Economics Bulletin

Volume 38, Issue 2

Optimal capital income taxation in the case of private donations to public goods

Shigeo Morita
Osaka University

Takuya Obara
Hitotsubashi University

Abstract

In this study, we investigate optimal nonlinear labor and capital income taxation and subsidies for contribution goods in a dynamic setting. We show that when individuals can contribute to a public good--even if additive and separable preference between consumption and labor supply is assumed and individuals differ only in earning ability--marginal capital income tax rate for low-income earners is not zero, indicating that the Atkinson-Stiglitz theorem does not hold.

We thank our supervisors Motohiro Sato and Yukihiro Nishimura for their invaluable comments, suggestions, and support. We also thank Yoshihiro Takamatsu for agreeing to be a discussant for the presentation of our paper and the participants at the Japanese Economic Association, June 19, 2016. Financial support from Hitotsubashi University is gratefully acknowledged.
Citation: Shigeo Morita and Takuya Obara, (2018) "Optimal capital income taxation in the case of private donations to public goods", *Economics Bulletin*, Volume 38, Issue 2, pages 921-939
Contact: Shigeo Morita - ngp024ms@student.econ.osaka-u.ac.jp, Takuya Obara - ed142003@g.hit-u.ac.jp.
Submitted: June 05, 2017.   **Published:** April 30, 2018.

1. Introduction

Optimal tax theory plays an important role in designing income redistribution policies and implementing public projects. For several years, researchers have explored aspects of the field: in particular, economists have been concerned with the question, "Should capital income be taxed?" This question arises from the fact that the government can reinforce redistributive policies by levying taxes on savings; however, this is a form of double taxation. Although the literature has discussed whether taxation on capital income is justified, it remains an ongoing research issue.

This study aims to investigate the desirability of capital income tax from the viewpoint of economic behaviors, specifically when individuals can contribute to public goods. The motivation stems from several empirical studies which find that a high tax rate on capital gains generally leads to taxpayers choosing charitable giving as strategies to avoid recognizing taxable gains. Indeed, Hood et al. (1977) find that Canada's 1971 Tax Reform, which introduced a 50% capital gain tax, brought about a decrease in individual charitable donations. More recently, Auten et al. (2002) estimate the price elasticity of donations, where price is a weighted average of the price of giving cash and appreciated properties. Their estimations show that the value of the price elasticity of donation is negative. These findings imply that individuals account for capital income tax when donating toward public goods. These suggest that individuals' charitable giving cannot be discussed independently of capital income taxation. Therefore, our study takes a step toward theoretically clarifying how capital income tax schemes should be designed when individuals can contribute to a public good.

Our analysis comprises a dynamic setting in which individuals live for two periods. We assume that in the first period, individuals can spend a part of their savings on donations to a charity. Conceptually, we regard the private donation to the public good as the charitable giving and utilize the framework of the optimal taxation in the presence of a public good. This setup is in line with Andreoni (1988), Saez (2004), and Diamond (2006). For simplicity, there are two types of individuals: high- and low-skilled individuals. The government designs three types of tax schedules: nonlinear taxes on labor and capital income and nonlinear subsidies for contributions to a public good. We demonstrate that although a utility function is represented by the preference that private goods are additively separable from leisure, the marginal tax on capital is zero for the high skilled but not low skilled when private contributions are made to public goods. The amount of donation to a public good differs between high- and low-skilled individuals, which affects the marginal rate of substitution between consumption in the first and second period. If a high-skilled individual's valuation of future consumption is higher than a low-skilled individual's one, the distortion on savings behavior for the latter relaxes the self-selection constraint for the former. Thus, the relationship between private consumption and donation to a public good plays an important role in characterizing the optimal tax rates for marginal capital income.

An important contribution to the literature on capital income taxation is Ordover and Phelps (1979). They examine optimal nonlinear taxation on income and savings in an overlapping generations model in the case of unobservable earnings ability. Their main conclusion states that if preferences are weakly separable between private goods and leisure, taxes on savings are redundant. This is consistent with the Atkinson and Stiglitz (1976) theorem. On the other hand, Saez (2002) investigates conditions necessary to obtain the Atkinson–Stiglitz theorem and shows that if individuals have heterogeneous taste in private consumption, the results are violated, even if the utility function is weakly separable between private goods and leisure.

The present study is closely related to the model explaining the desirability of capital income taxes on the basis of heterogeneous tastes for goods between high- and low-income earners, which stems from Saez (2002). The extant literature treats differentiation in taste based on initial endowments and discount rates as an assumption (Boadway et al. (2000), Cremer et al. (2001), Diamond and Spinnewijn (2011)). In contrast, we show that taste differentiation can also result from individuals' behavior without explicitly assuming additional characteristics. Thus, we present the desirability of capital income taxes by establishing the theoretical foundation that taste differentiation occurs. To the best of our knowledge, this study provides new evidence justifying capital taxation since private donations have not been considered in the context of individual behaviors in capital income taxation theory.

In our model, the government offers subsidies for contributions to public goods. Prior studies have investigated optimal tax policy assuming the presence of charitable giving (Andreoni (1988), Saez (2004)). The study most closely related in terms of tax treatment of private donations is Diamond (2006), who shows that the welfare-improving effect is achieved by introducing a subsidy on private donations toward a public good through nonlinear income taxes on labor. However, Diamond (2006) adopts a static model that does not allow the government to impose income taxes on capital and attempts to derive an optimal tax treatment formula. We extend the Diamond model as a two-period model to investigate the desirability of capital income taxes.

The remainder of this paper is organized as follows. Section 2 describes the framework of the basic model. Section 3 characterizes optimal tax formulas. Section 4 provides additional analyses. Section 5 concludes the paper.

2. Model

2.1 Environment

We consider an economy in which individuals live for two periods and work only in the first. There are two types of individuals: low-skilled individuals whose earning wage rate is w^1 and high-skilled ones with earning wage rate w^2, where $w^2 > w^1$. Their before-tax income is $y_i \equiv w^i l^i$, where l^i denotes the labor supply of type i individuals.

We consider a finite number of individuals as in Diamond (2006). The number of type i individuals is defined by π^i, which is a natural number greater than two and for now, is invariant.[1] The utility function of type i individuals is

$$U^i(c^i, x^i, G, l^i) = u(c^i, x^i, G) - v(l^i) \tag{1}$$

where c^i denotes consumption of a private good in the first period, x^i is consumption in the second period, and G is the amount of public good. Individuals can contribute to a public good; then, the aggregate amount of public good is

$$G = \sum_i \pi^i g^i + g^G \tag{2}$$

where g^i denotes type i's private donations to the public good and g^G the public donation to the public good. Contributions by a particular economic agent improve the service and then the amount of public good is an argument in the utility function of individuals.[2] The sub-utility function, $u(\cdot)$, is strictly increasing, concave, and twice differentiable and satisfies the Inada condition and $v(\cdot)$ is strictly increasing, convex, and twice differentiable.

Let s^i denote savings of type i individuals and r be the interest rate. The budget constraints type i individuals face can be written as follows:

$$c^i + s^i + g^i - \tau(g^i) = y^i - T(y^i) \tag{3}$$

$$s^i(1 + r) - \Phi(s^i r) = x^i \tag{4}$$

where $\tau(g^i)$ is a subsidy for private donations by type i individuals to a public good, $T(y^i)$ is an income tax payment, and $\Phi(rs^i)$ is the capital income tax payment, which respectively, are nonlinear functions of g^i, y^i, and rs^i. Individuals choose c^i, x^i, s^i, g^i, and l^i to maximize the utility function (equation (1)) subject to their budget constraints (equations (3) and (4)). Combining this with the first-order conditions yields

$$-u_c(c^i, x^i, G) + \{(1 + r) - r\Phi'(rs^i)\}u_x(c^i, x^i, G) = 0 \tag{5}$$

[1] Using a finite number of individuals, Piketty (1993) and Hamilton and Slutsky (2007) show that the first-best allocation can be achieved if an individual's tax schedule depends on the behavior of other individuals. Following traditional optimal taxation literature, the present study restricts an individual's tax schedule to a function of the value of his/her labor income, capital income, and private donation to a public good.

[2] We consider public goods financed by not only individuals but also the government such as health, education, and social services. According to Charitable Giving Statistics by National Philanthropic Trust, in the United States, individuals' charitable giving accounts for 71% of total giving and majority of donations are made to religious, educational, and healthcare organizations. A donation to religious organizations is a suitable example for the outcomes of Lemma 1 because the US government cannot contribute to them.

where $u_c(c^i, x^i, G) \equiv \frac{\partial u}{\partial c^i}$ denotes the marginal utility of consumption in the first period, $u_x(c^i, x^i, G) \equiv \frac{\partial u}{\partial x^i}$ is the marginal utility of consumption in the second period, and $\Phi'(rs^i) \equiv \frac{d\Phi}{drs^i}$ is the marginal capital income tax rate function corresponding to returns of savings rs^i.

The production sector utilizes labor and capital. Production technology exhibits constant return to scale. This means that each unit of effective labor $w^i l^i$ is required to produce one unit of private good and each unit of private good saved in the first period produces $(1+r)$ units of a private good in the second period.[3]

2.2 Planning problem

The objective of the government is represented by the following utilitarian social welfare function:

$$W = \sum_i \pi^i U^i(c^i, x^i, G, l^i) \tag{6}$$

The budget constraint for the government is

$$\sum_i \pi^i T(y^i) + (1+r)^{-1} \sum_i \pi^i \Phi(rs^i) = g^G + \sum_i \pi^i \tau(g^i) \tag{7}$$

Following Diamond (2006), we assume that there is no response of government budget constraints to a deviation from individuals' anticipated revealing strategies. Using the budget constraints that individuals face, equation (7) can be equivalently written as

$$\sum_i y^i \pi^i - \sum_i (c^i + \frac{x^i}{1+r} + g^i)\pi^i - g^G \geq 0 \tag{8}$$

The informational assumptions are conventional: the government can observe individuals' donation, labor income, and capital income, while their ability is never observable. We focus on the case in which the government attempts to redistribute from type-2 to type-1 individuals. This means the following incentive compatibility constraint is binding at the social optimum:

$$U^2(c^2, x^2, G, \frac{y^2}{w^2}) \geq U^2(c^1, x^1, \hat{G}, \frac{y^1}{w^2}) \tag{9}$$

[3]Pirttilä and Tuomala (2001) show that capital income taxation is justified when wages are endogenously determined and the relative wage rate is affected by the amount of savings. By contrast, we assume no general-equilibrium effects of input prices. This is because we never obtain the novel effect even if we endogenize input prices, that is, the optimal tax formula for capital income just involves the endogenous wage term proposed by Pirttilä and Tuomala (2001). Thus, for simplicity, our model can be seen as the two-period, partial equilibrium version of Pirttilä and Tuomala (2001) model. At the optimum, where only the government contributes to a public good, our model's outcome is consistent with that of their model.

where $\hat{G} \equiv G - g^2 + g^1$ denotes the aggregate level of a public good achieved when type-2 individuals mimic.

The social planning problem is to maximize the social welfare function (equation (6)), subject to the equations for a public good (equation (2)), resource constraints (equations (8)), and incentive compatibility constraints (equation (9)). The Lagrangean corresponding to this planning problem can be formulated as follows:

$$\mathcal{L} = W + \mu\Big[\sum_i g^i\pi^i + g^G - G\Big] + \gamma\Big[\sum_i y^i\pi^i - \sum_i (c^i + \frac{x^i}{1+r} + g^i)\pi^i - g^G\Big] \\ +\lambda\Big[U^2(c^2, x^2, G, \frac{y^2}{w^2}) - U^2(c^1, x^1, \hat{G}, \frac{y^1}{w^2})\Big] \tag{10}$$

where μ, γ, and λ are the Lagrange multipliers.

3. Characterizing optimal capital taxation

Here, we present the key features of our model's outcomes. The results imply that the government should design taxes on capital income such that it supplements the tax treatment of private donations to a public good.[4] Let

$$MRS^i_{cx} \equiv \frac{u_c(c^i, x^i, G)}{u_x(c^i, x^i, G)} \quad \text{and} \quad M\hat{R}S_{cx} \equiv \frac{u_c(c^1, x^1, \hat{G})}{u_x(c^1, x^1, \hat{G})}$$

denote the marginal rate of substitution between private consumption in the first and second period faced by type i individuals and the corresponding marginal rate of substitution that the mimicker faces. Combining the optimality condition regarding c^i and x^i yields the optimal capital income tax rate for type i individuals:

$$\frac{\Phi'(rs^1)}{1+r} = \frac{\lambda u_x(c^1, x^1, \hat{G})}{r\pi^1\gamma}\left(MRS^1_{cx} - M\hat{R}S_{cx}\right) \tag{11}$$

$$\Phi'(rs^2) = 0 \tag{12}$$

The derivation is presented in Appendix A. Equation (11) implies that the deviation of the optimal tax rate on capital income from the Atkinson–Stiglitz theorem depends on the term in the brackets on the right-hand side. These equations give the following proposition:

Proposition 1.

[4] In the present paper, we omit to characterize the optimal tax treatment of private donations for a public good. Morita and Obara (2016), which is the working paper version, derives the optimal tax treatment formula.

- *When a public good has a more complementary (substitutionary) relationship with the consumption good in the first period than in the second period, even if individual preferences can be separated between labor and consumption, the marginal capital income tax rate is positive (negative) for type-1 individuals and zero for the type-2 individuals.*

- *When a public good has no relationship with both the consumption good in the first period and in the second period, the marginal capital income tax rate is zero for both types of individuals.*

The result of Proposition 1 is crucially related to the difference between G and $\hat{G}$. At the optimum, the level of a public good is higher when a type-2 individual chooses a truth-telling strategy than a mimicking-one, that is, $G > \hat{G}$. As shown in the Appendix B, it is optimal that only type-2 individuals contribute to the public good, $g^1 = 0$, $g^2 > 0$, and $g^G = 0$.[5] This suggests that inducing type-2 individuals to contribute improves their level of social welfare from the allocation, where no one makes a private donation to public goods. This is consistent with Diamond (2006). At the optimum, the level of public good is higher when a type-2 individual chooses a truth-telling strategy than a mimicking one, that is, $G > \hat{G}$. Assuming that the public good has a stronger complementary relationship with the private good in the first period than in the second, the intertemporal marginal rate of substitution for the mimicker is lower than the corresponding marginal rate of substitution for the mimicked, that is, $MRS^1_{cx} > M\hat{R}S_{cx}$. In other words, the mimicker values the consumption in the second period more than the mimicked (type-1 individuals). This implies that distorting the capital income of type-1 individuals downward hurts the mimicker more than the mimicked and thus relaxes the incentive compatibility constraint. Therefore, the marginal capital income tax rate should be positive. Consequently, individuals' behavior in terms of private donations to a public good creates an informational advantage for the government. Note that this outcome depends on the assumption of finite number of individuals. This is because when there is an infinite population, the level of public good does not change even if a type-2 individual acts as a mimicker, that is, $G = \hat{G}$. This means that capital income taxation is redundant. On the other hand, equation (12) shows that the government should not distort type-2 individuals' saving behavior, making zero marginal capital income tax rate desirable.

4. Robustness

So far, we have assumed that the amount of savings are observable. In this section,

[5] The intuition is as follows: type-1 individuals do not have their incentive compatibility constraint tightened by private donation to a public good by type-2 individuals and thus inducing type-2 individuals to donate to a public good allows the government to reduce mimicker's utility due to $G > \hat{G}$, that is, it relaxes the binding incentive constraint.

we explore the robustness of our results by assuming that the amount of savings is unobservable, that is, the government is not allowed to employ nonlinear capital income taxes, and show that our main conclusion is robust as long as private donations to a public good are observable, otherwise it is ambiguous. We also analyze the model with the endogenous input prices.

4.1 Linear capital income taxation and nonlinear subsidies for private donations

First, we examine a case in which the government observes both private donations to a public good and labor income for each type, but is unable to observe capital income for each type. Therefore, the government can only levy linear tax on savings at rate t_s.

Following the traditional literatures in optimal income taxation, we decompose the individual's problem into two stages. First, each individual chooses the amount of labor supply and private donations to a public good, which determines disposable income $R^i \equiv y^i - T(y^i) - g^i + \tau(g^i)$, given nonlinear labor income taxes and subsidies for private donations. Second, disposable income is allocated into private consumption in the first and the second period. We suppose that individuals anticipate the outcome of the second stage at the first stage. In the second stage, type i individuals choose c^i and x^i to maximize the sub-utility function $u(c^i, x^i, G)$ subject to individual's budget constraint which is given by

$$c^i + s^i = R^i \tag{13}$$

$$s^i = q_s x^i \tag{14}$$

where $q_s \equiv 1/(1 + r(1 - t_s))$. The first-order condition is given by

$$\frac{u_c(c^i, x^i, G)}{u_x(c^i, x^i, G)} = \frac{1}{q_s} \tag{15}$$

On this occasion, we define the sub-indirect utility function for type i individuals as $V^i \equiv V(q_s, R^i, G) \equiv u(c_i^*, x_i^*, G)$, where $c_i^* \equiv c(q_s, R^i, G)$ denotes the optimal solution with respect to the consumption in the first period and $x_i^* \equiv x(q_s, R^i, G)$ the optimal solution with respect to the consumption in the second period.

The government chooses q_s, G, g^1, g^2, g^G, R^1, R^2, y^1, and y^2 to maximize the social welfare subject to the government's budget constraint, and the incentive compatibility constraint, which is expressed by

$$V(q_s, R^2, G) - v(\frac{y^2}{w^2}) \geq V(q_s, R^1, \hat{G}) - v(\frac{y^1}{w^2}) \tag{16}$$

where $\hat{V} \equiv V(q_s, R^1, \hat{G})$ indicates the mimicker's sub-indirect utility. Here, we define the optimal solution for the mimicker with respect to the consumption in the first and the second period as $\hat{c}^* \equiv c(q_s, R^1, \hat{G})$ and $\hat{x}^* \equiv x(q_s, R^1, \hat{G})$, respectively.[6]

The first-order conditions are shown in Appendix C. By rearranging these results, we can derive the optimal linear capital income tax rate as follows:

$$\frac{rt_s q_s}{1+r} = \frac{\tilde{\lambda}\frac{\partial \hat{V}}{\partial R^1}(\hat{x}^* - x_1^*)}{-\tilde{\gamma}\sum_i \pi^i \frac{\partial \tilde{x}_i^*}{\partial q_s}} \tag{17}$$

Equation (17) is consistent with the formula for optimal linear tax rate proposed by Edwards et al. (1994) and Nava et al. (1996). The deviation form the Atkinson-Stiglitz theorem relies on the numerator, that is, the difference of the demand for the consumption in the second period between the mimicker and the person being mimicked. Because only type-2 individuals contribute to a public good as with nonlinear capital tax instruments (see, Appendix D), that is, we have $G > \hat{G}$. Thus, the Atkinson-Stiglitz theorem is not valid. Therefore, even though the government cannot observe savings for each type, weak-separability of preferences between consumption and labor supply is not sufficient condition to make capital income taxation redundant. Note that when there is an infinite population, we obtain $G = \hat{G}$, and then $t_s = 0$.

4.2 Linear capital income taxation and linear subsidies for private donations

Next, this sub-section extends the model in which the government cannot observe both the amount of private donations to a public good and savings for each type. Therefore, the only nonlinear tax instrument is labor income taxation, and the government can only levy linear subsidies on private donations at rate t_g and linear tax on savings at rate t_s.

We now turn to the analysis of individual's behavior. In the second stage, type i individuals choose c^i, x^i, and g^i to maximize the sub-utility function $u(c^i, x^i, G)$ subject to individual's budget constraint which is given by

$$c^i + s^i + q_g g^i = y^i - T(y^i) \equiv R^i \tag{18}$$

$$s^i = q_s x^i \tag{19}$$

where $q_g \equiv 1 - t_g$ and $q_s \equiv 1/(1 + r(1 - t_s))$. The first-order conditions are given by

$$\frac{u_c(c^i, x^i, g^i + G_{-i})}{u_x(c^i, x^i, g^i + G_{-i})} = \frac{1}{q_s} \tag{20}$$

[6] We omit the first-order conditions with respect to G, y^1, and y^2 since we focus on the characterization of optimal linear capital income tax rates.

$$\frac{u_G(c^i, x^i, g^i + G_{-i})}{u_c(c^i, x^i, g^i + G_{-i})} = q_g \tag{21}$$

where $G_{-i} \equiv (\pi^i - 1)g^i + \pi^j g^j + g^G$, $i \neq j = 1, 2$. Here, let $c^i = c(q_s, q_g, R^i, G_{-i})$, $i = 1, 2$, be the best response function of the consumption in the first period, $x^i = x(q_s, q_g, R^i, G_{-i})$, $i = 1, 2$, the best response function of the consumption in the second period, and $g^i = g(q_s, q_g, R^i, G_{-i})$, $i = 1, 2$, the best response function of private donations to a public good. We can define the sub-indirect utility function for type i as $V^i \equiv V^i(q_g, q_s, R^1, R^2, g^G) \equiv u(c_i^*, x_i^*, g_i^* + G_{-i}^*)$ where the superscript (*) refers to the Nash equilibrium outcome. Under linear tax policy, the incentive constraint preventing high-type individuals from mimicking low-type ones is expressed by:

$$V^2(q_g, q_s, R^1, R^2, g^G) - v(\frac{y^2}{w^2}) \geq \hat{V}(q_g, q_s, R^1, R^2, g^G) - v(\frac{y^1}{w^2}) \tag{22}$$

where $\hat{V}(q_g, q_s, R^1, R^2, g^G)$ indicates the mimicker's sub-indirect utility. Here, we define the best response function for the mimicker with respect to the consumption in the first period as $\hat{c} \equiv c(q_g, q_s, R^1, \tilde{G}_{-2})$, the best response function for the mimicker with respect to the consumption in the second period as $\hat{x} \equiv x(q_g, q_s, R^1, \tilde{G}_{-2})$, and the best response function for mimickers with respect to the private donation to a public good as $\hat{g} \equiv g(q_g, q_s, R^1, \tilde{G}_{-2})$, where $\tilde{G}_{-2} = \pi^1 \tilde{g}^1 + (\pi^2 - 1)\tilde{g}^2 + g^G$ and $\tilde{g}^i$ is type i's private donations in the presence of the mimicker. Here, we define $\hat{G} \equiv \hat{g} + \tilde{G}_{-2}$ as the aggregate amount of public good faced by the mimicker. Note that it is not necessarily that $g_1^* = \tilde{g}_1^*$ or $g_2^* = \tilde{g}_2^*$ holds since these realize as a Nash equilibrium in contrast with the case in which the government can observe private donations to a public good and thus design the allocation for each type.

To sum up, the government chooses q_s, q_g, R^1, R^2, g^G, y^1, and y^2 to maximize the social welfare subject to the government's budget constraint, and the incentive compatibility constraint.[7] Solving the planning problem yields:

$$\begin{pmatrix} -t_g \\ \frac{rt_s q_s}{1+r} \end{pmatrix} = -\frac{1}{\bar{\gamma}}\Delta^{-1} \begin{pmatrix} \sum_{i\neq j=1,2} \pi^i u_G^i \left(\frac{\partial G_{-i}^*}{\partial q_g} + \frac{\partial G_{-i}^*}{\partial R^i} g_i^* + \frac{\partial G_{-i}^*}{\partial R^j} g_j^* \right) \\ \sum_{i\neq j=1,2} \pi^i u_G^i \left(\frac{\partial G_{-i}^*}{\partial q_s} + \frac{\partial G_{-i}^*}{\partial R^i} g_i^* + \frac{\partial G_{-i}^*}{\partial R^j} g_j^* \right) \end{pmatrix} - \frac{\bar{\lambda}\hat{u}_c}{\bar{\gamma}}\Delta^{-1} \begin{pmatrix} \hat{g} - g_1^* \\ \hat{x} - x_1^* \end{pmatrix}$$
$$- \frac{\bar{\lambda}}{\bar{\gamma}}\Delta^{-1} \begin{pmatrix} u_G^2 \left(\frac{\partial G_{-2}^*}{\partial q_g} + \frac{\partial G_{-2}^*}{\partial R^1} g_1^* + \frac{\partial G_{-2}^*}{\partial R^2} g_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_g} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} g_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} g_2^* \right) \\ u_G^2 \left(\frac{\partial G_{-2}^*}{\partial q_s} + \frac{\partial G_{-2}^*}{\partial R^1} x_1^* + \frac{\partial G_{-2}^*}{\partial R^2} x_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_s} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} x_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} x_2^* \right) \end{pmatrix} \tag{23}$$

[7] Following the argument in footnote 6, we omit the first-order conditions with respect to g^G, y^1, and y^2.

where Δ^{-1} is the inverse matrix of Δ which denotes the 2×2 matrix as follows:

$$\Delta \equiv \begin{pmatrix} \sum_i \pi^i \frac{\partial g_i^*}{\partial q_g} + \sum_i \pi^i \frac{\partial g_i^*}{\partial R^i} g_i^* + \sum_{i \neq j = 1,2} \pi^i \frac{\partial g_i^*}{\partial R^j} g_j^* & \sum_i \pi^i \frac{\partial x_i^*}{\partial q_g} + \sum_i \pi^i \frac{\partial x_i^*}{\partial R^i} g_i^* + \sum_{i \neq j = 1,2} \pi^i \frac{\partial x_i^*}{\partial R^j} g_j^* \\ \sum_i \pi^i \frac{\partial g_i^*}{\partial q_s} + \sum_i \pi^i \frac{\partial g_i^*}{\partial R^i} x_i^* + \sum_{i \neq j = 1,2} \pi^i \frac{\partial g_i^*}{\partial R^j} x_j^* & \sum_i \pi^i \frac{\partial x_i^*}{\partial q_s} + \sum_i \pi^i \frac{\partial x_i^*}{\partial R^i} x_i^* + \sum_{i \neq j = 1,2} \pi^i \frac{\partial x_i^*}{\partial R^j} x_j^* \end{pmatrix}$$

When the government cannot observe private donations to a public good for each types, the optimal tax formula consists of three terms. The first and second terms reflect well known effects respectively: the Pigouvian and non-Pigouvian elements discussed by Cremer et al. (1998).[8] The third term is the novel term, which comes from the different impact of the response of the other individuals to the perturbation in parameters to the mimicker and the high-skilled agent who does not behave as a mimicker. This is an additional information for the government to relax the binding incentive constraint. The third term stems from the assumption of a finite population since individuals' private donations can affect the total amount of a public good. If there is an infinite population, the term vanishes, which is consistent with Cremer et al. (1998).

The condition to restore Atkinson–Stiglitz theorem crucially depends on private donations to a public good among individuals. If $g_1^* = \tilde{g}_1^* = g_2^* = \tilde{g}_2^*$, we have $G_{-1}^* = \tilde{G}_{-2}$, and then $g_1^* = \hat{g}^*$ and $x_1^* = \hat{x}$. In addition, since it causes $G = \hat{G}$, the third term in the right hand side vanishes. Therefore, Atkinson–Stiglitz theorem remains. However, in contrast with the observability of private donations to a public good, we cannot analytically compare the amount of private donations among individuals since the government cannot directly control their private donations. As the same with the previous sub-section, if there is an infinite population, no person donates to a public good since the private donation does not affect the aggregate amount of public good. Thus, Atkinson–Stiglitz theorem is valid.

5. Concluding Remarks

This study is largely relevant to debates on the desirability of capital income taxes. Since Ordover and Phelps (1979) seminal work, a large body of literature has accumulated on whether capital income taxes are required from the viewpoint of heterogeneous

[8]If the public good is additively separable in the utility function, the marginal utility of the public good of a low-skilled individual coincides with that of a high-skilled individual, that is, $u_G \equiv u_G^1 = u_G^2$. In this case, the Pigouvian term is simplified as follows.

$$\begin{aligned} \begin{pmatrix} -t_g \\ \frac{rt_s q_s}{1+r} \end{pmatrix} = & -\frac{u_G}{\bar{\gamma}} \begin{pmatrix} \pi^1 + \pi^2 - 1 \\ 0 \end{pmatrix} - \frac{\bar{\lambda}\hat{u}_c}{\bar{\gamma}} \begin{pmatrix} \hat{g} - g_1^* \\ \hat{x} - x_1^* \end{pmatrix} \\ & - \frac{\bar{\lambda}}{\bar{\gamma}} \begin{pmatrix} u_G \left(\frac{\partial G_{-2}^*}{\partial q_g} + \frac{\partial G_{-2}^*}{\partial R^1} g_1^* + \frac{\partial G_{-2}^*}{\partial R^2} g_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_g} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} g_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} g_2^* \right) \\ u_G \left(\frac{\partial G_{-2}^*}{\partial q_s} + \frac{\partial G_{-2}^*}{\partial R^1} x_1^* + \frac{\partial G_{-2}^*}{\partial R^2} x_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_s} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} x_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} x_2^* \right) \end{pmatrix} \end{aligned} \tag{24}$$

tastes in private consumptions, even though the utility function is weakly separable between private goods and leisure. For instance, Boadway et al. (2000), Cremer et al. (2001), and Diamond and Spinnewijn (2011) consider a multidimensional heterogeneity setting in which individuals differ in not only earning abilities but also other characteristics such as initial endowments (bequest or inheritance) and discount rates, which are assumptions. By contrast, this study provides additional economic rationale for capital income taxes from the viewpoint of economic behavior that is, in reality, individuals deduct charitable contributions. Under the standard optimal tax approach, we show that the government should design taxes on capital income to supplement its redistribution policy when individuals can contribute to a public good. This persists even if the additive and separable preference between consumption and labor supply is satisfied and individuals differ in only earning abilities.

The theoretical contribution of this paper is as follows. Although we show that Atkinson–Stiglitz theorem breaks down as a result of heterogeneous preferences, as in the case of Saez (2002), we justify capital income taxes by clarifying the source of heterogeneity on the basis of individual behavior and not assumptions. It is worth noting that our justification is based on the assumption there is a finite population. Pirttilä and Tuomala (1997) examine the commodity taxation on an externality-generating good and nonlinear taxation on labor income under the condition of an infinite population. They show that the optimal tax formula reflects the two types of terms, that is, the externality internalizing effect and the influence through the incentive compatibility constraint. When the preference is assumed to be additive and separable between consumption and leisure, policy outcomes in their paper are consistent with standard Pigouvian taxes. However, under the setting of a finite population as our study, the corresponding tax formula includes the novel term, the interaction between the Pigouvian term and the self-selection term, which is the right hand side of equation (11). This is true even if the additive and separable preference is assumed.

Our paper derives a condition according to which capital income of low-income earners should be taxed or not. This policy implication depends on the shape of utility function, in particular, the signs of $\frac{\partial^2 u}{\partial c^i \partial G}$ and $\frac{\partial^2 u}{\partial x^i \partial G}$. This is an important issue of empirical study.

References

Andreoni, J., 1988. Privately provided public goods in a large economy: The limits of altruism. *Journal of Public Economics* 35, 57–73.

Atkinson, A. B., Stiglitz, J. E., 1976. The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics* 6, 55–75.

Auten, G. E., Sieg, H., Clotfelter, C. T., 2002. Charitable giving, income, and taxes: An analysis of panel data. *American Economic Review* 92, 371–382.

Boadway, R., Marchand, M., Pestieau, P., 2000. Redistribution with unobservable bequests: A case for taxing capital income. *Scandinavian Journal of Economics* 102, 253–267.

Caputo, M. R., 1996. The envelope theorem and comparative statics of Nash equilibria. *Games and Economic Behavior* 13, 201–224.

Cremer, H., Gahvari, F., Ladoux, N., 1998. Externalities and optimal taxation. *Journal of Public Economics* 70, 343–364.

Cremer, H., Pestieau, P., Rochet, J.-C., 2001. Direct versus indirect taxation: The design of the tax structure revisited. *International Economic Review* 42, 781–800.

Diamond, P. A., 2006. Optimal tax treatment of private contributions for public goods with and without warm glow preferences. *Journal of Public Economics* 90, 897–919.

Diamond, P. A., Spinnewijn, J., 2011. Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy* 3, 52–76.

Edwards, J., Keen, M., Tuomala, M., 1994. Income tax, commodity taxes and public good provision: A brief guide. *FinanzArchiv* 51, 472–497.

Hamilton, J., Slutsky, S., 2007. Optimal nonlinear income taxation with a finite population. *Journal of Economic Theory* 132, 548–556.

Hood, R. D., Martin, S. A., Osberg, L., 1977. Economic determinants of individuals charitable donations in Canada. *The Canadian Journal of Economics* 10, 653–669.

Morita, S., Obara, T., 2016. Optimal capital income taxation in the case of private donations to public goods, Discussion papers in Economics and Business, Graduate School of economics and Osaka School of International Public Policy (OSSIP), Osaka University, No.16-21.

Nava, M., Schroyen, F., Marchand, M., 1996. Optimal fiscal and public expenditure policy in a two-class economy. *Journal of Public Economics* 61, 119–137.

Ordover, J. A., Phelps, E. S., 1979. The concept of optimal taxation in the overlapping-generations model of capital and wealth. *Journal of Public Economics* 12, 1–26.

Piketty, T., 1993. Implementation of first-best allocations via generalized tax schedules. *Journal of Economic Theory* 61, 23–41.

Pirttilä, J., Tuomala, M., 1997. Income tax, commodity tax and environmental policy. *International Tax and Public Finance* 4, 379–393.

Pirttilä, J., Tuomala, M., 2001. On optimal non-linear taxation and public good provision in an overlapping generations economy. *Journal of Public Economics* 79, 485–501.

Saez, E., 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics* 83, 217–230.

Saez, E., 2004. The optimal treatment of tax expenditures. *Journal of Public Economics* 88, 2657–2684.

Appendix A

The first order conditions associated with c^1, x^1, c^2, and x^2 are

$$\frac{\partial \mathcal{L}}{\partial c^1} = \pi^1 u_c(c^1, x^1, G) - \gamma\pi^1 - \lambda u_c(c^1, x^1, \hat{G}) = 0 \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial x^1} = \pi^1 u_x(c^1, x^1, G) - \frac{\gamma\pi^1}{(1+r)} - \lambda u_x(c^1, x^1, \hat{G}) = 0 \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial c^2} = \pi^2 u_c(c^2, x^2, G) - \gamma\pi^2 + \lambda u_c(c^2, x^2, G) = 0 \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial x^2} = \pi^2 u_x(c^2, x^2, G) - \frac{\gamma\pi^2}{(1+r)} + \lambda u_x(c^2, x^2, G) = 0 \tag{A.4}$$

Combining equation (A.1) with equation (A.2) yields:

$$\pi^1\{u_c(c^1, x^1, G) - (1+r)u_x(c^1, x^1, G)\} = \lambda\{u_c(c^1, x^1, \hat{G}) - (1+r)u_x(c^1, x^1, \hat{G})\} \tag{A.5}$$

Combining equation (5) with equation (A.5) yields:

$$(\pi^1 u_x(c^1, x^1, G) - \lambda u_x(c^1, x^1, \hat{G}))r\Phi'(rs^1) = \lambda u_x(c^1, x^1, \hat{G})(\frac{u_c(c^1, x^1, G)}{u_x(c^1, x^1, G)} - \frac{u_c(c^1, x^1, \hat{G})}{u_x(c^1, x^1, \hat{G})}) \tag{A.6}$$

Substituting equation (A.2) into the term in the brackets of the left hand side, we obtain equation (11). Similarly, combining equation (A.3) with equation (A.4) yields:

$$-\pi^2\{u_c(c^2, x^2, G) - (1+r)u_x(c^2, x^2, G)\} = \lambda\{u_c(c^2, x^2, G) - (1+r)u_x(c^2, x^2, G)\} \tag{A.7}$$

This can be rewritten as follows:

$$(\pi^2 + \lambda)u_x(c^2, x^2, G)r\Phi'(rs^2) = 0 \tag{A.8}$$

Equation (A.3) implies that $\pi^2+\lambda$ is positive. Then, equation (A.8) implies that $\Phi'(rs^2)$ is zero. □

Appendix B

Differentiating $\mathcal{L}$ with respect to g^G, g^1, and g^2 implies

$$\frac{\partial \mathcal{L}}{\partial g^G} = -\gamma + \mu \tag{B.1}$$

$$\frac{\partial \mathcal{L}}{\partial g^1} = -\gamma\pi^1 - \lambda u_G(c^1, x^1, \hat{G}) + \mu\pi^1 \tag{B.2}$$

$$\frac{\partial \mathcal{L}}{\partial g^2} = -\gamma\pi^2 + \lambda u_G(c^1, x^1, \hat{G}) + \mu\pi^2 \tag{B.3}$$

If equation (B.1) is equal to zero, equation (B.3) is as follows:

$$\frac{\partial \mathcal{L}}{\partial g^2} = \lambda u_G(c^1, x^1, \hat{G}) > 0 \tag{B.4}$$

In this case, the optimal solution does not exist because of diverging. Therefore, at the optimum, we must have $\frac{\partial \mathcal{L}}{\partial g^G} < 0$ and $g^G = 0$ to satisfy Kuhn-Tucker conditions.

Given this condition, from equation (B.2), no contribution to a public good of type 1 individuals is optimal, that is, $g^1 = 0$. On the other hand, the private donation to a public good of type 2 individuals is not zero because the second term in equation (B.3) is sufficiently larger than the sum of the first and third term by the Inada condition when g^2 is close to zero given $g^1 = g^G = 0$. Therefore, g^2 is positive. In addition, g^2 is an interior solution. As g^2 is close to infinity, $\frac{\partial \mathcal{L}}{\partial g^2}$ converges to $-\gamma\pi^2 + \mu\pi^2$ which is negative. This implies that g^2 must not be corner solution at the optimum.

□

Appendix C

The corresponding Lagrangian is formulated as follows:

$$\begin{aligned}\tilde{\mathcal{L}} &= \tilde{W} + \tilde{\mu}\left[\sum_i g^i\pi^i + g^G - G\right] \\ &+ \tilde{\gamma}\left[\sum_i \pi^i(y^i - g^i - R^i) + \frac{1}{1+r}\sum_i \pi^i(q_s(1+r)-1)x(q_s, R^i, G) - g^G\right] \\ &+ \tilde{\lambda}\left[V(q_s, R^2, G) - v(\frac{y^2}{w^2}) - V(q_s, R^1, \hat{G}) + v(\frac{y^1}{w^2})\right]\end{aligned} \tag{C.1}$$

where $\tilde{\gamma}$, $\tilde{\lambda}$, and $\tilde{\mu}$ are the Lagrange multipliers. The first order conditions associated with q_s, R^1, and R^2 are

$$\frac{\partial \tilde{\mathcal{L}}}{\partial q_s} = \sum_i \pi^i \frac{\partial V^i}{\partial q_s} + \tilde{\gamma} \sum_i \pi^i \left(x_i^* + \frac{q_s(1+r)-1}{1+r} \frac{\partial x_i^*}{\partial q_s} \right) + \tilde{\lambda} \left(\frac{\partial V^2}{\partial q_s} - \frac{\partial \hat{V}^2}{\partial q_s} \right) = 0 \quad \text{(C.2)}$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial R^1} = \pi^1 \frac{\partial V^1}{\partial R^1} - \tilde{\gamma}\pi^1 + \tilde{\gamma}\pi^1 \frac{q_s(1+r)-1}{1+r} \frac{\partial x_1^*}{\partial R^1} - \tilde{\lambda} \frac{\partial \hat{V}^2}{\partial R^1} = 0 \quad \text{(C.3)}$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial R^2} = \pi^2 \frac{\partial V^2}{\partial R^2} - \tilde{\gamma}\pi^2 + \tilde{\gamma}\pi^2 \frac{q_s(1+r)-1}{1+r} \frac{\partial x_2^*}{\partial R^2} + \tilde{\lambda} \frac{\partial V^2}{\partial R^2} = 0 \quad \text{(C.4)}$$

We now combine these constraints by taking

$$\frac{\partial \tilde{\mathcal{L}}}{\partial q_s} + \sum_i \frac{\partial \tilde{\mathcal{L}}}{\partial R^i} x_i^* \quad \text{(C.5)}$$

From the Roy's identity and the Slutsky decomposition, we can get the following relationships:

$$\frac{\partial V^i}{\partial q_s} = -\frac{\partial V^i}{\partial R^i} \cdot x_i^* \quad \text{(C.6)}$$

$$\frac{\partial x_i^*}{\partial q_s} = \frac{\partial \tilde{x}_i^*}{\partial q_s} - \frac{\partial x_i^*}{\partial R^i} \cdot x_i^* \quad \text{(C.7)}$$

where $\tilde{x}_i^*$ indicates the compensated demand function of type i individuals for the consumption in the second period. Using Roy's identity and Slutsky decomposition, this gives

$$\tilde{\gamma} \frac{q_s(1+r)-1}{1+r} \sum_i \pi^i \frac{\partial \tilde{x}_i^*}{\partial q_s} + \tilde{\lambda} \frac{\partial \hat{V}}{\partial R^1} (\hat{x}^* - x_1^*) = 0 \quad \text{(C.8)}$$

Thus, we can obtain equation (17). □

Appendix D

Differentiating $\tilde{\mathcal{L}}$ with respect to g^G, g^1, and g^2 implies

$$\frac{\partial \tilde{\mathcal{L}}}{\partial g^G} = -\tilde{\gamma} + \tilde{\mu} \quad \text{(D.1)}$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial g^1} = -\tilde{\gamma}\pi^1 - \tilde{\lambda} u_G(q_s, R^1, \hat{G}) + \tilde{\mu}\pi^1 \quad \text{(D.2)}$$

$$\frac{\partial \tilde{\mathcal{L}}}{\partial g^2} = -\tilde{\gamma}\pi^2 + \tilde{\lambda} u_G(q_s, R^1, \hat{G}) + \tilde{\mu}\pi^2 \tag{D.3}$$

Following the same proof as in Appendix B, we conclude that $g^G = g^1 = 0$ and $g^2 > 0$.
□

Appendix E

Using Theorem 1 of Caputo (1996), we can get the following relationships.

$$\frac{\partial V^i}{\partial q_g} = -\phi g_i^* + u_G^i \frac{\partial G_{-i}^*}{\partial q_g} \tag{E.1}$$

$$\frac{\partial V^i}{\partial q_s} = -\phi x_i^* + u_G^i \frac{\partial G_{-i}^*}{\partial q_s} \tag{E.2}$$

$$\frac{\partial V^i}{\partial R^i} = \phi + u_G^i \frac{\partial G_{-i}^*}{\partial R^i} \tag{E.3}$$

$$\frac{\partial V^i}{\partial R^j} = u_G^i \frac{\partial G_{-i}^*}{\partial R^j}, \ i \neq j \tag{E.4}$$

where let ϕ be the Lagrange multiplier with respect to individual's budget constraint.

The corresponding Lagrangian is formulated as follows:

$$\begin{aligned} \bar{\mathcal{L}} = \bar{W} + \bar{\lambda}\Big[V^2(q_g, q_s, R^1, R^2, g^G) - v(\frac{y^2}{w^2}) - \hat{V}(q_g, q_s, R^1, R^2, g^G) + v(\frac{y^1}{w^2}) \Big] \\ + \bar{\gamma}\Big[\sum_i \pi^i (y^i - R^i) + \sum_i \pi^i (q_g - 1) g_i^* + \frac{1}{1+r} \sum_i \pi^i (q_s(1+r) - 1) x_i^* - g^G \Big] \end{aligned} \tag{E.5}$$

where $\bar{\gamma}$ and $\bar{\lambda}$ are the Lagrange multipliers. The first order conditions associated with q_s, q_g, R^1, and R^2 are as follows:

$$\begin{aligned} \frac{\partial \bar{\mathcal{L}}}{\partial q_s} = \sum_i \pi^i \frac{\partial V^i}{\partial q_s} + \bar{\lambda}\Big[\frac{\partial V^2}{\partial q_s} - \frac{\partial \hat{V}}{\partial q_s} \Big] \\ + \bar{\gamma}\Big[\frac{1}{1+r} \sum_i \pi^i \Big((1+r) x_i^* + (q_s(1+r) - 1) \frac{\partial x_i^*}{\partial q_s} \Big) + \sum_i \pi^i (q_g - 1) \frac{\partial g_i^*}{\partial q_s} \Big] = 0 \end{aligned} \tag{E.6}$$

$$\begin{aligned} \frac{\partial \bar{\mathcal{L}}}{\partial q_g} = \sum_i \pi^i \frac{\partial V^i}{\partial q_g} + \bar{\lambda}\Big[\frac{\partial V^2}{\partial q_g} - \frac{\partial \hat{V}}{\partial q_g} \Big] \\ + \bar{\gamma}\Big[\sum_i \pi^i \Big(g_i^* + (q_g - 1) \frac{\partial g_i^*}{\partial q_g} \Big) + \frac{1}{1+r} \sum_i \pi^i (q_s(1+r) - 1) \frac{\partial x_i^*}{\partial q_g} \Big] = 0 \end{aligned} \tag{E.7}$$

$$\begin{aligned}\frac{\partial \bar{\mathcal{L}}}{\partial R^1} &= \pi^1 \frac{\partial V^1}{\partial R^1} + \pi^2 \frac{\partial V^2}{\partial R^1} + \bar{\lambda}\Big[\frac{\partial V^2}{\partial R^1} - \frac{\partial \hat{V}}{\partial R^1}\Big] \\ &+ \bar{\gamma}\Big[-\pi^1 + \pi^1(q_g-1)\frac{\partial g_1^*}{\partial R^1} + \pi^2(q_g-1)\frac{\partial g_2^*}{\partial R^1} \\ &+ \frac{1}{1+r}\pi^1(q_s(1+r)-1)\frac{\partial x_1^*}{\partial R^1} + \frac{1}{1+r}\pi^2(q_s(1+r)-1)\frac{\partial x_2^*}{\partial R^1}\Big] = 0\end{aligned} \tag{E.8}$$

$$\begin{aligned}\frac{\partial \bar{\mathcal{L}}}{\partial R^2} &= \pi^1 \frac{\partial V^1}{\partial R^2} + \pi^2 \frac{\partial V^2}{\partial R^2} + \bar{\lambda}\Big[\frac{\partial V^2}{\partial R^2} - \frac{\partial \hat{V}}{\partial R^2}\Big] \\ &+ \bar{\gamma}\Big[-\pi^2 + \pi^2(q_g-1)\frac{\partial g_2^*}{\partial R^2} + \pi^1(q_g-1)\frac{\partial g_1^*}{\partial R^2} \\ &+ \frac{1}{1+r}\pi^2(q_s(1+r)-1)\frac{\partial x_2^*}{\partial R^2} + \frac{1}{1+r}\pi^1(q_s(1+r)-1)\frac{\partial x_1^*}{\partial R^2}\Big] = 0\end{aligned} \tag{E.9}$$

We now combine these constraints by taking

$$\frac{\partial \bar{\mathcal{L}}}{\partial q_g} + \sum_i \frac{\partial \bar{\mathcal{L}}}{\partial R^i} g_i^* \tag{E.10}$$

and

$$\frac{\partial \bar{\mathcal{L}}}{\partial q_s} + \sum_i \frac{\partial \bar{\mathcal{L}}}{\partial R^i} x_i^* \tag{E.11}$$

These give

$$\begin{aligned}&\sum_{i\neq j=1,2} \pi^i \Big[\frac{\partial V^i}{\partial q_g} + \frac{\partial V^i}{\partial R^i} g_i^* + \frac{\partial V^i}{\partial R^j} g_j^*\Big] + \bar{\gamma}(q_g-1)\Big[\sum_i \pi^i \frac{\partial g_i^*}{\partial q_g} + \sum_i \pi^i \frac{\partial g_i^*}{\partial R^i} g_i^* + \sum_{i\neq j=1,2} \pi^i \frac{\partial g_i^*}{\partial R^j} g_j^*\Big] \\ &+ \bar{\gamma}\frac{q_s(1+r)-1}{1+r}\Big[\sum_i \pi^i \frac{\partial x_i^*}{\partial q_g} + \sum_i \pi^i \frac{\partial x_i^*}{\partial R^i} g_i^* + \sum_{i\neq j=1,2} \pi^i \frac{\partial x_i^*}{\partial R^j} g_j^*\Big] \\ &+ \bar{\lambda}\Big[\frac{\partial V^2}{\partial q_g} + \frac{\partial V^2}{\partial R^1} g_1^* + \frac{\partial V^2}{\partial R^2} g_2^* - \frac{\partial \hat{V}}{\partial q_g} - \frac{\partial \hat{V}}{\partial R^1} g_1^* - \frac{\partial \hat{V}}{\partial R^2} g_2^*\Big] = 0\end{aligned} \tag{E.12}$$

and

$$\begin{aligned}&\sum_{i\neq j=1,2} \pi^i \Big[\frac{\partial V^i}{\partial q_s} + \frac{\partial V^i}{\partial R^i} g_i^* + \frac{\partial V^i}{\partial R^j} g_j^*\Big] + \bar{\gamma}(q_g-1)\Big[\sum_i \pi^i \frac{\partial g_i^*}{\partial q_s} + \sum_i \pi^i \frac{\partial g_i^*}{\partial R^i} x_i^* + \sum_{i\neq j=1,2} \pi^i \frac{\partial g_i^*}{\partial R^j} x_j^*\Big] \\ &+ \bar{\gamma}\frac{q_s(1+r)-1}{1+r}\Big[\sum_i \pi^i \frac{\partial x_i^*}{\partial q_s} + \sum_i \pi^i \frac{\partial x_i^*}{\partial R^i} x_i^* + \sum_{i\neq j=1,2} \pi^i \frac{\partial x_i^*}{\partial R^j} x_j^*\Big] \\ &+ \bar{\lambda}\Big[\frac{\partial V^2}{\partial q_s} + \frac{\partial V^2}{\partial R^1} x_1^* + \frac{\partial V^2}{\partial R^2} x_2^* - \frac{\partial \hat{V}}{\partial q_s} - \frac{\partial \hat{V}}{\partial R^1} x_1^* - \frac{\partial \hat{V}}{\partial R^2} x_2^*\Big] = 0\end{aligned} \tag{E.13}$$

Using equation from (E.1) to (E.4), (E.12) and (E.13) are transformed as

$$\sum_{i\neq j=1,2}\pi^i u_G^i\left[\frac{\partial G_{-i}^*}{\partial q_g}+\frac{\partial G_{-i}^*}{\partial R^i}g_i^*+\frac{\partial G_{-i}^*}{\partial R^j}g_j^*\right]$$
$$+\bar{\gamma}(q_g-1)\left[\sum_i\pi^i\frac{\partial g_i^*}{\partial q_g}+\sum_i\pi^i\frac{\partial g_i^*}{\partial R^i}g_i^*+\sum_{i\neq j=1,2}\pi^i\frac{\partial g_i^*}{\partial R^j}g_j^*\right]$$
$$+\bar{\gamma}\frac{q_s(1+r)-1}{1+r}\left[\sum_i\pi^i\frac{\partial x_i^*}{\partial q_g}+\sum_i\pi^i\frac{\partial x_i^*}{\partial R^i}g_i^*+\sum_{i\neq j=1,2}\pi^i\frac{\partial x_i^*}{\partial R^j}g_j^*\right]$$
$$+\bar{\lambda}\left[u_G^2\left(\frac{\partial G_{-2}^*}{\partial q_g}+\frac{\partial G_{-2}^*}{\partial R^1}g_1^*+\frac{\partial G_{-2}^*}{\partial R^2}g_2^*\right)-\hat{u}_G\left(\frac{\partial\tilde{G}_{-2}}{\partial q_g}+\frac{\partial\tilde{G}_{-2}}{\partial R^1}g_1^*+\frac{\partial\tilde{G}_{-2}}{\partial R^2}g_2^*\right)+\hat{u}_c(\hat{g}-g_1^*)\right]$$
$$=0 \tag{E.14}$$

and

$$\sum_{i\neq j=1,2}\pi^i u_G^i\left[\frac{\partial G_{-i}^*}{\partial q_s}+\frac{\partial G_{-i}^*}{\partial R^i}g_i^*+\frac{\partial G_{-i}^*}{\partial R^j}g_j^*\right]$$
$$+\bar{\gamma}(q_g-1)\left[\sum_i\pi^i\frac{\partial g_i^*}{\partial q_s}+\sum_i\pi^i\frac{\partial g_i^*}{\partial R^i}x_i^*+\sum_{i\neq j=1,2}\pi^i\frac{\partial g_i^*}{\partial R^j}x_j^*\right]$$
$$+\bar{\gamma}\frac{q_s(1+r)-1}{1+r}\left[\sum_i\pi^i\frac{\partial x_i^*}{\partial q_s}+\sum_i\pi^i\frac{\partial x_i^*}{\partial R^i}x_i^*+\sum_{i\neq j=1,2}\pi^i\frac{\partial x_i^*}{\partial R^j}x_j^*\right]$$
$$+\bar{\lambda}\left[u_G^2\left(\frac{\partial G_{-2}^*}{\partial q_s}+\frac{\partial G_{-2}^*}{\partial R^1}x_1^*+\frac{\partial G_{-2}^*}{\partial R^2}x_2^*\right)-\hat{u}_G\left(\frac{\partial\tilde{G}_{-2}}{\partial q_s}+\frac{\partial\tilde{G}_{-2}}{\partial R^1}x_1^*+\frac{\partial\tilde{G}_{-2}}{\partial R^2}x_2^*\right)+\hat{u}_c(\hat{x}-x_1^*)\right]$$
$$=0 \tag{E.15}$$

Using matrix notation, (E.14) and (E.15) can be rewritten as

$$\Delta\begin{pmatrix}-t_g\\ \frac{rt_sq_s}{1+r}\end{pmatrix}=-\frac{1}{\bar{\gamma}}\begin{pmatrix}\sum_{i\neq j=1,2}\pi^i u_G^i\left(\frac{\partial G_{-i}^*}{\partial q_g}+\frac{\partial G_{-i}^*}{\partial R^i}g_i^*+\frac{\partial G_{-i}^*}{\partial R^j}g_j^*\right)\\ \sum_{i\neq j=1,2}\pi^i u_G^i\left(\frac{\partial G_{-i}^*}{\partial q_s}+\frac{\partial G_{-i}^*}{\partial R^i}g_i^*+\frac{\partial G_{-i}^*}{\partial R^j}g_j^*\right)\end{pmatrix}-\frac{\bar{\lambda}\hat{u}_c}{\bar{\gamma}}\begin{pmatrix}\hat{g}-g_1^*\\ \hat{x}-x_1^*\end{pmatrix}$$
$$-\frac{\bar{\lambda}}{\bar{\gamma}}\begin{pmatrix}u_G^2\left(\frac{\partial G_{-2}^*}{\partial q_g}+\frac{\partial G_{-2}^*}{\partial R^1}g_1^*+\frac{\partial G_{-2}^*}{\partial R^2}g_2^*\right)-\hat{u}_G\left(\frac{\partial\tilde{G}_{-2}}{\partial q_g}+\frac{\partial\tilde{G}_{-2}}{\partial R^1}g_1^*+\frac{\partial\tilde{G}_{-2}}{\partial R^2}g_2^*\right)\\ u_G^2\left(\frac{\partial G_{-2}^*}{\partial q_s}+\frac{\partial G_{-2}^*}{\partial R^1}x_1^*+\frac{\partial G_{-2}^*}{\partial R^2}x_2^*\right)-\hat{u}_G\left(\frac{\partial\tilde{G}_{-2}}{\partial q_s}+\frac{\partial\tilde{G}_{-2}}{\partial R^1}x_1^*+\frac{\partial\tilde{G}_{-2}}{\partial R^2}x_2^*\right)\end{pmatrix} \tag{E.16}$$

Multiplying equation (E.16) by Δ^{-1} yields (23). If the public good is additively separable in the utility function, the marginal utility of the public good coincides between a low-skilled and a high-skilled individual. In this case, equation (E.16) reduces to

$$\Delta \begin{pmatrix} -t_g \\ \frac{rt_s q_s}{1+r} \end{pmatrix} = -\frac{u_G}{\bar{\gamma}}(\pi^1 + \pi^2 - 1)\delta_1 - \frac{\bar{\lambda}\hat{u}_c}{\bar{\gamma}} \begin{pmatrix} \hat{g} - g_1^* \\ \hat{x} - x_1^* \end{pmatrix}$$
$$- \frac{\bar{\lambda}}{\bar{\gamma}} \begin{pmatrix} u_G \left(\frac{\partial G^*_{-2}}{\partial q_g} + \frac{\partial G^*_{-2}}{\partial R^1} g_1^* + \frac{\partial G^*_{-2}}{\partial R^2} g_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_g} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} g_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} g_2^* \right) \\ u_G \left(\frac{\partial G^*_{-2}}{\partial q_s} + \frac{\partial G^*_{-2}}{\partial R^1} x_1^* + \frac{\partial G^*_{-2}}{\partial R^2} x_2^* \right) - \hat{u}_G \left(\frac{\partial \tilde{G}_{-2}}{\partial q_s} + \frac{\partial \tilde{G}_{-2}}{\partial R^1} x_1^* + \frac{\partial \tilde{G}_{-2}}{\partial R^2} x_2^* \right) \end{pmatrix} \tag{E.17}$$

where δ_1 indicates the first column vector of Δ. Multiplying equation (E.17) by Δ^{-1} yields (24).

□