UNIVERSITÉ TOULOUSE 1 CAPITOLE

THÈSE

Université Fédérale Toulouse Midi-Pyrénées

En vue de l'obtention du

DOCTORAT DE L'UNIVERSITE DE TOULOUSE

Délivré par l'Université Toulouse Capitole

École doctorale : **Sciences Economiques-Toulouse School of Economics**

Présentée et soutenue par

TSUGAWA SHUICHI

le 18 December 2018

Essays on public finance and publicly provided public good

Discipline : **Sciences Economiques**

Unité de recherche : **TSE-R (UMR CNRS 5314 – INRA 1415)**

Directeur de thèse : Professeur, Helmuth, Cremer, Docteur, Université Toulouse 1
Professeur, Jean-Marie, Lozachmeur, Docteur,
Université Toulouse 1

JURY

Rapporteurs Professeur, Pierre, PESTIEAU, Docteur, University of Liege
Professeur, Motohiro, SATO, Docteur, Hitotsubashi University

Suffragants Professeur, Francesca, BARIGOZZI, Docteur, University of Bologna
Professeur, Takuro, YAMASHITA, Docteur, Université Toulouse 1

Essays on public finance and publicly provided public good

Shuichi Tsugawa[1]

September 26, 2018

[1]Toulouse School of Economics; shuichi.tsugawa@ut-capitole.fr

Contents

Outline

This thesis deals with several theoretical subjects about optimal fiscal and government policy. It contains my four works about tax and other redistributive policy, starting with the general introductory survey as the first chapter.

Chapter 2 compares ad-valorem and specific taxation in models where a representative consumer with an exogenous income has both a quality and a quantity choice under perfect competition. In the setting, while ad-valorem tax causes income effect only, specific tax causes both income effect and substitution effect. Therefore, ad-valorem tax decreases consumer demand for both quality and quantity; on the other hand, specific tax decreases consumer demand for quantity. However, the sign of consumer demand for quality is ambiguous and is determined by the curvature of marginal utility on quantity. Additionally, using a constant elasticity of substitution (CES) utility function and a linear price function, we show that ad-valorem tax is superior to specific tax except for the Leontief preference under which the two forms of commodity taxes generate the same tax revenue. The substitution effect caused by specific tax disappears if the elasticity of substitution converges to zero.

In Chapter 3, We examine optimal taxation and public good provision by a government which takes reduction of envy into consideration as one of the constraints. We adopt the notion of extended envy-freeness proposed by Diamantaras and Thomson (1990), called λ-equitability. We derive the modified Samuelson rule at an optimum income tax, and show that, using a constant elasticity of substitution utility function, the direction of distorting the original Samuelson rule to relax λ envy free constraints is crucially determined by the elasticity of substitution. Furthermore, we numerically show that the level of public good increases (or decreases) in the degree of envy-freeness when the provision level is upwardly (or downwardly) distorted. Also, Chapter 4 covers the topic of public good provision under income transfer under that ethical constraint, but allows the social planner to set the surcharge fee for the purpose of excluding some agents whereas we simplify their income as exogenous one (or initial wealth). In this chapter, we study optimal public good provision and user fee in order to exclude some agents by Rawlsian or utilitarian government under lump-sum transfer, constrained by reduction of envy. In particular, we employ the exclusion technique used in Hellwig (2005), *i.e.*, the policymaker decides the level of provision and surcharge fee paid by those making access to it, as well as uniform transfer. Different from Hellwig (2005), we introduce heterogeneity in initial wealth for agents and the envy-free constraint with respect to their one, but not to their tastes for public good. In this setting, we derive the optimal provision level and user fee, and compared to those in Hellwig (2005), for Rawlsian government, the up-charge is lower than the one derived in Hellwig (2005) in order to reduce the envy.

Chapter 5 studies optimal nonlinear income tax schedule at symmetric equilibria at which two symmetric states (or tax authorities) compete in order to attract more taxpayers from the opposite. It is different from the existing papers that taxpayers' wage are endogenously determined by production technology. The optimal tax schedule embraces not only migration effect, but also trickle-down effect coming from endogenous wage, and the migration effect stimulates the trickle-down effect. Compared to previous works, the threat of emigration never disappears in marginal tax rate for high-skilled workers because emigration terms are embedded in the production and such factors have impacts on the productivities or their unit wages.

Acknowledgements

This thesis contains four research papers I wrote in doctoral track of Toulouse School of Economics. I would like to thank our supervisors Helmuth Cremer, Jean-Marie Lozachmeur for their guidance and support. Also, sincerely thanks Takuro Yamashita, Philippe Bontems, Takashi Kunimoto, Motohiro Sato, Jun-ichi Itaya, Pierre Boyer, Francǫis Maniquet, Kaname Miyagishima and Atsushi Yamagishi for invaluable comments and discussions, and Pierre Pestieau and Francesca Barigozzi for jury members of my thesis as well as valuable feedbacks. Finally, I wish to acknowledge my friend and co-author Takuya Obara with whom I investigate three of the four papers, and my family for their support and encouragement.

Chapter 1

General Introduction

This thesis consists of four independent chapters which are linked in contribution to the theory of public economics. Chapter 2 studies the problem of tax incidence and welfare comparison between ad-valorem tax and specific tax, and all the other chapters are based on the assumption that agents are privately informed about their preferences and wages, and use mechanism design theory. Chapter 3 and 4 study public good provision under nonlinear income tax by a government considering reduction of envy among taxpayers, and Chapter 5 uses common agency model under asymmetric information and studies nonlinear income tax under strategic situations, which mean that two symmetric tax authorities attract mobile workers with differential mobile costs in order to collect more taxes.

As said before, topics of this thesis are categorized into the followings: tax incidence, nonlinear income tax, public good provision (with or without exclusion) and tax competition. According to existing literatures, Suits and Musgrave (1953) is the canonical work of tax incidence, which considers ad-valorem and specific tax incidence on monopoly firm. The research question is to compare these two tax systems with respect to price, quantity and profit under the same tax revenues in monopoly market. A monopoly firm decides the quantity of product which determines its selling price per unit and production cost. Under plausible assumptions, they show that under plausible conditions, tax revenues by ad-valorem tax is greater than those by specific (or unit) tax if its firm chooses the same quantity under these two tax systems. The intuition is straightforward. Imagine switching from specific tax to ad-valorem tax where the firm produces the same output as under that unit tax. Output tends to increase with the change because such ad-valorem tax yields the quantity larger than that under specific tax. The price decreases as increasing the output, so the tax wedge between the producer and consumers under is also decreasing since the final price is decreasing in output. On the other hand, the tax wedge under specific tax is constant in output. Therefore, consumer surplus and profit under ad-valorem tax are superior to those under the other. In addition, increasing quantity leads to increasing before tax revenue until the monopoly firm earns the maximized profit without any taxes. Therefore, tax revenue by ad-valorem tax is also larger than that by unit tax. Delipalla and Keen (1992) also study the tax incidence similar to the above and achieve consistent result to it, but they include not only monopoly market, but also imperfect or perfect competitive market.

In addition, they compare ad-valorem tax, unit tax and the mix, but the tax authority still heavily rely on ad-valorem tax according to their results. As the extensions, there are papers which examine the tax incidence under imperfect competition, tax competition on capital income and two-sided market. For instance, see Skeath and Trandel

(1994) and Anderson et al. (2001) for imperfect competition, Lockwood (2004) for tax competition on capital income and Kind et al. (2008) for two-sided market.

Chapter 3 covers the topic of nonlinear income tax and public good provision, and Chapter 5 considers nonlinear income tax competition among symmetric two countries. Mirrlees (1971) is the foundational work of optimal tax that suggests a way to formalize the government's problem dealing explicitly with unobserved heterogeneity. The government can observe their labor income depending on ability and labor effort, but it can observe neither ability nor labor effort. The objective of the government is to maximize social welfare through income redistribution (i.e.,equity) without changing individual behaviors (e.g., labor responses) which affect tax revenues (i.e., efficiency). Contrary on the first-best environment, the unobserved heterogeneity implies that the government cannot employ lump-sum tax, which leads to distortions. The infeasibility of lump-sum tax stems from self-selection constraint to deter high-ability individuals from mimicking low-ability ones. According to the revelation principle and taxation principal which are the canonical mechanism design result, any optimal allocation can be achieved through the tax system inducing high-ability taxpayers to reveal their true types. Under the binding self-selection constraint, the government must rely on a distortionary nonlinear income tax to prevent high-ability from mimicking low-ability. Hence, lump-sum tax is not available. From the fact, since the government cannot implement the redistribution without distorting labor responses, equity and efficiency tradeoff occurs in the Mirrleesian framework. Here, the following question appears: how should the government design income tax schedules given equity and efficiency tradeoff?

Mirrlees (1971) characterizes the optimal marginal income tax rate formula under a continuum of individuals who differ in skill levels. The most famous property of optimal income tax schedules is that the optimal marginal tax rate must be nonnegative and equal zero at the top and the bottom. Stiglitz (1982) studies the same problem under two-class economy where all individuals have either high or low skill. This employs principal-agent model, and adopts incentive compatibility as self-selection constraint, and shows that at optimum, zero marginal tax rate for high-class and the positive one for low class when the government wants to redistribute collected taxes from high class to low class. Also, this paper derives all other cases: non-binding incentive constraint for all and binding incentive constraint for low class only, and it obtains that lump-sum tax is enough to achieve the optimal allocation in the former case whereas zero marginal tax rate for low-class and negative marginal tax rate for high-class implement the (constrained) best allocation in the latter case. However, these properties are not relevant to practical policy design, and many papers studying optimal fiscal policy with sufficient statistics approach emerged recently. Diamond (1998) and Saez (2001) decompose the first-order condition in terms of optimal tax rates into three terms, which is called ABC-formula in Diamond (1998), and clarifies three key parameters that determine optimal tax rates: the taxable income elasticity, the shape of the income distribution, and the social welfare weights. Therefore, the optimal marginal income tax formula is described by real economic factors which can be measured empirically.

The above studies assume that a government can set nonlinear income tax schedule, but in reality, such policy is difficult due to administrative complexity. So, it's natural to study the optimal structure of linear (commodity) taxation. The seminal contribution is the paper of Ramsey (1927) that consider a situation in which the government can only imposes taxes on commodities to collect a given amount of tax revenue. Ramsey (1927) show that commodity taxation on a good should be imposed in inverse proportion to the individual's elasticity of demand for the good. This means that commodities with inelastic demand are levied more heavily to minimize efficiency costs. Ramsey's formula is also formally identical to that obtained by Boiteux (1956) for the socially optimal pricing of a budget-constrained multiproduct monopoly, which is sometimes called the Ramsey-Boiteux formula. Corlett and Hague (1953) generalizes Ramsey's model for more general preferences. They consider a representative household that consumes two goods and leisure. The government can tax the two goods, but not leisure. The Corlett-Hague rule states that a good which is more complementary with leisure should be taxed at a higher rate. Diamond and Mirrlees (1971b) extends Ramsey's result to an economy with several consumers who differ in their wages, but not their preferences. In their analysis, we face a fundamental conflict between equity and efficiency: goods with the lowest price elasticity of demand should have high tax rates for the sake of efficiency, but they are also the ones with the lowest income elasticity of demand that should be taxed leniently on equity grounds. An important result from Diamond and Mirrlees (1971a) is the production efficiency theorem, which says that if the optimal commodity tax system is in place, and if pure profits are all taxed away,[3] production efficiency should apply for the economy as a whole.

Using the Mirrlees framework, Atkinson and Stern (1974) examines the optimal mixed taxation when the government can employ both linear commodity taxation and nonlinear income taxation, called mixed taxation. In the mixed tax framework, the most influential result is that of Atkinson and Stiglitz (1976): when consumer preferences are separable between goods and leisure, there is no need for differentiated commodity tax rates. There have been many papers verifying the robustness of the canonical result. For instance, Cremer et al. (2001) shows that, when individuals differ in skills and endowments, the Atkinson and Stiglitz Theorem is not robust because commodity taxation creates positive effect of relaxing incentive compatibility constraints. Saez (2002) shows that, under the heterogeneity in tastes for consumption goods, the Atkinson and Stiglitz Theorem cannot be adapted if high income earners have a relatively higher taste for this commodity. According to those results, individuals' heterogeneity plays an important role in breaking Atkinson-Stiglitz theorem. In another direction, Naito (1999) introduces public production sector and endogenous wages, and shows that it may be optimal for the tax authority to use as well a shadow price for skilled labor employed in the public production sector that is higher than its market price. Then, public firms would demand less skilled labor reducing its relative wage and thus contributes to redistributing incomes.

As mentioned above, most of papers about optimal labor income taxes assume that individuals' wages are exogenously fixed, but these volatilities are often seen in real world, and lots of studies try to adopt the endogenous wages. Stiglitz (1982) and Stern (1982) initiate its attempt in optimal income tax models, and obtain that taxpayers with high wage face negative marginal tax rate whereas those with low wage face negative marginal tax rate whose absolute value is larger than that without its endogeneity. Endogenous wages in optimal tax policies are related to analyses in Chapter 5.

Not only levying any taxes but also providing public goods are important policies for the government. Public goods will be under-provided by the market due to the free rider problem caused by properties of public goods: non-excludability and non-rivalry. From the reason, public economists are concerned with how the government should design the optimal provision of public goods. Samuelson (1954) suggests the criterion for providing public goods, which states that the optimal provision level equates the sum of marginal willingness to pay for the public good to the marginal cost of providing the public good, which is called the Samuelson rule. However, this criterion is derived under financing public good provision with non-distortionary taxation. Boadway and Keen (1993), which Chapter 3 and 4 are based on, integrates the theory of public goods provision and the theory of optimal income taxation. In their model, the role of the government is to mitigate income inequality through providing public good and income redistribution by considering heterogeneous individuals in terms of innate ability as in Mirrleesian economy. They show that the original Samuelson rule is valid if both private consumption and public good are weakly separable with leisure. Gaube (2005a) follows their work, and investigates conditions which increase or decrease the level of public provision which may not correspond to the distortion of provision rule. Bierbrauer (2009) and Bierbrauer (2014) study the similar model, but embed individuals' preference heterogeneity over public good and aggregate uncertainty about the preferences from the point of a policymaker and take their mimicking their preferences collectively into consideration. As another distortionary taxation, Atkinson and Stern (1974) and Gaube (2000) study a Ramsey economy, where individuals are identical and there is no lump-sum tax. They check Pigou (1947)'s claim: the existence of distortionary taxes was a reason for not ”carrying public expenditures as far” as would be done if we could apply the rule of equating marginal cost to marginal benefit of individuals. Apart from public good provision financed by nonlinear income taxation, there are many attempts to loose non-excludability in the context of public good provision. For instance, Moulin (1994), Schmitz (1997), Norman (2004) and Hellwig (2005a) study public good provision by a government which has the ability to exclude individuals from the enjoyment of a public good. In these papers, excludability is useful because, in bargaining under incomplete information with voluntary participation, the threat of individual exclusion makes participants more willing to contribute to the financing of the public good.

In Chapter 3 and 4, we assume that the government is constrained on reduction of envy among taxpayers. The notion of envy-free allocation introduced by Foley (1967), which means that all agents have no envy toward the others' consumption bundles, is important to consider fairness in a economy. The central question is when Pareto efficient and envy-free allocation exists. The fundamental conflict between efficiency and equity was first observed by Kolm (1971), and Tadenuma (2002) proposed two contrasting principles to construct social preference relations. Also, how much an agent prefers the other's bundle is also important, and cardinality is necessary to check the intensity of envy. The importance of cardinality was recognized by Feldman and Kirman (1974) and Varian (1976b) when they took the utility difference between two individuals as a cardinal measure of envy to construct social welfare functions. Introducing the cardinality, Chaudhuri (1986) and Diamantaras and Thomson (1989) define the weaker envy free notion, and Chapter 3 and 4 adopt it as one of constraints faced by the government.

Tax competition is also related to Chapter 5 in this thesis. Zodrow and Mieszkowski (1986) and Wilson (1986) study capital income tax competition in a economy where capital is mobile while workers are immobile. Following these works or ZMW model, most of the literatures about tax competition assume that there is at least one fixed factor. Apart from ZMW model, many papers have also studied optimal tax policy with mobile workers. Early contributions to this literature include Wilson (1980) for the case of linear taxation and Mirrlees (1982) for the case of nonlinear taxation. In much of this literature, potential emigrants choose between the best labor-consumption bundle available at home and some predetermined bundle or utility abroad. There is a potential conflict between a government's desire to tax on the basis of ability to pay and the possibility that more able individuals might emigrate to avoid high tax burdens. Hamilton and Pestieau (2005) consider a political economy model of competition between a large number of small countries when there are two skill types and only one of them is mobile.The objective function of each government is determined by majority rule, with the consequence that it wants to maximize the utility of the type of individual who is in the majority. Piaser (2007) studies nonlinear income tax competition by two benevolent governments in the presence of labor mobility, but in a model with only two skills, and Morelli et al. (2012a) also studies the same problem except for three types in a model. This enables us to study the government's choice of constitution: integrated one or independent one. Lehmann et al. (2014) also studies nonlinear income tax competition by two Rawlsian governments when workers are mobile and type spaces for wage as well as commuting cost are continuous.Contrary to these researches, Bierbrauer et al. (2013) assumes that workers are perfectly mobile without any costs, and study nonlinear income tax competition.

Each chapter in this dissertation contributes to the literature on theoretical subjects about tax and other redistributive policy. Chapter 2 compares ad-valorem and specific taxation in models where a representative consumer with an exogenous income has both a quality and a quantity choice under perfect competition. In the setting, while ad-valorem tax causes income effect only, specific tax causes both income effect and substitution effect. Therefore, ad-valorem tax decreases consumer demand for both quality and quantity; on the other hand, specific tax decreases consumer demand for quantity. However, the sign of consumer demand for quality is ambiguous and is determined by the curvature of marginal utility on quantity. Additionally, using a constant elasticity of substitution (CES) utility function and a linear price function, we show that ad-valorem tax is superior to specific tax except for the Leontief preference under which the two forms of commodity taxes generate the same tax revenue. The substitution effect caused by specific tax disappears if the elasticity of substitution converges to zero.

In Chapter 3, We examine optimal taxation and public good provision by a government which takes reduction of envy into consideration as one of the constraints. We adopt the notion of extended envy-freeness proposed by Diamantaras and Thomson (1989), called λ-equitability. We derive the modified Samuelson rule at an optimum income tax, and show that, using a constant elasticity of substitution utility function, the direction of distorting the original Samuelson rule to relax λ envy free constraints is crucially determined by the elasticity of substitution. Also, Chapter 4 covers the topic of public good provision under income transfer under that ethical constraint, but allows the social planner to set the surcharge fee for the purpose of excluding some agents whereas we simplify their income as exogenous one (or initial wealth). In this chapter, we study optimal public good provision and user fee in order to exclude some agents by Rawlsian or utilitarian government under lump-sum transfer, constrained by reduction of envy. In particular, we employ the exclusion technique used in Hellwig (2005a), *i.e.*, the policymaker decides the level of provision and surcharge fee paid by those making access to it, as well as uniform transfer. In this setting, we derive the optimal provision level and user fee, and compared to those in Hellwig (2005a), for Rawlsian government, the up-charge is lower than the one derived in Hellwig (2005a) in order to reduce the envy.

Chapter 5 studies optimal nonlinear income tax schedule at symmetric equilibria at which two symmetric states (or tax authorities) compete in order to attract more taxpayers from the opposite. It is different from the existing papers that taxpayers' wage are endogenously determined by production technology. The optimal tax schedule embraces not only migration effect, but also trickle-down effect coming from endogenous wage, and the migration effect stimulates the trickle-down effect. Compared to previous works, the threat of emigration never disappears in marginal tax rate for high-skilled workers because emigration terms are embedded in the production and such factors have impacts on the productivities or their unit wages.

Chapter 2

The Welfare Comparison of Ad-valorem Tax and Specific Tax with Both Quality and Quantity Choice of a Consumer

2.1 Introduction

In the field of public finance, there are many studies that compare social welfare under ad-valorem and specific taxation. The seminal contribution of Suits and Musgrave (1953) examines the impact of two different tax structures on welfare keeping the same tax revenue under a monopoly market. The author shows that an ad-valorem tax is superior to a specific tax. Some works that approach the comparison under imperfect competition support the conclusion (see Delipalla and Keen (1992), Skeath and Trandel (1994), Myles (1996), Denicoló and Matteuzzi (2000), Anderson et al. (2001) and Aiura and Ogawa (2013)).

However, these studies ignore the impact of the tax structure on the product quality selected by firms. If both forms of taxation affect the quality, wholly specific taxation can be optimal under perfect competition (see Kay and Keen (1983, 1991) and Delipalla and Keen (2006)). To cover the cost of improving quality, the increment in tax revenue under specific tax is less than the increment under ad-valorem tax because of the "multiplier effect."

Our aim in this paper is to compare the two forms of commodity taxation in terms of welfare when the consumer has both a quantity and a quality choice under perfect competition. Contrary to the aforementioned literature, we assume that a representative consumer, not firms, determines product quality provided under a perfect competition market. In other words, a given good/service is available to the consumer at different quality levels. In the setting, there are two margins of choice for the consumer: quality choice and quantity choice. At the quality choice margin, the consumer chooses a degree of excellence of the good/service. At the quantity choice margin, the consumer determines the amount of the good/service given the quality. Imagine a situation in which the consumer can choose not only the number of hours that his parents are at the elderly care center but also the quality of the facility. Both quality and quantity affect utility, but a crucial difference is that quality choice positively affects the unit price because it is quite natural that the higher quality he chooses, the more he pays the unit price. The basic structure of our model is close to the Bastani et al. (2016) model. The authors study optimal ad-valorem and specific taxation as well as nonlinear income taxation, but our setting and concern differ from Bastani et al. (2016). The study investigates the two forms of commodity taxation under nonlinear labor income taxes when individuals have different productivity, and there is asymmetric information between the policymaker and taxpayers with respect to individuals' productivity. Both consumption taxes play a crucial role in relaxing incentive constraints, thus, such taxes are necessary to implement the second-best allocation, which contradicts the canonical results provided by Atkinson and Stiglitz (1976). Contrary to their paper, we allow the government to levy taxes on such consumption only and assume homogeneous individuals, that is, there is no asymmetric information. Instead of studying the linkage between commodity taxes and relaxing incentive constraints on income taxes, we present the welfare comparison between an ad-valorem and a specific tax in the sense of taxpayer's utility and the response in the change of each tax rate using the comparative statics analysis. Additionally, **?** is related to our paper and studies consumer behavior when the consumer can choose both quality and quantity; however, the cited paper does not study tax incidence.

We find that specific taxation distorts the consumer's choice between quantity and quality, whereas introducing ad-valorem taxation does not affect the choice. In other words, while the marginal rate of substitution between a quantity and a quality includes the tax rate under specific tax, the corresponding marginal rate of substitution does not include the tax rate under an ad-valorem tax. The unit tax rate makes the consumer reluctant to purchase more, whereas he is willing to improve the quality to compensate for lost utility. Subsequently, assuming that a price function is linear and individual preference is expressed by the constant elasticity of the substitution (CES) utility function, we analytically show that an ad-valorem tax is superior to a specific tax in the presence of a substitution effect, if any. The difference in indirect utilities between ad-valorem tax and specific tax decreases as the elasticity of substitution decreases; the two forms of taxation are equivalent when consumers have Leontief preference. The findings imply that the substitution effect under a specific tax plays an important role in ad-valorem tax dominating a specific tax in terms of consumer welfare.

The next section describes our abstract model with assumptions guaranteeing the existence of an optimal solution for a representative consumer. Section 3 provides a comparative statics analysis under these two taxes. Using CES utility and linear price function, we study welfare comparison with numerical simulation as an example in section 4. We show most of our complex calculations to derive our results in appendices.

2.2 The Model

In this model, there is a single representative consumer with initial wealth I. He consumes only one type of good and can choose both quantity and quality. Let $y \in \mathbb{R}_{++}$ be a quantity and $\theta \in \mathbb{R}_{++}$ be a quality. The price for unit consumption is determined by quality via a strictly increasing and differentiable function $p(\theta)$. The consumer derives utility from quantity and quality, which is defined by $v(y, \theta)$. To guarantee an interior solution with respect to each input, we assume that v is twice differentiable, strictly increasing, and strictly concave in (y, θ).

The government can put tax on either the unit price $p(\theta)$ or the quantity y. If the consumer adopts an ad-valorem tax t^a, the budget constraint must be:

$$(1 + t^a)p(\theta)y \leq I, \tag{1}$$

while if the consumer chooses a specific tax t^s, the budget constraint must become

$$(p(\theta) + t^s)y \leq I. \tag{2}$$

The taxpayers maximize their own utility with respect to quantity y and quality θ given the budget constraint. From the first order conditions, we can derive the following ad-valorem and specific tax wedge:

$$\frac{v_y(y^a, \theta^a)}{v_\theta(y^a, \theta^a)} = \frac{p(\theta^a)}{p'(\theta^a)y^a} \tag{3}$$

$$\frac{v_y(y^s, \theta^s)}{v_\theta(y^s, \theta^s)} = \frac{p(\theta^s) + t^s}{p'(\theta^s)y^s} \tag{4}$$

where v_k denotes the derivative of v with respect to $k = y, \theta$, the subscript a the choice of an ad-valorem tax, and s is the choice of a specific tax. Given initial wealth I, tax rate t, and tax scheme $i = s, a$, their quantity choice function, quality choice function, and indirect utility function are defined as follows: $y(t^i, I)$, $\theta(t^i, I)$, and $V(t^i, I)$.

When the government employs an ad-valorem tax, the budget constraint faced by the consumer is:

$$t^a p(\theta^a)y^a \geq R \tag{5}$$

where R is an exogenous amount of public expenditure. On the other hand, if the consumer imposes a specific tax, the budget constraint is:

$$t^s y^s \geq R \tag{6}$$

2.3 Comparative Statics of (y, θ) on Tax Rate t Under Ad-valorem Tax and Specific Tax

How do an ad-valorem tax and a specific tax affect the consumer's demand for quantity y and quality θ? In this chapter, we study the sensitivity of their choice (y, θ) in response to each tax rate increase.

Let $|K|$ and $|M|$ be the determinants of a border Hessian matrix on the optimization problem under ad-valorem tax and specific tax. By the comparative statics of y and θ under ad-valorem tax, these derivatives are:

$$\frac{\partial \theta}{\partial t^a} = \frac{1}{|K|} \frac{v_\theta v_y p(\theta)}{\gamma^a} (1 - \rho - \varepsilon_y^{v_\theta}) \tag{7}$$

$$\frac{\partial y}{\partial t^a} = \frac{1}{|K|} \frac{v_\theta p(\theta) y}{\gamma^a} \Big(\frac{v_\theta}{y} (1 - \varepsilon_y^{v_\theta}) + v_y \{ \frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)} \} \Big) \tag{8}$$

where γ^a is the Lagrangian multiplier under ad-valorem tax, $\rho \equiv -\frac{v_{yy} y}{v_y}$ is the curvature of marginal utility on quantity y, $v_{\theta y}$ is the cross-derivative of v, and $\varepsilon_y^{v_\theta} \equiv \frac{v_{\theta y}}{v_\theta} y$ is the demand elasticity of v_θ. We would guarantee the second-order condition for this utility maximization under ad-valorem tax. Thus, we impose the following assumptions: $\varepsilon_y^{v_\theta} \geq 1$ and $p''(\theta) \geq 0$ (see Appendix A).

From Equation (7) and Equation (8), both demands for quality θ and quantity y decrease as ad-valorem tax increases; that is, $\frac{\partial \theta}{\partial t^a}(t^a, I^a) < 0$ and $\frac{\partial y}{\partial t^a}(t^a, I^a) < 0$ if the second-order condition is satisfied. The intuition is that as the ad-valorem tax increases, the consumer's disposable income decreases by $\frac{t^a}{1+t^a}$ percent. On the other hand, the marginal rate of substitution between quantity and quality is the same as that without imposing the tax. Hence, decreasing disposable income decreases the consumer's demands for both quantity and quality due to the income effect.

On the other hand, the comparative statics results under specific tax are:

$$\frac{\partial\theta}{\partial t^s} = \frac{1}{|M|}\frac{v_\theta v_y}{\gamma^s}(2-\rho-\varepsilon_y^{v_\theta}) \tag{9}$$

$$\frac{\partial y}{\partial t^s} = \frac{1}{|M|}\Big(-\frac{v_\theta v_{\theta y} y}{\gamma^s} + \frac{v_y v_\theta y}{\gamma^s}(\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)})\Big) \tag{10}$$

where γ^s is the Lagrangian multiplier under specific tax. Similar to the ad-valorem tax case, it is assumed that $\varepsilon_y^{v_\theta} \geq 1$ and $p''(\theta) \geq 0$ so that the second-order condition holds (see Appendix A).

From Equation (9) and Equation (10), while $\frac{\partial y}{\partial t^s}(t^s, I^s)$ is negative, the sign of $\frac{\partial\theta}{\partial t^s}(t^s, I^s)$ is ambiguous. Put differently, employing a specific tax induces the consumer to reduce his demand for quantity but perhaps to select better quality as opposed to that of the ad-valorem tax case. This is because the income effect and the substitution effect occur as shown in Equation (4). From Equation (9), whether the quality is improved depends on ρ, which indicates the extent to which the consumer wants to maintain the quantity. If ρ is larger (smaller) than 1, the bracket in Equation (9) may be negative (positive), so $\frac{\partial\theta}{\partial t^s}(t^s, I^s)$ is negative (positive).

The intuition is that when the elasticity of the consumer's marginal utility on quantity is greater than 1, the consumer is not willing to decrease quantity and debases the quality to cover his expenditure. In contrast, if ρ is sufficiently small, the consumer does not hesitate to decrease the quantity, whereas the consumer upgrades the quality to compensate for the utility loss caused by the decrease in quantity. Thus, whether the income effect dominates the substitution effect depends on ρ.

We summarize the above arguments in the following proposition.

Proposition 1. *Assume that each sufficient condition for the representative consumer's optimization problem is satisfied. Then,*

1. *Under an ad-valorem tax, both $\frac{\partial\theta}{\partial t^a}(t^a, I^a)$ and $\frac{\partial y}{\partial t^a}(t^a, I^a)$ are negative, that is,, both demand for quality θ and quantity y decrease in response to the tax increase.*

2. *Under specific tax, $\frac{\partial y}{\partial t^s}(t^s, I^s)$ is negative, which means that the demand decreases when the tax rate increases while the sign of $\frac{\partial\theta}{\partial t^s}(t^s, I^s)$ is determined by ρ. Particularly, the consumer's demand for quality decreases if $\rho > 1$.*

Here, we present a special case in which the second-order condition is satisfied, and the following example is also used in the next section. We assume that the price function is linear $p(\theta) = a\theta$ where $a > 0$ and individuals' preference is the constant elasticity of the substitution (CES) utility function expressed by $v(y, \theta) = (\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}}$, where $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$, and σ are measures of complementarity assuming that $\sigma \geq 1$. $\frac{1}{1+\sigma}$ is the elasticity of substitution. It is obvious that $p''(\theta) \geq 0$ is satisfied. Additionally, if $\sigma \geq 1$, $\varepsilon_y^{v_\theta} \geq 1$ under both ad-valorem tax and specific tax. Therefore, the results of proposition 1 hold in the environment. Moreover, we demonstrate that consumers retain quality or select lower quality under specific tax, that is, $\frac{\partial\theta}{\partial t^s} \leq 0$. The proofs are given in Appendix B, and the results are summarized in the next corollary.

Corollary 1. *Suppose that a representative consumer has CES utility function $v(y, \theta) =$*

$(\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}}$, *where* $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$, *and faces a linear price function* $p(\theta) = a\theta$. *Under both ad-valorem tax and specific tax, the consumer demands for both quality and quantity decrease in response to the tax rate increase.*

2.4 Welfare Comparison

This section examines whether the government should adopt an ad-valorem or a specific tax under the same tax revenue. For the remainder of the manuscript, we use the following environment that was used in the previous special case: the price function is $p(\theta) = a\theta$ where $a > 0$, and the individual's preference is the CES utility function given by $v(y, \theta) = (\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}}$ where $\alpha > 0, \beta > 0$, $\alpha + \beta = 1$ and $\sigma \geq 1$. As mentioned in section 3, the assumptions on the price function and the utility function ensure the second-order conditions for the maximization problem. Additionally, the elasticity of substitution $\frac{1}{1+\sigma}$ is between 0 and $\frac{1}{2}$. σ is a component of the elasticity of substitution and can be interpreted as the degree of complementarity. In the setting, we can state the following.

Proposition 2. *Assume that the price function is linear and the individual's preference is the CES utility function, that is,* $p(\theta) = a\theta$ *where* $a > 0$ *and* $v(y, \theta) = (\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}}$ *where* $\alpha > 0, \beta > 0$, $\alpha + \beta = 1$, *and* $\sigma \geq 1$. *If the tax revenue remains the same under ad-valorem and specific tax,*

1. *ad-valorem tax is superior to specific tax in terms of the consumer's welfare except for the elasticity of substitution* $\frac{1}{1+\sigma} = 0$*;*

2. *the difference in welfare, that is,* $v(y^a, \theta^a) - v(y^s, \theta^s)$ *decreases as* σ *increases, or*

the elasticity of substitution decreases;

3. *ad-valorem tax and specific tax are indifferent in terms of the welfare when* $\sigma = \infty$*, that is, the utility function is a Leontief preference* $v(y, \theta) = \min\{y, \theta\}$.

This is shown in Appendix C. The result is intuitive. As shown by Equation (3) and Equation (4), the marginal rate of substitution between quantity and quality includes the tax rate under specific tax; the corresponding marginal rate of substitution does not include the tax rate under an ad-valorem tax. In other words, the income effect only changes consumption choice under an ad-valorem tax (**Figure 2.1**), whereas the substitution effect as well as the income effect distort consumption choice under specific tax (**Figure 2.2**). This implies that if a substitution effect exists, specific taxation generates welfare losses since it distorts the individual's behavior relative to the case without any tax policy. Therefore, an ad-valorem tax is superior to a specific tax. The findings that an ad-valorem tax is superior to a specific tax does not depend on the specification of the utility function and the price function and the assumption that the government can impose either an ad-valorem or a specific tax. Indeed, we show the statement without specifying these functions when the government can employ both an ad-valorem and a specific tax simultaneously (see Appendix D).

Additionally, the increase in σ implies that the substitution effect becomes smaller. In other words, welfare losses caused by specific tax diminishes. Consequently, the welfare difference decreases and, particularly, when the utility function is Leontief preference, it is possible for the policymaker to achieve the same utility level via a specific tax as an ad-valorem tax. This is because there is no substitution effect under Leontief preference, which leads to the assumption that the consumption decision

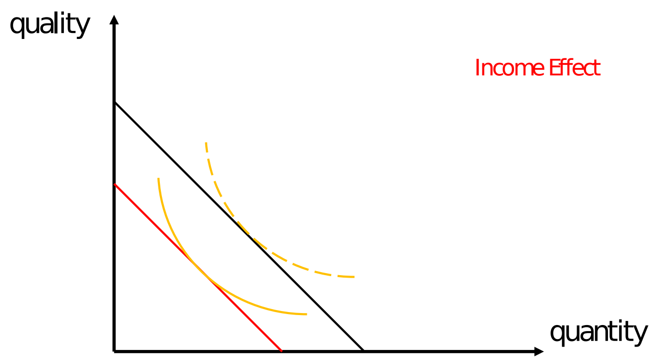

Figure 2.1: The effect of ad-valorem tax on a consumer's choice.

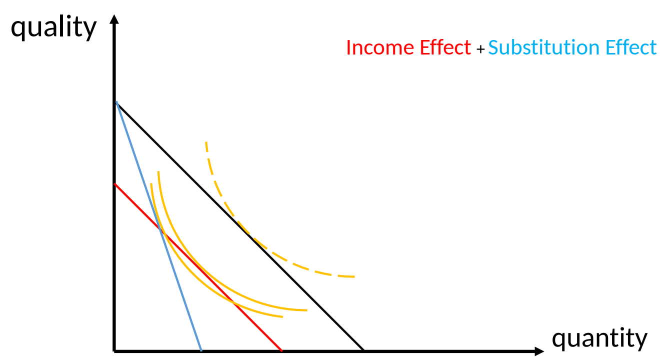

Figure 2.2: The effect of specific tax on a consumer's choice.

toward changing the tax rate is affected only by the income effect regardless of an ad-valorem tax and a specific tax. Our conclusion is illustrated in **Figure 2.3** for $R = 1$, $I = 6$, $a = 1$, $\alpha = \beta = 0.5$ and $\sigma \in [1, \infty)$.

Although **Figure 1** and **Figure 2** provide an illustrative explanation, these budget lines are not correct because the consumer's budget set is partially determined by a linear price function; therefore, each budget line should not be straight. However, it is sufficient to explain the mechanism of our result if the curvature of the budget line is

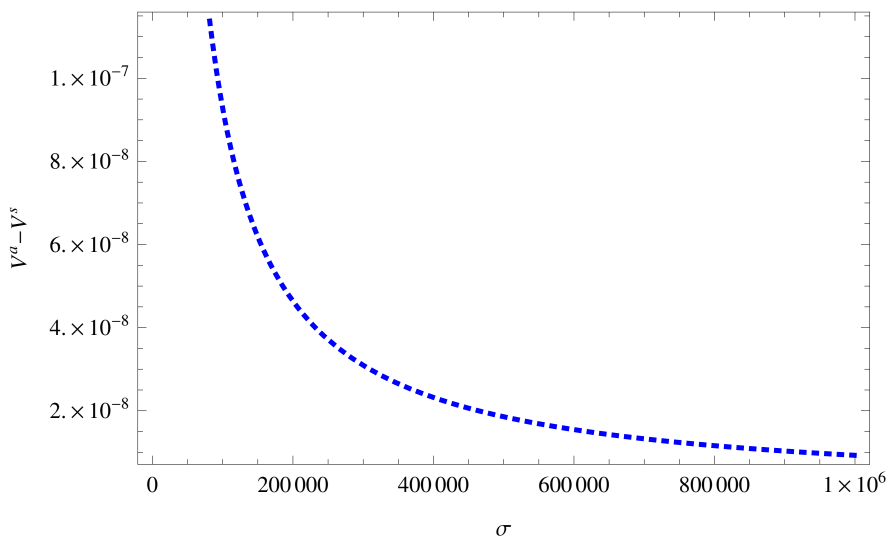

Figure 2.3: Simulations of $v(y^a, \theta^a) - v(y^s, \theta^s)$ with respect to σ.

smaller than that of the consumer's utility indifference curve, which is the condition to maximize utilities.

2.5 Concluding Remarks

This paper compares ad-valorem tax and specific tax from the perspective of a representative consumer's welfare when the consumer has both a quantity and a quality choice. Contrary to most papers investigating such a comparison of social welfare under imperfect competition, we allow for quality choice as well as a competitive environment. Using a linear price function for quality and CES utility function, we identify that the substitution effect distorts the consumer's optimal choice under specific tax, which leads to the consequence that ad-valorem tax dominates the consumer's choice in the sense of the consumer's utility. On the other hand, since the substitution effect vanishes under Leontief preferences, the two forms of tax are equivalent. Our main result is different from that of previous studies, which show that specific tax can be superior to ad-valorem tax for social welfare under quality choice by production sector, as in Kay and Keen (1983, 1991) and Delipalla and Keen (2006). These papers argue that the "multiplier effect" imposes excess burden on improving product quality. Therefore, specific tax can be superior to ad-valorem tax under firms' quality choices, but the representative consumer in our model can choose quality, so ad-valorem tax is like a lump-sum tax on initial wealth. Thus, only income effect distorts the consumer's decision. Roughly speaking, this explains why ad-valorem tax is superior to specific tax in our model unless the elasticity of substitution equals 0.

Although our model is simple in the sense that consumers focus on one taxable good, as the first step, we have priority over unveiling the interaction between preference and two tax schemes. Additionally, the result can be applied to a case in which the consumer faces liquidity constraints. For instance, in a housing choice, the consumer selects the location and scale, and the government often collects taxes from such consumption. Our result suggests that the government should levy taxes on the unit) land price if consumers have substitution between location and scale. Consequently, this study has implications for the understanding of the effect of two forms of commodity tax on consumer's quality choice.

The unsolved question in our paper is whether an ad-valorem tax is superior to a specific tax for many taxable goods. In this situation, the government faces an additional problem, which is the choice between a uniform or differentiated commodity tax from the standpoint of efficiency. If the government employs a uniform tax on all goods, we conjecture that our main conclusion holds because an ad-valorem tax still acts as a lump-sum tax while a specific tax distorts the consumer's choice between the quality and quantity of taxable goods. However, if the government employs differentiated indirect taxes, an ad-valorem tax distorts the consumer's choice compared to the model under one taxable good, which means that wholly ad-valorem tax may not be optimal. Atkinson and Stiglitz. (1972) is an important study that solves the question since the paper shows that a uniform tax is desirable under an exogenous income. If the findings hold even under the situation in which consumers can choose quantity and quality, wholly ad-valorem tax would be desirable even under many taxable goods if income is given as it is in our setting. Moreover, to clarify whether the desirability of an ad-valorem tax crucially depends on the assumption of an exogenous income, we will extend the model by allowing consumers to choose the amount of labor supply. This theme is left for future research.

Appendix A: Comparative statics under varying tax rates

Under the ad-valorem tax, individuals problem is formulated as follows.

$$\max_{y,\theta} \ v(y,\theta)$$

$$s.t. \ (1+t^a)p(\theta)y \leq I.$$

The corresponding Lagrangian is:

$$\mathcal{L} = v(y,\theta) + \gamma^a[I - (1+t^a)p(\theta)y]$$

The first-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial y} = v_y - \gamma^a(1+t^a)p(\theta) = 0$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = v_\theta - \gamma^a(1+t^a)p'(\theta)y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma^a} = I - (1+t^a)p(\theta)y = 0$$

The bordered Hessian matrix for this problem is as follows.

$$K = \begin{bmatrix} v_{yy} & -\gamma^a(1+t^a)p'(\theta)+v_{y\theta} & -(1+t^a)p(\theta) \\ -\gamma^a(1+t^a)p'(\theta)+v_{y\theta} & v_{\theta\theta}-\gamma^a(1+t^a)p''(\theta)y & -(1+t^a)p'(\theta)y \\ -(1+t^a)p(\theta) & -(1+t^a)p'(\theta)y & 0 \end{bmatrix}$$

Its determinant $|K|$ is:

$$|K| = 2\frac{v_y(v_\theta)^2}{y(\gamma^a)^2}(\varepsilon_y^{v_\theta}-1) - \left[\left(\frac{v_y}{\gamma^a}\right)^2 v_\theta\left\{\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)}\right\} + \left(\frac{v_\theta}{\gamma^a}\right)^2 v_{yy}\right]$$

Here, we assume that $\varepsilon_y^{v_\theta} \geq 1$ and $p''(\theta) \geq 0$ for utility maximization, that is, $|K|$ is positive. Therefore, it ensures the existence of the inverse matrix K^{-1} and then Cramer's rule is available. Then, the derivative of θ with respect to t^a is:

$$\frac{\partial \theta}{\partial t^a} = \frac{1}{|K|}\begin{vmatrix} v_{yy} & \gamma^a p(\theta) & -(1+t^a)p(\theta) \\ -\gamma^a(1+t^a)p'(\theta)+v_{y\theta} & \gamma^a p'(\theta)y & -(1+t^a)p'(\theta)y \\ -(1+t^a)p(\theta) & p(\theta)y & 0 \end{vmatrix}$$

$$= \frac{1}{|K|}\frac{v_y v_\theta p(\theta)}{\gamma^a}(1-\rho-\varepsilon_y^{v_\theta})$$

Also, the derivative of y can be found as follows:

$$\frac{\partial y}{\partial t^a} = \frac{1}{|K|}\begin{vmatrix} \gamma^a p(\theta) & -\gamma^a(1+t^a)p'(\theta)+v_{y\theta} & -(1+t^a)p(\theta) \\ \gamma^a p'(\theta)y & v_{\theta\theta}-\gamma^a(1+t^a)p''(\theta)y & -(1+t^a)p'(\theta)y \\ p(\theta)y & -(1+t^a)p'(\theta)y & 0 \end{vmatrix}$$

$$= \frac{1}{|K|}\frac{v_\theta p(\theta)y}{\gamma^a}\Big(\frac{v_\theta}{y}(1-\varepsilon_y^{v_\theta}) + v_y\{\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)}\}\Big)$$

On the other hand, individuals' problem under the specific tax is:

$$\max_{y,\theta}\ v(y,\theta)$$

$$s.t.\ (p(\theta)+t^s)y \leq I.$$

The corresponding Lagrangian is:

$$\mathcal{L} = v(y,\theta) + \gamma^s[I - (p(\theta)+t^s)y]$$

The first-order conditions are given by:

$$\frac{\partial\mathcal{L}}{\partial y} = v_y - \gamma^s(p(\theta)+t^s) = 0$$

$$\frac{\partial\mathcal{L}}{\partial\theta} = v_\theta - \gamma^s p'(\theta)y = 0$$

$$\frac{\partial\mathcal{L}}{\partial\gamma} = I - (p(\theta)+t^s)y = 0$$

The bordered Hessian matrix for this problem is:

$$M = \begin{bmatrix} v_{yy} & -\gamma^s p'(\theta) + v_{y\theta} & -(p(\theta)+t^s) \\ -\gamma^s p'(\theta) + v_{y\theta} & v_{\theta\theta} - \gamma^s p''(\theta)y & -p'(\theta)y \\ -(p(\theta)+t^s) & -p'(\theta)y & 0 \end{bmatrix}$$

Its determinant $|M|$ is:

$$|M| = 2\frac{v_y(v_\theta)^2}{y(\gamma^s)^2}(\varepsilon_y^{v_\theta}-1) - \Big[(\frac{v_y}{\gamma^s})^2 v_\theta\{\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)}\} + (\frac{v_\theta}{\gamma^s})^2 v_{yy})\Big]$$

Again, we assume that $\varepsilon_y^{v_\theta} \geq 1$ and $p''(\theta) \geq 0$ for utility maximization, that is, $|M|$ is positive. Using Cramer's rule,

$$\frac{\partial\theta}{\partial t^s} = \frac{1}{|M|}\begin{vmatrix} v_{yy} & \gamma^s & -(p(\theta)+t^s) \\ -\gamma^s p'(\theta) + v_{y\theta} & 0 & -p'(\theta)y \\ -(p(\theta)+t^s) & y & 0 \end{vmatrix}$$

$$= \frac{1}{|M|}\frac{v_\theta v_y}{\gamma^s}(2-\rho-\varepsilon_y^{v_\theta})$$

Similarly, we observe the derivative of y on t^s as follows.

$$\frac{\partial y}{\partial t^s} = \frac{1}{|M|}\begin{vmatrix} \gamma^s & -\gamma^s p'(\theta) + v_{y\theta} & -(p(\theta)+t^s) \\ 0 & v_{\theta\theta} - \gamma^s p''(\theta) y & -p'(\theta)y \\ y & -p'(\theta)y & 0 \end{vmatrix}$$

$$= \frac{1}{|M|}\Big(-\frac{v_\theta v_{\theta y} y}{\gamma^s} + \frac{v_y v_\theta y}{\gamma^s}(\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)})\Big)$$

□

Appendix B: A sufficient condition for the maximization problem

Assume $p(\theta) = a\theta$ and $v(y,\theta) = (\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}}$, where $a > 0$, $\alpha > 0$, $\beta > 0$, $\alpha + \beta = 1$, and $\sigma \geq 1$. In this case, $\varepsilon_y^{v_\theta}$ is given by:

$$\varepsilon_y^{v_\theta} = \alpha(\sigma+1)(\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-1} y^{-\sigma} \tag{B.1}$$

Before examining the second-order conditions under ad-valorem tax and specific tax, we suggest the first-order conditions in the setting from Equation (3) and Equation (4):

$$\alpha y^{-\sigma} = \beta\theta^{-\sigma} \tag{B.2}$$

$$y^\sigma = \frac{\alpha}{\beta}\frac{a\theta^{\sigma+1}}{a\theta + t^s} \tag{B.3}$$

Equation (B.2) is the first-order condition under ad-valorem tax and Equation (B.3) is one under specific tax.

We now turn to the analysis of the second-order conditions. First, we derive a sufficient condition under an ad-valorem tax. Substituting Equation (B.2) into Equation (B.1) yields:

$$\varepsilon_y^{v_\theta} = \frac{1}{2}(\sigma + 1) \tag{B.4}$$

Therefore, if $\sigma \geq 1$, $\varepsilon_y^{v_\theta} \geq 1$. On the other hand, substituting Equation (B.3) into Equation (B.1), $\varepsilon_y^{v_\theta}$ under specific tax can be rewritten as follows:

$$\varepsilon_y^{v_\theta} = (\sigma + 1)\frac{a\theta + t^s}{2a\theta + t^s} \tag{B.5}$$

This means that if $\sigma \geq \frac{a\theta}{a\theta+t^s}$, $\varepsilon_y^{v_\theta} \geq 1$. Note that $\frac{a\theta}{a\theta+t^s} < 1$ under $R > 0$.

To sum up, $\sigma \geq 1$ is a sufficient condition to yield a locally maximum solution under both ad-valorem and specific tax.

Next, we compute ρ under the setting. By the definition and Equation (B.1), it can be rewritten as follows:

$$\rho = -(\sigma + 1)\Big[\alpha(\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-1}y^{-\sigma} - 1\Big] = 1 - \varepsilon_y^{v_\theta} + \sigma \tag{B.6}$$

Substituting Equation (B.6) into Equation (9), it yields:

$$\frac{\partial\theta}{\partial t^s} = \frac{1}{|M|}\frac{v_\theta v_y}{\gamma^s}(1 - \sigma) \tag{B.7}$$

As a result, the sign of $\frac{\partial\theta}{\partial t^s}$ is determined by σ. If σ is lower than 1, it is positive since the substitutability is high. However, σ is equal to or greater than 1 for utility maximization, which means that the complementarity is high. Therefore, $\frac{\partial\theta}{\partial t^s}$ is non-positive under $\sigma \geq 1$.

□

Appendix C: The welfare comparison

Assume that the price function is linear and individual's preference is the CES utility function. First, we compute the optimal indirect utility function under ad-valorem tax. Substituting Equation (B.2) into Equation (1) yields:

$$y^a = \sqrt{\frac{I}{(1+t^a)a}\left(\frac{\alpha}{\beta}\right)^{\frac{1}{\sigma}}} \tag{C.1}$$

In addition, using Equation (B.2) and Equation (C.1), Equation (5) is rewritten as follows:

$$t^a = \frac{R}{I-R} \tag{C.2}$$

Combining Equation (C.1) and Equation (C.2) yields:

$$y^a = \sqrt{\frac{I-R}{a}\left(\frac{\alpha}{\beta}\right)^{\frac{1}{\sigma}}} \tag{C.3}$$

Thus, we can derive the optimal indirect utility function, using Equation (B.2) and then substituting Equation (C.3) into the CES utility function as follows:

$$\begin{aligned} v(y^a,\theta^a) &= (\alpha y^{-\sigma} + \beta\theta^{-\sigma})^{-\frac{1}{\sigma}} = (2\alpha)^{-\frac{1}{\sigma}} y^a \\ &= 2^{-\frac{1}{\sigma}}(\alpha\beta)^{-\frac{1}{2\sigma}} R^{\frac{1}{2}}\left(\frac{\frac{I}{R}-1}{a}\right)^{\frac{1}{2}} \end{aligned} \tag{C.4}$$

Here, we assume that $\frac{I}{R} > 1$ to avoid his bankruptcy. Next, we compute the optimal indirect utility function under specific tax. From Equation (6), we can get

$$y^s = \frac{R}{t^s} \tag{C.5}$$

Substituting Equation (C.5) into Equation (2) yields:

$$\theta^s = \frac{t^s(\frac{I}{R}-1)}{a} \tag{C.6}$$

Moreover, substituting Equation (C.6) into Equation (B.3) and using Equation (C.5) yields:

$$y^s = \Big(\frac{\alpha}{\beta}\Big)^{\frac{1}{2\sigma}}\Big[\frac{(\frac{I}{R}-1)^{\sigma+1}}{a^{\sigma}\frac{I}{R}}\Big]^{\frac{1}{2\sigma}}R^{\frac{1}{2}} \tag{C.7}$$

On the other hand, substituting Equation (C.5) into Equation (C.6) yields:

$$\theta^s = \frac{I-R}{ay^s} \tag{C.8}$$

Substituting Equation (C.7) into Equation (C.8) yields:

$$\theta^s = \frac{\frac{I}{R}-1}{a}\Big(\frac{\alpha}{\beta}\Big)^{\frac{-1}{2\sigma}}\Big[\frac{(\frac{I}{R}-1)^{\sigma+1}}{a^{\sigma}\frac{I}{R}}\Big]^{\frac{-1}{2\sigma}}R^{\frac{1}{2}} \tag{C.9}$$

Thus, we can derive the optimal indirect utility function, substituting Equation (C.7) and Equation (C.9) into the CES utility function as follows:

$$v(y^s,\theta^s) = (\alpha\beta)^{-\frac{1}{2\sigma}}R^{\frac{1}{2}}\Big(\frac{\frac{I}{R}-1}{a}\Big)^{\frac{\sigma+1}{2\sigma}}\Big(\frac{I}{R}\Big)^{\frac{-1}{2\sigma}}a^{\frac{1}{2\sigma}}\Big(2-\frac{R}{I}\Big)^{\frac{-1}{\sigma}} \tag{C.10}$$

Now, we compare the utility level under ad-valorem tax with one under specific tax. Using Equation (C.4) and Equation (C.10), the difference is

$$\begin{aligned}
&v(y^a,\theta^a) - v(y^s,\theta^s)\\
&\quad= (\alpha\beta)^{-\frac{1}{2\sigma}}R^{\frac{1}{2}}\Big[2^{-\frac{1}{\sigma}}\Big(\frac{\frac{I}{R}-1}{a}\Big)^{\frac{1}{2}} - \Big(\frac{\frac{I}{R}-1}{a}\Big)^{\frac{\sigma+1}{2\sigma}}\Big(\frac{I}{R}\Big)^{-\frac{1}{2\sigma}}a^{\frac{1}{2\sigma}}\Big(2-\frac{R}{I}\Big)^{\frac{-1}{\sigma}}\Big]\\
&\quad= (\alpha\beta)^{-\frac{1}{2\sigma}}R^{\frac{1}{2}}\Big(\frac{\frac{I}{R}-1}{a}\Big)^{\frac{1}{2}}\Big[\Big(\frac{1}{2}\Big)^{\frac{1}{\sigma}} - \Big(\frac{\frac{I}{R}-1}{\frac{I}{R}(2-\frac{R}{I})^2}\Big)^{\frac{1}{2\sigma}}\Big]\\
&\quad= (\alpha\beta)^{-\frac{1}{2\sigma}}R^{\frac{1}{2}}\Big(\frac{\frac{I}{R}-1}{a}\Big)^{\frac{1}{2}}\Big[\Big(\frac{1}{4}\Big)^{\frac{1}{2\sigma}} - \Big(\frac{\frac{I}{R}-1}{\frac{I}{R}(2-\frac{R}{I})^2}\Big)^{\frac{1}{2\sigma}}\Big]
\end{aligned} \tag{C.11}$$

Note that the second term in the bracket is smaller than $(\frac{1}{4})^{\frac{1}{2\sigma}}$. To show the fact, we define $f(\frac{I}{R})$ as follows:

$$\begin{aligned}
f(\frac{I}{R}) &\equiv \frac{I}{R}\Big(2-\frac{R}{I}\Big)^2 - 4\Big(\frac{I}{R}-1\Big)\\
&= 4\frac{R}{I}
\end{aligned} \tag{C.12}$$

Therefore, $f(\frac{I}{R})$ is positive. This implies that

$$\left(\frac{1}{4}\right)^{\frac{1}{2\sigma}} > \left(\frac{\frac{I}{R}-1}{\frac{I}{R}(2-\frac{R}{I})^2}\right)^{\frac{1}{2\sigma}} \tag{C.13}$$

Therefore, $v(y^a,\theta^a)-v(y^s,\theta^s)$ is positive except for $\sigma=\infty$. Moreover, $v(y^a,\theta^a)-v(y^s,\theta^s)$ is close to zero as σ goes to ∞ from Equation (C.11). In addition, $\frac{1}{\alpha\beta} > 1$ implies that $(\alpha\beta)^{-\frac{1}{2\sigma}}$ is decreasing in σ. With the bracket decreasing in σ, we can state that $v(y^a,\theta^a) - v(y^s,\theta^s)$ is decreasing in σ.

□

Appendix D: General case under both ad-valorem and specific taxes

Under both ad-valorem and specific taxes, individuals problem is formulated as follows.

$$\max_{y,\theta}\ v = v(y,\theta)$$

$$s.t.\ [(1+t^a)p(\theta)+t^s]y \le I.$$

The first-order conditions are given by:

$$\frac{\partial \mathcal{L}}{\partial y} = v_y - \gamma[(1+t^a)p(\theta)+t^s] = 0$$

$$\frac{\partial \mathcal{L}}{\partial \theta} = v_\theta - \gamma(1+t^a)p'(\theta)y = 0$$

$$\frac{\partial \mathcal{L}}{\partial \gamma} = I - [(1+t^a)p(\theta)+t^s]y = 0$$

The bordered Hessian matrix for this problem is:

$$Q = \begin{bmatrix} v_{yy} & -\gamma(1+t^a)p'(\theta) + v_{y\theta} & -[(1+t^a)p(\theta)+t^s] \\ -\gamma(1+t^a)p'(\theta) + v_{y\theta} & v_{\theta\theta} - \gamma(1+t^a)p''(\theta)y & -(1+t^a)p'(\theta)y \\ -[(1+t^a)p(\theta)+t^s] & -(1+t^a)p'(\theta)y & 0 \end{bmatrix}$$

Its determinant $|Q|$ is:

$$|Q| = 2\frac{v_y(v_\theta)^2}{y(\gamma)^2}(\varepsilon_y^{v_\theta} - 1) - \Big[(\frac{v_y}{\gamma})^2 v_\theta\{\frac{v_{\theta\theta}}{v_\theta} - \frac{p''(\theta)}{p'(\theta)}\} + (\frac{v_\theta}{\gamma})^2 v_{yy})\Big]$$

Here, we assume that $\varepsilon_y^{v_\theta} \geq 1$ and $p''(\theta) \geq 0$ to ensure the maximization problem. Therefore, since inverse matrix of Q exists, we can apply Cramer's rule and then obtain following results.

$$\frac{\partial\theta}{\partial t^a} - \frac{\partial\theta}{\partial t^s}p(\theta) = \frac{-1}{|Q|}\gamma p'(\theta)y[(1+t^a)p(\theta)+t^s]^2 \tag{D.1}$$

$$\frac{\partial y}{\partial t^s} - \frac{\partial y}{\partial t^s}p(\theta) = \frac{1}{|Q|}\gamma(1+t^a)(p'(\theta)y)^2[(1+t^a)p(\theta)+t^s] \tag{D.2}$$

The objective of the government is to maximize the indirect utility $V \equiv V(t^s, t^a) \equiv u(\theta^* \equiv \theta(t^s, t^a), y^* \equiv y(t^s, t^a))$ subject to the government's budget constraint $t^a p(\theta^*)y^* + t^s y^* = R$ by choosing t^a and t^s. The first order conditions are

$$\frac{\partial V}{\partial t^a} + \lambda\Big[p(\theta)y + t^a p'(\theta)\frac{\partial\theta}{\partial t^a}y + t^a p(\theta)\frac{\partial y}{\partial t^a} + t^s\frac{\partial y}{\partial t^a}\Big] = 0 \tag{D.3}$$

$$\frac{\partial V}{\partial t^s} + \lambda\Big[y + t^a p'(\theta)\frac{\partial\theta}{\partial t^s}y + t^a p(\theta)\frac{\partial y}{\partial t^s} + t^s\frac{\partial y}{\partial t^s}\Big] = 0 \tag{D.4}$$

Using the Roy's identity, Equation (D.3) and Equation (D.4) can be rewritten as

$$(\lambda - \alpha)p(\theta)y + \lambda\Big[t^a p'(\theta)\frac{\partial\theta}{\partial t^a}y + t^a p(\theta)\frac{\partial y}{\partial t^a} + t^s\frac{\partial y}{\partial t^a}\Big] = 0 \tag{D.5}$$

$$(\lambda - \alpha)y + \lambda\Big[t^a p'(\theta)\frac{\partial\theta}{\partial t^s}y + t^a p(\theta)\frac{\partial y}{\partial t^s} + t^s\frac{\partial y}{\partial t^s}\Big] = 0 \tag{D.6}$$

where $\alpha \equiv \frac{\partial V}{\partial I}$. Combining Equation (D.5) and Equation (D.6), we can get

$$\begin{aligned}\frac{t^a p(\theta^*)}{t^a p(\theta^*) + t^s} &= \frac{-p(\theta^*)\left[\frac{\partial y}{\partial t^s} - \frac{\partial y}{\partial t^s} p(\theta)\right]}{p'(\theta^*) y^* \left[\frac{\partial \theta}{\partial t^a} - \frac{\partial \theta}{\partial t^s} p(\theta)\right]} \\ &= \frac{(1 + t^a) p(\theta^*)}{(1 + t^a) p(\theta^*) + t^s}\end{aligned} \tag{D.7}$$

The second equality is derived by substituting Equation (D.1) and Equation (D.2). Note that the equality is satisfied only when t^s is zero. This is because $\frac{(1+t^a)p(\theta^*)}{(1+t^a)p(\theta^*)+t^s}$ is greater than $\frac{t^a p(\theta^*)}{t^a p(\theta^*)+t^s}$ if t^s is positive. Therefore, wholly ad-valorem taxation is optimal.

Chapter 3

Public good provision financed by nonlinear income tax under reduction of envy

3.1 Introduction

In an economy where agents have different skill levels, there are several ethical reasons to consider income redistribution. One such reason is the envy caused by income inequality. One agent envies another agent if he/she prefers the other's commodity bundle to his/her own. Income inequalities or many complaints among citizens lead to collective decision making out of the way. Recently, World Economic Forum (2017) reported that the income gap, one of the sources of envy, is a major driving force of polarized political outcomes. Further, as Bös and Tillmann (1985) noted,

> *the economic rationale for a minimization or reduction of envy by taxation is the following. Excessive envy in a society is an element of social disorder. Reducing envy in a society is a step towards increasing social harmony.* (p. 34)

Hence, reducing envy is not only a normative concept but also a relevant constraint for the politicians concerned with the harmony of society.

In the context of income taxation with endogenous labor supply, high-skilled agents cannot envy low-skilled ones because of the self-selection constraint. Conversely, low-skilled agents must envy high-skilled ones. While it is difficult to apply the original envy-free constraint presented by Varian (1974), we replace the weaker and cardinal criterion proposed by Diamantaras and Thomson (1989) to evaluate the intensity of envy, called λ envy-free, and examine the optimal policy schedule under not only self-selection or incentive compatibility, which extracts the true information on skill from each agent, but also constraints on the reduction of envy.

In this study, we investigate the optimal nonlinear income taxation with public good provision constrained on the reduction of envy, as well as the conventional constraints used by Boadway and Keen (1993). The objective of the government is to achieve a Pareto-efficient allocation, so that it maximizes low-class utility given the requirements for high-class utility, budget constraint, self-selection, and reduction of envy. In such situations, we derive the optimal provision rule of the public good, as well as the marginal income tax rate for each class. For the marginal income tax rate, we obtain the same results as Nishimura (2003b). Conversely, we derive the optimal provision rule following Boadway and Keen (1993), except for the distortion that arises from the λ envy-free constraint. This ethical constraint for the low class allows the policymaker to compare the marginal rate of substitution (MRS) for the high class with that for the λ high class and use the difference to relax that constraint. In particular, because changing the amount of private consumption for the high class implies changing that for the λ high-class λ times as much as for the original high class, the direction of the distortion is determined by whether the MRS is a step up or step down. To understand the provision rule, we use the constant elasticity of substitution (CES) utility function on private consumption and the public good, and show that the elasticity of substitution plays a key role in determining the sign. In addition, we conduct a numerical simulation to reveal the effect of λ on the public good provision. As extensions, we study the public good provision under mixed taxation, keeping the other settings constant.

The studies related to this research can be categorized into taxation with public good provision and optimal taxation under reduction of envy. The latter category may represent taxation in cases where a policymaker aims to satisfy fair distribution such as the maximin or Pigou-Dalton principle.[1] However, this study allows the policymaker to set the reduction of envy as one of the constraints, not as an objective. With regard to the optimal taxation for the reduction of envy, Nishimura (2003b) studies the optimal nonlinear income taxation under constraints on the reduction of envy, showing that the marginal income tax rate can increase only if leisure is a luxury. In addition, Nishimura (2003a) examines the optimal commodity taxation for the reduction of envy. Both these

[1] For instance, Fleurbaey and Maniquet (2006) derive the optimal income tax schedule in settings where the social planner maximizes the social index satisfying several axioms for fairness and inequality aversion. They characterize the social index as meeting several axioms before deriving the optimal policy.

studies adopt a particular envy-free notion, namely the λ envy-free of Diamantaras and Thomson (1989). While we adopt the same approach, our study introduces public good provision by the government.

On the optimal nonlinear income taxation with public good provision, Boadway and Keen (1993) show that a government provides a public good following the modified Samuelson rule, which embraces the self-selection term. This means that the policymaker reduces the provision level when the mimicker values the public good more than low-ability agents to redistribute more tax wealth. Nava et al. (1996) study the optimal nonlinear income and linear commodity taxation with pure public good provision, showing that the Samuelson rule is modified by two additional terms related to the self-selection constraint and revenue of indirect taxes, as well as that these terms disappear when the utility function is weakly separable between public and private goods (taken together) and leisure. Gaube (2005b) provides a sufficient condition for both a lower and higher level of public expenditure in the second best than in the first best based on Boadway and Keen (1993).

In the optimal tax literature, several theoretical studies have explored the effect of status or relative consumption (or income), that is, individuals' utilities depend on not only their own consumption of goods but also their relative standing in the society (e.g., Boskin and Sheshinski (1978), Oswald (1983), Seidman (1987), Persson (1995), Ireland (2001), Corneo (2002), Aronsson and Johansson-Stenman (2008), Balestrino (2009), Micheletto (2011), Kanbur and Tuomala (2013), Bruce and Peng (2018)). In particular, Aronsson and Johansson-Stenman (2008) and Micheletto (2011) examine the public good provision under the optimal nonlinear income tax in the presence of interdependence in individuals' utilities. Aronsson and Johansson-Stenman (2008) describe relative consumption as the difference between individual's own consumption and the average consumption in an economy, and show that the Samuelson rule should be upwardly distorted when leisure is weakly separable from private and public consumption. Additionally, Micheletto (2011) focuses only on the case where individuals care about the consumption of a richer group, according to evidence provided by Bowles and Park (2005). As the consumption of higher-income agents increases, it negatively affects the preferences of lower income agents, which is considered as "Veblen effect" in his paper. He also shows that the overprovision of the public good relative to the Samuelson rule is always optimal due to the Veblen effect if no self-selection constraints are binding. While these studies investigate the second-best allocation when agents have other-regarding preferences, we study the λ equitable allocation under the second-best environment when agents do not have other-regarding preferences.[2] We show that the λ envy-free approach proposes not only the case where the overprovision of the public good is optimal but also two novel cases where the underprovision of the public good is optimal and the Samuelson rule applies even if self-selection constraints are not binding, assuming a special form of the utility function.[3]

[2] To clarify how status effects should be reflected in the optimal provision rule of public goods expressed by equation (5), we will characterize the second-best Samuelson rule under the reduction of envy when agents have other-regarding preferences in future research. Velez (2016) and Nakada (2018) explore the equitable allocation of multiple indivisible goods and money among agents with other-regarding preferences. However, they do not analyze the second-best provision rule for public goods under the setting.

[3] This paper also belongs to the strand of optimal tax literature on poverty alleviation or income inequality, since we assume all individuals share the same utility function. While there are some papers on this topic (e.g., Besley and Kanbur (1988), Besley and Coate (1992, 1995), Kanbur et al. (1994), Pirttilä and Tuomala

The remainder of this paper is organized as follows. Section 2 examines the optimal provision rule for pure public goods under the reduction of envy and section 3 presents simple numerical examples. Section 4 extends the model to the case of linear commodity taxation and section 5 offers concluding remarks.

3.2 Optimal income taxation with public good provision for reduction of envy

We consider a two-class economy in which each agent ($i = H, L$) possesses an exogenous skill level w_i, where $w_H > w_L > 0$. There is a continuum of individuals with unit mass. Let $n_H \in (0, 1)$ denote the proportion of high-skilled individuals and the remaining $n_L = 1 - n_H$ the proportion of low-skilled ones. They earn their income by supplying labor, and their earnings are the product of the unit wage (or skill level) and amount of labor supply. The government collects taxes on their incomes, which can be scheduled nonlinearly. In addition, it provides a public good by using the collected taxes.

First, we assume three types of goods: consumption (or after-tax income) $c \in \mathbb{R}_+$, labor supply l, and public good $G \in \mathbb{R}_+$. We also assume that each worker provides at most $\bar{l}$ labor, meaning that he/she chooses supply level l to be between 0 and $\bar{l}$. Every agent shares an identical utility function, $U(c_i, G, l_i)$, and U is twice continuously differentiable, strictly concave, and strictly increasing in c and G and strictly decreasing

(2004), Kanbur et al. (2018)), they are different from our analysis because they consider non-welfarist's optimization, whereas we attempt to find the constrained Pareto-efficient allocation.

in l. Let Y be labor income. If agent i, with skill w_i, earns labor income Y_i, we can replace the expression with $U(c_i, G, \frac{Y_i}{w_i})$. To provide the public good, the government must incur production cost $\phi(G)$ with a strictly increasing, strictly convex, and twice continuously differentiable function. For all goods except the public good, a good with subscript i means one that agent i enjoys.

We assume that the government wants to achieve a constrained Pareto-efficient allocation. Specifically, we consider the problem of maximizing low-skilled utility subject to high-skilled agents having at least a given utility level, $\bar{u}$. The planner faces three other constraints. First, the government faces a resource constraint. Let $T : \mathbb{R} \to \mathbb{R}$ be the income tax function, and agent i's budget constraint is written as $c_i = w_i l_i - T(w_i l_i)$. Therefore, the government's resource constraint is

$$n_L T(w_L l_L) + n_H T(w_H l_H) = n_L(w_L l_L - c_L) + n_H(w_H l_H - c_H) \geq \phi(G). \tag{1}$$

Second, the policymaker cannot observe agents' skill directly but does know their earned income. Hence, we require that he/she resolves the information asymmetry problem, called the *self-selection* constraint. We formulate this as follows:

$$U(c_i, G, l_i) \geq U(c_j, G, \frac{w_j}{w_i} l_j), \tag{2}$$

for any $i, j = H, L$ with $i \neq j$. Finally, we impose an ethical constraint for reducing envy. The equity concept of no-envy faces a difficulty in the second-best situation, since the low-skilled agent always envies the high-skilled one, whereas the high-skilled agent never envies the low-skilled agent.[4] As a less demanding criterion of envy reduction,

[4] From the self-selection constraint for high-skilled agents, the following inequality holds: $U(c_H, G, l_H) \geq U(c_L, G, \frac{w_L}{w_H} l_L) > U(c_L, G, l_L)$. Therefore, the high-skilled agent never envies the low-skilled agent, which means that the envy-free constraint for the low-skilled agent is not satisfied.

we adopt the λ envy-free introduced by Diamantaras and Thomson (1989) and used by Nishimura (2003a,b) as a cardinal measure of the intensity of envy. The reason to employ cardinal concepts is that, according to Bös and Tillmann (1985), ordinal concepts are not useful because there is an invariant hierarchy of envy in the second-best analysis. Also, note that the Lagrangian expression of the optimization problem with the λ envy-free constraint (equation (4)) is similar to the social objective of Varian (1976a), who incorporates degrees of envy into the social objective, not constraint. However, λ envy-free is better in the sense that it is independent of the comparability and cardinality of utility functions (see Nishimura (2003b) for details).

Let λ_{ij} be a nonnegative real number, such that $U(c_i, G, l_i) = U(\lambda_{ij} c_j, G, \bar{l} - \lambda_{ij}(\bar{l} - l_j))$ when $U(c_i, G, l_i) \leq U(c_j, G, l_j)$ and $\lambda_{ij} \equiv 1$ when $U(c_i, G, l_i) > U(c_j, G, l_j)$. If λ_{ij} is unity, it is the no-envy case. When agent i envies agent j, the value of λ_{ij} represents the amount by which one would have to decrease j's bundle to stop agent i envying agent j. In other words, λ_{ij} indicates the intensity of envy. Assume that agent i compares his/her own bundle with the bundle containing the public good and a proportional contraction of agent j's consumption and leisure between points $(0, \bar{l})$ and (x_j, l_j). Let $\lambda \equiv \min_{ij} \lambda_{ij}$. Under the binding self-selection constraint, $\lambda = \lambda_{LH}$, since $\lambda_{LH} < 1$ and $\lambda_{HL} = 1$. An allocation is then λ envy-free if $U(c_i, G, l_i) \geq U(\lambda c_j, G, \bar{l} - \lambda(\bar{l} - l_j))$ for all i and j. We consider that the government is constrained by a given λ envy-free requirement:

$$U(c_i, G, l_i) \geq U(\lambda c_j, G, \bar{l} - \lambda(\bar{l} - l_j)), \tag{3}$$

for any $i, j = H, L$ with $i \neq j$.[5] Because a high-skilled agent never envies a low-skilled

[5]Nishimura (2000) presents the tax policy implications under the Pareto-efficient allocations that maximize λ as in Diamantaras and Thomson (1989). He also shows that envy is minimized at the leximin allo-

agent, we focus only on the λ envy-free constraint for the low-skilled agent.

Summarizing the above, the policymaker's optimization problem can be written as follows:

$$\max_{\{c_i,l_i\}_{i=L,H},G} U(c_L, G, l_L),$$

subject to

$$U(c_H, G, l_H) \geq \bar{u}$$

$$n_L(w_L l_L - c_L) + n_H(w_H l_H - c_H) \geq \phi(G)$$

$$U(c_i, G, l_i) \geq U(c_j, G, \frac{w_j}{w_i} l_j) \quad \text{where } i, j = H, L \text{ with } i \neq j$$

$$U(c_L, G, l_L) \geq U(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H)).$$

The Lagrangian is

$$\begin{aligned}
&\mathcal{L}(c_L, c_H, l_L, l_H, G; \gamma, \delta_r, \delta_{sH}, \delta_{sL}, \delta_e) = \\
&U(c_L, G, l_L) + \gamma\{U(c_H, G, l_H) - \bar{u}\} \\
&+\delta_r\{n_L(w_L l_L - c_L) + n_H(w_H l_H - c_H) - \phi(G)\} \\
&+\delta_{sH}\{U(c_H, G, l_H) - U(c_L, G, \frac{w_L}{w_H} l_L)\} + \delta_{sL}\{U(c_L, G, l_L) - U(c_H, G, \frac{w_H}{w_L} l_H)\} \\
&+\delta_e\{U(c_L, G, l_L) - U(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))\},
\end{aligned} \tag{4}$$

where $\gamma, \delta_r, \delta_{sH}, \delta_{sL}$, and δ_e are the Lagrangian multipliers associated with the first, second, third, fourth, and fifth constraints, respectively.[6] Note that this problem is almost the same as that of Boadway and Keen (1993), but we incorporate the λ envy-free con-

cation that maximizes the utility of the low-skilled agent. By contrast, this study examines the second-best Pareto-efficient allocations corresponding to various λ, as in Nishimura (2003a,b).

[6] Nishimura (2003b) demonstrates that the second-best frontier with the λ envy-free constraint gradually shrinks as λ increases. Indeed, as long as δ_e is positive, $\mathcal{L}$ decreases λ.

straint. Appendix A shows the first-order conditions with respect to the Lagrangian. Hereafter, we focus on the redistributive cases only: $\delta_{sL} = 0$ and $\delta_{sH} > 0$.

3.2.1 Marginal income tax rate

We derive the marginal income tax rate for each type in the same way as Nishimura (2003b). Let $U_a^i \equiv \partial U(c_i, G, l_i)/\partial a_i$, $\hat{U}_c \equiv \partial U(c_L, G, \frac{w_L}{w_H} l_L)/\partial c_L$, $\hat{U}_l \equiv \partial U(c_L, G, \frac{w_L}{w_H} l_L)/\partial(\frac{w_L}{w_H} l_L)$, and $\bar{U}_a \equiv \partial U(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))/\partial(\lambda a_H)$, where $i = H, L$ and $a = c, l$. The next lemma provides the marginal income tax rates.

Lemma 1. *Under the redistributive cases when $\delta_{sL} = 0$ and $\delta_{sH} > 0$,*

1. *The marginal income tax rate at the bottom is*

$$T'(w_L l_L) = \frac{\delta_{sH} \hat{U}_c}{\delta_r} \left[MRS^L(y, c) - \hat{MRS}(y, c) \right] > 0,$$

where $MRS^L(y, c) = -\frac{1}{w_L} \frac{U_l^L}{U_c^L}$ and $\hat{MRS}(y, c) = -\frac{1}{w_H} \frac{\hat{U}_l}{\hat{U}_c}$.

2. *The marginal income tax rate at the top is*

$$T'(w_H l_H) = \frac{\lambda \delta_e \bar{U}_c}{\delta_r w_H} \left[MRS_{lc}^H - \bar{MRS}_{lc} \right],$$

where $MRS_{lc}^H \equiv -\frac{U_l^H}{U_c^H}$ is the MRS for l_H measured by c_H and $\bar{MRS}_{lc} \equiv -\frac{\bar{U}_l}{\bar{U}_c}$ is the MRS measured at $(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))$.

This lemma is consistent with Nishimura (2003b).[7] Because of the self-selection constraint for agents with high skill, the marginal income tax rate for low-skilled agents must be positive, as shown by Stiglitz (1982). Conversely, the marginal income tax

[7] Nishimura (2003b) also examines these marginal income tax rates when the self-selection constraint for low-skilled workers is binding.

rate on the top is different from the standard result presented by Stiglitz (1982), since the term that represents the effect of the λ envy-free constraint appears.[8] Nishimura (2003b) shows that if the income elasticity of leisure is greater (less) than 1, MRS^H_{lc} is greater (less) than $M\bar{R}S_{lc}$, which means that the marginal income tax rate on the top must be positive (negative).[9] Of course, if $MRS^H_{lc} = M\bar{R}S_{lc}$, it must be zero. Moreover, if the equitability constraint does not bind (i.e., $\delta_e = 0$), then it must be zero.

3.2.2 Provision rule of the public good

This section presents the public good provision rule at the optimum. As per Boadway and Keen (1993), the optimal provision rule includes the self-selection term, which plays an important role in income redistribution. If the mimicker places more weight on the public good based on private consumption than the mimicked one with low skill, the government should reduce its production and transfer the tax revenue to low-class agents. In addition, to relax the λ envy-free constraint, the government increases or decreases the amount. For instance, if the evaluation of the public good for the private good at the λ-scaled bundle $(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))$ is higher than that the high-skilled agent receives, then he/she must reduce the provision level to redistribute more income.

Let $U^i_G \equiv \partial U(c_i, G, l_i)/\partial G$, $\hat{U}_G \equiv \partial U(c_L, G, \frac{w_L}{w_H} l_L)/\partial G$, and $\bar{U}_G \equiv \partial U(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))/\partial G$, where $i = H, L$. Formally, we can derive the optimal rule with respect to

[8] Note that the difference in the MRS between consumption and labor, not the efficiency-unit labor, between the envying and envied agent is useful information for the government, since the λ envy-free constraint allows us to consider a proportional decrease of the envied agent's bundle.

[9] According to the definition of Nishimura (2003b), if the income elasticity of leisure is greater (less) than 1, leisure is called a luxury (necessity).

the public good provision in the next proposition.

Proposition 3. *Under the optimal nonlinear income taxation with the λ envy-free and self-selection constraint, the optimal provision rule is characterized by*

$$\sum_{i=H,L} n_i MRS^{i}_{Gc} + \frac{\delta_{sH}}{\delta_r}\hat{U}_c(MRS^{L}_{Gc} - \hat{MRS}_{Gc}) + \frac{\lambda\delta_e}{\delta_r}\bar{U}_c(MRS^{H}_{Gc} - \frac{1}{\lambda}\bar{MRS}_{Gc}) = \phi'(G), \quad (5)$$

where $MRS^{i}_{Gc} \equiv \frac{U^i_G}{U^i_c}$ is type i's MRS for G measured by c_i, $\hat{MRS}_{Gc} \equiv \frac{\hat{U}_G}{\hat{U}_c}$ is the mimicker's MRS between c and G, and $\bar{MRS}_{Gc} \equiv \frac{\bar{U}_G}{\bar{U}_c}$ is the MRS measured at $(\lambda c_H, G, \bar{l} - \lambda(\bar{l} - l_H))$.

The first term is the sum of agent i's MRS for public good G measured by private consumption c_i and the second term the effect of the incentive constraint. The third term is a novel one, which reflects the effect on the λ envy-free constraint and whose implication is similar to that of the incentive constraint. Because λ distorts the consumption/leisure bundle for the envying agent, this term may not be zero. To relax the λ envy-free constraint, the government changes the provision level of the public good and makes room to improve welfare. We suggest an intuitive interpretation of the third term. Starting from the original Samuelson rule, consider the following redistribution. The government imposes an additional tax liability MRS^{i}_{Gc} on type-i individuals to increase G. The tax reform does not change the welfare of type-i individuals or the government's budget. The valuation of G of the envying agent is expressed by $\frac{1}{\lambda}\bar{MRS}_{Gc}$. If $MRS^{H}_{Gc} > \frac{1}{\lambda}\bar{MRS}_{Gc}$, the third term in equation (5) suggests that the original Samuelson rule should be upwardly shifted. This implies that an increase in G mitigates the intensity of envy for low-type agents because the tax liability of the envied agent is larger than that of the envying agent and, then, the difference between their utilities is reduced. Therefore, the upward distortion relaxes the λ envy-free constraint for low-type individuals.

Boadway and Keen (1993) show that the original Samuelson rule for the public good provision is replicated when each agent's preference is represented by $U(H(c,G),l)$, namely c and G are weakly separable with l in the utility function. In this case, while the second bracket on the left-hand side is zero, it is ambiguous whether the third bracket is zero.

3.2.3 A special case: CES utility function

This subsection derives the direction of the distortion due to the binding λ envy-free constraint on the provision rule by using a concrete utility function. To examine the direction of the distortion, we assume that the utility function is expressed by $U(H(c,G),l)$ and $H(c,G)$ is the CES functional form: $H(c,G) = (\alpha c^{\rho} + \beta G^{\rho})^{\frac{1}{\rho}}$, where $\rho \leq 1$. If ρ converges to zero, $H(c,G)$ converges to the Cobb–Douglas expression (i.e., $H(c,G) = c^{\alpha}G^{\beta}$). In this case, the round bracket can be represented by

$$MRS^{H}_{Gc} - \frac{1}{\lambda}M\bar{R}S_{Gc} = (1-\lambda^{-\rho})\left(\frac{\beta}{\alpha}\right)\left(\frac{c_H}{G}\right)^{1-\rho}.$$

$1-\lambda^{-\rho}$ determines the sign and the elasticity of substitution $\frac{1}{1-\rho}$ plays a crucial role since $\lambda < 1$. If $\frac{1}{1-\rho} \in (0,1)$, then the direction of the distortion is positive; otherwise, that direction is negative except for $\frac{1}{1-\rho} = 1$. If $\frac{1}{1-\rho} = 1$, the bracket equals zero and, thus, the third and second terms disappear. To sum up, the next corollary describes the direction of the distortion on the provision rule.

Corollary 2. *Assume that all agents have the following utility function:* $H(c,G) =$

$(\alpha c^{\rho} + \beta G^{\rho})^{\frac{1}{\rho}}$. The optimal provision rule distorts

- *Downwardly if the elasticity of substitution $\frac{1}{1-\rho} \in (1, +\infty)$ or $\rho = 1$;*
- *Upwardly if the elasticity of substitution $\frac{1}{1-\rho} \in (0, 1)$.*

In addition, the rule coincides with the Samuelson rule if the elasticity of substitution equals one.

If the elasticity of substitution is above one, the government increases private consumption for the high type by decreasing the provision level of the public good. The elasticity of substitution means the variation of the ratio between private consumption and the public good ($\frac{c_i}{G}$) when MRS^{i}_{Gc} changes. Hence, when the elasticity of substitution is above one, an decrease of the ratio due to a proportional contraction of private consumption for high type allows the MRS between private consumption and public good to decrease by less than the proportional decrease in the corresponding MRS. This means that $M\bar{R}S_{Gc} > \lambda MRS^{H}_{Gc}$, which is equivalently to stating that the corresponding MRS for envying agents (i.e., $\frac{1}{\lambda}M\bar{R}S_{Gc}$) is greater than the corresponding MRS for envied agent (i.e., MRS^{H}_{Gc}). That is, envying agents value the public good more than envied agent. Thus, it is desirable for the government that the amount of the public good decreases and private consumption for high type increases. Also, the argument is symmetric for the opposite case where the elasticity of substitution is below one.

The validity of the Samuelson rule under $\rho = 0$ stems from that $H(\cdot)$ is homothetic in c. In this case, a proportional decrease of the envied agent's consumption implies that the MRS decreases proportionally as c decreases. Therefore, $MRS^{H}_{Gc} = \frac{1}{\lambda}M\bar{R}S_{Gc}$ holds.[10]

[10]Consider that $H(\cdot)$ is homogeneous of degree k in c. The third term can be rewritten as follows:

3.2.4 Remarks: λ envy-free constraint

We make two comments about key constraint (3) in our model, the λ envy-free constraint: (i) the form of the λ envy-free constraint; (ii) the interaction between the self-selection constraint for the high-type and the λ envy-free constraint for the low-type.

First, we explain why the intensity of envy λ is not applied to the amount of the public good in the λ envy-free constraint. In our model, we consider that agent i compares two bundles: his/her own bundle, (x_i, G, l_i), and the bundle containing the public good and a proportion λ of agent j's consumption and leisure, $(\lambda c_j, G, \bar{l} - \lambda(\bar{l} - l_j))$. Remember that the classic concept of no-envy is not useful here due to the self-selection constraint. If no agent envies any other agent's bundle including the public good and the proportional contraction of the consumption and leisure, the allocation satisfies the λ envy-free constraint. This means that the utility of one agent does not increase, even if the government allocates the bundle of any other agent consisting of the public good and λ-scaled consumption and leisure to the agent. From this viewpoint, if one shrinks the amount of the public good in the λ envy-free constraint, the government needs to be able to implement the λ proportion of the public good to prevent one agent from envying any other agent. However, since all individuals share the same amount of the public good (i.e., public goods are non-rivalrous) provided by the government in this economy, such an allocation is infeasible. Therefore, it is inconsistent to impose the intensity of envy λ on the amount of the public good. Hence, we employ the version of

---

$\frac{\delta_e}{\delta_r}\lambda\bar{U}_c\left(\frac{H_G(G,c_H)}{H_{c_H}(G,c_H)} - \frac{H_G(G,\lambda c_H)}{\lambda H_{c_H}(G,\lambda c_H)}\right) = \frac{\delta_e}{\delta_r}\lambda\bar{u}^H_c\left(\frac{H_G(G,c_H)}{H_{c_H}(G,c_H)} - \frac{\lambda^k H_G(G,c_H)}{\lambda\times\lambda^{k-1}H_{c_H}(G,c_H)}\right) = 0$. As such, we obtain $MRS^H_{Gc} = \frac{1}{\lambda}M\bar{R}S_{Gc}$.

the λ envy-free constraint given by equation (3).

If we were to rewrite the constraint as:

$$U(c_i, G, l_i) \geq U(\lambda c_j, \lambda G, \bar{l} - \lambda(\bar{l} - l_j)),$$

then the optimal Samuelson rule is:

$$\sum_{i=H,L} n_i MRS^i_{Gc} + \frac{\delta_{sH}}{\delta_r} \hat{U}_c(MRS^L_{Gc} - M\hat{R}S_{Gc}) + \frac{\lambda\delta_e}{\delta_r} \bar{U}_c(MRS^H_{Gc} - M\tilde{R}S_{Gc}) = \phi'(G),$$

where $M\tilde{R}S_{Gc} \equiv \frac{\tilde{U}_G}{\tilde{U}_c}$, $\tilde{U}_G \equiv \partial U(\lambda c_H, \lambda G, \bar{l} - \lambda(\bar{l} - l_H))/\partial(\lambda G)$, and $\tilde{U}_c \equiv \partial U(\lambda c_H, \lambda G, \bar{l} - \lambda(\bar{l} - l_H))/\partial(\lambda c_H)$. If we assume that the utility function is described by $U(H(c, G), l)$, the second term disappears. Also, if the function $H(\cdot)$ is homothetic, the marginal rate of substitution between G and c is constant on the path of the λ-contraction of the envied agent's allocation, that is, $MRS^H_{Gc} = M\tilde{R}S_{Gc}$. Therefore, a sufficient condition for the original Samuelson rule is that $H(\cdot)$ is homothetic. For example, if $H(\cdot)$ is the CES functional form used in subsection 2.3, the original Samuelson rule is desirable.

Second, we clarify the interaction between self-selection constraints and λ envy-free constraints. Here, we focus on the case where the government faces the second-best environment, that is, the self-selection constraint for the high type must be binding. Choosing any plausible utility function and profile of wages, let us consider a set Ω consisting of all allocations satisfying the three constraints (the constraint that high-type agents have at least a given level of utility, resource constraint, and self-selection constraint for the high type) are binding and the λ envy-free constraint for the low-type is slack. Each element is denoted by $\omega \equiv \{(c^\omega_L, G^\omega, \ell^\omega_L), (c^\omega_H, G^\omega, \ell^\omega_H)\} \in \Omega$. However, any allocation ω is not implementable at the optimum as λ increases. Specifically, any allocation ω violates the λ envy-free constraint for the low-type for any $\lambda > \bar{\lambda}$, where

$\bar{\lambda} \equiv \sup_{\omega \in \Omega} \lambda^{\omega}$ and λ^{ω} denotes a threshold so that $U(c_L^{\omega}, G^{\omega}, l_L^{\omega}) = U(\lambda^{\omega} c_H^{\omega}, G^{\omega}, \bar{l} - \lambda^{\omega}(\bar{l} - l_H^{\omega}))$. This is because, when we fix any $\lambda > \bar{\lambda}$, $U(c_L^{\omega}, G^{\omega}, l_L^{\omega}) < U(\lambda c_H^{\omega}, G^{\omega}, \bar{l} - \lambda(\bar{l} - l_H^{\omega}))$ holds for any ω from the fact that $U(\lambda c_H^{\omega}, G^{\omega}, \bar{l} - \lambda(\bar{l} - l_H^{\omega}))$ is an increasing function in λ. Additionally, λ^{ω} is well defined from the intermediate value theorem because, since $U(\lambda c_H^{\omega}, G^{\omega}, \bar{l} - \lambda(\bar{l} - l_H^{\omega}))$ is continuously increasing in λ, the utility of low-type agents is lower than for the high-type ones because of the self-selection constraint for the high-type, and is greater than $U(0, G, \bar{l})$ due to the interior solutions of c_L and l_L. Here, assuming that $\bar{\lambda}$ is below one, we address that if there exists at least one feasible allocation in the optimization problem at $\lambda \in (\bar{\lambda}, 1)$, the allocation allows not only the above three constraints but also the λ envy-free constraint for the low-type to be binding. In other words, the following three cases do not appear at the optimum in the interval $(\bar{\lambda}, 1)$: (i) the self-selection constraint for the high-type is binding and the λ envy-free constraint for the low-type is slack; (ii) both the self-selection constraint for the high-type and the λ envy-free constraint for the low-type are slack; and (iii) the self-selection constraint for the high-type is slack and the λ envy-free constraint for the low-type is binding. Obviously, the first case does not occur due to the definition of $\bar{\lambda}$. Next, if the second case is realized, the government can implement the first-best allocation. Therefore, it contradicts the fact that we focus on the case where the government cannot implement the first-best Pareto efficient allocation. Finally, we denote by $\omega^0 \equiv \{(c_L^{\omega^0}, G^{\omega^0}, \ell_L^{\omega^0}), (c_H^{\omega^0}, G^{\omega^0}, \ell_H^{\omega^0})\}$ an allocation in the third case, and assume it is the optimal allocation at $\lambda \in (\bar{\lambda}, 1)$. For any $\lambda \leq \bar{\lambda}$, note that the self-selection constraint for the high-type is not affected and the λ envy-free constraint for the low-type is slackened. That is, allocation ω^0 is implementable as the first-best Pareto efficient allocation for any $\lambda \leq \bar{\lambda}$. Thus, it contradicts that the first-best Pareto efficient allocation is not implementable. As such, although an allocation ω is implementable at the optimum under any $\lambda \leq \bar{\lambda}$, the government needs to seek allocations that both the self-selection constraint for the high-type and the λ envy-free constraint for the low-type are binding over the interval $(\bar{\lambda}, 1)$, which means that the optimal level of the public good is derived from equation (5).[11]

So far, we have assumed that the self-selection constraint is always binding. However, if the government can implement the first-best Pareto efficient allocation, we cannot ignore the case where the λ envy-free constraint for the low-type is only binding. In other words, it is not necessarily that the binding self-selection constraint for the high-type is a necessary condition for the binding λ envy-free constraint for the low-type. If the self-selection constraint for the high-type is slack and the λ envy-free constraint for the low-type is binding, the original Samuelson rule should be modified only due to the effect of the third term on the left-hand side of equation (5).

[11] Although we consider the Pareto frontier when the government is constrained by a given λ equitability requirement as with Nishimura (2003a,b), there are the following unsolved questions left for future research. First, while λ^{ω} is below one due to the self-selection constraint for the high-type (see footnote 4), we cannot conclude that $\bar{\lambda}$ is below one from the definition of $\bar{\lambda}$. Second, even though $\bar{\lambda}$ is below one, it is unclear whether there exists at least one feasible allocation in the optimization problem at $\lambda \in (\bar{\lambda}, 1)$. In the present paper, instead of clarifying these questions analytically, we numerically provide some examples that the government can implement allocations so that both constraints are binding (see Tables 1 and 2).

3.3 Numerical examples

The previous section did not refer to the extent to which the λ envy-free requirement affects the amount of the public good. Therefore, this section presents a quantitative analysis of the amount of the public good. First, we examine the impact of the λ envy-free requirement on the utility level of a low-skilled agent and the amount of the public good. Next, we present the sensitivity of the amount of the public good with respect to the changes in the parameter value expressing the intensity of envy, namely λ.

In the simulation, we make the following assumptions. First, not only for the sake of simplicity but also to focus on the effect of the λ envy-free requirement on the provision level of public good, we assume that the functional form of the utility is $U(c, G, l) = H(c, G) - v(l)$, where the sub-utility function $H(\cdot)$ takes the CES form with $\alpha = \beta = 0.5$ and the disutility of labor $v(\cdot)$ takes an isoelastic form: $v(\ell_i) = \ell_i^{1+1/e}/(1 + 1/e)$ with $e > 0$. According to the empirical estimates (e.g., Chetty et al. (2011)), we set $e = 2$. Second, Fang (2006) and Goldin and Katz (2007) estimate that the college wage premium is approximately 60%. We normalize low-type individuals' parameter w_L to equal one and, thus, that of the high-type individuals is assumed to be $w_H = 1.6$. Third, according to an OECD (2010) report, approximately one-quarter of all adults have attained tertiary education. Therefore, we assume that 25% of individuals are high-skilled workers. In other words, we set $n_H = 0.25$ and $n_L = 0.75$. Finally, we assume that $\bar{u}$ is unity and the cost function takes the following form, being strictly increasing and strictly convex: $\phi(\cdot) = G^2$.

We suggest numerical examples for two cases: $\rho = 1$ and $\rho = -1$. Table 1 presents the case in which the original Samuelson rule is downwardly distorted. First, the utility level of the low-skilled agent decreases as λ increases, since the second-best frontier decreases. Second, the provision level decreases when the λ envy-free constraint is binding, which supports the results of Corollary 1. Third, the provision level decreases as λ increases. The intuition is that the government reinforces the income redistribution by distorting the provision level to alleviate the intensified envy. Table 2 describes the case in which the original Samuelson rule is upwardly distorted. As with the results in Table 1, low-skilled utility decreases as λ increases. However, in contrast to the results in Table 1, the provision level increases under the reduction of envy and increases much more as λ increases. That is, the government mitigates the intensified envy by the public good provision rather than income redistribution.

3.4 Mixed taxation

Here, we examine the optimal provision rule for public goods when the government employs not only labor income but also commodity taxes. We assume that the government can only levy linear commodity taxes, since it cannot observe individuals' consumption levels.

Again, we define the identical utility function of agent i as $U(c_i, x_i, G, l_i)$, where c_i is a numéraire commodity and x_i another commodity. The producer price of commodity x is constant and normalized to unity for simplicity. While the government cannot impose any taxes on the numéraire good, it imposes proportional commodity tax t on x_i. For simplicity, we assume that $n_H = n_L = 1$, which does not affect the tax schedule crucially. The other notations are the same as in the previous section.

Following Mirrlees (1976) and Jacobs and Boadway (2014), we decompose individual optimization into two stages. In the first stage, each agent chooses the amount of labor supply given nonlinear income taxes, which allows us to determine disposable income $R_i \equiv w_i l_i - T(w_i l_i)$. In the second stage, each agent expenses his/her disposable income to consume a numéraire and another commodity. We assume that individuals anticipate the outcome for the second stage in the first stage. Now, we formally analyze individuals' problem. In the second stage, given $\{p, R_i, G, l_i\}$, agent i chooses c_i and x_i to maximize utility $U(c_i, x_i, G, l_i)$ subject to budget constraint $c_i + px_i = R_i$, where $p \equiv 1 + t$ is the consumer price with respect to another commodity. The first-order conditions with respect to c_i and x_i yield

$$\frac{U_x^i}{U_c^i} = p. \tag{6}$$

The maximization problem in the second stage yields conditional commodity demands with respect to a numéraire and another commodity denoted by $c_i^* \equiv c(p, R_i, G, l_i)$ and $x_i^* \equiv x(p, R_i, G, l_i)$, respectively. As a result, substituting these solutions into the utility function yields a conditional indirect utility function, $V_i \equiv V(p, R_i, G, l_i) \equiv U(c_i^*, x_i^*, G, l_i)$. Let V_p^i, V_R^i, V_G^i, and V_l^i be the partial derivatives of V_i with respect to p, R_i, G, and l, respectively. From Roy's identity and the Slutsky decomposition, we can obtain the following relationship:

$$-\frac{V_p^i}{V_R^i} = x_i^*, \tag{7}$$

$$\frac{\partial x_i^*}{\partial p} = \frac{\partial \tilde{x}_i}{\partial p} - \frac{\partial x_i^*}{\partial R_i} \cdot x_i^*, \tag{8}$$

$$\frac{\partial x_i^*}{\partial G} = \frac{\partial \tilde{x}_i}{\partial G} + \frac{\partial x_i^*}{\partial R_i} \frac{V_G^i}{V_R^i}, \tag{9}$$

$$\frac{\partial c_i^*}{\partial p} = \frac{\partial \tilde{c}_i}{\partial p} - \frac{\partial c_i^*}{\partial R_i} \cdot x_i^*, \tag{10}$$

$$\frac{\partial c_i^*}{\partial G} = \frac{\partial \tilde{c}_i}{\partial G} + \frac{\partial c_i^*}{\partial R_i} \frac{V_G^i}{V_R^i}, \tag{11}$$

where $\tilde{c}_i$ and $\tilde{x}_i$ indicate the compensated conditional demands of individual i for the numéraire and the taxable good, respectively.

In the first stage, each agent chooses the amount of labor supply to maximize conditional indirect utility V_i subject to $R_i = w_i l_i - T(w_i l_i)$. The first-order condition is given by

$$-\frac{V_l^i}{w_i V_R^i} = -\frac{U_l^i}{w_i U_c^i} = 1 - T'(w_i l_i). \tag{12}$$

As above, the government faces the budget constraint, self-selection constraint to prevent high-skilled workers from mimicking low-skilled ones, and λ-equitability constraint for reducing envy. We respectively formulate these as follows:

$$\sum_{i=H,L} [w_i l_i - R_i + (p-1)x_i^*] \geq \phi(G), \tag{13}$$

$$V(p, R_H, G, l_H) \geq V(p, R_L, G, \frac{w_L}{w_H} l_L) \equiv \hat{V}, \tag{14}$$

$$V(p, R_L, G, l_L) \geq U(\lambda c_H^*, \lambda x_H^*, G, \bar{l} - \lambda(\bar{l} - l_H)) \equiv \bar{V}. \tag{15}$$

To sum up, the restricted Pareto optimization problem to the government is given by

$$\max_{\{R_i, l_i\}_{i=L,H}, p, G} V(p, R_L, G, l_L),$$

subject to

$$V(p, R_H, G, l_H) \geq \bar{u}$$

$$\sum_{i=H,L} [w_i l_i - R_i + (p-1)x_i^*] \geq \phi(G)$$

$$V(p, R_H, G, l_H) \geq V(p, R_L, G, \frac{w_L}{w_H} l_L)$$

$$V(p, R_L, G, l_L) \geq U(\lambda c_H^*, \lambda x_H^*, G, \bar{l} - \lambda(\bar{l} - l_H)).$$

The Lagrangian is

$$\begin{aligned} \mathcal{L}(p, R_L, R_H, l_L, l_H, G; \mu, \gamma, \delta, \eta) =& V(p, R_L, G, l_L) + \mu\{V(p, R_H, G, l_H) - \bar{u}\} \\ &+ \gamma\{\sum_{i=H,L} [w_i l_i - R_i + (p-1)x_i^*] - \phi(G)\} \\ &+ \delta\{V(p, R_H, G, l_H) - V(p, R_L, G, \frac{w_L}{w_H} l_L)\} \\ &+ \eta\{V(p, R_L, G, l_L) - U(\lambda c_H^*, \lambda x_H^*, G, \bar{l} - \lambda(\bar{l} - l_H))\}, \end{aligned} \tag{16}$$

where μ, γ, δ and η are the Lagrangian multipliers corresponding to the constraints, respectively. Appendix B shows the first-order conditions with respect to the Lagrangian.

Before analyzing the provision rule for public goods, it is useful to explore the optimal linear commodity tax rate. Let $\hat{V}_p$, $\hat{V}_R$, and $\hat{V}_G$ be the partial derivatives of $\hat{V}$ with respect to p, R, and G. The linear commodity tax rate is characterized by the following proposition.

Proposition 4. *Assume that the allocations are restricted by reduction of envy. The optimal commodity tax rate under the nonlinear labor income tax and public good provision is given by*

$$t \sum_{i=H,L} \frac{\partial \tilde{x}_i}{\partial p} = \frac{\delta}{\gamma} \hat{V}_R (x_L^* - \hat{x}) + \frac{\lambda \eta}{\gamma} \Big[\bar{U}_c \frac{\partial \tilde{c}_H}{\partial p} + \bar{U}_x \frac{\partial \tilde{x}_H}{\partial p} \Big], \tag{17}$$

where $\hat{x} \equiv x(p, R_L, G, \frac{w_L}{w_H} l_L)$ *is the mimicker's demand for another commodity and* $\bar{U}_r$ *(*$r = c, x$*) the derivative of* r *at the* λ*-scaled bundle* $(\lambda c_H^*, \lambda x_H^*, G, \bar{l} - \lambda(\bar{l} - l_H))$.

The first term on the right-hand side is the self-selection effect; if the agent's utility is separable between the commodity part and labor supply term, then it must disappear. We see this effect frequently in existing studies on mixed taxation. However, the second term on the right-hand side is the original part for reducing envy, as seen in Nishimura (2003a,b). Each term between the brackets is the inner product of the marginal utility of the low-skilled agent and substitution effect of the compensated demand, which reflects the reduction of envy by discouraging consumption by the high-skilled agent because of taxation. Moreover, the second term on the right-hand side can be rewritten as:[12]

$$\frac{\lambda\eta}{\gamma}\Big[\bar{U}_c\frac{\partial\tilde{c}_H}{\partial p} + \bar{U}_x\frac{\partial\tilde{x}_H}{\partial p}\Big] = \frac{\partial\tilde{x}_H}{\partial p}\bar{U}_c\Big[\frac{\bar{U}_x}{\bar{U}_c} - \frac{U_x^H}{U_c^H}\Big] \equiv \frac{\partial\tilde{x}_H}{\partial p}\bar{U}_c\Big[M\bar{R}S_{cx} - MRS_{cx}\Big]. \tag{18}$$

If the envying agent prefers the taxable good to the numéraire more than the envied agent, namely $M\bar{R}S_{cx} > MRS_{cx}$, it is taxed more heavily. This term remains even if the utility function is weakly separable between the public and private goods (taken together) and leisure, namely $U(H(c_i, x_i, G), l_i)$, while the first term on the right-hand side of equation (17) disappears. To replicate the Atkinson and Stiglitz (1976) theorem (hereafter, A-S theorem), we assume the following functional form: $H(f(c_i, x_i), G)$, where $f(\cdot)$ is homothetic. In this case, the second term on the right-hand side of equation (18) disappears, which means that commodity taxation is superfluous. The sufficient condition to hold the A-S theorem is slightly different from that in Nishimura (2003a,b), since we impose an additional restriction, namely weak separability between all private consumption and the public good.

[12]By using individuals' budget constraint $c_i + px_i = R_i$, the following relationship holds: $\frac{\partial\tilde{c}_H}{\partial p} = -p\frac{\partial\tilde{x}_H}{\partial p}$.

We now characterize the optimal provision rule for the public good. The optimal rule with respect to public good provision can be derived as in the next proposition.

Proposition 5. *Under linear commodity tax in addition to nonlinear income tax, the optimal provision rule taking reduction of envy into account is characterized by*

$$\sum_{i=H,L} \frac{V_G^i}{V_R^i} + \frac{\delta}{\gamma}\hat{V}_R\left[\frac{V_G^L}{V_R^L} - \frac{\hat{V}_G}{\hat{V}_R}\right] - \frac{\eta}{\gamma}\left[\bar{U}_c \frac{\partial \lambda \tilde{c}_H}{\partial G} + \bar{U}_x \frac{\partial \lambda \tilde{x}_H}{\partial G} + \bar{U}_G\right] = \phi'(G) - t \sum_{i=H,L} \frac{\partial \tilde{x}_i}{\partial G}, \quad (19)$$

where $\bar{V}_k$ *is the derivative of* $\bar{V} = U(\lambda c_H^*, \lambda x_H^*, G, \bar{l} - \lambda(\bar{l} - l_H))$ *with respect to* $k = G, R$.

On the left-hand side, the first term amounts to the sum of the evaluation for public good G based on marginal utility for disposable income R and the second term is the self-selection effect. The remaining part corresponds to the λ-equitability effect, which is different from that of Nava et al. (1996). This part consists of two effects. The first is the indirect effect, which is the inner product of the marginal utility of the low-skilled agent and the substitution effect of the compensated demand; this reflects the reduction of envy by discouraging consumption by the high-skilled agent because of the provision of the public good. The second is the direct effect, which reduces envy by decreasing the amount of the public good. On the right-hand side, the first term is the marginal cost of the public good and the second term is analogous to Nava et al. (1996), which means that the impact on indirect tax revenue increases the provision level through the compensated effects on consumption for the change in the level.

When can we apply the original Samuelson rule in this case? The third term on the left-hand side of equation (19) can be manipulated to yield

$$-\frac{\eta}{\gamma}\Big[\bar{U}_c\frac{\partial\lambda\tilde{c}_H}{\partial G}+\bar{U}_x\frac{\partial\lambda\tilde{x}_H}{\partial G}+\bar{U}_G\Big]=\frac{\lambda\eta}{\gamma}\bar{U}_c\Big[\frac{U_G^H}{U_c^H}-\frac{1}{\lambda}\frac{\bar{U}_G}{\bar{U}_c}\Big]+\frac{\lambda\eta}{\gamma}\frac{\partial\tilde{x}_H}{\partial G}\bar{U}_c\Big[\frac{U_x^H}{U_c^H}-\frac{\bar{U}_x}{\bar{U}_c}\Big]$$

$$\equiv\frac{\lambda\eta}{\gamma}\bar{U}_c\Big[MRS_{Gc}-\frac{1}{\lambda}M\bar{R}S_{Gc}\Big]+\frac{\lambda\eta}{\gamma}\frac{\partial\tilde{x}_H}{\partial G}\bar{U}_c\Big[MRS_{cx}-M\bar{R}S_{cx}\Big]. \tag{20}$$

Following the analysis above, if the agent's utility is expressed by $U(H(c_i,x_i,G),l_i)$, then the second term on the left-hand side of equation (19), which is the self-selection term, disappears. In addition, if function H meets the following functional form: $H(f(c_i,x_i),G)$, where $f(\cdot)$ is homothetic, the second term on the right-hand side of equation (20) must disappear, since $MRS_{cx} = M\bar{R}S_{cx}$ holds. At the same time, the second term on the right-hand side of equation (19) also disappears since t is zero, as shown above. Therefore, as in the analysis without linear commodity tax, whether to deviate from the original Samuelson rule depends on the first term on the right-hand side of equation (20).

As in subsection 2.3, we investigate the direction of the distortions when the utility function takes the CES form and has weak separability between labor and the other variables. Let the utility function be $H = (\alpha f(\cdot)^\rho + \beta G^\rho)^{\frac{1}{\rho}}$, where $\rho \leq 1$ and $f(\cdot)$ is homothetic. In this setting, the first term on the right-hand side of equation (20) can be rewritten as

$$MRS_{Gc}-\frac{1}{\lambda}M\bar{R}S_{Gc}=(1-\lambda^{-\rho})\frac{\beta G^{\rho-1}}{\alpha f(\cdot)^{\rho-1}f_c(\cdot)}.$$

Therefore, whether the original Samuelson condition is valid depends crucially on the elasticity of substitution. As with the result of Corollary 1, if the elasticity of substitution is above (below) one, the optimal provision rule is downwardly (upwardly) distorted, although the original Samuelson condition holds when the elasticity of substitution equals one.[13]

3.5 Conclusion

In this study, we analyze the optimal policy for income taxation with public good provision by a government concerned with the ethical constraint, namely the reduction of envy. As the new constraint, we use the λ-equitability by Diamantaras and Thomson (1989). To provide the public good, we then derive the optimal provision rule, as well as the marginal income tax rate in the optimal policy. Although the income tax part is the same as the results of Nishimura (2003a,b), the modified provision rule includes the effect of reducing envy, which is different from the modified Samuelson rule in Boadway and Keen (1993). To relax the ethical constraint, we adjust the amount of the provided public good to compare the evaluation of low-skilled agents with that at the referred commodity bundle. For instance, if an agent with the envied bundle places more weight on the public good than the low-skilled agent, he/she must decrease the provision level to use more tax income for redistribution. Furthermore, by using CES utility for the public good and private consumption, we show that if the elasticity of substitution is above (below) one, the original Samuelson condition is downwardly

[13] In general, if H is homogeneous of degree j in f on $H = H(f(c_i, x_i), G)$ and $f(\cdot)$ is homothetic under weak separability between labor and the other variables, the original Samuelson rule holds. The first term on the right-hand side of equation (20) can be rewritten as $\frac{\lambda\eta}{\gamma}\bar{U}_c\left(\frac{H_G(f(c_H,x_H),G)}{H_c(f(c_H,x_H),G)f_c(c_H,x_H)} - \frac{H_G(f(\lambda c_H,\lambda x_H),G)}{\lambda H_c(f(\lambda c_H,\lambda x_H),G)f_c(\lambda c_H,\lambda x_H)}\right) = \frac{\lambda\eta}{\gamma}\bar{U}_c\left(\frac{H_G(f(c_H,x_H),G)}{H_c(f(c_H,x_H),G)f_c(c_H,x_H)} - \frac{\lambda^{kj}H_G(f(c_H,x_H),G)}{\lambda\times\lambda^{kj-1}H_c(f(c_H,x_H),G)f_c(c_H,x_H)}\right) = 0$. Therefore, the λ-equitability term disappears and the original Samuelson rule is replicated.

(upwardly) distorted. However, the original rule is valid if the elasticity of substitution is one. As an extension, we add a taxable consumption good and linear commodity tax, and study both the optimal tax rate and the provision rule of the public good.

There are two policy implications from our model. First, when paying attention to the reduction of envy, the government must deal with the envied λ-scale bundle relative to the original bundle. Consequently, the government decreases the public good provision when the elasticity of substitution between private consumption and the public good is below one. In other words, a change in the ratio of their marginal utility is sensitive to variations in the ratio of these volumes. The second implication is that the public good provision increases much more or decreases much less as the intensity of envy increases. Since increasing the degree tightens the envy-free constraint, policymakers cannot use the other redistribution scheme; instead, they must reinforce the distorted direction of public good provision.

In Scandinavian countries, citizens place more weight on egalitarianism. Hence, governments or tax authorities should pay attention to the overall coverage of social security to reduce inequalities. They rely on the services provided by national systems, and the shares of social security are always relatively high in their budgets. According to OECD (2015), the 2015 Gini indexes of these countries (e.g., Sweden, Norway, Finland, and Denmark) where people place more weight on egalitarianism such as social justice were lower than the OECD average, while the government spending in 2015 on social protection in these countries was higher than the OECD average (see OECD (2017)). On the other hand, residents in the United States believe that most services must be accessible on the market and, thus, that the government's budget should decrease; they do not think it is necessary to make access to social security universal. In reality, President Trump has cut the budgets for social security and Medicare despite such fiscal policies breaking the major promises to his target voters (i.e., the poor).

Several tasks are left to future research. First, because we derive only the modified Samuelson rule, there is room to derive the provision level in general cases, divided into types of taxpayers' utility functions. Second, future studies could investigate the upper bound of intensity λ for the binding envy-free constraint. Intuitively, there are three regions for the degree of envy λ: nonbinding constraint, binding constraint, and violating constraint regions. Finally, related to the exogenous index λ, it would be interesting to conduct comparative statics of public good provision for λ analytically.

Appendix A

Assume that $\delta_{sH} > 0$ and $\delta_{sL} = 0$. Differentiating Lagrangian (4) with respect to c_L, c_H, l_L, l_H and G,

$$\frac{\partial \mathcal{L}}{\partial c_H} = (\gamma + \delta_{sH})U_c^H - \delta_r n_H - \delta_e \lambda \bar{U}_c = 0, \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial c_L} = (1 + \delta_e)U_c^L - \delta_r n_L - \delta_{sH} \hat{U}_c = 0, \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial l_H} = (\gamma + \delta_{sH})U_l^H + \delta_r n_H w_H - \delta_e \lambda \bar{U}_l = 0, \tag{A.3}$$

$$\frac{\partial \mathcal{L}}{\partial l_L} = (1 + \delta_e)U_l^L + \delta_r n_L w_L - \delta_{sH} \frac{w_L}{w_H} \hat{U}_l = 0, \tag{A.4}$$

$$\frac{\partial \mathcal{L}}{\partial G} = (\gamma + \delta_{sH})U_G^H + (1 + \delta_e)U_G^L - \delta_r \phi'(G) - \delta_e \bar{U}_G - \delta_{sH} \hat{U}_G = 0. \tag{A.5}$$

Rearranging (A.1) and (A.3) yields the optimal marginal income tax rate at the top. Conversely, we can derive the marginal income tax rate on the bottom by combining equations (A.2) and (A.4). The provision rule for the public good is obtained by substituting equations (A.1) and (A.2) into (A.5). □

Appendix B

Differentiating Lagrangian (20) with respect to p, R_L, R_H, and G,

$$\frac{\partial \mathcal{L}}{\partial p} = (1+\eta)V_p^L + (\mu+\delta)V_p^H - \delta\hat{V}_p - \eta\bar{V}_p + \gamma\sum_{i=H,L}[x_i^* + (p-1)\frac{\partial x_i^*}{\partial p}] = 0, \quad \text{(B.1)}$$

$$\frac{\partial \mathcal{L}}{\partial R_H} = (\mu+\delta)V_R^H - \eta\bar{V}_R - \gamma + \gamma(p-1)\frac{\partial x_H^*}{\partial R_H} = 0, \quad \text{(B.2)}$$

$$\frac{\partial \mathcal{L}}{\partial R_L} = (1+\eta)V_R^L - \delta\hat{V}_R - \gamma + \gamma(p-1)\frac{\partial x_L^*}{\partial R_L} = 0, \quad \text{(B.3)}$$

$$\frac{\partial \mathcal{L}}{\partial G} = (1+\eta)V_G^L + (\mu+\delta)V_G^H - \delta\hat{V}_G - \eta\bar{V}_G + \gamma\sum_{i=H,L}(p-1)\frac{\partial x_i^*}{\partial G} - \gamma\phi'(G) = 0. \quad \text{(B.4)}$$

Equations (B.1), (B.2), and (B.3) give

$$\frac{\partial \mathcal{L}}{\partial p} + \sum_i \frac{\partial \mathcal{L}}{\partial R_i}x_i^* = 0. \quad \text{(B.5)}$$

By using equations (7), (8), (10), and $\hat{x} = -\frac{\hat{V}_p}{\hat{V}_R}$, equation (B.5) can be transformed into equation (17). In addition, by substituting equations (B.2) and (B.3) into (B.4) and using equations (9) and (11), we can derive equation (19).

□

Table 3.1: Numerical examples under $\rho = 1$

	Low-skill utility	G
Case I		
Second best without the λ envy-free constraint	0.228045	0.499994
Case Ⅱ		
Second best with the λ envy-free constraint (λ=0.89)	0.225014	0.497525
Case Ⅲ		
Second best with the λ envy-free constraint (λ=0.91)	0.182929	0.489892
Case Ⅳ		
Second best with the λ envy-free constraint (λ=0.92)	0.0894965	0.450079

Table 3.2: Numerical examples under $\rho = -1$

	Low-skill utility	G
Case I		
Second best without the λ envy-free constraint	-0.406546	0.781823
Case Ⅱ		
Second best with the λ envy-free constraint (λ=0.785)	-0.406572	0.78252
Case Ⅲ		
Second best with the λ envy-free constraint (λ=0.79)	-0.428375	0.802776
Case Ⅳ		
Second best with the λ envy-free constraint (λ=0.792)	-0.481632	0.8229

Chapter 4

Envy-free pricing for excludable public good

4.1 Introduction

All over the world, income inequality becomes the biggest problem which leads to chaotic society unless reducing it, and the solution is redistributive policy by a government. There are several ways to redistribute collected incomes from rich ones to poor ones. For instance, the government levies taxes on workers' incomes, and transfers the wealth from rich to poor. Another is to provide public services which are useful for everyone but to which those with low income would have more limited access if the government did not provide them. With reducing complaints between members in society, the policymaker sets the optimal policy for income redistribution and implementing such public projects.

Why should policymakers take reduction of envy into consideration? There are two different reasons: widespread social justice like in Scandinavian countries and an element of social disorder which they should remove. For the first reason, all of Scandinavian countries like Sweden are welfare states, their social justice is egalitarianism. Recently, these social systems were partially reformed, but such societal norms remain ingrained in these countries. For the latter, as Bös and Tillmann (1985) noted:

> *the economic rationale for a minimization or reduction of envy by taxation is the following. Excessive envy in a society is an element of social disorder. Reducing envy in a society is a step towards increasing social harmony.* (p. 34)

However, this is not only the normative concept, but also an important issue which the whole world confronts. As seen in reality, Brexit or other electoral consequences like Donald Trump elected as the president of the United States reveal anti-globalism, and some of specialists claim that one of the reasons is to remove envy of the poor to the rich ones.[1]

In an economy where agents have different initial wealth, there are several ethical reasons to consider redistribution, and these are the background for mitigating inequality. One of them is envy. An agent envies the other agent if he prefers the other's commodity bundle to his own. We call envy-free allocation where there is no envy for every agent. It is difficult to apply the original envy-free constraint, but we replace the

[1] For example, according to World Economic Forum (2017), income gap is one of major sources bringing about polarized political outcomes.

weaker and cardinal criterion proposed by Diamantaras and Thomson (1989), called λ-equitability, and examine the optimal policy schedule under reduction of envy constraint.

In this paper, we analyze public good provision by public sector with surcharge fee and lump-sum transfer. Our analysis starts from Hellwig (2005b). In providing public good without exclusion and rivality, the government sets the optimal policy constraint on λ envy free as well as tax revenue. Before that, we check the differential transfer system depending on their income reported by themselves. However, we find that the only scheme for income transfer is uniform one constrained on information revelation. Under these cases, we examine the optimal provision rule of public goods with use exclusion and surcharge under the reduction of envy, assuming that individuals have additive and separable preferences and differ in both preferences for public goods and initial wealth or income. The consideration of λ envy free affects not only the amount of public goods but also the level of user fees. Additionally, we numerically simulate our model under Rawlsian government with the envy-free constraints and several intensities of envy, which support our analytical results and obtain new findings about comparative statics of the level of public provision about the index.

Related Literatures

We list papers related to this project categorized into optimal policy under reduction of envy and public good provision with use exclusion. With regard to optimal policy for reduction of envy, Nishimura (2003b) studies optimal nonlinear income taxation under constraints about reduction of envy, which shows that the marginal tax rate increases only if leisure is luxury. Also, Nishimura (2003a) examines optimal commodity taxation for reduction of envy. Both papers adopt particular envy-free notion suggested by Diamantaras and Thomson (1989), and we follow this manner, but our paper focus on public good provision and user fee by the government, and do not include endogenous labor supply but exogenous different incomes. Tsugawa and Obara (2017) incorporates endogenous labor supply like his two works, and we take pure public good provision by the policymaker into consideration as our novelty. In that paper, we show that such envy free constraint distorts the provision level, and find that the intensity of envy affects the amount of provision by numerical simulation. Lloret-Batlle and Jayakrishnan (2016) and Lloret-Batlle and Jayakrishnan (2017) study optimal pricing scheme of traffic system which addresses fairness by employing envy-free constraint.

As to public good provision with use exclusion, Hellwig (2005b) allows a policymaker to exclude agents who value a public good less than the surcharge set by her, which shows that a utilitarian government sets zero surcharge at optimum since revenues from increasing the surcharge is dominated by social welfare for reducing the fee, and that the revenue effect becomes stronger as the government is more risk-averse since it is better to utilize more surcharge fees for redistribution. Norman (2004) also examines excludable public good provision as mechanism design problems, but the difference from Hellwig (2005b) is that the government assigns the access right to each agent based on his reported preference while this paper allows her to set only the additional fee and the taxpayers to decide whether to make use of her services.[2]

[2]As well as those papers, for instance, Hellwig (2010) and Fang and Norman (2010) study excludable public good provision under asymmetric information as mechanism design problem.

This remainder of this paper is organized as follows. Section 2 analyzes the optimal provision rule for pure public goods and the optimal pricing rule for user fees under the reduction of envy, and the next section provides the numerical simulation. Section 4 offers concluding remarks.

4.2 The model

We consider an two-class economy in which each agent ($i = H, L$) possesses initial wealth Y_i, where $Y_H > Y_L > 0$. Without loss of generality, the population of each agent is equal to one, and for each group, they have tastes for public good θ which is distributed over $[\underline{\theta}, \overline{\theta}]$ with cumulative distributive function F where $0 = \underline{\theta} < \overline{\theta} < \infty$. The incomes are private information for agents, and the government collects taxes on their income nonlinearly. In addition, she provides public good by collected taxes and surcharges their usage in order to compensate the public expenditure, so agents who face the admission fee can decide whether to make access to the public good. At first, we assume three kinds of goods, consumption (or after-tax income) $c \in \mathbb{R}_+$, labor supply l and public good $G \in \mathbb{R}_+$. Every agent shares the identical utility function $U(c_i, G)$, and we restrict attention to that $U(c_i, G) = \theta G + c_i$.

$$U(c_i, G) = \theta G + c_i^j \tag{1}$$

In providing public good, the government must incur its production cost $\phi(G)$ with an increasing, strictly convex, and differentiable function. For all goods except for public good, a good with subscript i means the one which agent i enjoys.

Following Hellwig (2005b), we consider the maximization of the social welfare function, using Bergson-Samuelson social welfare function as her objective. It is obvious that she faces the resource constraint. Let $T : \mathbb{R} \to \mathbb{R}$ be the income tax function, and agent i's budget constraint is written as $c_i = Y_i - T(Y_i)$. So, the government's income tax revenue is

$$T(Y_L) + T(Y_H) = Y_L - c_L + Y_H - c_H. \tag{2}$$

We allow the government to provide public good with use-exclusion. Put it differently, it can impose user fees on those who enjoy public goods. Indicator B represents individuals who obtain the benefits from a public good and indicator NB individuals who are excluded from a public good. With those indicators and (1), their utility U_i^j is described by

$$U_i^j = \mathbf{1}(j) \cdot \theta G + c_i^j. \tag{3}$$

where $i = H, L \; and \; j = B, NB$ and $\mathbf{1}(j)$ is the characteristic function as follows:

$$\mathbf{1}(j) = \begin{cases} 1 & \text{if } i = B \\ \\ 0 & \text{if } i = NB \end{cases} .$$

The assumption on the utility function is useful to avoid the multidimensional heterogeneity problem. Also, the government imposes admission fees p on those who access to a public good with wage rate. Private consumption that type i individuals enjoy is given by $c_i^j = Y_i - \mathbf{1}(i) \cdot p - T(Y_i)$.

4.2.1 Extensive Margin

Individuals decide whether or not to get access to a public good. Individuals with type (θ, Y_i) obtain utility $\theta G + c_i^B$ if they have access to a public good, and utility c_i^{NB} if they are excluded. Therefore, they choose access to a public good if and only if

$$\theta \geq \frac{c_i^{NB} - c_i^{B}}{G} = \frac{p}{G} \equiv \hat{\theta} \tag{4}$$

where $\hat{\theta}$ is interpreted as the net gain from being excluded from a public good. We derive the equality in equation (27) using individual's budget constraint. Equation (27) means that, if public goods preferences of individuals are greater (lower) than the threshold $\hat{\theta}$, they (do not) access to a public good. Moreover, it is rewritten as:

$$p = \hat{\theta} G \tag{5}$$

That is, type i individuals prefer to make use of public good paying admission fees p if the benefit θG that they draw from the enjoyment of a public good exceeds p.

4.2.2 The Government

The budget constraint of the government takes the following form:

$$\begin{aligned} & 2\int_{\hat{\theta}}^{\bar{\theta}} p f(\theta) d\theta + \sum_i T(Y_i) = \phi(G) \\ \Leftrightarrow \quad & 2(1 - F(\hat{\theta}))\hat{\theta} G + \sum_i [Y_i - c_i^{NB}] = \phi(G) \end{aligned} \tag{6}$$

The first term is the aggregate revenue from admission fees. The second term represents the aggregate revenue from income taxes. She produces the public good, using these revenues. Following Hellwig (2005b), the social welfare function used in our analysis is Bergson-Samuelson criterion which is represented as follows:

$$
\begin{aligned}
\mathcal{W} \equiv & \int_{\hat{\theta}}^{\bar{\theta}} W(\theta G + c_H^B) f(\theta) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W(c_H^{NB}) f(\theta) d\theta \\
& + \int_{\hat{\theta}}^{\bar{\theta}} W(\theta G + c_L^B) f(\theta) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W(c_L^{NB}) f(\theta) d\theta \\
& = \int_{\hat{\theta}}^{\bar{\theta}} W(\theta G - \hat{\theta} G + c_H^{NB}) f(\theta) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W(c_H^{NB}) f(\theta) d\theta \\
& + \int_{\hat{\theta}}^{\bar{\theta}} W(\theta G - \hat{\theta} G + c_L^{NB}) f(\theta) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W(c_L^{NB}) f(\theta) d\theta
\end{aligned} \tag{7}
$$

where W is an increasing and concave function, that is, $W' \geq 0$ and $W'' \leq 0$.

In the second best environment, the government cannot observe initial wealth which are private information of individuals. By revelation principle, it suffices to induce individuals reveal their true types for initial endowment to maximize the objective of the government. Since the government can observe whether or not to access to a public good, they cannot mimic ones in the other group. Therefore, we consider only the incentive constraint within each group, and have only to check the followings:

$$
Y_H - T(Y_H) \geq Y_H - T(Y_L) \quad \text{and} \quad Y_L - T(Y_L) \geq Y_L - T(Y_H). \tag{8}
$$

Obviously, the government cannot set differential income taxes dependent on their revelation, in other words, $T(Y_H) = T(Y_L) = T$. Also, the government takes envy with respect to individuals' utilities into consideration. Introducing the concept of Diamantaras and Thomson (1989), the government must implement an allocation satisfying the following λ envy free constraint within each group[3]:

$$
Y_L - T \geq \lambda(Y_H - T) \tag{9}
$$

[3]Of course, we must satisfy λ envy-free constraint for high-class, but for any λ, the envy-free constraint for type H is satisfied whenever that for type L is satisfied.

for NB group, and

$$Y_L - p - T \geq \lambda(Y_H - p - T) \tag{10}$$

for B group. The inequality (9) holds if (10) is true, so it is enough to check inequality (10) for this ethical requirement. Furthermore, rearranging (10),

$$Y_L - \lambda Y_H \geq (1 - \lambda)(T + p).$$

If $\lambda > \frac{Y_L}{Y_H}$, (10) cannot hold vacuously since the right-hand side is non-negative, so we assume that $\lambda \leq \frac{Y_L}{Y_H}$.

To sum up, the government chooses the policy $\{T, \hat{\theta}, G\}$ to maximize the social welfare function (7) subject to the government's budget constraint (6), and the λ envy-free constraint (10). The corresponding Lagrangian is

$$\begin{aligned} \mathcal{L}(T, G, \hat{\theta}; \gamma, \eta) = &\mathcal{W} \\ &+\gamma\{2(1 - F(\hat{\theta}))\hat{\theta}G + 2T - \phi(G)\} \\ &+\eta\{Y_L - \lambda Y_H - (1 - \lambda)(T + p)\} \end{aligned} \tag{11}$$

where γ and η are Lagrangian multipliers corresponding to constraints individually. Using the first-order conditions with respect to the Lagrangian, we characterize the optimal provision rule for public goods for the reduction of envy.

Proposition 6. *Under nonlinear income tax, the optimal provision rule taking the reduction of envy into account is characterized by:*

$$\frac{\sum_{i=L,H} \int_{\hat{\theta}}^{\overline{\theta}} (\theta - \hat{\theta}) W'(\theta G - \hat{\theta} G + Y_i - T) f(\theta) d\theta}{\gamma} + 2(1 - F(\hat{\theta}))\hat{\theta} + \frac{\eta}{\gamma}\hat{\theta}(\lambda - 1) = \phi'(G) \tag{12}$$

where

$$\gamma = \frac{\sum_{i=H,L}\left[\int_{\hat{\theta}}^{\overline{\theta}} W'(\theta G - \hat{\theta} G + Y_i - T) f(\theta) d\theta + \int_{\underline{\theta}}^{\hat{\theta}} W'(Y_i - T) f(\theta) d\theta\right]}{2} + \frac{\eta(1 - \lambda)}{2} \tag{13}$$

The provision rule (12) is analogous to that derived in Tsugawa and Obara (2017) except for use exclusion. in the left hand, the first term is the sum of marginal rate of substitution between income and public good, the second term side represents the marginal benefit due to the increase in revenue from user fees, and the third term expresses the marginal loss caused by the argument of envy owing to the increase of user fees. These terms distort the level of public goods upwardly or downwardly, and the effect depends on the level of user fees. Therefore, we examine the level of user fees. Combining (A.1) with (A.2), the formula is characterized by:

$$\underbrace{2\hat{\theta}Gf(\hat{\theta}) + \frac{\eta}{\gamma}G(1-\lambda)F(\hat{\theta})}_{efficiency\ loss} = \underbrace{G(1-F(\hat{\theta}))\sum_i g^i_{NB} - GF(\hat{\theta})\sum_i g^i_B}_{equity\ gain} \tag{14}$$

where, $g^i_{NB} \equiv \frac{\int_{\underline{\theta}}^{\hat{\theta}} W'(Y_i - T)f(\theta)d\theta}{\gamma}$ and $g^i_B \equiv \frac{\int_{\hat{\theta}}^{\bar{\theta}} W'(\theta G - \hat{\theta}G + Y_i - T)f(\theta)d\theta}{\gamma}$. These are the marginal social welfare weight for individuals with wealth level i in group j, and measure the relative value of the government that gives an additional 1\$ to type ij individuals.

Equation (14) implies an *equity − efficiency tradeoff*. A small reform, such that p increases, distorts the decision making on the extensive margin. That is, individuals with lower preferences for a public good tend to hope the exclusion of public goods. Therefore, revenues from user fees decrease, which is expressed by the first term in the left hand side. On the other hand, the right hand side is the net welfare gains from the redistribution between groups, and the first and second terms describe the government's redistributive tastes for each group. If the government prefers to redistribute from group B to NB, that is, the first term in the right-hand side is greater than the second term, user fess are charged to increase revenues and raise consumption levels.

From equation (14), the equity gain crucially depends on the assumption on the social welfare function. We suggest two polar cases: utilitarian and Rawlsian. If the social welfare function is utilitarian, the equity gain is zero because the government is not interested in the redistribution. This means that efficiency loss is always larger than equity gain if the government imposes user fees. Therefore, $p = 0$ is optimal. On the other hand, when the objective of the government is Rawlsian, the equity gain which is equal to $G(1-F(\hat{\theta}))$ remains because the aim is to improve the utility of type-L agents.[4] That is, the level of p is determined by the comparison between the loss and the gain. However, not only interior solution but also a trivial case ($p = 0$ and $G = 0$) satisfies equation (14). In the section, we focus only on the analytical result under the interior solution, and compare the welfare level under the interior solution with that under the trivial case in the numerical section.

On the occasion, we characterize the optimal provision rule for public goods under utilitarian and Rawlsian as follows:

$$\mathbb{E}(\theta) + \frac{\eta}{\gamma}\lambda\left(\mathbb{E}(\theta) - \frac{\mathbb{E}(\theta)}{\lambda}\right) = \phi'(G) \tag{15}$$

$$(1 - F(\hat{\theta}))\hat{\theta} + \frac{\eta}{\gamma}\hat{\theta}(\lambda - 1) = \phi'(G) \tag{16}$$

where, $\mathbb{E}(\theta) \equiv \int_0^{\bar{\theta}} \theta f(\theta) d\theta$. Equation (15) is the optimal provision rule under utilitarian and the result suggests that under-provision is optimal, which is consistent with Proposition 1. On the other hand, equation (16) is that under Rawlsian. The envy term distorts downwardly the level of public goods.

[4] The objective of the Rawlsian government is to maximize the worst-off individuals' utility, i.e., $Y_L - T$.

	Social welfare (= x_L^{NB})	$\hat{\theta}$	G	x_H^{NB}
Case I				
$\lambda = .3$	0.605106	0.962194	0.925822	1.60511
Case Ⅱ				
$\lambda = .4$	0.661546	0.939104	0.881917	1.66155
Case Ⅲ				
$\lambda = .5$	0.737788	0.903638	0.816464	1.73779

Table 4.1: Simulation results under Rawlsian

4.3 Numerical examples

To assess our theoretical results in terms of the provision of excludable public goods, we present numerical example. The objective is to clarify the effect of λ, the degree of envy, to the amount of public goods and the level of the user fee under Rawlsian social welfare function.

In the simulation, we set the following assumptions. First, we assume that public goods preferences θ are distributed as the uniform function $F(\theta) = \theta/\overline{\theta}$ on the interval $[0, \overline{\theta} = 1]$. Second, individuals are categorized into $Y_L = 1$ and $Y_H = 2$. Third, the cost function for providing public good is denoted by G^2. Finally, we assume that $\pi_H = \pi_L = 0.5$ and $\lambda = 0.3, 0.4$ and 0.5.

We can check those results in Table 1. According to that, as λ increases, both user fee and the level of public provision decrease. In addition, the threshold of whether individuals utilize public goods or not decreases when the intensity of envy λ increases.

4.4 Conclusion

In this paper, we analyze optimal public good provision by a government when she is concerned with ethical constraint, reduction of envy, and able to utilize both lump-sum transfer and user fee as a way of exclusion. As the new constraint, we adopt λ envy free constraint borrowed from Diamantaras and Thomson (1989), and derive the optimal provision rule with user fee when the government has Bergson-Samuelson social welfare function under constraints on reduction of envy for agents with low income. Also, we conduct numerical simulation for optimal policy by Rawlsian government with different levels of the degree of envy λ. Our model employs the similar technique of introducing use exclusion to Hellwig (2005b), so we do not consider endogenous labour incomes, but heterogeneous initial wealth, which is different from Hellwig (2005b) as well as that ethical constraint.

The level of provision is determined by trade-off between efficiency loss and equity gain. Note that this is parallel to Hellwig (2005b) except for the nouveau part coming from λ envy free constraint. As the user fee increases, the efficiency loss also increases or λ-equitability constraint becomes tighter. In addition to these policies in general, we examine the two extreme cases: utilitarian or Rawlsian policymaker. In utilitarian case, the optimal surcharge must be 0 since the equity gain also equals 0 while public good is always provided. On the other hand, in Rawlsian case, there are two possible cases: the provision is implemented with positive user fee or no provision. Analytically, we cannot find which is better, but we simulate the proficiency numerically. According to our simulation result, Rawlsian government provides public good with user fee excluding agents who evaluate it less. Additionally, the level of public provision and user fee also decreases as the intensity of envy increases. Such results are parallel to those in numerical simulation part of our other work Tsugawa and Obara (2017).

In the end, there are two policy implications in our model. First of all, in paying attention to reduction of envy, the government must deal with the envied λ-scale bundle relative to the original bundle. Second, in Rawlsian policymaker's mitigating envy, she utilizes use exclusion. Finally, we think It interesting to remove the implicit assumption that λ-equitability binds. As λ sufficiently small, this constraint will not bind, so studying the problem as well as the comparative statics on λ may be good for future works.

Appendix

$$\frac{\partial \mathcal{L}}{\partial T} = -\sum_{i=H,L}\Big[\int_{\hat{\theta}}^{\bar{\theta}} W'(\theta G-\hat{\theta}G+Y_i-T)f(\theta)d\theta+\int_{\underline{\theta}}^{\hat{\theta}} W'(Y_i-T)f(\theta)d\theta\Big]+2\gamma+\eta(\lambda-1) = 0 \tag{A.1}$$

$$\frac{\partial \mathcal{L}}{\partial \hat{\theta}} = -G\sum_{i=L,H}\int_{\hat{\theta}}^{\bar{\theta}} W'(\theta G-\hat{\theta}G+Y_i-T)f(\theta)d\theta+2\gamma G[(1-F(\hat{\theta}))-\hat{\theta}f(\hat{\theta})]+G\eta(\lambda-1) = 0 \tag{A.2}$$

$$\frac{\partial \mathcal{L}}{\partial G} = \sum_{i=L,H}\int_{\hat{\theta}}^{\bar{\theta}}(\theta-\hat{\theta})W'(\theta G-\hat{\theta}G+Y_i-T)f(\theta)d\theta+\gamma 2(1-F(\hat{\theta}))\hat{\theta}-\gamma\phi'(G)+\eta\hat{\theta}(\lambda-1) = 0 \tag{A.3}$$

By rearranging (A.1) and (A.3), we obtain (13) and (12) respectively.

Chapter 5

Income tax competition with endogenous wages

5.1 Introduction

In the context of optimal income taxation, a government levies income taxes to earners with high ability more progressively than those with the low, and the reason is that she collects salaries from richer taxpayers and provides those poorer with collected taxes. However, in the literatures, we assume that all citizens cannot emigrate from the state or country because of their complaints about her tax policies whereas, in reality, many high skilled workers are demanded in their fields all over the world, and have chances to work in other places. Even though there are several obstacles to migrate from home to the other like language skills, the moving costs have gradually but steadily decreased for a long time by technology progress and internationalization.

From the above reasons, we should consider the optimal fiscal policy under the threat of migration as tax competitions by competing governments to attract more taxpayers from outside as well as to keep their citizens. We must care for the change of not only the number of populations paying taxes to an administration, but also wages for their labor supply there. If two governments compete with each other, and one of the two attracts workers by reducing their tax burden, the population must increase and at the same time, their wages vary responding to the overall fraction of labor supply or demand. Since the marginal tax rate at optimal income tax partially depends on the workers' wage, it's more useful to consider the setup of competitive optimal taxation with their unit wage determined endogenously. Also, in reality, such demographic changes and the structure of wages interact, and both of those have causal connection to fiscal policies. On the other hand, past works ignore the endogeneity, but it is natural extension to incorporate endogenous wages and its general equilibrium effect into tax competition model.

To sum up, this paper sheds light on two aspects which existing works ignore; tax competition in the framework of optimal nonlinear income tax and wages determined by the overall productivity. In order to compare these two effects on tax policies, we study the four cases: non-competing governments with or without endogenous wages, competition by two identical governments with fixed wages and endogenous ones. Except for the last, we review previous works in step with our novel setting, mobility of workers and endogenous wage. In our new setup, we show that the optimal income tax schedule embraces not only migration effect, but also trickle-down effect coming from endogenous wage, and the marginal tax rate depends on these two complementarily. In the tax rate for high-skilled workers in our new setup, the impact of emigration never disappears whereas that does in fixed wage cases. Why such difference arises is that such terms are built into production technology as the part of populations even though they vanish at symmetric equilibria.

Related literature

Much attention has been devoted to the analysis of optimal income taxation with mobile labor under strategic competitions between governments. For instance, Piaser (2007) studies nonlinear income tax competition by two benevolent governments in the presence of labor mobility, but in a model with only two skills. In contrast to contributions about tax competition like Wilson (1992) and Simula and Trannoy (2010)[1], Piaser (2007) adopts common agency model under asymmetric information based on competitive nonlinear pricing model with random participation like Rochet and Stole (2002) to study nonlinear income tax competition by competing two governments. Our model is relevant to this, but the main difference is that we take into consideration that each worker's unit wage is endogenously determined by the overall productivity, and the proportion depends on the ratio of populations in each class. Morelli et al. (2012b) also studies nonlinear income tax competition by competing two tax authorities, and they consider three classes of workers. Their main question is whether each class is

[1]With Simula and Trannoy (2012), they analyze income tax competition and its migration effect on marginal tax rate in the context of optimal income tax model with participation constraints for workers. Krause (2017) also examine a simple dynamic nonlinear income tax model without commitment in this framework.

in favor of unified tax authority or competitive tax authority, and they show that the highest class always prefers the competitive one while the lowest class always prefers the unified one. In addition, they show that the preferences of the middle class depend on the initial conditions in terms of the distribution of abilities, the relative power of the various classes, and mobility costs.

Lehmann et al. (2014) also studies nonlinear income tax competition by competing two governments strategically. Compared to the above literatures, they consider more general setting: continuous skill distribution and asymmetric two authorities in terms of the populations. They derive the optimal marginal income tax rates at the equilibrium, extending the Diamond (1998) and Saez (2001) formula and show that the level and the slope of the semi elasticity of migration are crucial to derive the shape of optimal marginal income tax analytically. Following the framework of competitive nonlinear pricing with random participation, they solve the problem by tax perturbation approach.

Bierbrauer et al. (2013) investigates strategic tax competition problem, but differently assume that all workers are perfectly mobile and that governments maximize the average utility of their residents while some of the papers assume those maximizing their utility with the lowest skill. Their results also support race-to-the-bottom thesis, which means that at the equilibrium, the highest skilled workers receive the benefits funded by taxes on lower skilled ones. We also refer Lipatov and Weichenrieder (2015) and Tóbiás (2016) as recent developments in nonlinear income tax competition by strategically competing tax authorities. Lipatov and Weichenrieder (2015) analyzes subgame-perfect Nash equilibrium of the game between two governments and two taxpayer populations, and Tóbiás (2016) considers more than two competing governments and obtains the converse result of race-to-the-bottom result.

Likewise, my model stands in the line of optimal income taxation with wages endogenously determined by the proportion of workers. Stiglitz (1982) studies Pareto efficient income taxation with endogenous wage, and derives the marginal tax rate at optimum. Sachs et al. (2016) generalizes the wage distribution into the continuous one, and studies the incidence and the optimal design of nonlinear income taxes in a economy of Mirrlees (1971). Micheletto (2004) sheds light on mixed taxation in this framework, and Gaube (2005b) shows that production efficiency is violated in the optimum with both non-linear and linear income taxation if and only if a distortionary tax schedule is implemented.

In what follows, next section introduces basic framework of the study, and Section 3 addresses the optimal tax schedule without endogenous wage and mobility of workers as benchmark. Section 4 and 5 add the structures of the mobility and wages determined by overall productivity independently, and Section 6 derives the optimal tax schedule with both mobile workers and endogenous wages.

5.2 The model

In this model, we consider an economy consisting of two states, indexed by $i = 1, 2$. Taxpayers live in the either state i, and each state has one population of taxpayers. Let p_i be the proportion of taxpayers with high ability for state i. In addition, they have locational preferences that are uniformly distributed between $[0, 1]$, and the distribution is independent from ability. Taxpayers who leave from their own state to the other pay

k per unit, so if they are located at m, and move to the other state, they must pay $(1-m)k$ as their cost of moving.

Taxpayers consume numeraire b and incur disutility from labor supply l. Assume that their utility are separable between consumption and labor supply, so they have the utility $U(b,l) = u(b) - l$ where u is strictly increasing and strictly concave in b. They earn their income from labor supply, and each state government can levy taxes only on their labor income. Let y be the labor income, the product of ability and labor supply; for instance, if they earn y with ability w, their labor supply is $\frac{y}{w}$. With their productivity (or unit wage) w and before-tax income y, their utility can be rewritten as

$$U(b,y,w) = u(b) - \frac{y}{w}.$$

In our model, it is assumed that there are two types of their productivities: w_H and w_L with $w_H > w_L$.

Tax schedule and timing of tax competition game

Each state government i sets nonlinear income tax $T_i(y)$. We assume that each state cannot discriminate the tax burden between their native state. If no taxpayer moves from their own state, and comes from the other state, government i faces the following budget constraint;

$$p_i\{\underbrace{y_i(w_H) - B_i(w_H)}_{=T_i(y_i(w_H))}\} + (1-p_i)\{\underbrace{y_i(w_L) - B_i(w_L)}_{=T_i(y_i(w_L))}\} \geq 0 \tag{1}$$

where $y_i(w_a)$ is the labor income for taxpayers with ability a and $B_i(w_a)$ is the labor income for taxpayers with ability a for $a = H, L$.

We introduce the tax competition game between two states. It is assumed that all taxpayers can move, so once they decide which state to live in, they also decide the state where they pay all taxes. First of all, in the environment of competitive mechanism design, it is not without loss of generality to restrict attention to direct contracts[2]. Hereafter, for simplicity, we assume that the two governments are symmetric or identical, so $p_1 = p_2 = p > 0$.

The timing of the tax competition game is as follows: in the first stage, each state i chooses its menu of contracts $T_i(\cdot)$. We assume that the two states have no asymmetric information, so each tax authority knows the distribution of workers and that of its commuting cost. Next, each citizen decides one state where he lives given these contracts. Citizens who move from their own state incur moving costs, so they maximize their utilities given the expense. Applying revelation principle in this setting, each government equivalently decides the menu of allocations: $\{Y_i(w), B_i(w)\}_w$. Therefore, we can summarize the timeline as follows:

1. Each state chooses its menu $\{Y_i(w), B_i(w)\}_w$,
2. Given these contracts, each citizen decides the location where he lives,
3. Allocations are determined according to the menu of contracts announced at time 1.

In order to describe the migration decision, we denote citizen's rent in their home with w_a by v_a, and her rent in emigrating and staying at the other state by v_a^*. The

[2]According to Martimort and Stole (2002), revelation principle can be applied as long as we consider only pure strategy equilibria.

proportion m of taxpayers in one state stay at home where

$$m \leq \min\{1, 1 + \frac{v_a - v_a^*}{k}\}.$$

It is obvious for taxpayers with w_a to stay at her home state if $v_a \geq v_a^*$. Otherwise, the citizens may move to another state, and for ability $a = H, L$, the proportion of paying taxes to state i is

$$m_i(w_a) = 1 + \frac{v_a - v_a^*}{k} \tag{2}$$

because m is uniformly distributed along Hoteling line $[0, 1]$. Including their migration decision, each authority i faces the modified budget constraint:

$$pm_i(w_H)(Y_i(w_H) - B_i(w_H)) + (1 - p)m_i(w_L)(Y_i(w_L) - B_i(w_L)) \geq 0 \tag{3}$$

where $m_i(w_a)$ is the fraction of taxpayers with w_a who pay taxes to state i. Also, the government extracts true information from the taxpayers, and because of single-crossing property with respect to w, he/she only sticks to downward incentive constraint:

$$u(B_i(w_H)) - \frac{Y_i(w_H)}{w_H} \geq u(B_i(w_L)) - \frac{Y_i(w_L)}{w_H}. \tag{4}$$

With or without strategic tax competition, each government can offer the above incentive constrained contract only. Given tax schedule T_1 and T_2, taxpayer with ability w_a , working at i, solves the following problem about before-tax income

$$\max_y \ u(y - T_i(y)) - \frac{y}{w_a}.$$

So, the marginal tax rate at y is determined by

$$(1 - T_i'(y)) \times \underbrace{u'(y - T_i(y))}_{=u'(B_i(w_a))} = \frac{1}{w_a}$$

$$\Leftrightarrow (1 - T_i'(y)) = \frac{1/w_a}{u'(B_i(w_a))} \equiv MRS_{yb}$$

where MRS_{yb} is the marginal rate of substitution between before-tax income y and numeraire or pre-tax income b.

5.3 Benchmark 1: optimal tax schedule under exogenous wage and without labor mobility

In this section, we derive the optimal income tax schedule under exogenous wage, but we do not allow all taxpayers to move from their home state. State i maximizes the taxpayers' utility for low ability: v_L subject to the resource constraint (1) and downward incentive constraint (4). So, we solve the following Lagrangian:

$$\begin{aligned} & u(B_i(w_L)) - \frac{Y_i(w_L)}{w_L} \\ & +\gamma\Big[p(Y_i(w_H) - B_i(w_H)) + (1-p)(Y_i(w_L) - B_i(w_L))\Big] \\ & +\lambda\Big[u(B_i(w_H)) - \frac{Y_i(w_H)}{w_H} - u(B_i(w_L)) + \frac{Y_i(w_L)}{w_H}\Big] \end{aligned} \tag{5}$$

where γ and λ are the multipliers of constraints (1) and (4) respectively. Differentiating it with respect to $Y_i(w_H)$, $Y_i(w_L)$, $B_i(w_H)$ and $B_i(w_L)$, we can obtain the FOCs:

$$B_i(w_H) : \lambda u'(B_i(w_H)) - \gamma p = 0$$

$$B_i(w_L) : (1-\lambda)u'(B_i(w_L)) - \gamma(1-p) = 0$$

$$Y_i(w_H) : -\lambda\frac{1}{w_H} + \gamma p = 0$$

$$Y_i(w_L) : -\frac{1}{w_L} + \lambda\frac{1}{w_H} + \gamma(1-p) = 0$$

Thus, we derive the following properties of allocation $\{B_i(w_a), Y_i(w_a)\}_{a=H,L}$ as the following proposition:

Lemma 2. *If taxpayers have no mobility, and fixed unit wage, and the government has Rawlsian preference, then the allocation* $\{B_i(w_a), Y_i(w_a)\}_{a=H,L}$ *satisfies*

- $u'(B_i(w_H)) = \frac{1}{w_H} \Leftrightarrow T_i'(Y_i(w_H)) = 0$*: no distortion on top,*
- $u'(B_i(w_L)) = \frac{\gamma(1-p)}{1-\lambda} = \frac{1}{1-\lambda}(\frac{1}{w_L} - \lambda\frac{1}{w_H}) > \frac{1}{w_L} \Leftrightarrow T_i'(Y_i(w_L)) > 0$ *: disposable income is distorted from the first best.*
- $B_i(w_H) > B_i(w_L)$*: disposable income for* w_H *is higher than that for* w_L*.*

This result is analogous to that in Stiglitz (1982)[3]. In order to find the degree of distortion, we derive the multipliers from the above FOCs as follows:

$$\gamma = \frac{1}{w_L} \quad \text{and} \quad \lambda = p \times \frac{w_H}{w_L}.$$

By this,

$$u'(B_i(w_L)) = \frac{1}{w_L} \times \frac{1-p}{1-p\frac{w_H}{w_L}} \equiv \frac{1}{w_L} \times \frac{1-p}{1-\mathcal{F}p} \quad \text{and} \quad T_i'(Y_i(w_L)) = \frac{p(\mathcal{F}-1)}{1-p}$$

where $\mathcal{F}$ is the ratio of unit wages $\frac{w_H}{w_L} > 1$. So, the marginal tax rate at $Y_i(w_L)$ must increase as $\mathcal{F}$ increases. In addition, the tax burden for taxpayers with high ability is larger than that for those with low ability.

Lemma 3. *If taxpayers have no mobility, and fixed unit wage, and the government has Rawlsian preference, then*

- *the marginal tax rate* $T_i'(Y_i(w_L))$ *increases as the wage ratio* $\mathcal{F}$ *increases.*

[3]Here, we only take downward incentive constraint into consideration. Of course, we may incorporate upward incentive constraint, but it is sufficient to focus only on downward incentive constraint because the government has Rawlsian social preference. Stiglitz (1982) discusses about incentive constraints in detail.

- $T^i_H > T^i_L$ *where* $T^i_a = Y_i(w_a) - B_i(w_a)$.

for $a = H, L$ *and* $i = 1, 2$

Proof. It remains to show the last part. Suppose, on the contrary, that $T^i_H \leq T^i_L$. Since the incentive constraint binds,

$$u(B_i(w_H)) - u(B_i(w_L)) = \frac{1}{w_H}(Y_i(w_H) - Y_i(w_L)) \leq \frac{1}{w_H}(B_i(w_H) - B_i(w_L)).$$
$$\Leftrightarrow u(B_i(w_H)) - \frac{B_i(w_H)}{w_H} \leq u(B_i(w_L)) - \frac{B_i(w_L)}{w_H}.$$

However, this contradicts the fact that $B_i(w_H) = \arg\max_B\{u(B) - \frac{B}{w_H}\}$ and $B_i(w_H) > B_i(w_L)$ (from Lemma 1). Hence, $T^i_H > T^i_L$. $\square$

Also, if $T^i_H < 0$, the government budget constraint is violated, so we must have $T^i_L < 0 < T^i_H$. When taxpayers are not allowed to move from their home country to the other, the government levies taxes on rich ones, and gives subsidies to poor ones.

5.4 Benchmark 2: optimal tax schedule under exogenous wage and with labor mobility

Next, we check the optimal tax schedule with their labor mobility under exogenous wage. In this case, we stick to the conditions that $v_a \leq v^*_a$ for any tax authority and wage level a. Also, since we cling to symmetric equilibria in tax competition games, we leave out the index of tax authority i.

The government solves the following maximization problem:

$$\max_{\{B(w_a),Y(w_a)\}_{a=H,L}} u(B(w_L)) - \frac{Y(w_L)}{w_L} \quad \text{subject to}$$

$$pm(w_H)(Y(w_H) - B(w_H)) + (1-p)m(w_L)(Y(w_L) - B(w_L)) \geq 0 \quad \text{and} \tag{6}$$

$$u(B(w_H)) - \frac{Y(w_H)}{w_H} - u(B(w_L)) + \frac{Y(w_L)}{w_H} \geq 0.$$

Note that the derivatives of fraction of which populations pay taxes to that tax authority with respect to disposable income $B(w_j)$ and before-tax income $Y(w_j)$ $(j = L, H)$ are

$$\frac{\partial m(w_j)}{\partial B(w_j)} = \frac{u'(B(w_j))}{k}, \quad \text{and}$$

$$\frac{\partial m(w_j)}{\partial Y(w_j)} = -\frac{1}{w_j k}.$$

The FOCs are:

$$B(w_H) : \lambda u'(B(w_H)) - \gamma p\{m(w_H) - \frac{u'(B(w_H))}{k}(Y(w_H) - B(w_H))\} = 0$$

$$B(w_L) : (1-\lambda)u'(B(w_L)) - \gamma(1-p)\{m(w_L) - \frac{u'(B(w_L))}{k}(Y(w_L) - B(w_L))\} = 0$$

$$Y(w_H) : -\lambda\frac{1}{w_H} + \gamma p\{m(w_H) - \frac{1}{w_H k}(Y(w_H) - B(w_H))\} = 0$$

$$Y(w_L) : -\frac{1}{w_L} + \lambda\frac{1}{w_H} + \gamma(1-p)\{m(w_L) - \frac{1}{w_L k}(Y(w_L) - B(w_L))\} = 0$$

Again, since these governments have the same proportion of skilled and unskilled taxpayers, and they have no tax revenue requirement, it's enough to restrict attention on symmetric equilibrium outcome. Let $\{B(w_a), Y(w_a)\}_{a=H,L} = \{B_a^m, Y_a^m\}_{a=H,L}$ and $v_a = v_a^* = u(B_a^m) - \frac{Y_a^m}{w_a}$ for any ability a. Hence, at symmetric equilibria, $m(w_a) = 1$ for $a = H, L$. Above the FOCs, we obtain $u'(B_H^m) = \frac{1}{w_H}$, and by similar argument in Lemma 2, we can show that $T_L^m < 0 < T_H^m$ where $T_a^m = Y_a^m - B_a^m$ for $a = H, L$.

The properties of allocation under such conditions can de described as the following lemma:

Lemma 4. *If taxpayers are mobile, and have their fixed unit wage, and the government has Rawlsian preference, then the allocation* $\{B_i(w_a), Y_i(w_a)\}_{a=H,L} \equiv \{B_a^m, Y_a^m\}_{a=H,L}$ *for each state i satisfies*

- $u'(B_H^m) = \frac{1}{w_H} \Leftrightarrow T_i'(Y_i(w_H)) = 0$ *for all* $i = 1, 2$*: no distortion on top,*
- $u'(B_L^m) = \frac{\gamma(1-p)}{(1-\lambda)+\gamma(1-p)\frac{T_L^m}{k}} \Leftrightarrow T_i'(Y_i(w_L)) = \frac{\lambda(1-\mathcal{F}^{-1})}{\gamma(1-p)w_L} > 0$ *for all* $i = 1, 2$*: upward distortion on top.*

In order to find progressivity of the tax schedule, we obtain from those FOCs:

$$\gamma = \frac{1}{w_L - (\mathcal{F}^{-1}pT_H^m + (1-p)T_L^m)/k} < \frac{1}{w_L}$$
$$\lambda = \frac{pw_H}{w_L - (\mathcal{F}^{-1}pT_H^m + (1-p)T_L^m)/k} - \frac{T_H^m}{k} < p\mathcal{F}$$

for all $k > 0$, and as k goes to infinity, γ and λ get the same ones as those in Benchmark 1. Therefore, the tax schedule comes close to one in Benchmark 1 when emigration cost is so high that few ones can exit from home country and work in another one, and according to the multipliers γ and λ, each tax authority attracts taxpayers with high skill and incurs more burden to those with low-skill as k decreases.

5.5 Benchmark 3: Optimal tax schedule under endogenous wage and without labor mobility

From now on, we introduce aggregate production technology to decide the wage level endogenously. Here, we assume that the each tax authority has one technology respectively, but that these two are identical. Let $F(Z_L, Z_H)$ be the production technology for each state where Z_a is the total labor supply for ability a, and it is assumed that

F is strictly increasing, strictly concave and constant return to scale, so $F(\alpha Z_L, \alpha Z_H) = \alpha F(Z_L, Z_H)$ for all $(Z_L, Z_H) \in \mathbb{R}^2_+$ and $\alpha > 0$, and $F_1(Z_L, Z_H), F_2(Z_L, Z_H) > 0$, $F_{11}(Z_L, Z_H) < 0$ and $F_{22}(Z_L, Z_H) < 0$ for any $(Z_L, Z_H) \in \mathbb{R}^2_+$.

For each class a, wage level per unit w_a is determined by the marginal product for labor supply: $w_L = F_1(Z_L, Z_H)$ and $w_H = F_2(Z_L, Z_H)$. Since F is constant return to scale, $F(Z_L, Z_H) = Z_L F(1, \frac{Z_H}{Z_L})$, and let $f(\frac{Z_H}{Z_L}) \equiv F(1, \frac{Z_H}{Z_L})$. Using f, the wage level for each class is determined by

$$F_1(Z_L, Z_H) = f(z) - f'(z) \times z$$

$$F_2(Z_L, Z_H) = f'(z).$$

where z is the ratio of aggregate labor supply between two classes, $z = \frac{Z_H}{Z_L}$.

As benchmark case, we derive the optimal tax schedule without labor mobility. Each wage level is determined by the marginal product for total labor supply, but taxpayers are not allowed to move from their home state to the other. In describing the self-selection constraint, taxpayers with high ability w_H who mimic the low ones work for earning $w_L Q_L$ where Q_a is the labor supply for each worker with ability a, so the amount of their labor supply is:

$$Q_L \times \frac{w_L}{w_H} = Q_L \times \frac{f(n) - nf'(n)}{f'(n)} \equiv Q_L \phi(n)$$

where $n = \frac{pQ_H}{(1-p)Q_L}$ is the ratio of total labor supply between two classes, and $\phi(\cdot)$ is the unit wage or productivity ratio. In this economy, the amount of total consumption $pB_H + (1-p)B_L$ comes from the production $F((1-p)Q_L, pQ_H)$. Thus, the resource constraint is written by

$$F((1-p)Q_L, pQ_H) - pB_H - (1-p)B_L \geq 0.$$

In order to obtain the optimal tax schedule, the government i solves the following maximization problem:

$$\max_{\{B_a,Q_a\}_a} u(B_L) - Q_L \quad \text{subject to}$$
$$u(B_H) - Q_H - u(B_L) + Q_L \times \phi(\frac{pQ_H}{(1-p)Q_L}) \geq 0 \tag{7}$$
$$F((1-p)Q_L, pQ_H) - pB_H - (1-p)B_L \geq 0$$

Let λ and γ be the Lagrangian multipliers of self-selection constraint and resource constraint. The FOCs are:

$$\begin{aligned}
&B_L : (1-\lambda)u'(B_L) - \gamma(1-p) = 0\\
&B_H : \lambda u'(B_H) - \gamma p = 0\\
&Q_L : -1 + \lambda(\phi(\frac{pQ_H}{(1-p)Q_L}) - \frac{pQ_H}{(1-p)Q_L}\phi'(\frac{pQ_H}{(1-p)Q_L})) + \gamma(1-p)F_1((1-p)Q_L, pQ_H) = 0\\
&Q_H : \lambda(-1 + \frac{p}{1-p}\phi'(\frac{pQ_H}{(1-p)Q_L})) + \gamma pF_2((1-p)Q_L, pQ_H) = 0.
\end{aligned} \tag{8}$$

Rearranging this, the disposable income for high ability B_H and the marginal tax rate for type H is determined by

$$u'(B_H) = \frac{1 - \frac{p}{1-p}\phi'(\frac{pQ_H}{(1-p)Q_L})}{F_2((1-p)Q_L, pQ_H)} < \frac{1}{F_2((1-p)Q_L, pQ_H)}$$
$$\Leftrightarrow T'(Y_H) = 1 - \frac{1}{1 - \frac{p}{1-p}\phi'(\frac{pQ_H}{(1-p)Q_L})} < 0 \quad \text{where } Y_H \text{ is the before-tax labour income,}$$

as $\phi'(\cdot) > 0$. The variation of wage ratio weighted by population ratio play a crucial role in disposable income B_H and the marginal tax rate. If $\frac{p}{1-p}\phi'$ increases, disposable income B_H also increases, and this result is different from fixed wage cases. Similarly, those with type H face negative marginal tax rate, and the absolute value increases as the proportion increases. Also, pre-tax income for low ability B_L is derived from the following equation:

$$
\begin{aligned}
u'(B_L) &= \frac{1}{1-\lambda} \times \frac{1-\lambda\phi(\frac{pQ_H}{(1-p)Q_L})(1-\sigma)}{F_1((1-p)Q_L, pQ_H)} \\
&> \frac{1}{1-\lambda} \times \frac{1-\lambda\phi(\frac{pQ_H}{(1-p)Q_L})}{F_1((1-p)Q_L, pQ_H)} > \frac{1}{F_1((1-p)Q_L, pQ_H)} \\
\Leftrightarrow T'(Y_L) &= \frac{\lambda-\lambda\phi(\frac{pQ_H}{(1-p)Q_L})(1-\sigma)}{1-\lambda\phi(\frac{pQ_H}{(1-p)Q_L})(1-\sigma)} > 0
\end{aligned}
$$

where $\sigma = \frac{pQ_L}{(1-p)Q_H}\frac{\phi'(\frac{pQ_H}{(1-p)Q_L})}{\phi(\frac{pQ_H}{(1-p)Q_L})} = \frac{d\ln\phi}{d\ln n}$ is the elasticity of substitution and Y_L is the before-tax labour income. The disposable income B_L depends on the elasticity σ and wage ratio ϕ. As the wage ratio or the elasticity increases, B_L decreases. Again, the general equilibrium effect as well as self-selection effect distort the disposable allocation from the first best. Summarizing the argument, we state the following lemma about properties of the optimal allocation.

Lemma 5. *If taxpayers are not mobile but their wage is determined by marginal productivity of aggregate production, and the government has Rawlsian preference, then the allocation* $\{B_a^e, Q_a^e\}_{a=H,L}$ *for each state i satisfies*

- $u'(B_H^e) = \frac{1-\frac{p}{1-p}\phi'(\frac{pQ_H}{(1-p)Q_L})}{F_2((1-p)Q_L,pQ_H)} < \frac{1}{F_2((1-p)Q_L,pQ_H)}$*: upward distortion caused by the variation of wage ratio,*

- $u'(B_L^e) = \frac{1}{1-\lambda} \times \frac{1-\lambda\phi(\frac{pQ_H}{(1-p)Q_L})(1-\sigma)}{F_1((1-p)Q_L,pQ_H)} > \frac{1}{F_1((1-p)Q_L,pQ_H)}$*: downward distortion caused by the wage ratio and the elasticity of substitution*

- $T'(Y_H) < 0$ *and* $T'(Y_H) > 0$*: marginal tax rate for H is negative and positive for L,*

where $\sigma = \frac{d\ln\phi}{d\ln n}$ *is the elasticity of substitution for wage ratio.*

5.6 Optimal tax schedule under endogenous wage and with labor mobility

We derive the optimal tax policy with endogenous wage and labor mobility. In the same manner as Section 4, we assume that taxpayers are able to move from their own state to the other, and they incur the cost $(1-x)k$ in leaving from their state where x is their locational preference between $[0,1]$ which is uniformly distributed, and $k \in \mathbb{R}_{++}$ is the marginal cost for emigration. Also, their wage is determined by the marginal product in the country where they work. So, two assumptions in models discussed above are combined.

Here, we assume that these two states are symmetric: they have the same population for each class, and the same technology F. Each government solves the following optimization problem:

$$\begin{aligned}
&\max_{\{B_a,Q_a\}_{a=H,L}} u(B_L) - Q_L \quad \text{subject to} \\
&F((1-p)m_L Q_L, pm_H Q_H) - pm_H B_H - (1-p)m_L B_L \geq 0 \quad \text{and} \\
&u(B_H) - Q_H - u(B_L) + \phi(\frac{pm_H Q_H}{(1-p)m_L Q_L})Q_L \geq 0.
\end{aligned} \tag{9}$$

The Lagrangian is:

$$\begin{aligned}
&u(B_L) - Q_L + \lambda\{u(B_H) - Q_H - u(B_L) + \phi(\frac{pm_H Q_H}{(1-p)m_L Q_L})Q_L\} \\
&+\gamma\{F((1-p)m_L Q_L, pm_H Q_H) - pm_H B_H - (1-p)m_L B_L\}
\end{aligned} \tag{10}$$

At symmetric equilibria, $m_L = m_H = 1$ because their utilities between two states are equal for each class. Differentiating it with respect to B_H, B_L, Q_H and Q_L and rearranging the FOCs[4], we obtain properties of the optimal allocation, and those are summa-

[4]The FOCs and its derivations are left in Appendix.

rized as the following theorem.

Theorem 1. *If taxpayers are mobile, their wage is determined by marginal productivity of aggregate production, and the two symmetric governments have Rawlsian preference, then the allocation* $\{B_a^{me}, Q_a^{me}\}_{a=H,L}$ *satisfies*

- $$u'(B_H^{me}) = \frac{1}{F_2((1-p)Q_L, pQ_H) + \frac{\lambda}{\gamma}\frac{1}{1-p}\phi'(\frac{pQ_H}{(1-p)Q_L})} < \frac{1}{F_2((1-p)Q_L, pQ_H)}$$
$$\Leftrightarrow T'(Y_H^{me}) < 0 \tag{11}$$

  upward distortion caused by the variation of wage ratio and negative marginal tax rate for H.

- $$u'(B_L^{me}) = \frac{1}{F_1((1-p)Q_L, pQ_H) + \frac{\lambda}{\gamma(1-p)}(\phi(\frac{pQ_H}{(1-p)Q_L}) - 1 - \frac{pQ_H}{(1-p)Q_L}\phi'(\frac{pQ_H}{(1-p)Q_L}))}$$
$$> \frac{1}{F_1((1-p)Q_L, pQ_H)} \quad \textit{and} \quad T'(Y_L^{me}) > 0 \tag{12}$$

  downward distortion caused by the wage ratio and the elasticity of substitution.

According to Theorem 1, we can find that trickle down effect caused by general equilibrium distorts the optimal allocation from the first best, which is similar to the labor-immobile cases.

Also, we can see that the labor mobility results in the distortion. Rearranging the FOC with respect to Q_H, we obtain that

$$\frac{\lambda}{\gamma} = p \times \frac{F_2((1-p)Q_L, pQ_H) - \frac{1}{k}T_H^{em}}{1 - \phi'(\frac{pQ_H}{(1-p)Q_L})\frac{p}{1-p} + \frac{Q_H}{k}\phi'(\frac{pQ_H}{(1-p)Q_L})\frac{p}{1-p}}$$

where $T_H^{em} = F_2((1-p)Q_L, pQ_H)Q_H - B_H$. Because of the self-selection constraint, $T_H^{em} > 0$, following the argument in Lemma 2. On top of that, $\phi'(\frac{pQ_H}{(1-p)Q_L}) > 0$. Therefore, $\frac{\lambda}{\gamma}$ distorted downwardly as k increases, and at the extreme cases, if k goes to infinity, $u'(B_H^{em})$ becomes identical to $u'(B_H^e)$. It is different from fixed-wage cases that the degree of labor mobility effect k depends on endogenous wage effect $\phi'(\cdot)$. According to the ratio of these multipliers, the denominator contains $\frac{\phi'(\cdot)}{k}$, so both marginal cost of moving k and the sensitivity of wage ratio toward that of total labor supply $\phi'(\cdot)$ play an important role in allocation B_H^{em}. The intuition is as follows; if the marginal cost k decreases, more taxpayers living in one state come to the other state, and it also increases the wage ratio $\phi(\cdot)$ due to decreasing the marginal productivity for skilled labor supply. In a nutshell, emigration terms m_H and m_L are built into the production technology F. Even though $m_L = m_H = 1$ at symmetric equilibria, any marginal exit effects do not disappear. The argument can be applied to $u'(B_L^{me})$ or $T'(Y_L^{em})$. If low-skilled taxpayers are more likely to move, then the number of them increases while the unit wage decreases. So, the wage ratio $\phi(\cdot)$ decreases, and it distorts their disposable income downwardly. By the same token, the marginal tax rate increases as k increases, which lead to that $\frac{\lambda}{\gamma}$ decreases.

Summarizing the discussions, the next proposition claims the effects of moving cost k and general equilibrium or trickle-down $\phi'(\cdot)$ on the marginal tax rates.

Proposition 7. *According to the tax schedules derived in Theorem 1,*

- *About $T'(Y_H^{em})$, the marginal tax rate comes close to 0 as emigration cost k increases or the threat of their emigration becomes weaker. In an analogous way, the tax rate for low-class $T'(Y_L^{em})$ decreases as that threat becomes weaker.*

5.7 Conclusion

In this paper, we investigate nonlinear income tax competition between two competing governments with mobile high-skilled or low-skilled workers, and derive the tax schedules at symmetric equilibria, in particular, focusing on the marginal tax rate for each class. Our novelty is to introduce not only mobility for all taxpayers, but also endogenous wages determined by marginal productivities at national level, and in this setting, we obtain the marginal tax rate for each skill. Compared to tax schemes for labor mobility but fixed wages, we show that the equilibrium marginal tax rate which high-skilled workers face contains migration threat effect k as well as general equilibrium one, so the sign must be negative and the slope be affected by marginal emigration cost k. Also, as the ones become immobile, the marginal tax rates for all classes become closer to those for models with endogenous wages but not embracing labor mobility. In a similar fashion, the marginal tax rates near those with exogenous wages and incorporating taxpayers' emigration when the general equilibrium effect or the variation of wage ratio with regard to the fraction of workers within one country $\phi'(\cdot)$ decreases. Intuitively, since emigration effects are embedded with each tax authority's production technology as well as the resource constraint, the marginal cost for moving k does not completely vanish in the marginal tax rate for high-class. On top of that, such emigration terms within productivity lead to the impact that marginal effects of changes in their wages become stronger as threats of their leaving increase.

In all benchmark cases, we review models and its results studied in existing literatures, which are beneficial for analyzing our new model in comparison to them. Without general equilibrium effect, high-skilled workers never face non-zero marginal tax rates while low-skilled ones always face positive marginal tax rates, and the progressivity of taxation depends on how low their emigration cost is and how large the variation of wage ratio toward the fraction of taxpayers and the wage disparity are. If such pay gaps and changes are high, the tax rates for low class must increase, and their disposable incomes must decrease. Due to trickle-down and migration effects, each tax authority is reluctant to collect taxes from high skilled workers in order to attract them; instead, she cannot redistribute lots of collected incomes from high class to low class. Especially, in our new kind of model, their threat of relocation, which does not appear in the marginal tax rate for high skilled workers in the setup with exogenous wages, emerges in the tax rate for those.

In the end, we leave several several future works we have to unveil. At first, we find the directions of marginal tax rates in our novel setting, but never completely clarify the impacts of general equilibrium part and moving cost on the tax schedules. In order to figure out such unsolved questions, we should conduct more analytical studies and numerical simulations like papers studying optimal taxation with general equilibrium effect[5]. Another is to generalize our setups. For instance, like Lehmann et al. (2014) and Lipatov and Weichenrieder (2015), we have the space to introduce asymmetricity of two tax authorities about demographic composition and production technology. Though such generalization makes much more difficult to solve the model, we must obtain more findings about equilibrium outcomes of tax competitions which are closer to events in reality. Considering asymmetric settings as well as symmetric settings, characterizing all equilibria in such tax competitions is also nice for future research.

[5] For example, see Jacobs (2013) about numerical simulation.

Apart from our model, nonlinear income tax competition model can be applied to tax authorities' competition problems for attracting international firms or entrepreneurs[6].

Appendix: FOCs under endogenous wage and labor mobility

First of all, differentiating Lagrangian (10) with respect to B_H, B_L, Q_H and Q_L, we obtain the following four FOCs:

$$\begin{aligned}
B_H :&\lambda(u'(B_H) + \frac{p\frac{u'(B_H)}{k}Q_H}{(1-p)Q_L}Q_L\phi'(\frac{pQ_H}{(1-p)Q_L})) \\
&+ \gamma[p\frac{u'(B_H)}{k}Q_H F_2((1-p)Q_L, pQ_H) - p(1 + \frac{u'(B_H)}{k}B_H)] = 0 \\
B_L :&(1-\lambda)u'(B_L) - \lambda\frac{pQ_H}{(1-p)Q_L}Q_L\frac{u'(B_L)}{k}\phi'(\frac{pQ_H}{(1-p)Q_L}) \\
&+ \gamma[(1-p)\frac{u'(B_L)}{k}Q_L F_1((1-p)Q_L, pQ_H) - (1-p)(1 + \frac{u'(B_L)}{k}B_L)] = 0 \\
Q_H :&\lambda(-1 + \phi'(\frac{pQ_H}{(1-p)Q_L})\frac{p(1-\frac{Q_H}{k})}{(1-p)Q_L}Q_L) \\
&+ \gamma\{p(1-\frac{Q_H}{k})F_2((1-p)Q_L, pQ_H) + p\frac{B_H}{k}\} = 0 \\
Q_L :& -1 + \lambda(\phi(\frac{pQ_H}{(1-p)Q_L}) - \frac{pQ_H}{(1-p)Q_L}\phi'(\frac{pQ_H}{(1-p)Q_L})(1-\frac{Q_L}{k})) \\
&+ \gamma\{(1-p)(1-\frac{Q_L}{k})F_1((1-p)Q_L, pQ_H) + (1-p)\frac{B_L}{k}\} = 0.
\end{aligned}$$

In order to derive $T'(Y_H^{me}) < 0$ and $T'(Y_L^{me}) > 0$, we multiply the first two equations by $u'(B_H)$ and $u(B_L)$ respectively, and rearrange those with the remaining two equations. Also, for the purpose of obtaining $\frac{\lambda}{\gamma}$, we utilize the FOCs with respect to Q_H and Q_L.

[6]For instance, see Olsen and Osmundsen (2001) and Boyer and Kempf (2017).

Bibliography

Aiura, H., Ogawa, H., 2013. Unit tax versus ad valorem tax: A tax competition model with cross-border shopping. Journal of Public Economics 105, 30–38.

Anderson, S. P., De Palma, A., Kreider, B., 2001. The efficiency of indirect taxes under imperfect competition. Journal of Public Economics 81 (2), 231–251.

Aronsson, T., Johansson-Stenman, O., 2008. When the Joneses' consumption hurts: Optimal public good provision and nonlinear income taxation. Journal of Public Economics 92 (5), 986–997.

Atkinson, A. B., Stern, N. H., 1974. Pigou, Taxation and Public Goods. The Review of Economic Studies 41 (1), 119.

Atkinson, A. B., Stiglitz., J. E., 1972. The structure of indirect taxation and economic efficiency. Journal of Public economics 1.1, 97–119.

Atkinson, A. B., Stiglitz, J. E., 1976. The design of tax structure: Direct versus indirect taxation. Journal of Public Economics 6 (1-2), 55–75.

Balestrino, A., 2009. Tax avoidance, endogenous social norms, and the comparison income effect, CHILD Working Paper No.15.

Bastani, S., Blomquist, S., Micheletto, L., 2016. Optimal commodity taxation with varying quality of goods. Research in Economics 70, 89–100.

Besley, T., Coate, S., 1992. Workfare versus welfare: incentive arguments for work requirements in poverty-alleviation programs. *American Economic Review* 82, 249–261.

Besley, T., Coate, S., 1995. The design of income maintenance programmes. *Review of Economic Studies* 62 (2), 187–221.

Besley, T., Kanbur, R., 1988. Food subsidies and poverty alleviation. *Economic Journal* 98, 701–719.

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

Bierbrauer, F. J., 2014. Optimal tax and expenditure policy with aggregate uncertainty. American Economic Journal: Microeconomics 6 (1 D), 205–257.

Boadway, R., Keen, M., 1993. Public Goods, Self-Selection and Optimal Income Taxation. International Economic Review 34 (3), 463–478.

Bös, D., Tillmann, G., 1985. An envy tax: theoretical principles and applications to the German surcharge on the rich. *Public Finance* 40, 35–63.

Boskin, M. J., Sheshinski, E., 1978. Optimal redistributive taxation when individual welfare depends upon relative income. *Quarterly Journal of Economics* 92, 589–601.

Bowles, S., Park, Y., 2005. Emulation, inequality, and work hours: was Thorsten Veblen right? *Economic Journal* 115, 397–412.

Boyer, P. C., Kempf, H., 2017. Regulatory arbitrage and the efficiency of banking regulation. Journal of Financial Intermediation.

Bruce, D., Peng, L., 2018. Optimal taxation in the presence of income-dependent relative income effects. *Social Choice and Welfare*, 1–23.

Chaudhuri, A., 1986. Some implications of an intensity measure of envy. Social Choice and Welfare 3 (4), 255–270.

Chetty, R., Guren, A., Manoli, D., Weber, A., 2011. Are micro and macro labor supply elasticities consistent? A review of evidence on the intensive and extensive margins. *American Economic Review* 101, 471–475.

Corlett, W. J., Hague, D. C., 1953. Complementarity and the excess burden of taxation. The Review of Economic Studies, 21–30.

Corneo, G., 2002. The efficient side of progressive income taxation. *European Economic Review* 46, 1359–1368.

Delipalla, S., Keen, M., 1992. The comparison between ad valorem and specific taxation under imperfect competition. Journal of Public Economics 49 (3), 351–367.

Delipalla, S., Keen, M., 2006. Product quality and the optimal structure of commodity taxes. Journal of Public Economic Theory 8, 547–554.

Denicoló, V., Matteuzzi, M., 2000. Specific and ad valorem taxation in asymmetric cournot oligopolies. International tax and public finance 7, 335–342.

Diamantaras, D., Thomson, W., 1989. A refinement and extension of the no-envy concept. Economics Letters 30 (2), 103–107.

Diamond, P. A., 1998. Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. American Economic Review 88 (1), 83–95.

Diamond, P. A., Mirrlees, J. A., 1971a. Optimal Taxation and Public Production I: Production Efficiency. American Economic Review 61 (1), 8–27.

Diamond, P. A., Mirrlees, J. A., 1971b. Optimal Taxation and Public Production II: Tax Rules. The American Economic Review 61 (3), 261–278.

Fang, H., 2006. Disentangling the college wage premium: estimating a model with endogenous education choices. *International Economic Review* 47, 1151–1185.

Fang, H., Norman, P., 2010. Optimal provision of multiple excludable public goods. American Economic Journal: Microeconomics 2 (4), 1–37.

Feldman, A., Kirman, A., 1974. Fairness and Envy. American Economic Review 64 (6), 995–1005.

Fleurbaey, M., Maniquet, F., 2006. Fair income tax. *The Review of Economic Studies*, 55–83.

Foley, D. K., 1967. Resource allocation and the public sector.

Gaube, T., 2000. When do distortionary taxes reduce the optimal supply of public goods? Journal of Public Economics 76 (2), 151–180.

Gaube, T., 2005a. Financing public goods with income taxation: Provision rules vs. provision level. International Tax and Public Finance 12 (3), 319–334.

Gaube, T., 2005b. Income Taxation, Endogenous Factor Prices and Production Efficiency. Scandinavian Journal of Economics 107 (2), 335–352.

Goldin, C., Katz, L. F., 2007. The race between education and technology: the evolution of u.s. educational wage differentials, NBER Working Paper, No. 12984.

Hamilton, J., Pestieau, P., 2005. Optimal income taxation and the ability distribution: Implications for migration equilibria. International Tax and Public Finance 12 (1), 29–45.

Hellwig, M. F., 2005a. A utilitarian approach to the provision and pricing of excludable public goods. Journal of Public Economics 89 (11-12), 1981–2003.

Hellwig, M. F., 2005b. A utilitarian approach to the provision and pricing of excludable public goods. *Journal of Public Economics* 89, 1981–2003.

Hellwig, M. F., 2010. Utilitarian mechanism design for an excludable public good. Economic theory 44 (3), 361–397.

Ireland, N. J., 2001. Optimal income tax in the presence of status effects. *Journal of Public Economics* 81, 193–212.

Jacobs, B., 2013. Optimal redistributive tax and education policies in general equilibrium. International Tax and Public Finance 20 (2), 312–337.

Jacobs, B., Boadway, R., 2014. Optimal linear commodity taxation under optimal non-linear income taxation. *Journal of Public Economics* 117, 201–210.

Kanbur, R., Keen, M., Tuomala, M., 1994. Optimal non-linear income taxation for the alleviation of income-poverty. *European Economic Review* 38, 1613–1632.

Kanbur, R., Paukkeri, T., Pirttilä, J., Tuomala, M., 2018. Optimal taxation and public provision for poverty reduction. *International Tax and Public Finance* 25, 64–98.

Kanbur, R., Tuomala, M., 2013. Relativity, inequality, and optimal nonlinear income taxation. *International Economic Review* 54 (4), 1199–1217.

Kay, J. A., Keen, M. J., 1983. How should commodities be taxed? European Economic Review 23, 339–358.

Kay, J. A., Keen, M. J., 1991. Product quality under specific and ad valorem taxation. Public Finance Quarterly 19, 371–385.

Kind, H. J., Koethenbuerger, M., Schjelderup, G., 2008. Efficiency enhancing taxation in two-sided markets. Journal of Public Economics 92 (5-6), 1531–1539.

Kolm, S.-C., 1971. Justice et equite (cepremap, paris).

Krause, A., 2017. On redistributive taxation under the threat of high-skill emigration. Social Choice and Welfare 48 (4), 845–856.

Lehmann, E., Simula, L., Trannoy, A., 2014. Tax me if you can! optimal nonlinear income tax between competing governments. Quarterly Journal of Economics 129 (4), 1995–2030.

Lipatov, V., Weichenrieder, A., 2015. Welfare and labor supply implications of tax competition for mobile labor. Social Choice and Welfare 45 (2), 457–477.

Lloret-Batlle, R., Jayakrishnan, R., 2016. Envy-minimizing pareto efficient intersection control with brokered utility exchanges under user heterogeneity. Transportation Research Part B: Methodological 94, 22–42.

Lloret-Batlle, R., Jayakrishnan, R., 2017. Envy-free pricing for collaborative consumption of supply in transportation systems. Transportation Research Procedia 23, 772–789.

Lockwood, B., 2004. Competition in unit vs. ad valorem taxes. International Tax and Public Finance 11 (6), 763–772.

Martimort, D., Stole, L., 2002. The Revelation and Delegation Principles in Common Agency Games. Econometrica 70 (4), 1659–1673.

Micheletto, L., 2004. Optimal Redistributive Policy with Endogenous Wages. FinanzArchiv / Public Finance Analysis 60 (2), 141–159.

Micheletto, L., 2011. Optimal nonlinear redistributive taxation and public good provision in an economy with veblen effects. Journal of Public Economic Theory 13 (1), 71–96.

Mirrlees, J., 1976. Optimal tax theory: a synthesis. *Journal of Public Economics* 6, 327–358.

Mirrlees, J. A., 1971. An Exploration in the Theory of Optimal Taxation. Review of Economic Studies 38 (2), 175–208.

Mirrlees, J. A., 1982. Migration and optimal income taxes. Journal of Public Economics 18 (3), 319–341.

Morelli, M., Yang, H., Ye, L., 2012a. Competitive nonlinear taxation and constitutional choice. American Economic Journal: Microeconomics 4 (1), 142–175.

Morelli, M., Yang, H., Ye, L., 2012b. Competitive nonlinear taxation and constitutional choice. American Economic Journal: Microeconomics 4 (1), 142–175.

Moulin, H., 1994. Serial Cost-Sharing of Excludable Public Goods. Source: The Review of Economic Studies 61 (2), 305–325.

Myles, G. D., 1996. Imperfect competition and the optimal combination of ad valorem and specific taxation. International tax and public finance 3, 29–44.

Naito, H., 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. Journal of Public Economics 71 (2), 165–188.

Nakada, S., 2018. Envy-free allocation of indivisible goods with money and externalities. *Economics Bulletin* 38, 52–59.

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.

Nishimura, Y., 2000. Envy minimization in the optimal tax context, Paper presented at the Fifth International Conference of Society of Social Choice and Welfare, University of Alicante, Spain.

Nishimura, Y., 2003a. Optimal commodity taxation for reduction of envy. *Social Choice and Welfare* 21, 501–527.

Nishimura, Y., 2003b. Optimal non-linear income taxation for reduction of envy. *Journal of Public Economics* 87, 363–386.

Norman, P., 2004. Efficient mechanisms for public goods with use exclusions. Review of Economic Studies 71 (4), 1163–1188.

OECD, 2015. How's life? 2015 measuring well-being.

OECD, 2017. General government spending.

Olsen, T. E., Osmundsen, P., 2001. Strategic tax competition; Implications of national ownership. Journal of Public Economics 81 (2), 253–277.

Oswald, A. J., 1983. Altruism, jealousy and the theory of optimal non-linear taxation. *Journal of Public Economics* 20, 77–87.

Persson, M., 1995. Why are taxes so high in egalitarian societies? *Scandinavian Journal of Economics* 97, 569–580.

Piaser, G., 2007. Labor mobility and income tax competition. In: International taxation handbook: policy, practice, standards and regulation. CIMA Publishing, Ch. 4, pp. 73–94.

Pirttilä, J., Tuomala, M., 2004. Poverty alleviation and tax policy. *European Economic Review* 48 (5), 1075–1090.

Rochet, J.-C., Stole, L. A., 2002. Nonlinear Pricing with Random Participation. Review of Economic Studies 69 (1), 277–311.

Sachs, D., Tsyvinski, A., Werquin, N., 2016. Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68 (1), 205–229.

Saez, E., 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. Journal of Public Economics 83 (2), 217–230.

Samuelson, P., 1954. The Pure Theory of Public Expenditure. Review of Economics and Statistics 36 (4), 387–389.

Seidman, L. S., 1987. Relativity and efficient taxation. *Southern Economic Journal* 54, 463–474.

Simula, L., Trannoy, A., 2010. Optimal income tax under the threat of migration by top-income earners. Journal of Public Economics 94 (1-2), 163–173.

Simula, L., Trannoy, A., 2012. Shall we keep the highly skilled at home? The optimal income tax perspective. Social Choice and Welfare 39 (4), 751–782.

Skeath, S. E., Trandel, G. A., 1994. A Pareto comparison of ad valorem and unit taxes in noncompetitive environments. Journal of Public Economics 53 (1), 53–71.

Stern, N., 1982. Optimum taxation with errors in administration. Journal of Public Economics 17 (2), 181–211.

Stiglitz, J. E., 1982. Self-selection and Pareto efficient taxation. Journal of Public Economics 17 (2), 213–240.

Suits, D. B., Musgrave, R. A., 1953. AD VALOREM AND UNIT TAXES COMPARED. Quarterly Journal of Economics 67 (4), 598–604.

Tadenuma, K., 2002. Efficiency first or equity first? Two principles and rationality of social choice. Journal of Economic Theory 104 (2), 462–472.

Tóbiás, Á., 2016. Income redistribution in open economies. Journal of Public Economics 134, 19–34.

Tsugawa, S., Obara, T., 2017. Public good provision financed by nonlinear income tax under reduction of envy.

Varian, H., 1974. Equity, envy, and efficiency. *Journal of Economic Theory* 9, 63–91.

Varian, H., 1976a. Two problems in the theory of fairness. *Journal of Public Economics* 5, 249–260.

Varian, H. R., 1976b. Two problems in the theory of fairness. Journal of Public Economics 5 (3-4), 249–260.

Velez, R. A., 2016. Fairness and externalities. *Theoretical Economics* 11, 381–410.

Wilson, J. D., 1980. The effect of potential emigration on the optimal linear income tax. Journal of Public Economics 14 (3), 339–353.

Wilson, J. D., 1986. A theory of interregional tax competition. Journal of Urban Economics 19 (3), 296–315.

Wilson, J. D., 1992. Optimal Income Taxation and International Personal Mobility. The American Economic Review 82 (2), 191–196.

World Economic Forum, 2017. The global risks report 2017 12th edition.

Zodrow, G. R., Mieszkowski, P., 1986. Pigou, Tiebout, property taxation, and the underprovision of local public goods. Journal of Urban Economics 19 (3), 356–370.