrease in public good users from region A mitigates net marginal congestion cost $\phi_{N_A} - \phi_{N_B}$. As a result, we can express the participation effect denoted by dP as follows:

$$dP = -(E - (\phi_{N_A} - \phi_{N_B}))f(\hat{\theta})d\hat{\theta}$$

Moreover, using $dE = d\hat{\theta} \cdot \Delta G$ obtained from equation (9), gives us

$$dP = -\frac{E - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta})dE$$

Therefore, the participation effect exhibits a net efficiency loss from imposing additional taxes. Second, a small perturbation that uniformly increases additional taxes E affects tax revenues from income taxes from region A without behavioral responses and the net mechanical effect denoted by dM is measured as follows:

$$dM = \int_{\underline{w}}^{\overline{w}} (1 - g_A(x)) f_A(x) dx \times dE$$

Rearranging this and then substituting equation (A.13) in Appendix A yields

$$\begin{aligned} dM &= \frac{1}{\gamma}\left[\gamma(1 - F(\hat{\theta})) - \int_{\underline{w}}^{\overline{w}} \int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta\right] \times dE \\ &= \left[N_A \int_{\underline{w}}^{\overline{w}} g_B(w) f_B^c(w) f(w) dw - N_B \int_{\underline{w}}^{\overline{w}} g_A(w) f_A^c(w) f(w) dw\right] \times dE \end{aligned}$$

The increase in additional taxes E amounts to revenue $N_A dE$, which increases the level of private consumption by $N_A dE$ units. Therefore, the first term on the right-hand side is the welfare gain from an increase in the private consumptions of individuals in region B. On the other hand, although the tax burdens of individuals in region A decrease $N_A dE$ units, the level of private consumptions decreases $N_B dE$ units since they are levied dE. As a result, the second term on the right-hand side represents the welfare loss from the decrease in the private consumptions of individuals in region A. That is, the term on the right-hand side is interpreted as the net welfare gain from redistribution. In sum, we must have $dP + dM = 0$ at the optimum, which leads to equation (15). Put differently, equation (15) implies an *equity − efficiency tradeoff*.

3.2 Special cases for the Ramsey inverse elasticity formula

The determinants for whether additional taxes should be charged above the marginal congestion cost are the correlation between θ and w and the government's redistributive tastes, as seen in equation (15). However, we do not know the direction of the optimal tax policy in general since the formula is complicated.

Here, we assume that θ and w are independently distributed. In this case, equation (15) is transformed as follows:

$$\frac{E - (\phi_{N_A} - \phi_{N_B})}{E} = \frac{N_B}{\eta} \int_{\underline{w}}^{\overline{w}} (g_B(w) - g_A(w)) f(w) dw \tag{16}$$

If social welfare is a strictly concave function as in equation (6), the sign of the equation is positive because $g_B(w) - g_A(w)$ is positive for any labor productivities given the concavity of the social welfare function. In other words, the redistribution from region A to B causes net welfare gains. Therefore, we can summarize the statement as follows:

Corollary 1. *If θ and w are independently distributed, the level of lump-sum transfer exceeds the net marginal congestion cost.*

4. Differentiated marginal tax rates

In this section, we examine the effect of introducing the differentiation of marginal income tax rates between two regions. If the government is able to design differential income tax schedules with differentiated marginal tax rates, it faces with the problem of choosing $V_i(w)$, $\ell_i(w)$ for $i = A, B$, and $\hat{\theta}(w)$ to maximize social welfare function (6) subject to budget constraint (5), incentive constraints (7), and participation constraint (4). Therefore, the optimization problem is formulated as follows:

$$\max_{V_i(w),\ell_i(w),\hat{\theta}(w)} \quad \mathcal{W} \quad s.t. \quad V_i'(w) = \frac{\ell_i(w)}{w} v'(\ell_i(w)), \quad \hat{\theta}(w)\Delta G + V_A(w) = V_B(w) \quad and$$
$$\int_{\underline{w}}^{\overline{w}} T_A(z_A(w)) f_A(w) dw + \int_{\underline{w}}^{\overline{w}} T_B(z_B(w)) f_B(w) dw = \phi(G_A, N_A) + \phi(G_B, N_B) \tag{17}$$

The corresponding Lagrangian is

$$\mathcal{L} = \mathcal{W} + \gamma \left[\int_{\underline{w}}^{\overline{w}} T_A(z_A(w)) f_A(w) dw + \int_{\underline{w}}^{\overline{w}} T_B(z_B(w)) f_B(w) dw - \phi(G_A, N_A) - \phi(G_B, N_B) \right]$$
$$+ \sum_{i=A,B} \int_{\underline{w}}^{\overline{w}} \lambda_i(w) \left[\frac{\ell_i(w)}{w} v'(\ell_i(w)) - V_i'(w) \right] dw + \int_{\underline{w}}^{\overline{w}} \mu(w) \left[\hat{\theta}(w)\Delta G + V_A(w) - V_B(w) \right] dw \tag{18}$$

where γ is the Lagrangian multiplier on the resource constraint, $\lambda_i(w)$ is the co-state variable associated with the incentive constraint in region i, and $\mu(w)$ is the co-state variable associated with the participation constraint. The first-order conditions are given in Appendix B and the optimal marginal income tax rate for each region is derived by rearranging them.

Proposition 2. *The optimal marginal income tax rate for each region is characterized by*

$$\frac{T_A'(z_A(w))}{1 - T_A'(z_A(w))} = \left[1 + \frac{1}{\epsilon_A} \right] \cdot \frac{1}{w f_A(w)} \cdot \int_w^{\overline{w}} \left[(1 - g_A(x)) f_A^c(x) - \Phi(x) \right] f(x) dx \tag{19}$$

$$\frac{T_B'(z_B(w))}{1 - T_B'(z_B(w))} = \left[1 + \frac{1}{\epsilon_B} \right] \cdot \frac{1}{w f_B(w)} \cdot \int_w^{\overline{w}} \left[(1 - g_B(x)) f_B^c(x) + \Phi(x) \right] f(x) dx \tag{20}$$

where $\Phi(x) \equiv \frac{T_A(z_A(x)) - T_B(z_B(x)) - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(x)|x)$

These formulas describe the optimal differentiated marginal income tax rate for each region. This result is consistent with those in the existing literature (Kleven et al. (2006, 2009), Kessing et al. (2015)). In contrast with the optimal tax rate on the basis of an immutable tag, the novel effect $\Phi(\cdot)$ appears, which negatively (positively) works for tax rates on individuals in region A (region B) if $\Phi(\cdot)$ is positive. The term consists of two terms: $T_A(z_A(w)) - T_B(z_B(w))$ and $\phi_{N_A} - \phi_{N_B}$. The first term expresses the additional tax revenue obtained by inducing individuals to access region A. Thus, if this term is positive, the government intends to decrease (increase) the marginal tax rates in region A (region B) to attract people to the region. The second term is the net marginal congestion cost $\phi_{N_A} - \phi_{N_B}$ for the government, which differs from the previous literature examining variable categories. If the government decreases the marginal tax rate in region A as an incentive to live in region A, efficiency loss occurs in region A owing to the congestion cost. On the other hand, efficiency gain occurs in region B due to population outflow. This mechanism reflects the net marginal congestion cost $\phi_{N_A} - \phi_{N_B}$. If it is positive, this implies that the government increases (decreases) the marginal tax rates in region A (region B) to mitigate the congestion cost in total. As a result, even if the government can extract additional tax revenue inducing individuals to access region A, it must determine income tax schedules while taking account of the congestion cost.

The social welfare criterion affects the differentiation of the marginal income tax rate. If the government has distributional concerns, government redistributive tastes for region A is estimated to be lower than those for region B because the utility of individuals in region A is higher than that of individuals in region B. That is, from the concavity of the social welfare function, we have $g_B > g_A$. Therefore, the government intends to redistribute income from region A to region B by imposing more income taxes on individuals in region A.

Furthermore, it is shown that the property of optimal marginal income tax rates under tax systems with tagging obtained by Cremer et al. (2010) holds.

Corollary 2. *(i) The optimal marginal income tax rate with non-differentiation incurred by individuals with labor productivity w is bracketed by the optimal marginal income tax rate with differentiation incurred by individuals with labor productivity w in each region:*

$$\frac{T'(z(w))}{1-T'(z(w))} = \frac{T'_A(z_A(w))}{1-T'_A(z_A(w))}\frac{f_A(w)}{f(w)} + \frac{T'_B(z_B(w))}{1-T'_B(z_B(w))}\frac{f_B(w)}{f(w)} \tag{21}$$

(ii) If $T'_h(z_h(w)) > T'_j(z_j(w))$, $T'_h(z_h(w)) > T'(z(w)) > T'_j(z_j(w))$, where $h, j = A, B$ and $h \neq j$

The first result is that the marginal income tax rate with non-differentiation coincides with a weighted average of the marginal income tax rate for individuals in each region. The second result is that, if applying differentiated marginal tax rates, higher marginal income tax rates is imposed on one and a lower marginal income tax rate on the other compared to the marginal income tax rate with non-differentiation. While Cremer et al. (2010) present these results under Rawlsian preferences, Corollary 2 indicates that their findings hold under the Bergson-Samuelson criterion.

4.1 Direct proof of optimal differentiated marginal tax rate

We present an intuitive interpretation of formulas in Proposition 2 by characterizing optimal marginal nonlinear income tax rates by means of direct derivation as in Saez (2001). We consider a situation in which the government marginally increases the marginal income tax rates for individuals in region A whose income levels are distributed over $[z_A, z_A + dz_A]$, denoted by dT'_A. This small tax reform causes the following three effects: mechanical, behavioral, and participation effect.

4.1.1 Mechanical effect

The rise in marginal income tax rates increases tax receipts without behavioral responses. Since individuals in region A with labor productivity above w_{z_A} must pay the additional payment $dT'_A \times dz_A$, the added net tax receipts amount to

$$\varrho_M^A \equiv \int_{w_{z_A}}^{\overline{w}} (1 - g_A(x)) f_A(x) dx \times dT_A \tag{22}$$

where w_{z_A} is the ability of individuals in region A who earn labor income z_A.

4.1.2 Behavioral effect

The change in the marginal income tax rate distorts decision making in terms of labor supply. If the marginal income tax rates increase, the tax base decreases by the reduction of labor supply. Thus, a decrease in tax receipts occurs due to behavioral responses. To measure this effect, we rearrange the change in labor income owing to a small change in the marginal income tax rates, denoted by $\frac{dz_A}{dT'_A}$ as follows:

$$dz_A = -\frac{z_A}{1 - T'_A(z_A)} \times \epsilon_A \times dT'_A \tag{23}$$

Substituting equation (23) with $dT_A(z_A) = T'_A(z_A) dz_A$ yields

$$dT_A(z_A) = -T'_A(z_A) \frac{z_A}{1 - T'_A(z_A)} \times \epsilon_A \times dT'_A \tag{24}$$

Let ϱ_B^A be the total reduction of tax receipts from region A brought about by a behavioral effect. Thus, ϱ_B^A is equal to $dT_A(z_A) \times f_A(w_{z_A}) d\hat{w}$ because individuals whose skill levels are within the interval $[w_{z_A}, w_{z_A} + d\hat{w}]$ are affected by the change in marginal tax rates. Given that $d\hat{w} = \frac{dz_A}{(1+\epsilon_A)\ell}$ is derived using equation (2), we have

$$\varrho_B^A = -\frac{T_A'(z_A)}{1 - T_A'(z_A)} \times \frac{\epsilon_A}{1+\epsilon_A} \times w_{z_A} f_A(w_{z_A}) \times dT_A \tag{25}$$

4.1.3 Participation effect

Unlike in the traditional literature examining differential income taxation on immutable categories, our model considers a variable category as a tag. The increase in marginal tax rates induces individuals in region A with $x \geq w_{z_A}$ such that the number of switchers amounts to $f(\hat{\theta}(x), x)d\hat{\theta}(x)$ to drop out of the access to region A. Because their payments change from $T_A(z_A)$ to $T_B(z_B)$, the government's revenue decreases by $T_A(z_A) - T_B(z_B)$ units. In addition, the decrease in the number of individuals in region A alleviates the congestion cost in region A and augments that in region B, measured by $\phi_{N_A} - \phi_{N_B}$. Therefore, a net effect on tax revenues is equal to $-(T_A(z_A) - T_B(z_B)) + (\phi_{N_A} - \phi_{N_B})$. Using $d\hat{\theta}(x) \cdot \Delta G = dT_A$, the total effect on tax receipts is as follows:

$$\varrho_P^A \equiv -\int_{w_{z_A}}^{\overline{w}} \frac{T_A(z_A) - T_B(z_B) - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(x), x)dx \times dT_A$$

As a whole, the three effects need to be offset at the optimum, and accordingly, we have $\varrho_M^A + \varrho_B^A + \varrho_P^A = 0$. Rearranging this, we can obtain the optimal marginal income tax rate in region A in Proposition 2.

Using a similar method, the optimal marginal income tax rate in region B in Proposition 2 is characterized, where the participation effect is the opposite since the increase in marginal tax rates in region B induces individuals to access region A.

In the traditional literature, a tagged group as an immutable category depends on the mechanical effect and the behavioral effect. However, if a tagged group is a variable category, the change in the tax system due to differential income taxes distorts decision making on the extensive margin. Hence, the government takes account of the participation effect on tax revenues when differentiating income taxes.

4.2 Tax perturbation method: welfare gains introducing differentiated marginal tax rates

Beginning from the tax system with non-differentiated marginal tax rates at the optimum, we examine how differentiation of marginal tax rates should be introduced. Similar to Kleven et al. (2006, 2009), we consider a little bit of tax reform at any labor productivity w as depicted in Figure 1. The tax reform is decomposed into two com-

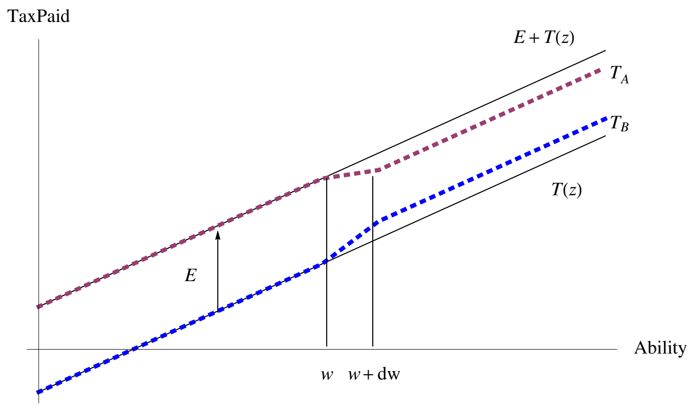

Figure 1: Small tax reform perturbation

ponents. Above labor productivity w, we decrease income taxes on people accessing region A and increase income taxes on people accessing region B. Let dT_A^a and dT_B^a be the small tax reform for each region above labor productivity w. Here, we assume that θ and w are independently distributed. We numerically examine the implication of correlation between two characteristics in section 5. Let the change in income taxes on each segment be inversely proportional to the population on the segment; in other words, $dT_A^a = -\frac{dT}{1-F_A(w)}$ and $dT_B^a = \frac{dT}{F_A(w)}$. Therefore, the tax reform is revenue neutral.

The tax reform causes three effects. First, an implementation of the tax reform induces individuals with labor productivity above w to access region A. The effect is associated with participation responses. Above w, while individuals accessing region A provide the government with additional revenue $T_A - T_B$, they cause efficiency loss by the amount of the net marginal congestion effect $\phi_{N_A} - \phi_{N_B}$. The number of switchers due to the tax reform amounts to the size of $|f(\hat{\theta}(w))f(w)d\hat{\theta}(w)|$, which is an absolute value. Therefore, a net effect at labor productivity w is measured by $(T_A - T_B - (\phi_{N_A} - \phi_{N_B}))|f(\hat{\theta}(w))f(w)d\hat{\theta}(w)|$. Moreover, given that $d\hat{\theta}(x) \cdot \Delta G = dT_A^a - dT_B^a$, the total effect associated with participation responses denoted by dP is expressed as follows:

$$dP = \int_w^{\overline{w}} \frac{T_A - T_B - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(x))f(x)dx \times \left(\frac{1}{1-F_A(w)} + \frac{1}{1-F_B(w)}\right)dT \tag{26}$$

As we begin with tax systems with non-differentiated marginal tax rates, we have $T_A = T_B$ and constant threshold $\hat{\theta}$. Using the assumption of independence between θ and w and substituting equation (15), equation (26) can be rewritten as follows:

$$dP = \int_{\underline{w}}^{\overline{w}} (g_B(w) - g_A(w))f(w)dw \tag{27}$$

The sign of dP is positive, implying that the government can collect more tax revenues through the response that individuals participate in region A owing to the tax reform.

Second, a small perturbation that changes the tax burden on each segment directly affects tax revenues without behavioral responses and the net mechanical effect denoted by dM is measured as follows:

$$dM \equiv \int_{w}^{\overline{w}} (1 - g_B(x)) f_B(x) dx \times dT_B^a + \int_{w}^{\overline{w}} (1 - g_A(x)) f_A(x) dx \times dT_A^a \tag{28}$$

Since we start from tax systems with non-differentiated marginal tax rates, we have the constant threshold $\hat{\theta}$. By the assumption of independence, dM is transformed as follows:

$$dM = \frac{1}{1 - F(w)} \int_{w}^{\overline{w}} (g_A(x) - g_B(x)) f(x) dx \times dT \tag{29}$$

The sign of dM is negative, which means that the mechanical effect caused by the tax reform decreases tax revenues.

Finally, the tax reform associated with the change in marginal income tax rates affects an individual's behaviors with respect to labor supply with labor productivity around w. The decrease in tax rates for individuals in region A increases tax receipts through the promotion of labor responses and the increase in tax rates for individuals in region B reduces tax receipts through the distortion of labor responses. As with the derivation of behavioral effects in subsection 4.1.2, we describe the effect on $[w, w + dw]$ in each region, denoted by dB_A^a and dB_B^a.

$$dB_A^a \equiv -\frac{T_A'}{1 - T_A'} \frac{e}{1 + e} w f_A(w) \times dT_A^a \tag{30}$$

$$dB_B^a \equiv -\frac{T_B'}{1 - T_B'} \frac{e}{1 + e} w f_B(w) \times dT_B^a \tag{31}$$

Since we start from tax systems with non-differentiated marginal tax rates, we have $T_A' = T_B'$ and constant threshold $\hat{\theta}$. Therefore, by the assumption of independence, these behavioral effects cancel out.

Here, we denote the total welfare effect by dW, which is the sum of the effects above. As a result, if θ and w are independently distributed, the total welfare effect of introducing differentiated marginal tax rates starting from tax systems with non-differentiated marginal tax rates is as follows:

$$\begin{aligned} dW &= dP + dM \\ &= \Bigg[\underbrace{\int_{\underline{w}}^{\overline{w}} (g_B(w) - g_A(w)) f(w) dw}_{Participation\ Effect} + \underbrace{\frac{1}{1 - F(w)} \int_{w}^{\overline{w}} (g_A(x) - g_B(x)) f(x) dx}_{Net\ Mechanical\ Effect}\Bigg] \times dT \end{aligned} \tag{32}$$

The direct welfare effect represents the trade-off between the positive effect due to participation responses and the negative effect associated with the mechanical effect. The first term expresses welfare gain (the tax reform enables the government to reinforce the redistributive tax system, inducing individuals to access region A from the decrease in tax burdens) and the second term reflects welfare loss (the tax reform weakens redistribution by decreasing total tax receipts). As shown in Appendix C, dW is positive if the first derivative of the social welfare function is strictly convex such as the constant rate of risk aversion (CRRA) form $W = V^{1-\pi}/(1-\pi)$, where π measures the government's taste for redistribution. This means that the welfare gain owing to the participation effect exceeds the welfare loss caused by the mechanical effect, and thus, $dW > 0$. In sum, the government can enhance social welfare by implementing the tax reform (Figure 1) and the following statement holds.

Proposition 3. *If θ and w are independently distributed and the first derivative of the social welfare function is strictly convex, starting from the tax system with non-differentiated marginal tax rates, the social welfare increases by introducing differentiated marginal tax rates, such that the marginal tax rate on individuals who access region A decreases and the marginal tax rate on individuals who access region B increases for any labor productivity w.*

However, we cannot assess whether this result holds even if it allows for the correlation between θ and w. To confirm how income tax schemes at the optimum are differentiated in various situations, we exercise numerical simulations in section 5.

5. Numerical examples

To illustrate our results at the optimum, we now exercise a simulation. The objective is to (i) confirm that, if preferences for public goods and labor productivity are independently distributed, the tax perturbation analysis in section 4.2 is consistent with tax reforms implemented at the optimum while checking the robustness with respect to alternative parameters (ii) examine the impact of the correlation between two characteristics on the differentiation of marginal tax rates at the optimum, and (iii) contrast the marginal tax rate of an immutable tag with that of a variable tag under the correlation between two characteristics.

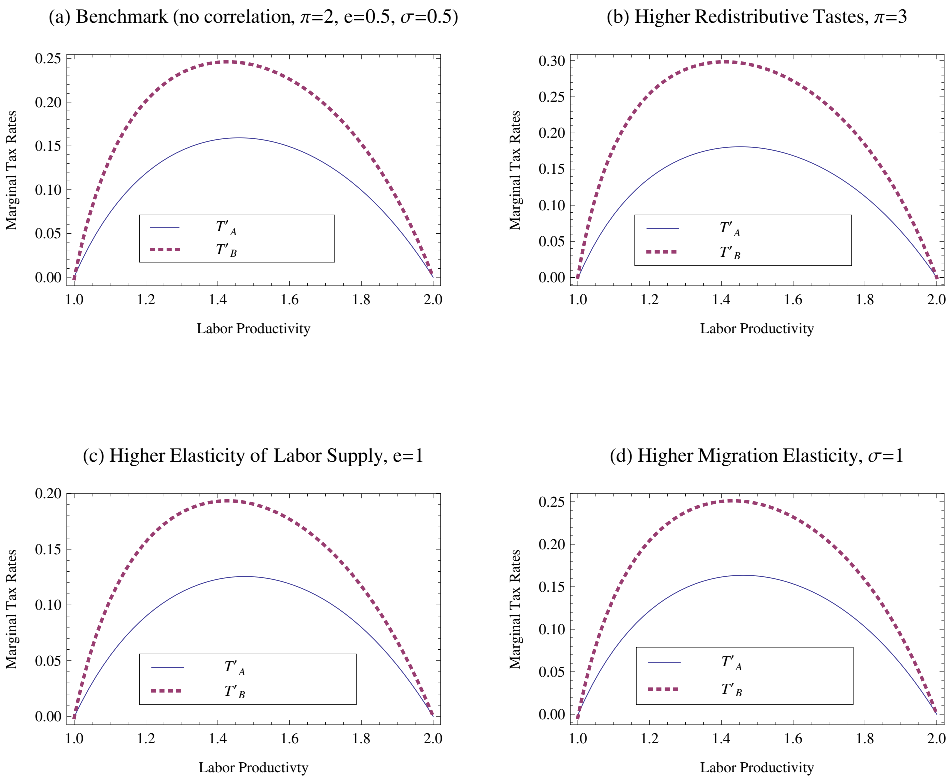

Figure 2: Benchmark simulations and sensitivity with alternative parameters

In the simulation, we set the following assumptions. First, we assume that the Bergson-Samuelson criterion is CRRA form. Second, we assume that the disutility of labor $v(\cdot)$ takes the following functional form: $v(\ell_i) = \ell_i^{1+1/e}/(1+1/e)$, where $e > 0$. In this case, $e = \epsilon_A = \epsilon_B$, and thus, the elasticity of labor supply with respect to the net-of-tax wage rate ϵ_i is constant. Third, following by Kleven et al. (2006, 2009), public goods preferences θ are distributed as the power function $F(\theta) = (\theta/\overline{\theta})^\sigma$ with the density function $f(\theta) = \sigma \cdot \theta^{\sigma-1}/\overline{\theta}^\sigma$ on the interval $[\underline{\theta} = 0, \overline{\theta} = 2.5]$, where σ indicates the constant migration elasticity $\hat{\theta}(w)f(\hat{\theta}(w))/F(\hat{\theta}(w))$ in region B. Fourth, we assume that a cumulative distribution function of labor productivity w is a truncated Pareto distribution with parameter $a = 2$ over $[\underline{w} = 1, \overline{w} = 2]$, expressed by $F(w) = [1 - (\underline{w}/w)^a]/[1-(\underline{w}/\overline{w})^a]$. Fifth, we consider $G_A = 0.8$ and $G_B = 0.05$ and assume that the cost function for the public good takes the following functional form: $\phi(G_i, N_i) = G_i^2 N_i$. In the benchmark simulation, we consider $\pi = 2$, $e = 0.5$, and $\sigma = 0.5$ and that the preference for a public good and labor productivity are independently distributed.

We plot optimal marginal tax rates T'_A and T'_B in each figure. Figure 2(a) is our benchmark simulation and describes their results. The marginal income tax rate on individuals enjoying a higher quality public good is lower than that on individuals enjoying a lower quality public good, which is in line with Proposition 3. Moreover,

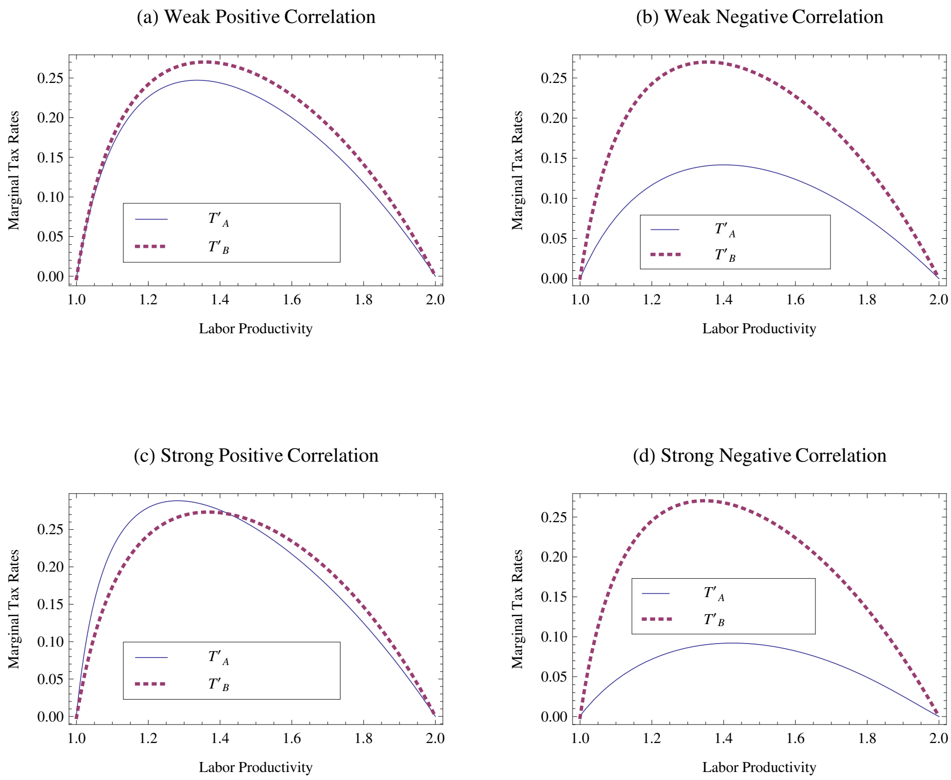

Figure 3: Simulations with positive or negative correlation

we present the sensitivity of optimal policies with respect to changes in the parameter values around the benchmark simulation. First, we increase the redistributive taste π from 2 to 3, the result of which is depicted in Figure 2(b). We find that all marginal tax rates increase to reinforce the redistribution. Second, we increase the elasticity of labor supply ϵ from 0.5 to 1, whose effect is described in Figure 2(c). As expected, all marginal tax rates decrease. Third, we increase migration elasticity σ from 0.5 to 1, the outcome of which is shown in Figure 2(d), and find that the increase in migration elasticity has a limited impact on the marginal tax rate. Nevertheless, the marginal tax rate in region A remains lower. Thus, as long as preferences for public goods and labor productivity are independently distributed, the tax reform in our tax perturbation method is implemented at the optimum, regardless of the sensitivity of alternative parameter values.

Here, we examine the implication of introducing a positive and negative correlation between the preference for a public good and labor productivity. We introduce a positive correlation by considering $\overline{\theta}$ as a increasing function of w and a negative correlation $\overline{\theta}$ as a decreasing function of w, as in Kleven et al. (2006, 2009). First, we consider the case in which the preference for a public good is weakly correlated with labor productivity ($\overline{\theta} = 1 + 0.5w$). As shown in Figure 3(a), the level of marginal tax rates

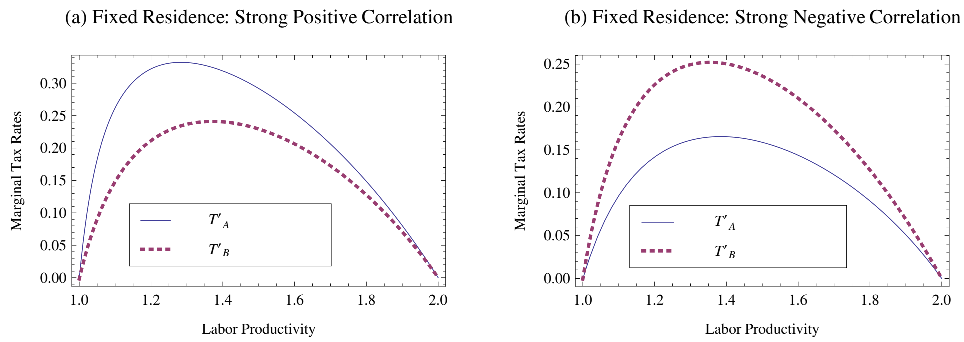

Figure 4: Simulations with labor immobility

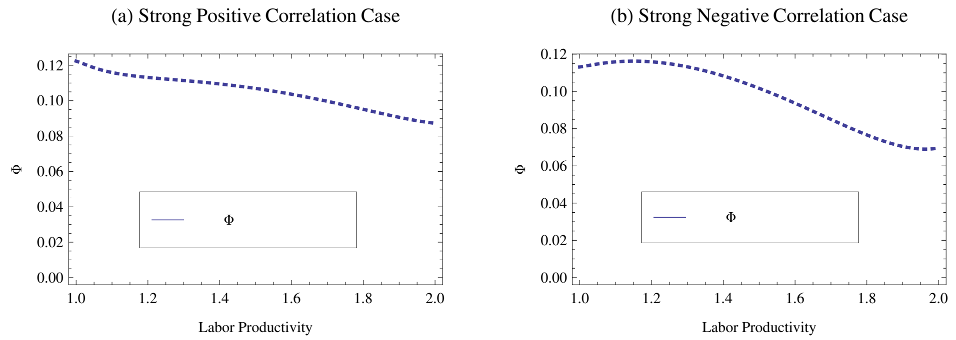

Figure 5: Simulations with respect to $\Phi(\cdot)$

in the weak positive correlation case is higher than the independent case to reinforce the income redistribution because the inequalities between categories are more serious. In contrast, as depicted in Figure 3(b), the level of marginal tax rates in the weak negative correlation case ($\overline{\theta} = 3.5 - 0.5w$) is lower than the independent case because the inequalities are mitigated. The fact that the marginal tax rate in region A is lower remains even though weak correlation is allowed. This is consistent with the findings of Kleven et al. (2006, 2009), who numerically demonstrate that introducing a positive or negative correlation between two characteristics does not overturn the tax perturbation results. Next, we present a situation in which the preference for a public good is strongly correlated with labor productivity. Figure 3(c) shows that the marginal tax rates in region A in a strong positive correlation case ($\overline{\theta} = 1.5w$) can be higher than those in region B. Therefore, it is not necessary that the effect of a positive correlation does not overturn the theoretical results obtained from the tax perturbation analysis. Undoubtedly, a strong negative correlation ($\overline{\theta} = 4.5 - 1.5w$) does not affect the relationship between T'_A and T'_B, as described in Figure 3(d).

Finally, we investigate how the correlation between two characteristics should be reflected in the optimal tax system depending on whether the category is immutable or variable. We apply an income distribution that is endogenously generated from the result in Figures 3(c) and 3(d) when calibrating the marginal tax rate without labor mobility. Figure 4(a) (Figure 4(b)) depicts the optimal differentiated marginal tax rate in the positive correlation case (negative correlation case) with immobile labor. These outcomes imply that the marginal tax rate on the region comprising a higher proportion of individuals with high ability is greater. Note that labor mobility decreases (increases) the marginal tax rate in region A (region B), regardless the correlation. As shown in Figures 5(a) and 5(b), $\Phi(\cdot)$, which appears when the labor is mobile, is always positive; this acts the marginal tax rate in region A (region B) as downward (upward) pressure, which in line with Proposition 2. Therefore, labor mobility weakens the differentiation of marginal tax rates on the basis of the positive correlation.[9]

6. Concluding Remarks

This study analyzes optimal nonlinear income taxes under spatial differences in terms of the quality of public goods when individuals have two types and determine labor supply along both intensive and extensive margins. The government can design differential income taxes on two regions, of which one has a higher quality public good. We show that the government's redistributive tastes and correlation between preferences for a public good and labor productivity are especially crucial in determining the shape of income tax schedules. In particular, if the preference for a public good and labor productivity are independently distributed and the first derivative of the social welfare function is strictly convex, the marginal income tax rate on individuals enjoying a higher quality public good is lower. This is because the decrease in tax burdens on the region in the presence of higher quality public goods leads to greater welfare gain from individuals being induced to access the region (participation effect) than the welfare loss caused by a decrease in tax receipts (mechanical effect). Moreover, we numerically find that the theoretical results are supported when the correlation between two characteristics is weak. However, the marginal tax rate on individuals enjoying a higher quality public good can be higher when the positive correlation between two characteristics is strong, although labor mobility weakens the differentiation of marginal tax rates on the basis of the positive correlation. Therefore, the present study offers implications for the optimal design of differential income tax schedules in the presence of a correlation between two characteristics, which is in contrast to Kleven et al. (2006, 2009).

Our findings have key implications for applied tax policies. If the aim of the government is to mitigate inequalities between regions in which income distributions slightly differ, the marginal income tax rate on individuals who enjoy higher quality public good services should be lower. However, if income stratification across regions is serious, the

[9] These results cannot be directly compared in general because the redistributive taste g_i can differ depending on whether labor is mobile. However, the simulation result suggests that the the effect of the change in $\Phi(\cdot)$ is crucial to decrease (increase) the marginal tax rate in region A (region B), regardless of the change in g_i.

differentiation of the marginal tax rates on the basis of the correlation is recommended, as shown in previous studies examining tagging on immutable categories corresponding to the present study, in which individuals do not vote with their feet. Therefore, our novel findings are that the government should put emphasis on the information in terms of serious income stratification across regions in designing income tax schedules.

Our findings can be further applied to the optimal tax and transfer program designed by the central government. Indeed, Boadway and Pestieau (2006) present the federal government with regions of different income distributions as an example. In particular, the results of this study will be useful when a supranational government such as the European Union is transfered the responsibility of redistributive taxation from national governments.

Appendix A: First-order conditions under the tax system with non-differentiated marginal tax rates

Proof of Proposition 1

Using integration by parts, $\int_{\underline{w}}^{\overline{w}} \lambda(w)V_B'(w)$ is transformed into $\lambda(\overline{w})V_B(\overline{w})-\lambda(\underline{w})V_B(\underline{w})-\int_{\underline{w}}^{\overline{w}} \lambda'(w)V_B(w)$. Applying this to the optimization problem with non-differentiated marginal tax rates, the corresponding Lagrangian is rewritten as follows:

$$\begin{aligned}\mathcal{L} = \hat{\mathcal{W}} + \gamma\left[\int_{\hat{\theta}}^{\overline{\theta}} \hat{\theta}\Delta G f(\theta)d\theta + \int_{\underline{w}}^{\overline{w}} T_B(z(w))f(w)dw - \phi(G_A, N_A) - \phi(G_B, N_B)\right] \\ + \int_{\underline{w}}^{\overline{w}} \lambda(w)\frac{\ell(w)}{w}v'(\ell(w))dw + \int_{\underline{w}}^{\overline{w}} \lambda'(w)V_B(w)dw - \lambda(\overline{w})V_B(\overline{w}) + \lambda(\underline{w})V_B(\underline{w})\end{aligned} \tag{A.1}$$

By the definition of indirect utilities, income taxes are expressed by $T_B(z(w)) = w\ell(w)-V_B(w) - v(\ell(w))$ and substitute this for (A.1) yields:

$$\begin{aligned}\mathcal{L} = \hat{\mathcal{W}} + \gamma\left[\int_{\hat{\theta}}^{\overline{\theta}} \hat{\theta}\Delta G f(\theta)d\theta + \int_{\underline{w}}^{\overline{w}} \Big(w\ell(w) - V_B(w) - v(\ell(w))\Big)f(w)dw - \left(\sum_{i=A,B} \phi(G_i, N_i)\right)\right] \\ + \int_{\underline{w}}^{\overline{w}} \lambda(w)\frac{\ell(w)}{w}v'(\ell(w))dw + \int_{\underline{w}}^{\overline{w}} \lambda'(w)V_B(w)dw - \lambda(\overline{w})V_B(\overline{w}) + \lambda(\underline{w})V_B(\underline{w})\end{aligned} \tag{A.2}$$

The first-order conditions associated with $V_B(w)$, $\ell(w)$, and $\hat{\theta}$ are as follows:

$$\frac{\partial \mathcal{L}}{\partial V_B(w)} = \int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W'(\theta G_B + V_B(w)) f(\theta, w) d\theta + \lambda'(w) - \gamma f(w) = 0 \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial \ell(w)} = \lambda(w) \left[\frac{v'(\ell(w))}{w} + \frac{\ell(w)}{w} v''(\ell(w)) \right] + \gamma \left[w - v'(\ell(w)) \right] f(w) = 0 \tag{A.4}$$

$$\frac{\partial \mathcal{L}}{\partial \hat{\theta}} = -\gamma \left[\hat{\theta} f(\hat{\theta}) - (1 - F(\hat{\theta})) \right] \Delta G - \int_{\underline{w}}^{\overline{w}} \int_{\hat{\theta}}^{\overline{\theta}} \Delta G \cdot W'(\theta G_A + V_A) f(\theta, w) d\theta dw + \gamma(\phi_{N_A} - \phi_{N_B}) f(\hat{\theta}) = 0 \tag{A.5}$$

$$\frac{\partial \mathcal{L}}{\partial V_B(\overline{w})} = -\lambda(\overline{w}) = 0, \quad \frac{\partial \mathcal{L}}{\partial V_B(\underline{w})} = \lambda(\underline{w}) = 0 \tag{A.6}$$

Integrating (A.3) between w and $\overline{w}$ and using the transversality condition (A.6) yields:

$$-\frac{\lambda(w)}{\gamma} = \int_{w}^{\overline{w}} \left[1 - \frac{\int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(x)) f(\theta|x) d\theta}{\gamma} - \frac{\int_{\underline{\theta}}^{\hat{\theta}} W'(\theta G_B + V_B(x)) f(\theta|x) d\theta}{\gamma} \right] f(x) dx \tag{A.7}$$

Rearranging equation (A.7) yields:

$$-\frac{\lambda(w)}{\gamma} = \int_{w}^{\overline{w}} \left[1 - \frac{\int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(x)) f(\theta|x) d\theta}{\gamma f_A^c(x)} f_A^c(x) - \frac{\int_{\underline{\theta}}^{\hat{\theta}} W'(\theta G_B + V_B(x)) f(\theta|x) d\theta}{\gamma f_B^c(x)} f_B^c(x) \right] f(x) dx \tag{A.8}$$

Therefore, by the definition of g_i and $\overline{g}$,

$$-\frac{\lambda(w)}{\gamma} = \int_{w}^{\overline{w}} \left[1 - g_A(x) f_A^c(x) - g_B(x) f_B^c(x) \right] f(x) dx = \int_{w}^{\overline{w}} (1 - \overline{g}(x)) f(x) dx \tag{A.9}$$

On the other hand, equation (A.4) is transformed as follows:

$$\lambda(w) \frac{v'(\ell(w))}{w} \left[1 + \frac{\ell(w)}{v'(\ell(w))} v''(\ell(w)) \right] + \gamma w \left[1 - \frac{v'(\ell(w))}{w} \right] f(w) = 0 \tag{A.10}$$

Substituting equation (2) and (3) and rearranging, (A.10) is rewritten as follows:

$$\frac{T'(z(w))}{1 - T'(z(w))} = -\left[1 + \frac{1}{\epsilon} \right] \frac{\lambda(w)}{\gamma} \frac{1}{w f(w)} \tag{A.11}$$

Finally, combining (A.9) and (A.11) yields (14).

We transform (A.5) using $E = \hat{\theta}\Delta G$ as follows:

$$\gamma \frac{E - (\phi_{N_A} - \phi_{N_B})}{E} \hat{\theta} f(\hat{\theta}) = \gamma(1 - F(\hat{\theta})) - \int_{\underline{w}}^{\overline{w}} \int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta dw \quad (A.12)$$

Furthermore, integrating (A.3) between $\underline{w}$ and $\overline{w}$ and using the transversality conditions (A.6) yields:

$$\gamma = \int_{\underline{w}}^{\overline{w}} \left[\int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W'(\theta G_B + V_B(w)) f(\theta, w) d\theta \right] dw \quad (A.13)$$

Substituting (A.13) into (A.12) and dividing by N_A, the following is obtained:

$$\begin{aligned} \frac{E - (\phi_{N_A} - \phi_{N_B})}{E} \frac{\hat{\theta} f(\hat{\theta})}{1 - F(\hat{\theta})} = & \frac{1}{N_A} \left[N_A \int_{\underline{w}}^{\overline{w}} \frac{\int_{\underline{\theta}}^{\hat{\theta}} W'(\theta G_B + V_B(w)) f(\theta|w) d\theta}{\gamma} f(w) dw \right. \\ & \left. - N_B \int_{\underline{w}}^{\overline{w}} \frac{\int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta|w) d\theta}{\gamma} f(w) dw \right] \end{aligned} \quad (A.14)$$

By the definition of g_i and η, we can rewrite (A.14) for (15).

Appendix B: First-order conditions under differentiated marginal income tax rates

Proof of Proposition 2

Using integration by parts, $\int_{\underline{w}}^{\overline{w}} \lambda_i(w) V_i'(w)$ is transformed into $\lambda_i(\overline{w}) V_i(\overline{w}) - \lambda_i(\underline{w}) V_i(\underline{w}) - \int_{\underline{w}}^{\overline{w}} \lambda_i'(w) V_i(w)$. Applying this to the optimization problem with differentiated marginal tax rates, the corresponding Lagrangian is rewritten as follows:

$$\begin{aligned} \mathcal{L} = & \mathcal{W} + \gamma \left[\int_{\underline{w}}^{\overline{w}} T_A(z_A(w)) f_A(w) dw + \int_{\underline{w}}^{\overline{w}} T_B(z_B(w)) f_B(w) dw - \left(\sum_{i=A,B} \phi(G_i, N_i) \right) \right] \\ & + \sum_{i=A,B} \left[\int_{\underline{w}}^{\overline{w}} \lambda_i(w) \frac{\ell_i(w)}{w} v'(\ell_i(w)) dw + \int_{\underline{w}}^{\overline{w}} \lambda_i(w) V_i(w) dw - \lambda_i(\overline{w}) V_i(\overline{w}) + \lambda_i(\underline{w}) V_i(\underline{w}) \right] \\ & + \int_{\underline{w}}^{\overline{w}} \mu(w) \left[\hat{\theta}(w) \Delta G + V_A(w) - V_B(w) \right] dw \end{aligned} \quad (B.1)$$

By the definition of indirect utilities, government's revenues from region i are expressed by $T_i(z_i(w)) = w\ell_i(w) - V_i(w) - v(\ell_i(w))$ and substitute this for (B.1) yields:

$$
\begin{aligned}
\mathcal{L} = \mathcal{W} + \gamma \Bigg[\sum_{i=A,B} \int_{\underline{w}}^{\overline{w}} \Big(w\ell_i(w) - V_i(w) - v(\ell_i(w)) \Big) f_i(w) dw - \Big(\sum_{i=A,B} \phi(G_i, N_i) \Big) \Bigg] \\
+ \sum_{i=A,B} \Bigg[\int_{\underline{w}}^{\overline{w}} \lambda_i(w) \frac{\ell_i(w)}{w} v'(\ell_i(w)) dw + \int_{\underline{w}}^{\overline{w}} \lambda_i(w) V_i(w) dw - \lambda_i(\overline{w}) V_i(\overline{w}) + \lambda_i(\underline{w}) V_i(\underline{w}) \Bigg] \\
+ \int_{\underline{w}}^{\overline{w}} \mu(w) \Big[\hat{\theta}(w) \Delta G + V_A(w) - V_B(w) \Big] dw
\end{aligned}
\tag{B.2}
$$

The first-order conditions associated with $V_i(w)$, $\ell_i(w)$, and $\hat{\theta}(w)$ are as follows:

$$
\frac{\partial \mathcal{L}}{\partial V_A(w)} = \int_{\hat{\theta}(w)}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta + \lambda_A'(w) - \gamma f_A(w) + \mu(w) = 0 \tag{B.3}
$$

$$
\frac{\partial \mathcal{L}}{\partial V_B(w)} = \int_{\underline{\theta}}^{\hat{\theta}(w)} W'(\theta G_B + V_B(w)) f(\theta, w) d\theta + \lambda_B'(w) - \gamma f_B(w) - \mu(w) = 0 \tag{B.4}
$$

$$
\frac{\partial \mathcal{L}}{\partial \ell_A(w)} = \lambda_A(w) \Bigg[\frac{v'(\ell_A(w))}{w} + \frac{\ell_A(w)}{w} v''(\ell_A(w)) \Bigg] + \gamma \Big[w - v'(\ell_A(w)) \Big] f_A(w) = 0 \tag{B.5}
$$

$$
\frac{\partial \mathcal{L}}{\partial \ell_B(w)} = \lambda_B(w) \Bigg[\frac{v'(\ell_B(w))}{w} + \frac{\ell_B(w)}{w} v''(\ell_B(w)) \Bigg] + \gamma \Big[w - v'(\ell_B(w)) \Big] f_B(w) = 0 \tag{B.6}
$$

$$
\frac{\partial \mathcal{L}}{\partial \hat{\theta}(w)} = -\gamma \Big[T_A(z_A(w)) - T_B(z_B(w)) - (\phi_{N_A} - \phi_{N_B}) \Big] f(\hat{\theta}(w), w) + \mu(w) \Delta G = 0 \tag{B.7}
$$

$$
\frac{\partial \mathcal{L}}{\partial V_i(\overline{w})} = -\lambda_i(\overline{w}) = 0, \; \frac{\partial \mathcal{L}}{\partial V_i(\underline{w})} = \lambda_i(\underline{w}) = 0 \; i = A, B \tag{B.8}
$$

Substituting (B.7) for (B.3) to delete $\mu(w)$ and dividing by γ yields:

$$
\begin{aligned}
\frac{\lambda_A'(w)}{\gamma} = f_A(w) - \frac{\int_{\hat{\theta}(w)}^{\overline{\theta}} W'(\theta G_A + V_A(w)) f(\theta, w) d\theta}{\gamma} \\
- \frac{T_A(z_A(w)) - T_B(z_B(w)) - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(w), w)
\end{aligned}
\tag{B.9}
$$

By the definition of g_i,

$$\frac{\lambda_A'(w)}{\gamma} = \left[(1-g_A(w))f_A^c(w) - \frac{T_A(z_A(w)) - T_B(z_B(w)) - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(w)|w)\right] f(w) \tag{B.10}$$

Integrating (B.10) between w and $\overline{w}$ and using the transversality condition (B.8) yields:

$$-\frac{\lambda_A(w)}{\gamma} = \int_w^{\overline{w}} \left[(1-g_A(x))f_A^c(x) - \frac{T_A(z_A(x)) - T_B(z_B(x)) - (\phi_{N_A} - \phi_{N_B})}{\Delta G} f(\hat{\theta}(x)|x)\right] f(x)dx \tag{B.11}$$

On the other hand, (B.5) is transformed as follows:

$$\lambda_A(w)\frac{v'(\ell_A(w))}{w}\left[1 + \frac{\ell_A(w)}{v'(\ell_A(w))}v''(\ell_A(w))\right] + \gamma w\left[1 - \frac{v'(\ell_A(w))}{w}\right] f_A(w) = 0 \tag{B.12}$$

Substituting equation (2) and (3) and rearranging,

$$\frac{T_A'(z_A(w))}{1 - T_A'(z_A(w))} = -\left[1 + \frac{1}{\epsilon_A}\right]\frac{\lambda_A(w)}{\gamma}\frac{1}{wf_A(w)} \tag{B.13}$$

Finally, combining (B.11) and (B.13), we can obtain equation (19). In the similar way, (20) is obtained.

Appendix C

Proof of Proposition 3

Let us start from the separable taxation with non-differentiated marginal income tax rates. By differentiating $g_B(w) - g_A(w)$ with respect to w, we can get the following:

$$g_B'(w) - g_A'(w) = \left[\frac{W''(\theta G_B + V_B(w))}{\gamma} - \frac{\int_{\hat{\theta}}^{\overline{\theta}} W''((\theta - \hat{\theta})\Delta G + G_B + V_B(w))f(\theta)d\theta}{\gamma(1 - F(\hat{\theta}))}\right] \cdot V_B'(w) \tag{C.1}$$

By the assumption, W'' is strictly increasing. Hence, $g_B' - g_A'$ is negative for any w. By using this fact, we can derive the following inequality for any w:

$$\frac{1}{F(w)}\int_{\underline{w}}^{w}(g_B(x) - g_A(x))f(w)dw > g_B(w) - g_A(w) > \frac{1}{1 - F(w)}\int_w^{\overline{w}}(g_B(x) - g_A(x))f(w)dw \tag{C.2}$$

Here, we rearrange the equation (32) as follows:

$$dW = \frac{1}{1-F(w)} \left[(1-F(w)) \int_{\underline{w}}^{w} (g_B(x)-g_A(x)) f(x) dx - F(w) \int_{w}^{\overline{w}} (g_B(w)-g_A(w)) f(w) dw \right] \times dT \tag{C.3}$$

Using (C.2), we can conclude that dW is positive.

References

Akerlof, G., 1978. The economics of "tagging" as applied to the optimal income tax, welfare programs, and manpower planning. American Economic Review 68, 8–19.

Bargain, O., Dolls, M., Fuest, C., Neumann, D., Peichl, A., Pestel, N., Siegloch, S., 2013. Fiscal union in europe? redistributive and stabilizing effects of a european tax-benefit system and fiscal equalization mechanism. Economic Policy 28, 375–422.

Bayer, P., McMillan, R., 2012. Tiebout sorting and neighborhood stratification. Journal of Public Economics 96, 1129–1143.

Bierbrauer, F., Brett, C., Weymark, J. A., 2013. Strategic nonlinear income tax competition with perfect labor mobility. Games and Economic Behavior 82, 292–311.

Blumkin, T., Sadka, E., Shem-Tov, Y., 2015. International tax competition: Zero tax rate at the top re-established. International Tax and Public Finance 22, 760–776.

Boadway, R., Pestieau, P., 2006. Tagging and redistributive taxation. Annals of Economics and Statistics 83-84, 123–147.

Buchanan, J. M., 1965. An economic theory of clubs. Economica 32, 1–14.

Cremer, H., Gahvari, F., Lazachmeur, J.-M., 2010. Tagging and income taxation: Theory and an application. American Economic Journal: Economic Policy 2, 31–50.

Diamond, P. A., 1998. Optimal income taxation: An example with u-shaped pattern of optimal marginal tax rates. American Economic Review 88, 83–95.

Gomes, R., Lazachmeur, J.-M., Pavan, A., 2014. Differential taxation and occupational choice, CESifo Working Paper 5054.

Immonen, R., Kanbur, R., Keen, M., Tuomala, M., 1998. Tagging and taxing: The optimal use of categorical and income information in designing tax/transfer schemes. Economica 65, 179–192.

Jacquet, L., Lehmann, E., der Linden, B. V., 2013. Optimal redistributive taxation with both extensive and intensive responses. Journal of Economic Theory 148, 1770–1805.

Kessing, S., Lipatov, V., Zoubek, J. M., 2015. Optimal taxation under regional inequality, CAGE Working Papers 253.

Kessing, S., Strozzi, C., 2016. The regional distribution of public employment: Theory and evidence. published online in Regional Studies.

Kleven, H. J., Kreiner, C. T., Saez, E., 2006. The optimal income taxation of couples, NBER Working Papers 12685.

Kleven, H. J., Kreiner, C. T., Saez, E., 2009. The optimal income taxation of couples. Econometrica 77, 537–560.

Lehmann, E., Simula, L., Trannoy, A., 2014. Tax me if you can! optimal nonlinear income tax between competing governments. Quarterly Journal of Economics 129, 1995–2030.

Mankiw, N. G., Weinzierl, M., 2010. The optimal taxation of height: A case study of utilitarian income redistribution. American Economic Journal: Economic Policy 2, 155–176.

McGuire, M., 1974. Group segregation and optimal jurisdictions. Journal of Political Economy 82, 112–132.

Mirrlees, J., 1971. An exploration in the theory of optimum income taxation. Review of Economic Studies 38, 175–208.

Morelli, M., Yang, H., Ye, L., 2012. Competitive nonlinear taxation and constitutional choice. American Economic Journal: Microeconomics 4, 142–175.

Rothchild, C., Scheuer, F., 2013. Redistributive taxation in the roy model. The Quarterly Journal of Economics 128, 623–668.

Rothchild, C., Scheuer, F., 2016. Optimal taxation with rent-seeking. Review of Economic Studies 83, 1225–1262.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68, 205–229.

Verdugo, G., 2016. Public housing magnets: Public housing supply and immigrants' location choices. Journal of Economic Geography 16, 237–265.

Viard, A. D., 2001. Optimal categorical transfer payments: The welfare economics of limited lump-sum redistribution. Journal of Public Economic Theory 3, 483–500.

Weinzierl, M., 2014. The promise of positive optimal taxation: Normative diversity and a role for equal sacrifice. Journal of Public Economics 118, 128–142.