Queen's Economics Department Working Paper No. 1178

Envy Minimization in the Optimal Tax Context

Yukihiro Nishimura
Yokohama National University and Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

8-2008

Envy Minimization in the Optimal Tax Context

Yukihiro Nishimura
Department of Economics, Yokohama National University, 79-3 Tokiwadai, Hodogaya-ku, Yokohama, 240-8501, Japan.*†

August 22, 2008

Abstract

This paper examines the second-best tax policy to minimize envy in the sense of Chaudhuri (1986) and Diamantaras and Thomson (1990). An allocation is λ-equitable if no agent prefers a proportion λ of any other agent's bundle. We study the allocations that maximize λ among the second-best Pareto efficient allocations. In the standard two-class economy with identical preferences, the Chaudhuri-Diamantaras-Thomson allocation coincides with the leximin allocation. In many-agent economies, it is possible to order the class of second-best Pareto efficient allocations graded by progressivity in the sense of Hemming and Keen (1983), with respect to the intensity of envy. Envy is then minimized in the most progressive tax system.

Key Words: Income Taxation, Envy
JEL Classification: D63, H21

1 Introduction

It has been a common belief among public finance researchers that there is no alternative to the utilitarian social welfare function and its generalized form including a welfaristic

*Tel: +81-45-339-3569. Fax: +81-45-339-3574. E-mail: ynishimu@ynu.ac.jp

†The author would like to express special thanks to Robin Boadway, Koichi Tadenuma, Walter Bossert, and William Thomson for constructive advice and encouragement. Comments and suggestions by Francois Maniquet, Oana Secrieru, Stefan Dodds, Stanley Winer and Toshihiro Ihori are also valuable. The usual disclaimer applies. Financial support from JSPS (Japan Society for the Promotion of Science) Postdoctoral Fellowships for Research Abroad, Grant-in-Aid for Young Scientists (B) from the Ministry of Education, Culture, Sports, Science and Technology, and Nomura Foundation for Social Science is gratefully acknowledged.

version of the Rawlsian maximin principle.[1] On the other hand, the social choice literature examined non-welfaristic equity criteria for resource allocation,[2] but the analysis was mainly devoted to the first-best environment without distortionary taxation. In this paper, we choose minimization of envy as the social objective, and analyze the policy implications in the second-best environment.

We define envy with respect to agents' bundle holdings. Agent i envies agent j if the former perceives the latter's bundle to be better than his own. Alternative motivations of the importance of envy are provided by several researchers. First, it suits the tradition of the new welfare economics, which avoids both comparability and measurability of individual welfare. Second, Varian (1974) and Thomson and Varian (1985) emphasized symmetry in the sense that each agent puts himself in the place of the other agent to evaluate the other agent's position on the same terms that he judges his own. Third, as Varian (1974) clarified, we do not regard envy as the propensity to view with hostility towards the fortune of others (Ralws (1971, p.531)),[3] but rather as a representation of Rawls' notion of resentment:[4] the existence of envy is clear-cut evidence that agents are being treated asymmetrically, and

[1]The studies on the informational basis for social-welfare judgment (e.g., d'Aspremont and Gevers (1977)) confirmed the modern public economists' presumption that interpersonal comparison is necessary for normative public decision making. For example, '[g]iven Arrow's theorem, [the impossibility of interpersonal welfare comparison] did not leave very much for welfare economics to do. Modern approaches, including optimal income tax theory, are firmly based on explicit interpersonal comparisons' (Tuomala (1990), p.24), and 'as the welfare economics literature has made clear, interpersonal comparisons are mandatory for policy prescriptions' (Boadway (2002), p.50).

[2]See, e.g., Thomson and Varian (1985) and Moulin and Thomson (1997) for an overview.

[3]Such an individual sentiment stems from interdependent preferences, but we do not assume such preferences in the present paper.

[4]'If we resent our having less than others, it must be because we think that their being better off is the result of unjust institutions' (Rawls (1971, p.533)).

reduction of such resentment is a valid social objective. Forth, a society with minimal or fewer envy reflects a policy-maker's desire for social harmony, if there are psychological and sociological reasons that the sentiments of envy exist. Bös and Tillmann (1985) said, '[t]he 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).

We consider a second-best framework originated by Mirrlees (1971), where an allocation has to satisfy self-selection (or incentive compatibility). We show in Section 3 that there arises a hierarchy of envy, and the low-skilled agents always envies the high-skilled agent. In order to pursue equitable and efficient tax policies based on envy, we use the notion by Chaudhuri (1986) and Diamantaras and Thomson (1990). An allocation is λ-equitable if no agent prefers a proportion λ of the bundle of any other agent. The value of λ measures the intensity of envy reduction: if λ is unity, it is the usual no-envy; any feasible allocation satisfies zero-equitability. Diamantaras and Thomson (1990) proposed to maximize λ among the Pareto efficient allocations (hereafter referred to as *the DT allocations*).[5] They showed the existence of such an allocation in a first-best environment. This paper, on the other hand, examines the tax policy implications of their criterion in a second-best environment.

The issue we consider is the amount of compensation given to the low-skilled agents in

[5] Chaudhuri's objective differs from Diamantaras-Thomson when there are more than three types of agents, which we will discuss in section 5.1.

the second-best DT allocations. A conflict between reducing envy and enhancing the low-skilled agents' welfare under second-best taxation is pursued in Nishimura (2003a) where the available policy for the government is linear income tax (Sheshinski (1972), Hellwig (1986), and Sandmo (1993)).[6] When the income elasticity of leisure is less than unity, the DT allocation may differ from the Rawlsian solution. Moreover, when preferences differ, there exists an economy in which the DT allocation is the Pareto efficient allocation which *minimizes* the welfare of the low-skilled agent. In this paper, we examine the same issue in the context of the nonlinear income taxation model (Mirrlees (1971), Stiglitz (1982)).[7] It turns out that the results are substantially different under nonlinear taxation model.

In Section 4 we start from a two-class economy, commonly dealt with in the optimal taxation literature, in which there are agents with high and low skill levels. We show that a conflict between reducing envy and enhancing the low-skilled agent's welfare discussed above disappears when a nonlinear income tax is available. Under identical preferences, envy reduction is consistent with enhancement of the low skill's welfare. The same conclusion also holds under different preferences among the allocations where the high-skilled agent pays the higher tax (Proposition 1).

In Sections 5 and 6 we extend our analysis to many-agent economies. Two kinds of

[6] Nishimura (2003a) also examined commodity taxes an economy with many commodities. The analysis in this paper can be extended to the multi-commodity setting. We discuss this in Section 7.

[7] Linear income tax has advantages of administrative simplicity and analytical tractability, and this policy tool has been studied for a long time in the optimal taxation literature. On the other hand, nonlinear income tax is the best policy tool to achieve any informationally feasible allocations under Mirrleesian informational framework (see Guesnerie (1995), Chapter 1).

tax changes, discussed in Lambert (1993, Chapter 9), are examined. First, the class of the second-best Pareto efficient allocations which satisfy the following *single crossing of the tax schedules* is examined: for any pair of allocations, the consumption-gross income schedules cross once, hence the distribution of the post-tax income or welfare is more equalized under one tax system than the other (Hemming and Keen (1983), Gans and Smart (1996)). An example is the set of Pareto efficient allocations among those supported by linear income taxes. We show that these allocations can be ordered with respect to the intensity of envy: the DT criterion chooses the most progressive taxation which maximizes the welfare of the lowest-skilled agent (Proposition 2). This result remains valid under different preferences (Propositions 4 and 5). On the other hand, when we consider 'double-crossing' tax reform, where tax cuts benefit only the middle-income agents, it may well be the case that envy is reduced under the tax system where both the highest skilled and the lowest skilled are worse-off (Proposition 3). These results show consistency as well as the difference between the DT criterion and the Rawlsian maximin principle, in highly policy-relevant circumstances such as single- versus double-crossing tax reforms and the amount of subsidies given to the hard-working poor.

2 The Model

Consider an n-agent economy ($n \geq 2$) with one consumption good, c, and labor, l. Let $N \equiv \{1, ..., n\}$. Each agent i ($i \in N$) has a preference represented by a binary relation R_i, 'at least as good as', defined over $(c, l) \in I\!R_+ \times [0, \bar{l}]$. The indifference relation and the strict preference relation associated with R_i are denoted by I_i and P_i, respectively. Each R_i is a continuous ordering which is increasing in the consumption good, decreasing in labor, and strictly convex. Each agent i is endowed with an exogenous skill level $w_i \geq 0$. The set of feasible allocations is $\{(c_i, l_i)_{i=1}^n \in (I\!R_+ \times [0, \bar{l}])^n | \sum_{i=1}^n c_i \leq \sum_{i=1}^n w_i l_i\}$. A feasible allocation is often denoted by $x = (x_i)_{i=1}^n$.

We assume that both consumption and leisure are normal goods. Formally, let $(\hat{c}_i(w, T), \hat{l}_i(w, T))$ be the welfare-maximizing bundle of agent i subject to the budget constraint $c = wl - T$. That is, $\hat{c}_i(w, T) = w\hat{l}_i(w, T) - T$ and $(\hat{c}_i(w, T), \hat{l}_i(w, T))\ R_i\ (c, l)$ for all (c, l) *s.t.* $c = wl - T$. Consumption and leisure are normal goods if, for all $i \in N$, $w \geq 0$, T^a and T^b such that $T^a < T^b$, $\hat{c}_i(w, T^a) \geq \hat{c}_i(w, T^b)$ and $\hat{l}_i(w, T^a) \leq \hat{l}_i(w, T^b)$ hold with at least one strict inequality.

The property of *income elasticity of leisure* is cricical in the following analysis:

Definition 1 *The income elasticity of leisure is greater than (resp. less than/ equal to) unity if, all $i \in N$, $w \geq 0$, T^a and T^b where $T^a < T^b$,*

$$\textit{if }\ \hat{c}_i(w, T^b) > 0,\ \hat{l}_i(w, T^a) > 0 \textit{ and } \hat{l}_i(w, T^b) < \bar{l}, \textit{then } \frac{\hat{c}_i(w, T^a)}{\bar{l} - \hat{l}_i(w, T^a)} < \ (\textit{resp.} > / =)\ \frac{\hat{c}_i(w, T^b)}{\bar{l} - \hat{l}_i(w, T^b)}.$$

Here, the income elasticity of leisure is defined in terms of 'full income', $w\bar{l} - T$.[8] Three

[8]When $\hat{c}_i(w, T)$ and $\hat{l}_i(w, T)$ are differentiable, $\frac{\hat{c}_i(w, T^a)}{\bar{l} - \hat{l}_i(w, T^a)} < \ (\textit{resp.} > / =)\ \frac{\hat{c}_i(w, T^b)}{\bar{l} - \hat{l}_i(w, T^b)}$ implies $\frac{w\bar{l} - T}{\bar{l} - \hat{l}_i(w, T^a)} \frac{\partial l_i}{\partial T} > \ (\textit{resp.} < / =)\ 1$.

benchmark cases are studied in the optimal taxation literature. First, Weymark (1986) and Boadway *et al.* (2000, 2002) assume that preferences are quasi-linear in leisure, in which the income elasticity of leisure is greater than unity. Second, the comprementary case where preferences are quasi-linear in consumption, in which the income elasticity of leisure is less than than unity, is studied by Phelps (1973) and Diamond (1998). Third, the case where the income elasticity of leisure is unity is called *homotheticity*. This is common in the quantitative optimal taxation studies (Mirrlees (1971), Stern (1976), Tuomala (1990), Tarkiainen and Tuomala (1999)).

It is often convenient to convert the preference R_i into $\tilde{R}_i$ over $(c, y) \in I\!R_+ \times [0, w_i\bar{l}]$ in the following way:

$$(c, y)\ \tilde{R}_i\ (\hat{c}, \hat{y}) \iff (c, y/w_i)\ R_i\ (\hat{c}, \hat{y}/w_i) \text{ when } w_i > 0,\ (c, 0)\ \tilde{R}_i\ (\hat{c}, 0) \iff (c, 0)\ R_i\ (\hat{c}, 0). \quad (1)$$

The indifference relation and the strict preference relation associated with $\tilde{R}_i$ are denoted by $\tilde{I}_i$ and $\tilde{P}_i$, respectively.

Seade (1982) showed that normality of consumption implies the following *agent monotonicity* (hereafter *AM*): at any point in the consumption-gross income space, the indifference curve of the low-skilled agent is steeper than that of the high-skilled agent:

Definition 2 *(Condition AM) For all agents i and j with $w_i > w_j$ and all $(\hat{c}, \hat{y}), (\tilde{c}, \tilde{y}) \in I\!R_+ \times [0, w_j\bar{l}]$, $(\hat{c}, \hat{y})\ \tilde{I}_i\ (\tilde{c}, \tilde{y})$ implies (i) $(\hat{c}, \hat{y})\ \tilde{P}_j\ (\tilde{c}, \tilde{y})$ if $(\hat{c}, \hat{y}) < (\tilde{c}, \tilde{y})$,*[9] *and (ii) $(\tilde{c}, \tilde{y})\ \tilde{P}_j\ (\hat{c}, \hat{y})$ if $(\hat{c}, \hat{y}) > (\tilde{c}, \tilde{y})$.*

[9] Vector inequality: given $(\hat{c}, \hat{y}), (\tilde{c}, \tilde{y}) \in I\!R^2_+$, $(\hat{c}, \hat{y}) > (\tilde{c}, \tilde{y})$ means $\hat{c} > \tilde{c}$ and $\hat{y} > \tilde{y}$.

The government can observe the gross income, $y_i \equiv w_i l_i$, of each agent i, but cannot observe either w_i or l_i. The income tax is used for redistribution:

Definition 3 *An income tax schedule* $T^x : \mathbb{R}_+ \to \mathbb{R}$ **supports** *allocation* $x = (c_i, l_i)_{i=1}^n$ *if (i)* $c_i = w_i l_i - T^x(w_i l_i)$*, (ii)* $(c_i, w_i l_i)\ \tilde{R}_i\ (y - T^x(y), y)$ *for all* i *and all* $y \in [0, w_i \bar{l}]$*, and (iii)* $l_i = 0$ if $w_i = 0$.

It is well-knwon that (i) and (ii) are equivalent to the following *self-selection* condition:[10]

$$\forall i,\ j \in N,\ (c_i, w_i l_i)\ \tilde{R}_i\ (c_j, w_j l_j) \text{ if } y_j \leq w_i \bar{l}, \tag{2}$$

which requires that agent i cannot be better off by choosing the gross income and consumption of agent j by working $w_j l_j / w_i$ hours.

Allocation $(c_i, l_i)_{i=1}^n$ is *first-best Pareto efficient* iff there is no other feasible allocation $(\tilde{c}_i, \tilde{l}_i)_{i=1}^n$ such that $(\tilde{c}_i, \tilde{l}_i)\ R_i\ (c_i, l_i)$ for all $i \in N$ with at least one agent with $(\tilde{c}_i, \tilde{l}_i)\ P_i\ (c_i, l_i)$. An allocation is *second-best Pareto efficient* iff it satisfies (2) and there is no other feasible allocation $(\tilde{c}_i, \tilde{l}_i)_{i=1}^n$, satisfying (2), such that $(\tilde{c}_i, \tilde{l}_i)\ R_i\ (c_i, l_i)$ for all $i \in N$ with at least one agent with $(\tilde{c}_i, \tilde{l}_i)\ P_i\ (c_i, l_i)$.

3 The Equity Objective

3.1 Hierarchy of Envy in the Second-best Environment

We consider envy with respect to agents' consumption-leisure bundles. Agent i envies agent j at allocation x iff $x_j\ P_i\ x_i$.[11] Allocation x satisfies no-envy iff $x_i\ R_i\ x_j$ for all i

[10] Condition (iii) is the utility-maximization condition for $i \in N$ such that $w_i = 0$.

[11] This formulation of envy does not require the observability of agents' labor supply. A reasonable supposition is that the agents, as well as the government, know the demographic data in the economy, i.e.,

and j. In the class of self-selective allocations, we observe a hierarchical structure of envy. First, the higher skilled agent *never* envies the lower skilled agent, since the self-selection implies:

$$\forall i,\ j \in N,\ (c_i, l_i)\ R_i\ (c_j, \frac{w_j}{w_i} l_j)\ R_i\ (c_j, l_j) \text{ if } w_i > 0 \text{ and } w_i \geq w_j. \tag{3}$$

On the other hand, if $w_i > w_j$ and $R_i = R_j$, then $(c_i, l_i)\ P_j\ (c_j, l_j)$ always holds, except when $l_j = 0$. Therefore, *in a standard economy in which all agents have identical preferences, at any envy-free and self-selective allocation, at most one agent supplies labor* (Bös and Tillmann (1985)). This result shows that no-envy is not useful in examining equitable and efficient taxation. Tillmann (2005) generalizes this result to economies where preferences differ. In a class of economies that include those we study in Section 6, this result still holds.[12] In order to pursue equitable and efficient tax policies based on envy, we need to formulate a less demanding criterion.

3.2 Reduction of Envy

To evaluate the intensity of envy, we will adopt a radial contraction measure advocated by Chaudhuri (1986) and Diamantaras and Thomson (1990).[13]

---

distribution of preferences and the skill level of the other agents, hence they correctly *infer* the other agents' labor supply as a best response to the given tax schedule.

[12] There are two remarks. First, Bös and Tillmann (1985) originally showed that, if the skill distribution is variable (the true economy is drawn from a skill distribution whose upper bound is infinity), the self-selective and envy-free allocation is 'no production'. Second, notice that no-envy and self-selection are disjoint conditions if skills are unequal. There are envy-free allocations which violate (2), and most of the self-selective allocations invite envy of the lower skilled as shown in (3).

[13] Several alternative concepts of no-envy are proposed in the social choice literature. Feldman and Kirman (1974) first advocated reducing envy based on a cardinal measure; their measure is the difference in utilities (see Bös and Tillmann (1985) for tax policy implications of their concept). However, their concept

Figure 1 around here.

Figure 1 illustrates the idea. At allocation x, agent i envies agent j. By continuity and monotonicity, the straight line connecting the points $(0, \bar{l})$ and $x_j = (c_j, l_j)$ intersects agent i's indifference curve passing through x_i exactly once. The point of intersection is denoted by $(\tilde{c}, \tilde{l})$, and thus $x_i \; I_i \; (\tilde{c}, \tilde{l})$. The ratio $\lambda^x_{ij} \equiv \dfrac{\bar{l} - \tilde{l}}{\bar{l} - l_j} = \dfrac{\tilde{c}}{c_j}$ measures the factor by which one would have to shrink j's bundle in order for i to stop envying j. The inverse of λ^x_{ij} indicates the extent to which agent i envies agent j at allocation x.

Formally, let $\lambda \in I\!R_+$ be a nonnegative real number, and let $\circ$ be the operation such that $\lambda \circ (c_i, l_i) \equiv (\lambda c_i, \bar{l} - \lambda(\bar{l} - l_i))$, which represents a proportional contraction of agent i's consumption-leisure bundle. Given allocation x, let $\lambda^x_{ij} \in I\!R_+$ be such that $x_i \; I_i \; (\lambda^x_{ij} \circ x_j)$ when $x_j \; R_i \; x_i$, and $\lambda^x_{ij} \equiv 1$ when $x_i \; P_i \; x_j$.[14] Let $\lambda^x \equiv \min_{i,j} \lambda^x_{ij}$.

Definition 4 *Allocation x is λ-equitable if $\lambda^x \geq \lambda$, or equivalently, $\forall i, \; j \in N, \; x_i \; R_i \; (\lambda \circ x_j)$.*

is not invariant with respect to monotonic transformations of the utility functions. The property does not suit the classic assumption of new welfare economics: social choice should be ordinally invariant. On other alternative concepts advocated by the mid 1980's (see Thomson and Varian (1985)), Tillmann (1984) and Bös and Tillmann (1985) show that most of them are not successful in addressing the Bös-Tillmann's result discussed above. On the concepts in economies with different consumptive talents (Arnsperger (1994), Section 6), we can easily see that the application of these criteria into the current environment also results in the Bös-Tillmann's result.

[14] Here, we modify the definition of Diamantaras and Thomson (1990). In the current environment, the λ-expansion of the leisure of the opponent may meet the upper bound $\bar{l}$, so that a value of λ^x_{ij} such that $x_i \; I_i \; (\lambda^x_{ij} \circ x_j)$ may not exist when $x_i \; P_i \; x_j$.

Diamantaras and Thomson propose to maximize λ^x in the set of Pareto efficient allocations (hereafter referred to as *the DT allocation*).[15] Our goal is to examine the tax policy implications of their concept, by seeking an allocation which maximizes λ^x in the set of the *second-best* Pareto efficient allocations.

4 Two-class Economies

We start with a two-class economy, commonly dealt with in the optimal taxation literature. Without loss of generality, we examine a two-agent economy ($n = 2$) with $w_2 > w_1$. We begin with the case of identical preferences, i.e., $R_1 = R_2 \equiv R$.

4.1 The Second-Best Pareto Optima

The purpose of this section is to illustrate the set of second-best Pareto efficient allocations. It is convenient to proceed with the preference $\tilde{R}_i$ defined in (1). Hereafter, 'the self-selection constraint of agent i against agent j ($i \neq j$)' is abbreviated to SS_i. The SS_i is *binding* at allocation $((c_1, l_1), (c_2, l_2))$ if $(c_i, w_i l_i)\ \tilde{I}_i\ (c_j, w_j l_j)$. Also, let *the income-consumption curve* of agent i (hereafter ICC_i^*) be the collection of the bundles $(c_i(T), y_i(T))$ defined with respect to $T \in I\!R$ such that $c_i(T) = y_i(T) - T$ and $(c_i(T), y_i(T))\ \tilde{R}_i\ (\tilde{c}, \tilde{y})$ for all $(\tilde{c}, \tilde{y})$ *s.t.* $\tilde{c} = \tilde{y} - T$. Namely,

$$ICC_i^* \equiv \{(c, y) \in I\!R_+ \times [0, w_i \bar{l}] | (c, y) = (c_i(T), y_i(T)) \text{ for some } T \in I\!R\}. \tag{4}$$

[15] This is in general different from Chaudhuri's (1986) concept. However, they are the same in a two-class economy with a hierarchy of envy. The difference in a many-agent economies will be discussed in Section 5.1.

Following the convention of the optimal taxation literature, we say that *the marginal income tax rate of agent i is zero* at a self-selective allocation $x = (c_j, l_j)_{j=1}^n$ if $(c_i, w_i l_i) \in ICC_i^*$.

We now introduce the following result in the optimal income taxation:

Lemma 1 *(Stiglitz (1982)) At any second-best Pareto efficient allocation, (i) at most one self-selection constraint is binding, and, (ii) if SS_j $(j \neq i)$ is not binding, the marginal income tax rate of agent i is zero.*

Figure 2 illustrates this result on (c, y)-space.

Figure 2 around here.

In each panel, the upward-sloping curves denoted by $\tilde{I}_i$ are indifference curves of agent i, namely, the sets $\tilde{I}(a, \tilde{R}_i) \equiv \{b \in I\!R_+ \times [0, w_i \bar{l}] | b \ \tilde{I}_i \ a\}$ corresponding to bundle a. By agent monotonicity (AM), for all a, $\tilde{I}(a, \tilde{R}_1)$ is steeper than $\tilde{I}(a, \tilde{R}_2)$. The dotted curves, denoted by ICC_i^*, are the income-consumption curves of agent i $(i = 1, 2)$. By strict convexity, $(c_i(T), y_i(T))$ is unique for all T. By normality, ICC_i^* is downward sloping.

Panel (a) depicts allocations where the marginal income tax rates of agents are zero. From the second welfare theorem, every first-best Pareto efficient allocation corresponds to allocation $G(T) \equiv ((c_1(-T), y_1(-T)/w_1), (c_2(T), y_2(T)/w_2))$ for $T \in I\!R$. For example, $G(0)$ is the laissez-faire allocation, which trivially satisfies (2). By Tillmann (1989, Proposition 1.1.1), there is a continuum of first-best Pareto efficient allocations which satisfy (2), i.e., there is a nondegenerate interval $[\underline{T}, \overline{T}] \subset I\!R$ such that $G(T)$ is self-selective for all

$T \in [\underline{T}, \overline{T}]$. Let $A \equiv (A_1, A_2) \equiv G(\underline{T})$ and $B \equiv (B_1, B_2) \equiv G(\overline{T})$. Any point along A_1B_1 corresponds to a bundle $(c_1(-T), y_1(-T)/w_1)$ constituting $G(T)$, $T \in [\underline{T}, \overline{T}]$, and there is a corresponding bundle $(c_2(T), y_2(T)/w_2)$ for agent 2 along A_2B_2.

Panel (b) illustrates allocations in which SS_2 binds. Since $G(T)$, $T > \overline{T}$ does not satisfy SS_2, a transfer of income to agent 1 greater than $\overline{T}$ through income taxation is possible only if agent 1's consumption-gross income deviates from ICC_1^*. Stiglitz (1982) showed that, at any second-best Pareto efficient allocation where agent 1 receives a subsidy greater than $\overline{T}$, the bundle of agent 1 has to be southwest of ICC_1^*, and that of agent 2 belongs to ICC_2^*.[16] In panel (b), the bundle of agent 2 belongs to B_2D_2, and there is a corresponding bundle for agent 1 on B_1D_1 (for example, (B_1', B_2') is an allocation with $B_1' \; \tilde{I}_2 \; B_2'$). From $D \equiv (D_1, D_2)$ in the panel where $l_1 = 0$, no further transfer to agent 1 is possible. As we increase the transfer to agent 1 from allocation B to D, agent 2's welfare apparently decreases. On the other hand, due to the increased distortions, the welfare of agent 1 may or may not increase by increasing the transfer to him. Therefore, some allocations in panel (b) may be second-best Pareto inefficient.

Panel (c) illustrates allocations where the transfer to agent 1 is lower than $\underline{T}$. In such allocations, SS_1 is binding,[17] and the bundle of agent 1 belongs to ICC_1^*. From $C \equiv (C_1, C_2)$, where $w_2 l_2 = w_1 \bar{l}$ in the panel, no further transfer to agent 2 is possible. As

[16] In the conventional terminologies, downward distortion is imposed on agent 1 who faces the positive marginal income tax rate, and no distortion is imposed on agent 2.

[17] For an illustrative purpose, we will depict the case that $w_2 l_2 \leq w_1 \bar{l}$ at A_2 so that agent 1 can supply the gross income $w_2 l_2$ at allocation A (in other words, there exists $T \in I\!R$ such that $(c_1(-T), y_1(-T)) \; \tilde{I}_1 \; (c_2(T), y_2(T))$). Otherwise, there is no second-best Pareto efficient allocation where SS_1 binds.

in panel (b), some allocations in panel (c) may not be second-best Pareto efficient.

Panel (d) illustrates the set of second-best Pareto efficient allocations. Along B_1D_1, agent 1's welfare is maximized at E_1. This means that any bundle along E_1D_1 gives a lower welfare to agent 1 than E_1, so that the corresponding allocations are not second-best Pareto efficient. An analogous explanation applies for F_2. The allocation which maximizes the welfare of agent 1 among the second-best Pareto efficient allocations is called *the leximin allocation*.

4.2 Variation in λ^x in the Non-Regressive Area

Let us evaluate the allocations with respect to Chaudhuri-Diamantaras-Thomson's measure defined in Section 3. As discussed earlier, when preferences are identical, $\lambda_{21}^x = 1$ and $\lambda_{12}^x \leq 1$ for any self-selective allocation (with the second inequality strict except at allocation D in Figure 2). Therefore, we can identify $\lambda^x = \lambda_{12}^x$.

Let *the non-regressive area* be the Pareto efficient subset of the area where SS_1 is not binding, corresponding to the Pareto efficient subset of allocations in panels (a) and (b) in Figure 2. We now show that the higher is the transfer to agent 1, the higher λ^x is in this area.

Lemma 2 *Consider a two-class economy with identical preferences. In the non-regressive area, λ^x is maximized at the leximin allocation.*

Proof: As shown in Figure 2, SS_1 is not binding in the non-regressive area. By Lemma 1, the bundle of agent 2 belongs to ICC_2^*. Consider any pair of allocations $x = ((c_1, l_1), (c_2, l_2))$

and $z = ((c'_1, l'_1), (c'_2, l'_2))$ in the non-regressive area with $x_1 \; P \; z_1$ and $z_2 \; P \; x_2$. As $(c_2, w_2 l_2), (c'_2, w_2 l'_2) \in ICC^*_2$ and $z_2 \; P \; x_2$, normality implies $c'_2 \geq c_2$ and $l_2 \geq l'_2$ with at least one strict inequality. Therefore, $(\lambda^z \circ z_2) \; P \; (\lambda^z \circ x_2)$. Combining with $x_1 \; P \; z_1 \; I \; (\lambda^z \circ z_2)$, we have $\lambda^x > \lambda^z$. *Q.E.D.*

4.3 Variation in $\boldsymbol{\lambda^x}$ in the Regressive Area

For those whose concern is the redistribution from the high-skilled agent to the low-skilled agent, Lemma 2 is sufficient to conclude that envy is minimized at the leximin allocation. However, the maximization of λ^x does not necessarily imply redistribution from the high-skilled agent to the low-skilled agent. In fact, the change in λ^x at allocations in panel (c) of Figure 2 (hereafter *the regressive area*) would be analyzed when preferences are differentiable. The following is proved in an appendix:

Lemma 3 *Consider a two-class economy with identical preferences. Suppose that preferences are twice continuously differentiable. Then λ^x increases as we increase the welfare of the low-skilled agent among the second-best Pareto efficient allocations.*

4.4 Different Preferences

We now consider the case with $n = 2$ and $w_2 > w_1$ where preferences differ, i.e., R_1 and R_2 differ. Brito *et al.* (1990, Corollary 1) show that Lemma 1.(i) and (ii) still hold under different preferences. Therefore, as in the case of identical preferences, we can divide the second-best Pareto efficient allocations into the three cases: (a) no self-selection constraint is binding; (b) SS_2 is binding; (c) SS_1 is binding. We now examine the pattern of λ^x_{12} in cases (a) and (b) (the non-regressive area). Applying Proposition 3 of Brito *et al.* (1990), the marginal income tax rate of agent 2 is zero: that is, as is depicted in Figure 2, agent 2's bundle belongs to ICC_2, which is downward-sloping on (c, y)-space if consumption and leisure are normal goods. Therefore, Lemma 2 is extended to the case of different preferences: *in the non-regressive area,* λ^x_{12} *is increasing when* $\lambda^x_{12} < 1$ *as we increase the welfare of the low-skilled agent.* Brito *et al.* (1990, Proposition 4) also show that the marginal income tax rate of the highest tax payer is zero, so that the non-regressive area includes every second-best Pareto efficient allocation where the high-skilled agent pays the higher tax.

Lemma 2, Lemma 3 and the above discussion are summarized as a following proposition:

Proposition 1 *Consider a two-class economy.*

1. Among the second-best Pareto efficient allocations where the high-skilled agent pays the higher tax, λ^x *increases when* $\lambda^x_{12} < 1$*, as we increase the welfare of the low-skilled agent.*

2. If preferences are identical and twice continuously differentiable, then λ^x *increases as we increase the welfare of the low-skilled agent among the second-best Pareto efficient allocations.*

Proposition 1 contrasts Nishimura (2003a) in which the available policy for the government is *linear income tax.* Formally, consider a class of allocations where there exist some α and

β such that:[18]

$$\forall i \in N,\ l_i = \hat{l}_i((1-\beta)w_i, -\alpha),\ c_i = \alpha + (1-\beta)w_i l_i \text{ and } \sum_{i=1}^{n} \beta w_i l_i = \alpha n. \tag{5}$$

He showed that, (i) when preferences are identical, envy reduction is consistent with enhancement of the welfare of the low-skilled agent when the income elasticity of leisure is greater than or equal to unity (Nishimura (2003a, Proposition 1)); (ii) when the income elasticity of leisure is less than unity, the DT allocation may differ from the leximin allocation (Nishimura (2003a, Proposition 2)); (iii) when preferences differ, there exists an economy in which the DT allocation is the Pareto efficient allocation which *minimizes* the welfare of the low-skilled agent (Nishimura (2003a, Proposition 3)). In contrast, our Proposition 1.1 is valid in the case of different preferences as well. Also, Proposition 1.2 is valid when the income elasticity of leisure is less than unity. Therefore, a conflict between envy reduction and compensation to the low-skilled agent depends on whether the government's tax possibilities are restricted or not.

5 Many-Agent Economies

We now consider many-agent economies with identical preferences, R. Suppose that the agents' skill levels are ordered with $w_i \leq w_{i+1}$ for all $i = 1, ..., n-1$ with at least two strict inequalities.

[18] Namely, the tax that supports the allocation x is $T^x(y) = \beta y - \alpha$, which consists of the single marginal income tax rate, β, and the uniform basic income, α. This is the simplest form of redistributive income taxation which is widely studied by Sheshinski (1972), Hellwig (1986) and Sandmo (1993) and others.

5.1 Identification of the Envying/Envied Agent

In discussing the Diamantaras-Thomson allocations, we first need to identify the pair of agents i^* and j^* such that $\lambda^x_{i^*j^*} = \min_{i,j} \lambda^x_{ij}$ for each allocation x. Because of the hierarchical structure discussed in Section 3, it is natural to conjecture that λ^x_{1n} is the smallest among λ^x_{ij}'s. This is true when the income elasticity of leisure is greater than or equal to unity.

Lemma 4 *Consider an n-agent economy with identical preferences. For any self-selective allocation x, $\lambda^x_{1j} \leq \lambda^x_{ij}$ for all i and j. Moreover, if the income elasticity of leisure is greater than or equal to unity, $\lambda^x = \lambda^x_{1n}$. When the income elasticity of leisure is less than unity, $\lambda^x \neq \lambda^x_{1n}$ may happen at some self-selective allocation x.*

Even if we omit counting the psychic compensation that the higher skilled agent obtains from the lower skilled agent's misfortune (namely, greater than one λ^x_{ij}), there arise $n(n-1)/2$ pairs of envy (the lower skilled envying the higher skilled). Indeed, it is hard to characterize the optimal allocation on the basis of Chaudhuri's (1986) additive criterion.[19] On the other hand, Lemma 4 shows that it is easy to identify the target for minimization

[19]Chaudhuri's measure is the minimization of the sum of $1/\lambda^x_{ij}$'s for all pairs of i, j which are greater than 1:

$$\min \sum_{\{(i,j)|\lambda^x_{ij}<1\}} \left(\frac{1}{\lambda^x_{ij}} - 1 \right).$$

For λ-equitability, $\lambda^x \equiv \min_{i,j} \lambda^x_{ij}$ matters. On the other hand, with this equity criterion, it is the sum of envies that matters.

of envy with respect to the Diamantaras-Thomson (1990) criterion.[20]

5.2 Variation in λ^x When Single-Crossing Applies

We consider two kinds of binary comparisons, which are typically considered in tax reform discussion, according to Lambert (1993). First, we examine the case where the consumption-gross income schedules induced by the tax schedule that supports each allocation cross exactly once, which is called *single-crossing* by Hemming and Keen (1983). The case of double-crossing, where tax cuts benefit only the middle-income agents, is examined in the next subsection.

The single-crossing is defined with respect to the income tax schedules that support allocations (recall Definition 3). Formally:

Definition 5 *For self-selective allocations x and z, x* **single-crosses** *z if there exist $T^x(\cdot)$ and $T^z(\cdot)$ that support x and z respectively, such that there exists y^* such that $T^x(y) \leq T^z(y)$ for all $y \leq y^*$ and $T^x(y) \geq T^z(y)$ for all $y \geq y^*$.*

Notice that, under AM, at any self-selective allocation $x = (c_i, l_i)_{i=1}^n$, $w_n l_n \geq w_{n-1} l_{n-1} \geq \ldots \geq w_1 l_1$ and $c_n \geq c_{n-1} \geq \ldots \geq c_1$ hold. Therefore, the implication of single-crossing is that

[20] This point relates to Rawls' (1974) argument that 'informational requirement' is an advantage of the maximin solution (compared to utilitarianism). 'Once the least-favored group is identified, it may be relatively easy to determine which policies are to their advantage. By comparison it is much more difficult to know what maximizes average utility. ... In application [the maximization of the average utility] leaves so much to judgment that some may reasonably claim that the gains of one group outweigh the losses of another, while others may equally reasonably deny it' (Rawls (1974), p.343). One may argue that the identification of the least-favored is as hard as that of cardinal comparison of the utility gains and losses among alternative allocations. On the other hand, λ-equitability gives a systematic way of identifying the least favored. The advantage vis-a-vis Chaudhuri's measure with respect to computation is similar to that of maximin vis-a-vis utilitalianism.

the tax schedules are ordered according to their progressivity, in the sense that, starting from a critical agent, every higher income earner (hence higher skilled agent under AM) is worse off, and every lower income earner is better off under one tax schedule than the other. The following lemma can be shown, applying Gans and Smart (1996), Proposition 2.

Lemma 5 *Consider an n-agent economy with identical preferences. Consider self-selective allocations x and z where x single-crosses z. If $x_i \; P \; z_i$ (resp. $z_i \; P \; x_i$) for some $i \in N$, then $x_j \; R \; z_j$ (resp. $z_j \; R \; x_j$) for all $j \in N$ with $w_j < w_i$ (resp. $w_j > w_i$).*

In addition to single-crossing, we impose either of the following two requirements. First, agent n, the highest-skilled agent (the highest income earner under AM), bears the highest tax burden among all the agents. This is a minimal requirement of 'ability to pay'.[21] A second alternative is homotheticity of preferences, which is common in the quantitative optimal taxation studies (Mirrlees (1971), Stern (1976), Tuomala (1990)). In the class of second-best Pareto efficient allocations which satisfies the above requirements, we can show the following:

Proposition 2 *Consider an n-agent economy with identical preferences.*

1. If the second-best Pareto efficient allocations x and z satisfy: (a) x single-crosses z, and (b) the highest income earner pays the highest tax payment, then $\lambda^x_{1n} \geq \lambda^z_{1n}$.

[21]Under the welfaristic framework, this property is satisfied if the social welfare function satisfies the redistributive property (Guesnerie and Seade (1982)); that is, the redistribution from the higher skilled agent to the lower skilled agent is desirable as long as the self-selection constraint is satisfied.

2. If preferences satisfy homotheticity, and if the self-selective allocations x and z satisfy: (a) x single-crosses z, and (b) $x_i \ P_i \ z_i$ and $z_j \ P_j \ x_j$ for some $i \in N$ and $j \in N$, then $\lambda^x_{1n} \geq \lambda^z_{1n}$.

Proof: 1. By Lemma 5, $x_1 \ R \ z_1$ and $z_n \ R \ x_n$. By Proposition 4 of Brito *et al.* (1990), agent n's bundle belongs to ICC^*_n defined in (4). The rest of the derivation is the same as that of Lemma 2.

2. By Lemma 5, $x_1 \ R \ z_1$ and $z_n \ R \ x_n$. Under homotheticity, the latter implies that $(\lambda^z_{1n} \circ z_n) \ R \ (\lambda^z_{1n} \circ x_n)$. We then conclude that $x_1 \ R \ z_1 \ I \ (\lambda^z_{1n} \circ z_n) \ R \ (\lambda^z_{1n} \circ x_n)$, so that $\lambda^x_{1n} \geq \lambda^z_{1n}$. *Q.E.D.*

It is well-knwon in the optimal taxation literature that the analytical features of the second-best Pareto optimal allocations are highly complex. However, Proposition 2 provides a simple and useful way to order the class of second-best Pareto efficient allocations with respect to the intensity of envy. Lemma 4 shows that, when the income elasticity of leisure is greater than or equal to unity, $\lambda^x = \lambda^x_{1n}$ at any second-best Pareto efficient allocation. Therefore, the DT allocation coincides with the leximin allocation in the range of allocations ordered by progressivity. Notice that, although Hemming and Keen (1983) showed that the single-crossing relation is not transitive, we can order the welfares of agent 1 and λ^x_{1n} among the allocations satisfying the single-crossing: if x is more progressive than x' and x' is more progressive than z, then $x_1 \ R \ x'_1 \ R \ z_1$ and $\lambda^x_{1n} \geq \lambda^{x'}_{1n} \geq \lambda^z_{1n}$ among the allocations satisfying (b).

When the income elasticity of leisure is less than unity, there are two potential reasons that envy reduction may not justify a progressive tax system. First, when $\lambda^x \neq \lambda^x_{1n}$ as shown in Lemma 4, the envied agent may be an agent with middle income, not the one with the highest income, so that progressive taxation that possibly benefits the middle income may not reduce envy. Second, when a more progressive taxation reduces the labor supply (increases leisure) of the envied agent, it may have a negative effect on envy. Intuitively, when the income elasticity of leisure is less than unity, agents prefer leisure more relative to consumption as they become poorer. See Nishimura (2003a), pp. 507-508 and pp. 519-520 in the case of the linear income tax.

Proposition 2 can also hold in economies with a continuum of agents with bounded highest skill. We make use of the result of 'zero marginal income tax rate at the highest income' (Phelps (1973), Sadka (1976) and Seade (1977)) which holds under an assumption that the highest-skilled agent bears the highest tax burden.[22]

The case of linear income taxation is a special case of single-crossing.[23] Nishimura (2003a, Proposition 1) showed that envy reduction is consistent with higher progressivity when the income elasticity of leisure is greater than or equal to unity. Same result is obtained in the case of non-linear income tax.

[22]Consider an economy with a continuum of agents with identical preferences represented by R. The agent's skill level w belongs to an interval $[\underline{w}, \overline{w}]$ according to a CDF $F(w)$. A skill distribution is bounded if $\overline{w} < \infty$. At any second-best Pareto efficient allocation, the marginal income tax rate at the highest income level is non-positive (see, e.g., Seade (1977), Theorem 1 and Ebert (1992), equations (P1*), (P3*), (P4**). If it is negative (by the binding second-order incentive constraint), an agent with lower skill level bears higher tax burden, which is a contradiction. Diamond (1998) showed that this result is no longer valid when the skill distribution is unbounded ($\overline{w} = \infty$).

[23]Straight lines cross at most once. If the lines do not cross, one allocation Pareto dominates another.

5.3 Variation in λ^x When Single-Crossing Does Not Apply

In general, however, there are second-best Pareto efficient allocations which do not single cross, and tax reform sometimes entails tax cuts on the middle-income only, increasing the tax burden and the welfare loss of the top and the bottom. Regarding envy, we can find an allocation where the Diamantaras-Thomson measure is higher at an allocation than the leximin allocation when we allow double-crossing. We now show the following:

Proposition 3 *There exists an economy in which a self-selective allocation x has a higher λ^x than the leximin allocation and moreover it is Pareto noncomparable with the leximin allocation.*

Proof: Consider the following three-agent economy: $w_1 = 1,\ w_2 = 3,\ w_3 = 5$, preferences R can be represented by a common utility function $u(c, l) = c(\bar{l} - l)$, $\bar{l} = 1$.

The leximin allocation, $L = (c_i^L, l_i^L)_{i=1}^3$, is as follows:

$c_1^L = 0.72847,\ l_1^L = 0,\ c_2^L = 1.05457,\ l_2^L = 0.30923,\ c_3^L = 2.07234,\ l_3^L = 0.58554,\ \lambda^L = 0.92094.$

On the other hand, there is the following allocation $x = (c_i^x, l_i^x)_{i=1}^3$:

$c_1^x = 0.73361,\ l_1^x = 0.01361,\ c_2^x = 1.03669,\ l_2^x = 0.29556,\ c_3^x = 2.06500,\ l_3^x = 0.58700,\ \lambda^x = 0.92113.$

$(c_i^L, l_i^L)\ P\ (c_i^x, l_i^x)\ (i = 1,\ 3)$ and $(c_2^x, l_2^x)\ P\ (c_2^L, l_2^L)$, so that x is Pareto noncomparable with L. *Q.E.D.*

Here, both L and x satisfy condition (b) of Proposition 2.1 and 2.2, but, in order for $T^L(\cdot)$ and $T^x(\cdot)$ to support L and x respectively, $T^L(w_1 l_1^L) < T^x(w_1 l_1^L)$, $T^L(w_2 l_2^x) >$

$T^x(w_2 l_2^x)$ and $T^L(w_3 l_3^L) < T^x(w_3 l_3^L)$ have to hold. As $w_1 l_1^L < w_2 l_2^x < w_3 l_3^L$, $T^L(\cdot)$ and $T^x(\cdot)$ cannot single-cross. Envy reduces (i.e., $\lambda^z > \lambda^L$) at the expense of the welfares of *both* the envying agent and the envied agent (i.e., $(c_1^L, l_1^L)\ P\ (c_1^z, l_1^z)$ and $(c_3^L, l_3^L)\ P\ (c_3^z, l_3^z)$).

6 Two-Dimensional Characteristics

The optimal taxation literature typically assumes identical preferences. However, in this section we show that the result of Proposition 2 also holds in economies with different preferences. Let $\mathcal{R}$ be a class of preferences satisfying the properties mentioned in Section 2, and $\mathcal{W} \subset [0, \infty)$ be a class of agents' possible skill levels. Let $\underline{w}$, $\overline{w} \in \mathcal{W}$ denote the lowest and the highest skill levels, respectively. The characteristics of an agent are determined by preference $R \in \mathcal{R}$ and the skill level $w \in \mathcal{W}$. An economy with the above characteristics is called *a two-dimensional economy*.

The following two assumptions are employed. First, there is an agent whose preference for leisure is the lowest at any point in (c, l)-space. Formally, there exists $R^* \in \mathcal{R}$ such that, for all $(\hat{c}, \hat{l}), (\tilde{c}, \tilde{l}) \in \mathbb{R}_+ \times [0, \overline{l}]$ and $R \in \mathcal{R}$, $(\hat{c}, \hat{l})\ I^*\ (\tilde{c}, \tilde{l})$ implies: (i) $(\hat{c}, \hat{l})\ P\ (\tilde{c}, \tilde{y})$ if $(\hat{c}, \hat{l}) < (\tilde{c}, \tilde{l})$, and (ii) $(\tilde{c}, \tilde{l})\ P\ (\hat{c}, \hat{l})$ if $(\hat{c}, \hat{l}) > (\tilde{c}, \tilde{l})$. That is, $I((c, l), R^*)$ is flatter than $I((c, l), R)$ in (c, l)-space, which means that an agent with preference R^* has the lower propensity to consume leisure than one with R. Second, there exist agents $i,\ j \in N$ such that $(R_i, w_i) = (R^*, \underline{w})$ and $(R_j, w_j) = (R^*, \overline{w})$. These conditions are weaker than ones employed in the optimal tax studies by Sandmo (1993), Tarkiainen and Tuomala (1999) and

Tillmann (2005). They consider an economy in which agents can be ordered with respect to their preferences for leisure (stronger than the first assumption), and the distribution of $(R_i, w_i),\ i \in N$ has full support in $\mathcal{R} \times \mathcal{W}$ (stronger than the second assumption).

Issues addressed by Roemer (1996, 1998), Roemer *et alii.* (2003), Fleurbaey and Maniquet (1998, 2006, 2007), Bossert *et al.* (1999), and Schokkaert *et al.* (2004) with important policy relevance are the choice among the tax schemes that enhance the welfare of the lowest-skilled agents, and the amount of compensation given to agent i with $(R_i, w_i) = (R^*, \underline{w})$ ('the hard-working poor'), consistent with notions of fairness. Below we discuss these issues regarding envy-minimization under two cases: (i) $\underline{w} = 0$ and (ii) $\underline{w} > 0$.

6.1 The Case of $\underline{w} = 0$

We first examine the case where there are agents whose skill levels are zero. The following result is proved in an appendix.

Proposition 4 *Consider a two-dimensional economy. Suppose that the income elasticity of leisure is greater than or equal to unity.*

1. if the second-best Pareto efficient allocations x and z satisfy: (a) x single-crosses z, and (b) the highest income earner pays the highest tax payment, then $\lambda^x \geq \lambda^z$.

2. if preferences satisfy homotheticity, and if the self-selective allocations x and z satisfy: (a) x single-crosses z, and (b) $x_i\ P_i\ z_i$ and $z_j\ P_j\ x_j$ for some $i \in N$ and $j \in N$, then $\lambda^x \geq \lambda^z$.

3. if the allocations x and z are supported by linear income taxes and x single-crosses z, then $\lambda^x \geq \lambda^z$.

As Proposition 2, this proposition can also hold in economies with a continuum of agents with bounded highest skill.

Proposition 4 shows that, in the class of allocations graded by progressivity, the DT allocation coincides with an allocation that maximizes *the lowest income*. Justification for maximizing the lowest income in the case of $\underline{w} = 0$ is given by Fleurbaey and Maniquet (1998, 2007). Fleubaey and Maniquet (1998) advocated the following *compensation* axiom for a comparison of two allocations: allocation x is preferred to allocation z if, for all $i \in N$,

$$z_i \; P_i \; x_i \Rightarrow \exists j \in N, \; R_i = R_j, \; x_j \; P_i \; z_j \text{ and } x_i \; R_j \; x_j. \tag{6}$$

This has a resembrance to Hammond's (1976) equity axiom.

If x and z are self-selective, (6) implies $w_i > w_j$. If a tax schedule that supports each allocation crosses once, then x single-crosses z (x is more progressive). Therefore, the compensation axiom is consistent with increasing progressivity in the sense of Hemming-Keen (1983), which is in turn consistent with reducing envy. Flauebaey and Maniquet (1998, Proposition 1 and 2007, Theorem 1) also showed that the maximization of the lowest income is consistent with an extension of the leximin principle based on the notions of responsibility and compensation. Basically, when the lowest skill level is zero, maximizing the welfare or an opportunity set of the lowest-skilled agents is equivalent to maximizing the lowest income, which is also consistent with reducing envy in the sense of Diamantaras-Thomson.

6.2 The Case of $\underline{w} > 0$

A conflict between reducing envy of agent i and that of j with $w_i = w_j = \underline{w}$ and $R_i \neq R_j$ arises when $\underline{w} > 0$. The preferred tax rates of these agents, as well as the tax rates that minimize the envy of these agents differ. With regard to envy minimization, the following result can be shown. As a benchmark, we assume homothetic and differentiable preferences, where R is represented by a unitility function $u(c, l, R)$. We also restrict our attention to the linear income taxation[24] where there exist some α and β that satisfy (5). Let $x(\beta) \equiv (c_i(\beta), l_i(\beta))_{i=1}^n$ be such an allocation,[25] and let $v(\beta, w_i, R_i) \equiv u(c_i(\beta), l_i(\beta), R_i)$.

Proposition 5 *Consider a two-dimensional economy. Suppose that, for all $R \in \mathcal{R}$, there exist agents $i,\ j \in N$ such that $(R_i, w_i) = (R, \underline{w})$ and $(R_j, w_j) = (R, \overline{w})$. Suppose also that preference R is homothetic and represented by a twice-continuously differentiable unitility function $u(c, l, R)$ for all $R \in \mathcal{R}$. At any allocation $x(\beta)$ where $l_i(\beta) > 0$ for $i \in N$ such that $(R_i, w_i) = (R^*, \underline{w})$, the following holds:*

$$\frac{\partial v(\beta, \underline{w}, R^*)}{\partial \beta} \geq 0 \Rightarrow \frac{\partial v(\beta, \underline{w}, R)}{\partial \beta} > 0 \text{ for all } R \in \mathcal{R} \text{ and } \frac{\partial \lambda^{x(\beta)}}{\partial \beta} > 0. \tag{7}$$

[24] In the literature of fair income taxation mentioned above, most researches focus on linear income taxation.

[25] Define $\alpha(\beta)$ implicitly by $\alpha(\beta)n - \sum_{i=1}^n \beta w_i \hat{l}_i((1-\beta)w_i, -\alpha(\beta)) = 0$. Then $l_i(\beta) \equiv \hat{l}_i((1-\beta)w_i, -\alpha(\beta))$, $c_i(\beta) \equiv \alpha(\beta) + (1-\beta)w_i l_i(\beta)$.

By the same logic as in Lemma 5, in the range where the increasing tax rates enhances the welfare of the hard-working poor, it also enhances the welfare of *any* agent with the lowest skill who has higher preferences for leisure. Proposition 5 shows that this is also consistent with reducing envy in the sense of Diamantaras-Thomson. Therefore, in this range, envy reduction is consistent with increasing the welfare of any agent with the lowest skill.

However, there are some cases where envy reduction entails the decrease of the welfare of the lowest-skilled agents, i.e., $\dfrac{\partial v(\beta, \underline{w}, R)}{\partial \beta} < 0$, $R \in \mathcal{R}$ is compatible with $\dfrac{\partial \lambda^{x(\beta)}}{\partial \beta} > 0$. First, the proof of Proposition 5 in an appendix shows that, if the uncompensated labor-supply curves are upward-sloping ($\dfrac{\partial \hat{l}_i(w,T)}{\partial w} \geq 0$ for all $i \in N$, $w \geq 0$ and $T \in \mathbb{R}$), then $\dfrac{\partial \lambda^{x(\beta)}}{\partial \beta} > 0$ along the Pareto frontier, regardless of the sign of $\dfrac{\partial v(\beta, \underline{w}, R^*)}{\partial \beta}$. Second, when the full-support assumption is dropped, there is an economy in which the DT allocation is the Pareto efficient allocation which *minimizes* the welfare of the low-skilled agent (Nishimura (2003a), Proposition 3).

7 Concluding Remarks

This paper has examined the second-best tax policy to minimize envy in the sense of Chaudhuri (1986) and Diamantaras and Thomson (1990). We have first shown that the Diamantaras-Thomson allocations coincide with the leximin allocation when consumption and leisure are normal goods in the two-class economy (Proposition 1). We conjecture that the result may not hold when there is an inferior good, which we leave as an open question to be addressed in future work. This also relates to the point discussed by Dixit and Seade (1979) in the utilitarian case, who showed that there is a possibility of regressive tax policy when there is an inferior good. In many-agent economies, it turned out that the property of income elasticity of leisure is cricical. The assumption of homothetic preferences is a good benchmark case. We showed that, under homotheticity, the DT criterion selects the most progressive taxation in the sense of Hemming and Keen (1983) and Gans and Smart (1996) among those which can be ordered with respect to progressivity (Propositions 2, 4 and 5). However, the ranking with respect to single-crossing is partial, and we have shown that envy reduction may be justified by reducing welfare of both the highest- and the lowest-skilled agents (Proposition 3). Given that the analytical features of the second-best tax schedules are highly complex, our results provide simple and useful policy prescriptions for reduction of envy, in policy-relevant circumstances such as single- versus double-crossing tax reforms and the amount of subsidies given to the hard-working poor.

Our results can be generalized to an economy with many commodities in the following way. When a fully non-linear tax policy is feasible (where the self-selection condition is the necessary and sufficient condition for tax implementability), the result of Proposition 1.1 holds. When the feasible policies are restricted to a linear commodity taxation and non-linear income taxation (Edwards *et al.* (1994), Guesnerie (1995)), (i) the results of Proposition 1.1, Proposition 2 and Proposition 4 hold when R_i can be represented by $u_i(c_i, l_i) = u_i(f(c_i), l_i)$ for all i, where $f : \mathbb{R}_+^k \to \mathbb{R}_+$ is common for all individuals (a generalized form of weak separability discussed in Atkinson and Stiglitz (1976)), and (ii) the result of Proposition 1.2 holds under identical homotheticity or $u_i(c_i, l_i) = u(f(c_i), l_i)$ for all i with $f : I\!R^k_+ \to I\!R_+$ bing homothetic. The proofs are straightforward extensions of those of Propositions 1, 2 and 4. Generalizations of these results would be left for future researches.

The present paper is part of a larger project in welfare economics to address the equity-efficiency trade-off. As a complementary approach to this paper, Nishimura (2003b) characterized the shape of non-linear income taxes and commodity taxes when the government is concerned not just with social welfare but also with the reduction of envy, based on Varian's (1976) generalized social welfare function and Tadenuma's (2002) Equity-first Principle. No-envy and its alternatives have been one of the most celebrated concepts in the social choice literature. More broadly, the analysis of tax policy implications based on non-welfaristic value judgments would deepen the discussion of the normative analysis of taxation. It is also an important field of the application of the theory of distributive justice.

Appendix

Proof of Lemma 3

By Lemma 2, it is sufficient to show that λ^x decreases along the regressive area (the Pareto efficient allocations in panel (c) of Figure 2), as the transfer to agent 1 decreases (or equivalently, the allocation is less favorable to agent 1).

Let $u(c,l)$ be a utility function that represents R. In the rgressive area, allocation $x = ((c_l, l_1), (c_2, l_2))$ is characterized by: (i) $c_1 = w_1 l_1 - T_1$ (where T_1 is the tax payment by agent 1), (ii) $c_2 = w_2 l_2 + T_1$ (fiscal budget balance), (iii) $u(c_1, l_1) = u(c_2, w_2 l_2 / w_1)$ (SS_1 is binding), and (iv) $u_c w_1 + u_l = 0$ (where $u_r \equiv \partial u / \partial r|_{(c,l)=(c_1,l_1)}$ ($r = c,\ l$)) or $c_1 = 0$ ($(c_1, w_1 l_1) \in ICC_1^*$).[26] Differentiation of (iii) taking account of (i), (ii) and (iv) yields:

$$\frac{u_l}{w_1} dT_1 = \hat{u}_c (w_2 dl_2 + dT_1) + \hat{u}_l \frac{w_2}{w_1} dl_2,$$

where $\hat{u}_r \equiv \partial u / \partial r|_{(c,l)=(c_2, \frac{w_2}{w_1} l_2)}$ ($r = c,\ l$).[27] Re-arranging, we obtain:

$$\frac{dl_2}{dT_1} = \frac{u_l / w_1 - \hat{u}_c}{\hat{u}_c + \hat{u}_l / w_1} \frac{1}{w_2}. \tag{8}$$

We now consider the change in λ^x by changing one tax parameter T_1, by differentiating $u(c_1, l_1) = u(\lambda^x \circ (c_2, l_2))$, taking account of (i)-(iv):

$$\frac{d\lambda^x}{dT_1} = \frac{u_l / w_1 - \bar{u}_c \lambda^x (w_2 dl_2 / dT_1 + 1) - \bar{u}_l dl_2 / dT_1}{\bar{u}_c c_2 - \bar{u}_l (\bar{l} - l_2)},$$

where $\bar{u}_r \equiv \partial u(\lambda^x \circ (c_2, l_2)) / \partial r$ ($r = c, l$). The denominator of the right hand side is positive. Substituting (8), the numerator is:

$$\begin{aligned} & \frac{u_l}{w_1} - \bar{u}_c \lambda^x \left(w_2 \frac{dl_2}{dT_1} + 1 \right) - \bar{u}_l \frac{\lambda^x dl_2}{dT_1} \\ &= \lambda^x \frac{\frac{u_l / w_1}{\lambda^x} (\hat{u}_c + \hat{u}_l / w_1) - \bar{u}_c (u_l / w_1 - \hat{u}_c + \hat{u}_c + \hat{u}_l / w_1) - \bar{u}_l (u_l / (w_1 w_2) - \hat{u}_c / w_2)}{\hat{u}_c + \hat{u}_l / w_1} \end{aligned} \tag{9}$$

[26] The condition of $c_1 = 0$ refers to the case of the corner solution. For other possibilities: (a) $l_1 = 0$ belongs to a non-regressive area (see Figure 2). (b) The possibility of $l_1 = \bar{l}$ and $w_2 l_2 > w_1 l_1$ are excluded from the regressive area (see footnote 17). Notice that $w_2 l_2 > w_1 l_1$ (e.g., Stiglitz (1982), p.230).

[27] The left hand side of the above equation is $u_c (w_1 dl_1 - dT_1) + u_l dl_1$. When $c_1 > 0$, we apply (iv). When $c_1 = 0$, $w_1 l_1 = T_1$, so that $dl_1 = dT_1 / w_1$.

Notice that $w_2 l_2 > w_1 l_1$ (e.g., Stiglitz (1982), p.230). By (iii) and (iv), the denominator of the right hand side of (9) is negative. The numerator is manipulated to yield:

$$\begin{aligned}
&\frac{u_l/w_1}{\lambda^x}\left(\hat{u}_c + \frac{\hat{u}_l}{w_1}\right) - \bar{u}_c\left(\frac{u_l}{w_1} + \frac{\hat{u}_l}{w_1}\right) - \bar{u}_l\left(\frac{u_l}{w_1 w_2} - \frac{\hat{u}_c}{w_2}\right)\\
&= \frac{u_l \hat{u}_l/w_1}{\lambda^x}\left(\frac{\hat{u}_c}{\hat{u}_l} + \frac{1}{w_1}\right) - \frac{u_l \bar{u}_l}{w_1}\left(\frac{\bar{u}_c}{\bar{u}_l} + \frac{1}{w_2}\right) + \bar{u}_c \hat{u}_c\left(\frac{-\hat{u}_l}{\hat{u}_c w_1} + \frac{\bar{u}_l}{\bar{u}_c w_2}\right) \qquad (10)
\end{aligned}$$

By construction, $-\hat{u}_l/\hat{u}_c + \bar{u}_l/\bar{u}_c > 0$, so that, combining with $w_1 < w_2$, $\bar{u}_c \hat{u}_c(-\hat{u}_l/(\hat{u}_c w_1) + \bar{u}_l/(\bar{u}_c w_2)) > 0$, i.e., the last term of (10) is positive. Consider now the first and the second terms. $\frac{u_l \hat{u}_l/w_1}{\lambda^x}(\hat{u}_c/\hat{u}_l + 1/w_1)$ is positive. If $\bar{u}_c/\bar{u}_l + 1/w_2 < 0$, the remaining term is also positive. Otherwise, $-\hat{u}_c/\hat{u}_l < -\bar{u}_c/\bar{u}_l \leq 1/w_2 < 1/w_1$,[28] so that $\hat{u}_c/\hat{u}_l + 1/w_1 > \bar{u}_c/\bar{u}_l + 1/w_2$. Also, $\hat{u}_l \leq \bar{u}_l$ under normality, and $\lambda^x < 1$. Combining these facts, (10) is positive.

By (9) and (10), we conclude that $d\lambda^x/dT_1 < 0$. *Q.E.D.*

Proof of Lemma 4

Before proving Lemma 4, we first introduce the following notations. Let $I(a, R) \equiv \{b \in I\!R_+ \times [0, \bar{l}] | bIa\}$ be an indifference curve in (c, l)-space, and $U(a, R) \equiv \{b \in I\!R_+ \times [0, \bar{l}] | bRa\}$ be the upper counter set, and $L(a, R) \equiv \{b \in I\!R_+ \times [0, \bar{l}] | aRb\}$ be the lower counter set. Let $\hat{I}((c, l), \lambda, R) \equiv \{(\hat{c}, \hat{l}) \in I\!R_+ \times [0, \bar{l}] | \exists\ (\tilde{c}, \tilde{l}) \in I((c, l), R) \text{ s.t. } \lambda\hat{c} = \tilde{c},\ \lambda(\bar{l} - \hat{l}) = \bar{l} - \tilde{l}\}$ be a homothetic expansion of $I((c, l), R)$ with proportion $\frac{1}{\lambda}$, and $\hat{L}((c, l), \lambda, R) \equiv \{(\hat{c}, \hat{l}) \in I\!R_+ \times [0, \bar{l}] | \exists\ (\tilde{c}, \tilde{l}) \in I((c, l), R) \text{ s.t. } \lambda\hat{c} \leq \tilde{c},\ \lambda(\bar{l} - \hat{l}) \leq \bar{l} - \tilde{l}\}$. The following lemma is helpful.

[28] Let $u_r^2 \equiv \partial u/\partial r|_{(c,l)=(c_2,l_2)}$ $(r = c,\ l)$. Then $-u_l^2/u_c^2 > w_2$ in the regressive area (see Figure 2). By normality, $-\hat{u}_l/\hat{u}_c \geq -u_l^2/u_c^2$, so that $-\hat{u}_l/\hat{u}_c > w_2$. By construction, $-\hat{u}_l/\hat{u}_c > -\bar{u}_l/\bar{u}_c$.

Lemma 6 *For any allocation* $x = (c_i, l_i)_{i=1}^n$ *and agents* j *and* k *with* $R_j = R_k \equiv R$, *if* $x_k\ R\ x_j$, *and the income elasticity of leisure is greater than or equal to unity, then* $\{(c,l)|c \leq c_k,\ l \leq l_k\} \cap L(x_k, R) \subset \{(c,l)|c \leq c_k,\ l \leq l_k\} \cap \hat{L}(x_j, \lambda^x_{jk}, R)$.

Proof: Let take any $(c^*, l^*) \in I(x_k, R)$ and let $\hat{w}$ be the lowest value of w such that $U((c^*, l^*), R) \subset \{(c,l)|c - wl \geq c^* - wl^*\}$, i.e., the lowest wage level generating the compensated demand (c^*, l^*). Let $(\hat{c}, \hat{l})$ be the intersection of $I(x_j, R)$ and the set $\{(c,l) \in I\!R_+ \times [0, \bar{l}]|(c,l) = (c_j(\hat{w}, T), l_j(\hat{w}, T))$ for some $T \in I\!R\}$. Let $(\tilde{c}, \tilde{l})$ be the intersection of $I(x_j, R)$ and the straight line connecting the point $(0, \bar{l})$ and (c^*, l^*). By continuity, monotonicity and normality of demand, the points $(\hat{c}, \hat{l})$ and $(\tilde{c}, \tilde{l})$ are uniquely determined. Also, since income elasticity of leisure is greater than or equal to unity, $\hat{c} \geq \tilde{c}$ and $\hat{l} \geq \tilde{l}$. Let $\tilde{w}$ be the lowest value of w such that $U((\tilde{c}, \tilde{l}), R) \subset \{(c,l)|c - wl \geq \tilde{c} - w\tilde{l}\}$. By strict convexity, we obtain $\tilde{w} \leq \hat{w}$ for all $(c^*, l^*) \in I(x_k, R)$, which derives the desired result. *Q.E.D.*

Proof of Lemma 4: Let $x = (c_i, l_i)_{i=1}^n$ be any self-selective allocation. It is straightforward to show that $\lambda^x_{1j} \leq \lambda^x_{ij}\ \forall\ i,\ j \in N$.

Consider first the case where the income elasticity of leisure is greater than or equal to unity. By AM, $c_i \leq c_n$ for all i. If $l_i \geq l_n$, apparently $\lambda^x_{1i} \geq \lambda^x_{1n}$. It is left to show that $\lambda^x_{1i} \geq \lambda^x_{1n}$ when $l_n \geq l_i$. By Lemma 6, $\{(c,l)|c \leq c_n,\ l \leq l_n\} \cap L((c_n, l_n), R) \subset \{(c,l)|c \leq c_n,\ l \leq l_n\} \cap \hat{L}((c_1, l_1), \lambda^x_{1n}, R)$. Combining with the observation that $\{(c_j, l_j)|w_j \leq w_i\} \subset L((c_i, l_i), R)$ for all i at any self-selective allocation ((3)), it follows that for any (c_i, l_i), $i = 2, ..., n-1$, with $l_i \leq l_n$, $(c_i, l_i) \in \hat{L}((c_1, l_1), \lambda^x_{1n}, R)$, which implies that $\lambda^x_{1i} \geq \lambda^x_{1n}$.

Consider next a three-agent economy where R is represented by the following utility function:

$$u(c,l) = \min\{c - s^1 l, c - s^2 l + (s^2 - s^1)\hat{l}^1, c - s^3 l + (s^3 - s^2)\hat{l}^2 + (s^2 - s^1)\hat{l}^1\} - \epsilon(l)^2,$$

where we set $\hat{l}^1 = 0.5,\ \hat{l}^2 = 0.9,\ s^1 = 0.1,\ s^2 = 2,\ s^3 = 4,\ w_1 = 0.2,\ w_2 = 2,\ w_3 = 2.1,$ $\epsilon = 0.000001$ and $\bar{l} = 1$. In this utility function, preferences are quasi-linear with respect to consumption, and thus the income elasticity of leisure is 0. Consider the competitive equilibrium allocation $x = (x_1, x_2, x_3) = ((0.1, 0.5), (1, 0.5), (1.89.0.9))$. This allocation is second-best Pareto efficient. In this economy, $\lambda^x = \lambda^x_{12} = 0.55000 < \lambda^x_{13} = 0.56769$. *Q.E.D.*

Proof of Lemma 5

Suppose $x{\equiv}(c_i, l_i)_{i=1}^n$ single-crosses z and $x_i\ P\ z_i$. Let $T^x(\cdot)$ and $T^z(\cdot)$ be income tax schedules that support x and z, respectively, such that there exists y^* such that $T^x(y) \leq T^z(y)$ for all $y \leq y^*$ and $T^x(y) \geq T^z(y)$ for all $y \geq y^*$. Let $h^x(y) \equiv y - T^x(y)$ and $h^z(y) \equiv y - T^z(y)$. By the weak axiom of revealed preference, $h^x(w_i l_i) > h^z(w_i l_i)$. Single-crossing implies that $h^x(y) \geq h^z(y)$ for all $y \leq w_i l_i$. By AM, $w_i l_i \geq w_j l_j$ for all $j \in N$ such that $w_j < w_i$. Since $T^x(\cdot)$ supports x, then $(c_j, w_j l_j)\ \tilde{R}_j\ (h^x(y), y)$ for all $y \leq w_j \bar{l}$, and thus $(c_j, w_j l_j)\ \tilde{R}_j\ (h^z(y), y)$ for all $y \leq \min\{w_j \bar{l}, w_i l_i\}$. If $w_i l_i > w_j \bar{l}$, we reach the desired conclusion of $x_j\ R\ z_j$. Suppose that $w_i l_i < w_j \bar{l}$. Let $\tilde{L}(a, \tilde{R}_i) \equiv \{b \in I\!R_+ \times [0, w_i \bar{l}] | a\ \tilde{R}_i\ b\}$. We then have $\{(h^z(y), y) | w_j \bar{l} \geq y \geq w_i l_i\} \subset \tilde{L}((c_i, w_i l_i), \tilde{R}_i)$ by $x_i\ P\ z_i$, and $\{(c, y) | c \geq$

$c_i,\ w_j\bar{l} \geq y \geq w_i l_i\} \cap \tilde{L}((c_i, w_i l_i), \tilde{R}_i) \subset \{(c,y)|c \geq c_i,\ w_j\bar{l} \geq y \geq w_i l_i\} \cap \tilde{L}((c_i, w_i l_i), \tilde{R}_j)$ by AM. Since $(c_j, w_j l_j)\ \tilde{R}_j\ (c_i, w_i l_i)$ by self-selection, we conclude that $(c_j, w_j l_j)\ \tilde{R}_j\ (h^z(y), y)$ for all $y \leq w_j\bar{l}$, and $x_j\ R\ z_j$, the desired conclusion. The same logic applies if $z_i\ P\ x_i$ and $w_j > w_i$. *Q.E.D.*

Proof of Proposition 4

Let $x = (x_k)_{k=1}^n$ be any self-selective allocation, and let $i,\ j \in N$ be such that $(R_i, w_i) = (R^*, \underline{w})$ and $(R_j, w_j) = (R^*, \overline{w})$. We will show that $\lambda^x = \lambda^x_{ij}$. To see this, first, for any $k \in N$, if $w_k = \underline{w} = 0$, then $x_k = (c, 0)$ for some c. Second, AM, (3) and the definition of R^* imply $x_j\ R^*\ x_k$ and $c_j \geq c_k$ for all $k \in N$. By an analogy to Lemma 6, $x_k \in \hat{L}(x_j, \lambda^x_{ij}, R^*)$ for all $k \in N$. Third, as $U(x_k, R_k) \subset U((c,0), R_k) \subset U((c,0), R^*)$ for all $k \in N$, if $x_m\ \tilde{R}_k\ x_k$ for some $m \in N$, then the straight line connecting x_m and $(0, \bar{l})$ crosses $I(x_k, R_k)$ at $U(x_i, R^*)$. As a result, $\lambda^x_{ij} \leq \lambda^x_{km}$ for all $k, m \in N$. We then have $\lambda^x = \lambda^x_{ij}$. The rest of the proof is the same as that of Proposition 1.3 in Nishimura (2003a) and Proposition 2. *Q.E.D.*

Proof of Proposition 5

Let $x = (x_k)_{k=1}^n$ be any self-selective allocation. Let take any $R \in \mathcal{R}$, and Let $i,\ j \in N$ be such that $R_i = R_j = R$, $w_i = \underline{w}$ and $w_j = \overline{w}$. We first show that $\min_{m,k \in N,\ R_k = R} \lambda^x_{km} = \lambda^x_{ij}$ for all $R \in \mathcal{R}$. When preferences are homothetic, $\hat{I}(x_i, \lambda^x_{ij}, R_i) = I(x_j, R_i)$ for all $i,\ j \in N$ such that $R_i = R_j$ and $w_i \leq w_j$, i.e., homothetic expansion coincides with the agent's indifference curve. From (3), $x_j\ R_j\ x_k$ for all $j,\ k \in N$ with $w_j \geq w_k$. It follows that, if $R_i = R_j,\ w_i = \underline{w}$ and $w_j = \overline{w}$, then $x_k \in \hat{L}(x_j, \lambda_{ij}^x, R_i)$ for all $k \in N$, so that $\lambda_{ij}^x \leq \lambda_{ik}^x$ for all $k \in N$. As in Lemma 4, $\lambda_{ik} \leq \lambda_{mk}$ for all $i, k, m \in N$ such that $R_i = R_m$ and $w_i \leq w_m$. Therefore, $\min_{m,k \in N,\ R_k = R} \lambda_{km}^x = \lambda_{ij}^x$ for all $R \in \mathcal{R}$. As a result, $\lambda^x = \min\{\lambda_{ij}^x | i,\ j \in N,\ R_i = R_j,\ w_i = \underline{w}$ and $w_j = \overline{w}\}$.

Define $\alpha(\beta)$ implicitly by $\alpha(\beta)n - \sum_{i=1}^n \beta w_i \hat{l}_i((1-\beta)w_i, -\alpha(\beta)) = 0$. Then one can show that $\dfrac{\partial v(\beta, w_k, R_k)}{\partial \beta} = \gamma_k \left(-w_k l_k + \dfrac{\partial \alpha}{\partial \beta}\right)$, where $\gamma_k > 0$ is agent k's marginal utility of income (that is, the Lagrange multiplier of $\max_{c,l} u(c, l, R)$ $s.t.$ $c = (1-\beta)w_k l + \alpha$). By assumption for the preference for leisure, $\hat{l}_i((1-\beta)\underline{w}, -\alpha) > \hat{l}_j((1-\beta)\underline{w}, -\alpha)$ for all $i,\ j \in N$ such that $R_i = R^*$ and $R_j = R,\ R \neq R^*$. Therefore, $\dfrac{\partial v(\beta, \underline{w}, R^*)}{\partial \beta} \geq 0 \rightarrow \dfrac{\partial v(\beta, \underline{w}, R)}{\partial \beta} > 0$ for all $R \in \mathcal{R}$.

Next, any utility function that represents a homothetic preference is such that $u(c, l, R) = \hat{u}(c, \bar{l} - l, R) = g(\phi^R(c, \bar{l} - l))$, $g : I\!R \rightarrow I\!R$ is an increasing function, and $\phi^R : I\!R_+ \times [0, \bar{l}] \rightarrow I\!R_+$ is homogeneous of degree 1. Then,

$$u(c, l, R_i) = u(\lambda \circ (\hat{c}, \hat{l}), R_i) \iff \phi^{R_i}(c, \bar{l} - l) = \lambda \phi^{R_i}(\hat{c}, \bar{l} - \hat{l}).$$

For the purpose of the proof, we only need to examine the case where $u(c, l, R) = \phi^R(c, \bar{l} - l)$.

Let take any $R \in \mathcal{R}$, and let $i,\ j \in N$ be such that $(R_i, w_i) = (R, \underline{w})$ and $(R_j, w_j) = (R, \overline{w})$. We now differentiate $v(\beta, \underline{w}, R) = u(\lambda_{ij}^x \circ (x_j), R)$ with respect to β. We then obtain:

$$\frac{\partial \lambda_{ij}^x}{\partial \beta} = \frac{\gamma_i(-w_i l_i + \partial \alpha / \partial \beta) - \lambda_{ij}^x \gamma_j \left(-w_j l_j + \partial \alpha / \partial \beta\right)}{\bar{u}_c^j c_j - \bar{u}_l^j (\bar{l} - l_j)}, \tag{11}$$

where $\bar{u}_r^j \equiv \partial u(\lambda_{ij}^x \circ (c_j, l_j), R)/\partial r$ $(r = c, l)$. The denominator of the right hand side of (11) is positive. We now check the numerator. By homogeneity of degree 1, $\gamma_k = \dfrac{v(\beta, w_k, R)}{(1-\beta)w_k\bar{l} + \alpha}$ $(k = i,\ j)$. Noting also that $v(\beta, w_i, R) = u(\lambda_{ij}^x \circ (x_j), R) = \lambda_{ij}^x v(\beta, w_j, R)$, the numerator of the right hand side of (11) is rearranged to yield:

$$\begin{aligned}
&\gamma_i\left(-w_i l_i + \frac{\partial\alpha}{\partial\beta}\right) - \lambda_{ij}^x \gamma_j \left(-w_j l_j + \frac{\partial\alpha}{\partial\beta}\right) = \\
&v(\beta, w_i, R)\left(\left(\frac{1}{Z_i} - \frac{1}{Z_j}\right)\left(-w_i l_i + \frac{\partial\alpha}{\partial\beta}\right) + \frac{1}{Z_j}(w_j l_j - w_i l_i)\right),
\end{aligned}$$

where $Z_k \equiv (1-\beta)w_k\bar{l} + \alpha$ $(k = i\ j)$. Notice that $Z_i < Z_j$, $-w_i l_i + \partial\alpha/\partial\beta \geq 0$ (with strict inequality if $R_i \neq R^*$), and $w_j l_j \geq w_i l_i$ by AM (with strict inequality if $l_j > 0$). Hence, the value of (11) is positive at all $i,\ j \in N$ such that $(R_i, w_i) = (R, \underline{w})$ and $(R_j, w_j) = (R, \overline{w})$. As a result, we obtain $\dfrac{\partial\lambda^{x(\beta)}}{\partial\beta} > 0$.

An alternative expression of the numerator of the right hand side of (11) is:

$$v(\beta, w_i, R)\left(\frac{(1-\beta)w_i w_j \bar{l}(l_j - l_i) + \alpha(w_j l_j - w_i l_i)}{Z_i Z_j} + \left(\frac{1}{Z_i} - \frac{1}{Z_j}\right)\frac{\partial\alpha}{\partial\beta}\right).$$

$\partial\alpha/\partial\beta > 0$ along the Pareto frontier. If $l_j \geq l_i$ (i.e., under $\dfrac{\partial\hat{l}_i(w,T)}{\partial w} \geq 0$), then the expression is positive, regardless of the sign of $\dfrac{\partial v(\beta, \underline{w}, R^*)}{\partial\beta}$. *Q.E.D.*

References

Arnsperger, C. (1994). Envy freeness and distributive justice. *Journal of Economic Surveys* 8, 155–185.

Atkinson, A. and J. Stiglitz (1976). The design of tax structure — direct versus indirect tax. *Journal of Public Economics* 6, 55–75.

Boadway, R. (2002). The role of public choice considerations in normative public economics. In S. L. Winer and H. Shibata (Eds.), *Political economy and public finance: the role of political economy in the theory and practice of public economics*, pp. 47–68. Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing Limited.

Boadway, R., K. Cuff, and M. Marchand (2000). Optimal income taxation with quasi-linear preferences revisited. *Journal of Public Economic Theory* 2, 435–460.

Boadway, R., M. Marchand, P. Pestieau, and M. del Mar Racionero (2002). Optimal redistribution with heterogeneous preferences for leisure. *Journal of Public Economic Theory* 4, 475–498.

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

Bossert, W., M. Fleurbaey, and D. Van de gear (1999). Responsibility, talent and compensation: a second-best analysis. *Review of Economic Design* 4, 35–55.

Brito, D., J. Hamilton, S. Slutsky, and J. Stiglitz (1990). Pareto efficient tax structures. *Oxford Economic Papers* 42, 61–77.

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

d'Aspremont, C. and L. Gevers (1977). Equity and the informational basis of collective choice. *Review of Economic Studies* 44, 199–209.

Diamantaras, D. and W.Thomson. (1990). A refinement and extension of the no-envy concept. *Economics Letters* 33, 217–222.

Diamond, P. (1998). Optimal income taxation: an example with a u-shaped pattern of optimal marginal tax rates. *American Economic Review* 88, 83–95.

Dixit, A. and J. Seade. (1979). Utilitarian versus egalitarian redistributions. *Economics Letters* 4, 121–124.

Ebert, U. (1992). A reexamination of the optimal nonlinear income tax. *Journal of Public Economics* 49, 47–73.

Edwards, J., M. Keen, and M. Tuomala (1994). Income tax, commodity tax and public good provision: a brief guide. *Finanzarchiv* 51, 472–487.

Feldman, A. and A. Kirman. (1974). Fairness and envy. *American Economic Review* 63, 995–1005.

Fleurbaey, M. and F. Maniquet (1998). Optimal income taxation: an ordinal approach. CORE Discussion Paper No.9865 http://www.core.ucl.ac.be/services/psfiles/dp98/DP9865.PDF.

Fleurbaey, M. and F. Maniquet (2006). Fair income tax. *Review of Economic Studies* 73, 55–83.

Fleurbaey, M. and F. Maniquet (2007). Help the low-skilled or let the hardworking thrive? A study of fairness in optimal income taxation. *Journal of Public Economic Theory* 9, 467–500.

Gans, J. and M. Smart (1996). Majority voting with single-crossing preferences. *Journal of Public Economics* 59, 219–237.

Guesnerie, R. (1995). *A contribution to the pure theory of taxation*. Econometric Society Monographs, Cambridge: Cambridge University Press.

Guesnerie, R. and J. Seade (1982). Nonlinear pricing in a finite economy. *Journal of Public Economics* 17, 157–179.

Hammond, P. (1976). Equity, Arrow's conditions and Rawls' difference principle. *Econometrica* 44, 793–804.

Hellwig, M. (1986). The optimal linear income tax revisited. *Journal of Public Economics* 31, 163–179.

Hemming, R. and M. Keen (1983). Single-crossing conditions in comparisons of tax progressivity. *Journal of Public Economics* 20, 373–380.

Lambert, P. (1993). *The distribution and redistribution of income: a mathematical analysis*. Manchester: Manchester University Press.

Mirrlees, J. (1971). An exploration in the theory of optimal taxation. *Review of Economic Studies* 38, 175–208.

Moulin, H. and W. Thomson (1997). Axiomatic analysis of resource allocation problems. In K. Arrow, A. Sen, and K. Suzumura (Eds.), *Social choice re-examined,* Volume 1, pp. 101–120. London: Macmillan and New York: St. Martin's Press.

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

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

Phelps, E. (1973). The taxation of wage income for economic justice. *Quarterly Journal of Economics* 87, 331–354.

Rawls, J. (1971). *A theory of justice*. Cambridge MA: Harvard University Press.

Rawls, J. (1974). Concepts for distributional equity: some reasons for the maximin principle. *American Economic Review, Papers and Proceedings* 64, 141–145.

Roemer, J. (1996). *Theories of distributive justice*. Harvard University Press.

Roemer, J. (1998). *Equality of opportunity*. Harvard University Press.

Roemer, J., R. Aaberge, U. Colombino, J. Fritzell, S. Jenkins, A. Lefranc, I. Marx, M. Page, E. Pommer, J. Ruiz-Castillo, M. S. Segundo, T. Tranaes, A. Trannoy, G. Wagner, and I. Zubiri (2003). To what extent do fiscal regimes equalize opportunities for income acquisition among citizens? *Journal of Public Economics* 87, 539–565.

Sadka, E. (1976). On income distribution, incentive effects and optimal income taxation. *Review of Economic Studies* 43, 261–268.

Sandmo, A. (1993). Optimal redistribution when tastes differ. *Finanzarchiv* 50, 149–163.

Schokkaert, E., D. V. de gaer, F. Vandenbroucke, and R. Luttens (2004). Responsibility sensitive egalitarianism and optimal linear income taxation. *Mathematical Social Sciences* 48, 151–182.

Seade, J. (1977). On the shape of optimal income tax schedule. *Journal of Public Economics* 43, 203–235.

Seade, J. (1982). On the sign of optimum marginal income taxation. *Review of Economic Studies* 49, 637–643.

Sheshinski, E. (1972). The optimal linear income tax. *Review of Economic Studies* 39, 297–302.

Stern, N. (1976). On the specification of models of optimum income taxation. *J.Pub.E.* 6, 123–162.

Stiglitz, J. (1982). Self-selection and pareto efficient taxation. *J.Pub.E.* 17, 213–240.

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

Tarkiainen, R. and M. Tuomala (1999). Optimal nonlinear income taxation with a two-dimensional population; a computational approach. *Computational Economics* 13, 1–16.

Thomson, W. and H. Varian. (1985). Theory of justice based on symmetry. In L. Hurwicz, D. Schmeidler, and H. Sonnenshein (Eds.), *Social goals and social organization*, pp. 107–129. Cambridge: Cambridge University Press.

Tillmann, G. (1984). Equity and taxation. In D. Bös, M. Rose, and C. Seidl (Eds.), *Beitrage zur neueren steuetheorie*, pp. 75–94. Berlin: Springer.

Tillmann, G. (1989). *Equity, incentives, and taxation.* Berlin: Springer-Verlag.

Tillmann, G. (2005). The equity-efficiency trade-off reconsidered. *Social Choice and Welfare* 24, 63–81.

Tuomala, M. (1990). *Optimal income tax and redistribution.* Oxford: Clarendon Press.

Varian, H. (1974). Equity, envy and efficiency. *Journal of Economic Theory* 4, 63–90.

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

Weymark, J. (1986). A reduced-form nonlinear tax problem. *Journal of Public Economics* 30, 199–217.

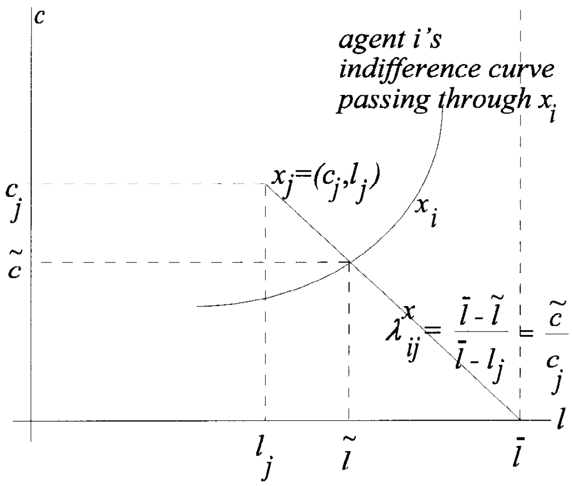

Figure 1: λ^x_{ij}

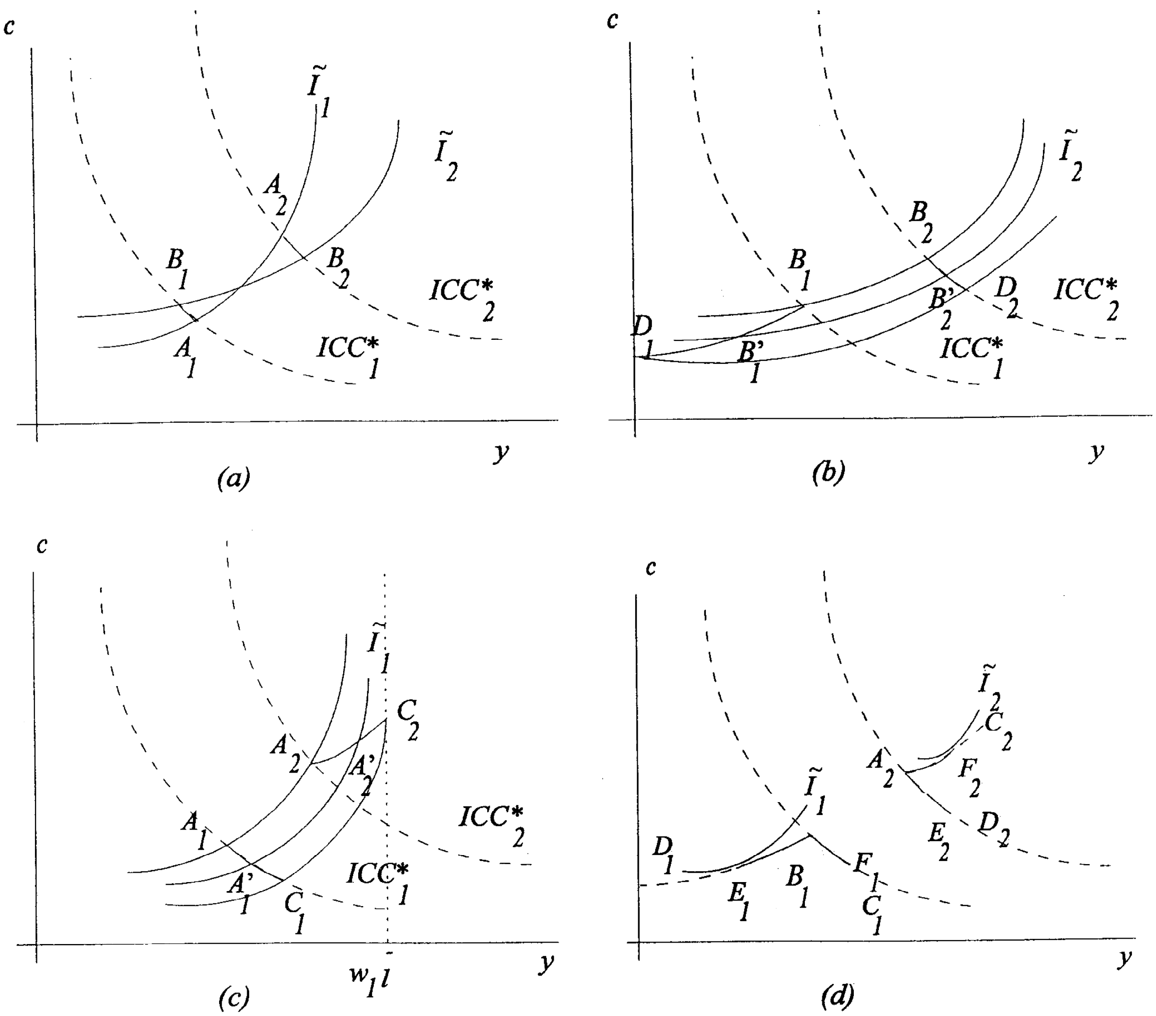

Figure 2: Characterization of the Second-best Pareto Optima