Optimality Conditions and Comparative Static Properties of Optimal Nonlinear Income Taxes Revisited

Laurent SIMULA
EHESS, PSE, GREQAM and IDEP*

March 12, 2007

Abstract

Comparative static properties of the optimal Mirrleesian nonlinear income tax are obtained for a finite population and quasilinear in consumption preferences. Contrary to Weymark (1987) who considers quasilinear in leisure preferences, the linearity with respect to the variable observed by the government and used as a tax base is lost. A reduced-form optimal income tax problem is derived, in which consumption levels are obtained as functions of gross incomes. The contribution of this new reduced form is twofold. First, the optimal allocation can be characterized geometrically in a simple way. Second, comparative static results with respect to individual productivities are easy to obtain.

Keywords: Optimal Tax, Income Tax, Comparative Statics.

*Affiliations: Ecole des Hautes Etudes en Sciences Sociales, Paris School of Economics, Research Group in Quantitative Economics of Aix-Marseille and Institute for Public Economics. Corresponding Address: Paris School of Economics, 48 boulevard Jourdan, 75014 Paris, France. Email: Laurent.Simula@ehess.fr
1. INTRODUCTION

This paper uses a reduced form of the optimal non-linear income tax problem to derive a geometric characterization of the social optimum in the gross-income/consumption space as well as comparative static properties. Following Guesnerie and Seade (1982) and Weymark (1987), it considers a finite population version of Mirrleses (1971)'s model. All individuals have the same preferences over consumption and leisure, but differ in skill levels. The government wants to redistribute income from the more to the less productive individuals. However, if the distribution of this parameter within the population is common knowledge, each agent’s productivity is private information. Accordingly, the government is restricted to setting taxes as a function of earnings and faces an adverse selection problem when designing the optimal income tax schedule.

The optimal tax structure is the product of different sort of interacting influences. It basically depends on the skill distribution (Diamond, 1998, Saez, 2001), on the government’s aversion to income inequality, reflected by the welfare weights in the social objective function, but also on the responsiveness of labour supply. In addition, the way in which all these influences interact is affected both by the incentive-compatibility constraints and the tax revenue constraint, which restrict the possibilities for income redistribution. Because of the complexity of the relationship between the optimal tax schedule and the set of underlying parameters, investigations must usually resort to numerical simulations (Tuomala, 1990). This is an unfortunate state of affairs because some features of the model are necessarily left somewhat obscure by such an approach, which is very useful since it allows the optimal tax rates to be quantified, but is not ideally suited for shedding light on the economic intuition behind the results.

In a pioneering paper, Weymark (1987) has derived a number of comparative static results of optimal non-linear income taxes for the case in which individual preferences are quasilinear in leisure. Their derivation uses a reduced form of the optimal tax problem, which does only involve the choice of the consumption good (Weymark, 1986b). This methodology has been adapted to obtain comparative static properties for a model in which the government both designs an optimal income tax and provides a public good optimally (Brett and Weymark, 2004). It has also been extended to analyse a two-period model in which income and savings are taxed optimally (Brett and Weymark, 2005). The assumption that individual preferences are quasilinear in leisure is maintained in both papers. The disutility of effort is therefore constant. In other words, when a price is varied, the change in individual consumption does only depend on the substitution effect while all income effects are absorbed by the labour supply. This preference specification has also been employed to derive an explicit solution to the optimal income tax problem with a continuous population (Lollivier and Rochet, 1983), the properties of which
have been investigated by Boadway, Cuff, and Marchand (2000). Hamilton and Pestieau (2005) have used it to derive some comparative static results with respect to individual productivity in an economy with two classes of agents where the government adopts a maximin or maximax objective function. Its tractability has been exploited by Ebert (1992) to provide a complete example in which different types of individuals are bunched together, establishing that the first-order approach to Mirrlees (1971)'s model can be misleading.

Quasilinear-in-leisure preferences offers technical advantages. They are indeed linear with respect to gross income, i.e. to the variable observed by the government and used as the tax base. The linearity with respect to individual productivity of the social objective in Weymark (1986b)'s reduced form stems from this utility specification and allows the reduced-form optimal income tax problem to have an explicit solution. Hence, assuming quasilinear-in-leisure preferences is somewhat more restrictive than assuming quasilinear-in-consumption preferences. Indeed, when considering the latter, the linearity with respect to the observable variable is lost, the social objective of the reduced-form is no longer linear in productivity and its maximization does not yield an explicit solution.

Although working with them is less tractable, quasilinear-in-consumption preferences are worth examining for at least three reasons. First, from the theoretical viewpoint, assuming that all income effects are absorbed by consumption is a more satisfying assumption. Otherwise, as is made clear in the continuous population framework, the optimal tax schedule does only depend on the skill distribution and on the social weights (Boadway, Cuff, and Marchand, 2000), but not on the labour response. The rent-extraction trade-off reflected by the income tax schedule is thus very specific. On the contrary, if the income effects on the labour supply are omitted, the optimal tax scheme basically depends on the elasticity of the labour supply as well. Second, most of the empirical studies, though not all, gives credence to small income effects relative to substitution effects as regards labour supply (Blundell, 1992, Blundell and MaCurdy, 1999). Accordingly, the case with no income effects on labour supply provides a useful benchmark, which has been theoretically studied by Atkinson (1990), Diamond (1998), Piketty (1997), Salanié (1998) or d'Autume (2000) and used in the numerical part of other papers (Saez, 2001). Third, the comparative static properties of the optimal non-linear income tax problem could differ significantly from those obtained under quasilinear-in-leisure preferences.

This paper derives a reduced-form optimal non-linear income tax problem involving only the allocation of gross incomes within the population. Consumption levels are thus obtained as a function of gross incomes. The reduced form clearly reflects the trade-off between equity and efficiency. Its firm order conditions receives a clear geometric interpretation. At any given gross-income/consumption bundle designed for a given individual, the angle of the Spence-Mirrlees single-crossing condition specifies to which extent the indifference curve, through this bundle,
of the nearest more productive individual must be flatter. At the social optimum, this angle is entirely determined by the cumulative social weights. The reduced form is then used to provide the comparative statics of the optimal tax schedule with respect to the marginal utility of money and the weights in the welfare function, as in Weymark (1987), but also to individual productivity. Varying the skill level of an individual alters the optimal allocation through three channels: it involves a local substitution effect, an incentive effect and an informational externality.

The paper is organized as follows. Section 2 sets up the model. Section 3 derives the reduced form of the optimal non-linear income tax problem and provides a geometric characterization of the optimal allocation. Section 4 examines the comparative statics of the solution to the optimal income tax problem. Section 5 concludes.

2. THE MODEL

The population consists of \( I \geq 2 \) individuals, indexed by \( i \in I := \{1, \ldots, I\} \). There are two goods, consumption and leisure. Person \( i \)'s consumption and labour supply are denoted \( x_i \) and \( \ell_i \), respectively. The economy is competitive, with constant-returns-to-scale technology, so person \( i \)'s wage rate is fixed and equal to his productivity \( \theta_i \). For convenience, only one person has a given productivity level. Individuals are thus indexed in terms of productivity. Without loss of generality, the vector of productivities \( \theta := (\theta_1, \ldots, \theta_I) \) is taken to be monotonically increasing,

\[
0 < \theta_1 < \ldots < \theta_I. \tag{1}
\]

An individual with productivity \( \theta_i \) working \( \ell_i \) units of time has gross income

\[
z_i := \theta_i \ell_i, \quad i \in I. \tag{2}
\]

All individuals have the same preferences over consumption and leisure, represented by the utility function \( U : \mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R} \),

\[
U(x_i, \ell_i) := \gamma x_i - v(\ell_i), \quad i \in I, \tag{3}
\]

where \( \gamma \in \mathbb{R}_{++} \) is the marginal utility of money. It is assumed that the disutility of labour \( v(\ell_i) \) is a \( \mathcal{C}^4 \)-function which satisfies \( v' > 0, v'' > 0, v''' > 0, v(0) = 0, v'(0) = 0, \) and \( v'(\overline{\ell}) \rightarrow \infty \) where \( \overline{\ell} \) is the time endowment of each individual.

By (2), the utility function (3) can be rewritten as \( U(x_i, \ell_i) = U(x_i, z_i/\theta_i) \). Individuals have
therefore personalized utility functions \( u : \mathbb{R}_+ \times \mathbb{R} \to \mathbb{R} \) in the gross-income/consumption space,

\[
u(x_i, z_i; \theta_i) := \gamma x_i - v\left(\frac{z_i}{\theta_i}\right), \quad i \in \mathcal{I}. \tag{4}\]

The marginal rate of substitution \( s(z_i; \theta_i) \) of the \( \theta_i \)-individual at the \((x_i, z_i)\)-bundle only depends on gross income, with

\[
s(z_i; \theta_i) := \frac{u'_z(x_i, z_i; \theta_i)}{u'_x(x_i, z_i; \theta_i)} = \frac{v'(z_i/\theta_i)}{\gamma \theta_i}, \quad i \in \mathcal{I}. \tag{5}\]

In particular, the higher is \( \gamma \), the flatter are the indifference curves and thus the lower is the increase in consumption required to compensate for an increase in gross income while keeping utility constant.

A social allocation specifies the consumption and gross income levels for each individual. It is represented by a vector \( a = (x, z) \in \mathbb{R}_+^I \times \mathbb{R}^I \), with \( x = (x_1, \ldots, x_I) \) and \( z = (z_1, \ldots, z_I) \). The tax policymaker knows the functional form of the utility function and the distribution of wages in the population. He is however unable to observe each individual’s productivity. As a result, he is restricted to setting taxes as a function of gross income \( z_i \). By the taxation principle, a non-linear income tax schedule is therefore a mapping

\[
\begin{aligned}
\theta & \quad \mapsto \quad a \\
\theta_i & \quad \mapsto \quad (x_i, z_i),
\end{aligned} \tag{6}
\]

which satisfies the self-selection constraints

\[
u(x_i, z_i; \theta_i) \geq u(x_j, z_j; \theta_i), \quad \forall (i, j) \in \mathcal{I}^2, \tag{7}\]

and the tax revenue constraint

\[
\sum_{i=1}^I z_i \geq \sum_{i=1}^I x_i. \tag{8}\]

An allocation \( a \) is production efficient if the budget balanced constraint (8) is binding.

The social welfare function \( W : \mathbb{R}_+^I \times \mathbb{R}^I \to \mathbb{R} \) is a weighted sum of individual utilities,

\[
W(a) := \sum_{i=1}^I \lambda_i u(x_i, z_i; \theta_i), \tag{9}\]

in which \( \lambda := (\lambda_1, \ldots, \lambda_I) \) are individual social weights. The tax policymaker’s taste for redistribution from the high to the low productive individuals is captured through the requirement that
the higher the individual productivity the less the weight in the social objective, i.e.

\[ 0 < \lambda_i < \ldots < \lambda_1. \tag{10} \]

As \( W(a) \) is homogeneous of degree one in \( \lambda \), the sum of the social weights can be normalized without loss of generality. It is convenient to define \( \Lambda(\theta_i) \) as the cumulative social weight of the \( i \) less productive individuals, and to set

\[ \Lambda(\theta_i) = I. \tag{11} \]

Consequently, admissible social weights belong to the set

\[ \mathcal{D} := \{ \lambda \mid (10) \text{ and } (11) \text{ are satisfied} \}. \tag{12} \]

The optimal nonlinear income tax problem can thus be formulated as follows:

**Problem 1** (Optimal Non-linear Income Tax Problem). For \((\theta, \gamma, \lambda) \in \mathbb{R}^2_+ \times \mathcal{D}, \) choose an allocation \( a \in \mathbb{R}^I_+ \times \mathbb{R}^I \) to maximize \( W(a) \) under the self-selection constraints (7) and the tax revenue constraint (8).

For fixed values of the parameters \((\theta, \gamma, \lambda)\), there is a unique solution to Problem 1. I denote by \( g^x : \mathbb{R}^2_+ \times \mathcal{D} \rightarrow \mathbb{R}^I_+ \) and \( g^z : \mathbb{R}^2_+ \times \mathcal{D} \rightarrow \mathbb{R}^I \) the functions which relate \((\theta, \gamma, \lambda)\) to the optimal consumption and gross income vectors for Problem 1 respectively, with

\[ g^x(\theta, \gamma, \lambda) := (g^x_1(\theta, \gamma, \lambda), \ldots, g^x_I(\theta, \gamma, \lambda)), \tag{13} \]

\[ g^z(\theta, \gamma, \lambda) := (g^z_1(\theta, \gamma, \lambda), \ldots, g^z_I(\theta, \gamma, \lambda)) \tag{14}. \]

The indirect utilities \( V_i(\theta, \gamma, \lambda) \) are thus obtained as

\[ V_i(\theta, \gamma, \lambda) := u(g^x_i(\theta, \gamma, \lambda), g^z_i(\theta, \gamma, \lambda); \theta_i), \quad i \in \mathcal{I}. \tag{15} \]

### 3. THE REDUCED-FORM PROBLEM

The optimal non-linear income tax problem involves two sets of control variables, gross income \( z \) and net income \( x \). It can however be transformed into a reduced-form problem in which the policymaker chooses only one of these variables. The reduced-form problem makes it easier to interpret the social value function and to derive comparative static results. For this purpose, Problem 1 is separated into two subproblems. In the first one, gross income is arbitrarily chosen within the set of incentive-feasible gross income levels \( \mathcal{Z} \).
Subproblem 1. Given a gross income vector $z \in \mathcal{Z}$ and the parameters $(\theta, \gamma, \lambda) \in \mathbb{R}_{++}^2 \times \mathcal{D}$, choose the consumption vector $x \in \mathbb{R}_+^I$ to maximize the social welfare function $W(a)$ subject to the self-selection constraints (7) and the tax revenue constraint (8).

Let $\mathcal{X}(z; \theta, \gamma, \lambda)$ be the set of maximizers. Then, if there is a unique consumption vector $x(z; \theta, \gamma, \lambda)$ in $\mathcal{X}(z; \theta, \gamma, \lambda)$, the solution in $z$ to Problem 1 is obtained as

$$\arg\max_{x \in \mathcal{X}} W(x(z; \theta, \gamma, \lambda), z).$$

So, the reduced-form problem can be stated as follows.

Subproblem 2. Given the parameters $(\theta, \gamma, \lambda) \in \mathbb{R}_{++}^2 \times \mathcal{D}$, choose $z \in \mathcal{Z}$ to maximize the social welfare function $W(x(z; \theta, \gamma, \lambda), z)$.

For this two-stage reasoning to hold, it remains to clarify why all implications of the self-selection constraints, except $z \in \mathcal{Z}$, are taken into account in Subproblem 1 and to establish that the function $x(z; \theta, \gamma, \lambda)$ is unique and differentiable.

3.1. Implications of the Self-Selection Constraints

The self-selection constraints (7) place structure on the solution to Problem 1. Indeed, incentive compatibility of the income tax schedule requires the indirect utility to increase at a specific rate and the gross income to be non-decreasing. These restrictions can be used to derive sufficient conditions under which an allocation $a \in \mathbb{R}_+^I \times \mathbb{R}^I$ satisfies the incentive-compatibility constraints (7). We proceed in two steps.

First, if an allocation $a \in \mathbb{R}_+^I \times \mathbb{R}^I$ satisfies (7), then gross income and net income must be non-decreasing in productivity, i.e.

$$(x_1, z_1) \leq \ldots \leq (x_I, z_I),$$

with $(x_{i-1}, z_{i-1}) \ll (x_i, z_i)$ if $(x_{i-1}, z_{i-1}) \neq (x_i, z_i)$, $i = 2, \ldots, I$. As a result, the solution in $z$ to Problem 1 must lie in the set $\mathcal{Z}$ defined as

$$\mathcal{Z} := \{z \in \mathbb{R}^I|0 \leq z_1 \leq \ldots \leq z_I\}.$$
Seade (1982), a simple monotonic chain to the left is an allocation \( a \in \mathbb{R}^I_+ \times \mathbb{R}^I \) such that
\[
\begin{align*}
u(x_{i+1}, z_{i+1}; \theta_{i+1}) &= u(x_i, z_i; \theta_{i+1}), \\ i &= 1, ..., I-1.
\end{align*}
\]
(19)

Given quasilinear-in-consumption preferences, the requirement (19) is equivalent to
\[
\begin{align*}
u(x_{i+1}, z_{i+1}; \theta_{i+1}) - u(x_i, z_i; \theta_i) &= v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right).
\end{align*}
\]
(20)

In words, the adjacent downward compatibility constraints are active for all \( i = 2, ..., I \).

**Proposition 1.** Let an allocation \( a \in \mathbb{R}^I_+ \times \mathbb{R}^I \) be a simple monotonic chain to the left and \( z \in \mathcal{Z} \). Then \( a \) satisfies (7).

**Proof.** Guesnerie and Seade (1982). \( \square \)

This pattern expresses a specific rent-extraction trade-off. Indeed, (20) expressed at which rate utility must be increased for the tax schedule to induce individual truth-telling. For each pair of adjacent productivity levels \( (\theta_i, \theta_{i+1}) \), this rate basically depends on
\[
R_{i+1} := v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right), \\ i = 1, ..., I-1,
\]
(21)

which may be regarded as the marginal rent the policymaker has to leave to the more productive individuals because of the informational externality. Consequently, (20) constitutes the discrete analogue of the first-order condition for incentive compatibility obtained in the models with a continuum of individuals.

### 3.2. Optimal Consumption Given Fixed Levels of Income

The properties of the solution in \( x \) to Subproblem 1 are now investigated. The next lemma establishes that there exist solutions to Subproblem 1 for all gross income vector \( z \in \mathcal{Z} \) and all \( (\theta, \gamma, \lambda) \in \mathbb{R}_{2+}^2 \times \mathcal{D} \). In addition, each of them is a simple monotonic chain to the left for which the tax revenue constraint (8) is binding.

**Lemma 1.** Given \( z \in \mathcal{Z} \) and all \( (\theta, \gamma, \lambda) \in \mathbb{R}_{2+}^2 \times \mathcal{D} \), there is at least one solution to Subproblem 1 and any allocation \( a = (x, z) \) where \( x^* \in \mathcal{X}^* (z; \theta, \gamma, \lambda) \), is a simple monotonic chain to the left which is production efficient.

**Proof.** See the Appendix. \( \square \)
The implications are twofold. First, combined with Proposition 1, Lemma 1 ensures that all implications of the incentive-compatibility constraints (7) are embedded in any solution to Subproblem 1, provided \( z \in \mathcal{Z} \). Second, the fact that \( a \) is a simple monotonic chain to the left gives rise to a specific consumption pattern. Indeed, by (19),

\[
x_i = x_{i-1} + \frac{1}{\gamma} \left[ v \left( \frac{z_i}{\theta_j} \right) - v \left( \frac{z_{i-1}}{\theta_j} \right) \right], \quad i = 2, \ldots, I, \tag{22}
\]

and so

\[
x_i = x_1 + \frac{1}{\gamma} \sum_{j=2}^I \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_j} \right) \right], \quad i = 2, \ldots, I. \tag{23}
\]

As any solution to Subproblem 1 is production efficient, by Lemma 1, the binding tax revenue constraint (8) can be substituted in

\[
\sum_{i=1}^I x_i = I x_1 + \frac{1}{\gamma} \sum_{j=2}^I \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_j} \right) \right] = I x_1 + \frac{1}{\gamma} \sum_{j=2}^I (I + 1 - j) \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_j} \right) \right]. \tag{24}
\]

This equation admits a unique solution in \( x_1 \). Substituting the latter in (23) and proceeding sequentially show that there is a unique consumption vector in \( \mathcal{Z}^* (z; \theta, \gamma, \lambda) \), which is independent of the social weights \( \lambda \) and inherits the differentiability properties of \( v \).

**Proposition 2.** Given \( z \in \mathcal{Z} \) and \( (\theta, \gamma, \lambda) \in \mathbb{R}^3_{++} \times \mathcal{D} \), the unique function solution to Subproblem 1 is twice continuously differentiable, defined by \( x^*: \mathcal{Z} \times \mathbb{R}^3_{++} \rightarrow \mathbb{R}^I_+ \) with

\[
x_1^* (z; \theta, \gamma) = \frac{1}{I} \left\{ \sum_{j=1}^I z_j - \frac{1}{\gamma} \sum_{j=2}^I (I + 1 - j) \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_j} \right) \right] \right\}, \tag{25}
\]

\[
x_i^* (z; \theta, \gamma) = x_1^* (z; \theta, \gamma) + \frac{1}{\gamma} \sum_{j=2}^i \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_j} \right) \right], \quad i = 2, \ldots, I. \tag{26}
\]

### 3.3. The Reduced Form

We can now take stock of the previous results to give a more compact formulation of Subproblem 2. For this purpose, it is convenient to introduce the new vector of social parameters

\[
\beta = (\beta_1, \ldots, \beta_I) \quad \text{with} \quad \beta_i := \Lambda (\theta_i) - i, \quad i \in \mathcal{I}. \tag{27}
\]

Because of (10) and (11), the graph of \( i \rightarrow \Lambda (\theta_i) \) is hump-shaped and above the 45°-line. Hence, \( \beta_i > 0 \) for all \( i = 1, \ldots, I - 1 \).

The parameters \( \beta_i \) summarize in a transparent way the redistributive taste of the government.
Since individual preferences are quasilinear in consumption, the marginal social benefit of increasing consumption of the $\theta_i$-individual is equal to $\lambda_i$. Hence, $\Lambda(\theta)$ corresponds to the gross social benefit of a marginal increase in consumption for each of the $i$ less productive individuals. Since giving of extra euro of consumption to each of these individuals reduces tax revenue by $i$ euros, $\beta_i$ is nothing but the net social benefit of marginally increasing the consumption of the $i$ less skilled individuals. Alternatively, it can be noted that all $\lambda_i$ would be equal if the government adopted pure utilitarianism as a social objective. In this case, $\beta_i = 0$ for every $i$. Consequently, the social parameters $\beta_i$ express the policymaker’s strict aversion to income inequality. That is why they are henceforth referred to as net cumulative social weights. Thank to them, the reduced-form optimal income tax problem can be written compactly as follows.

**Problem 2 (Reduced Form).** For $(\theta, \gamma, \lambda) \in \mathbb{R}_+^2 \times \mathcal{D}$, choose $z$ in $\mathcal{Z}$ so as to maximize the social objective function $W^* (z; \theta, \gamma, \lambda) : \mathcal{Z} \times \mathbb{R}_+^2 + \mathcal{D} \rightarrow \mathbb{R}$, with

$$W^* (z; \theta, \gamma, \lambda) := \sum_{i=1}^I \left[ \gamma z_i - v \left( \frac{z_i}{\theta_i} \right) \right] - \sum_{i=1}^I \beta_i R_{i+1}. \quad (28)$$

This problem is called a reduced form of Problem 1 because the optimal solution in gross income of the former, and the consumption pattern it generates through Proposition 2, are the optimal solutions of the latter.

**Proposition 3.** For all $(\theta, \gamma, \lambda) \in \mathbb{R}_+^2 \times \mathcal{D}$, the optimal solution to Subproblem 2 is $g^* (\theta, \gamma, \lambda)$, the gross income vector solution to Problem 1, and the optimal consumption vector for Problem 1 is

$$g^* (\theta, \gamma, \lambda) = x^* (g^* (\theta, \gamma, \lambda); \theta, \gamma). \quad (29)$$

**Proof.** See the Appendix.

An important implication of Proposition 8 is that the social allocation solution to the optimal income tax problem Problem 1 is a monotonic chain to the left. In consequence, the optimal tax schedule is not differentiable at each observed gross income level $z_i$. It is nevertheless possible to use the differentiability of the indifference curves in order to define implicit marginal tax rates. Since at the optimum only the adjacent downward self-selection constraints are binding, two implicit marginal tax rates are of particular interest at each observed gross income level $z_i$: the implicit marginal tax rate $T'(z_i; \theta_i)$ faced by the $\theta_i$-individual for whom the $(x^*_i (z_i; \theta_i, \gamma), z_i)$-bundle is designed by the policymaker, on the one hand, and the implicit marginal tax rate $T'(z_i; \theta_{i+1})$ the nearest more productive $\theta_{i+1}$-individuals would face if they were mimicking the
The implicit marginal tax rates allows us to get further understanding of the social objective function of the reduced-form optimal income tax problem \( W^* (z; \theta; \gamma; \lambda) \). Indeed, let \( z \) be a fixed gross income vector and consider that the gross income \( z_i \) of the \( \theta_i \)-individual is increased at the margin. As

\[
\frac{\partial W^* (z; \theta; \gamma; \lambda)}{\partial z_i} = \gamma T^0 (z_i; \theta_i) - \beta_i \frac{\partial R_{i+1}}{\partial z_i},
\]

by (28) and (30), the impact on social welfare is twofold. First, the \( \theta_i \)-individual pays \( T^0 (z_i; \theta_i) \) additional euros in taxes, which relaxes the tax revenue constraint (8). Since \( \gamma \) is the marginal utility of money, the positive effect in terms of welfare amounts to \( \gamma T^0 (z_i; \theta_i) \). Second, since the \( \theta_i \)-individual receives \( 1 - T^0 (z_i; \theta_i) \) extra euro of consumption, it becomes more attractive for all more skilled individuals to mimic the im. In order to restore individual truth-telling, the policymaker has to increase the marginal rent left to the more productive individuals in such a way that a new monotonic chain to the left is obtained. Holding all other gross income levels fixed, this can only be achieved if the consumption of each of the \( I - i \) more able workers is raised by \( \partial R_{i+1} / \partial z_i \) euros. The average cost of giving \( \partial R_{i+1} / \partial z_i \) extra euros of consumption to these \( I - i \) individuals is simply \( \beta_i \partial R_{i+1} / \partial z_i \). In consequence, the social optimum is obtained when the positive effect on social welfare offsets the negative one, provided an interior solution exists.

### 3.4. Characterization of the Social Optimum

Henceforth, the initial optimal allocations under scrutiny do not exhibit bunching and are such that every individual has a strictly positive gross income. In other words, the gross income levels \( z = g^\theta (\theta, \gamma, \lambda) \) lie in the interior of \( \mathcal{Z} \). \( \mathcal{Z}^0 := \{ z \in \mathbb{R}^I | 0 < z_1 < ... < z_I \} \). The optimal gross-income vector must therefore be such that (31) is equal to zero for every \( i \in I \). Using (30) and (21), the following alternative characterization is obtained.

**Proposition 4.** A gross-income vector \( z \in \mathcal{Z}^0 \) is socially optimal if and only if

\[
\alpha_i = \frac{1}{\beta_i}, \quad \forall i \in I,
\]

where

\[
\alpha_i (z; \theta; \gamma) = \frac{T^0 (g^\theta_i; \theta_i) - T^0 (g^\theta_{i+1}; \theta_i)}{T^0 (g^\theta_i; \theta_i)}, \quad i = 1, ..., I - 1.
\]
and $\alpha_i$ is an arbitrary number.

For a gross income level $g_i^z$, $\alpha_i$ tells us to which extent the indifference curves of the $\theta_{i+1}$-individual must be flatter than those of the $\theta_i$-individual. Geometrically, it thus corresponds to the angle depicted in Figure 1. The Spence-Mirrlees single-crossing condition is a restriction on the sign of the angle $\alpha_i$, which must be strictly positive. Here, this condition is automatically satisfied since individual preferences are quasilinear in consumption. The conditions for social optimality (32) introduce an additional restriction on $\alpha_i$: an allocation is socially optimal only if, at each observed gross income level $z_i$, the angle between the indifference curves of the $\theta_i$ and $\theta_{i+1}$-individuals is entirely determined by the exogenously given cumulative social weight $\beta_i$.

In the absence of bunching at the social optimum, $\alpha_i(z_i; \theta, \gamma)$ is a strictly increasing function of $z_i$. In other words, for fixed $\theta$ and $\gamma$, $z_i$ monotonically depends on $\beta_i$. In this case, it is possible to construct the optimal allocation geometrically. Thanks to Proposition 4, it is known that each $\alpha_i$ is independent of consumption. Therefore, in the first step, the tax revenue constraint can be ignored and the consumption level of the less productive individual arbitrarily set to $x_1$. Starting from zero, gross income is gradually increased until $\alpha_1 = 1/\beta_1$ and the bundle $(x_1, z_1)$ is determined. Then, gross income is increased along the indifference curve of the $\theta_2$-individual.
through the \((x_1, z_1)\)-bundle until the angle with the indifference curve of the \(\theta_i\)-individual is equal to \(1/\beta_2\). Proceeding recursively, a monotonic chain to the left \((x, z)\) is obtained. This allocation is incentive compatible, but not necessarily budget-balanced since \(x_1\) was chosen arbitrarily. That is why, in the second step, each \(x_i\) is varied by a same amount \(\epsilon\) so as to get a binding tax revenue constraint. The resulting allocation \((x+\epsilon, z)\), which is both incentive compatible and production efficient, is socially optimal.

Before going further and derive comparative static properties, it is instructive to compare our results with those in Weymark (1986a,b, 1987). In these papers, the quasilinear-in-leisure utility function \(u(x_i, z_i; \theta_i) = h(x_i) + \gamma z_i\) is replaced by its monotone transform \(\tilde{u}(x_i, z_i; \theta_i) = \theta_i h(x_i) - \gamma z_i\) in order to sum \(\tilde{u}(x_i, z_i; \theta_i)\) over all \(i \in I\), get

\[ \sum_{i=1}^I \tilde{u}(x_i, z_i; \theta_i) = \sum_{i=1}^I h(x_i) - \gamma \sum_{i=1}^I z_i, \]  

(34)

and replace \(\sum_{i=1}^I z_i\) by \(\sum_{i=1}^I x_i\). In consequence, skill-normalized social weights \(\tilde{\lambda}_i := \lambda_i/\theta_i\) are used in the social objective \(\sum_{i=1}^I \tilde{\lambda}_i \tilde{u}(x_i, z_i; \theta_i)\). The first-order conditions of the reduced-form problem involve therefore skilled-normalized cumulative social weights \(\sum_{j=1}^i \tilde{\lambda}_j\) instead of \(\Lambda_i\).

So, the impact of the policymaker’s taste for redistribution is less transparent. However, since the reduced-form is linear with respect to gross income, the variable observed by the policymaker, the first-order conditions involve \(h'(x_i)\) for every \(i \in I\). An explicit expression of the consumption level \(x_i\) is thus obtained as \(h'\) is invertible. Weymark (1986a) employs this expression to examine in which cases bunching occurs. In the present framework, the loss of the linearity in gross income prevents us to provide clear-cut results as regards bunching.

4. COMPARATIVE STATIC PROPERTIES

Besides providing a geometric interpretation of the optimality conditions, the reduced-form makes it possible to derive comparative static results of the optimal income tax allocation. As previously, the initial optimal allocations under scrutiny do not exhibit bunching and are such that every individual has a strictly positive gross income.

Since the disutility of labour \(v\) is \(\mathcal{C}^2\), the implicit function theorem implies that \(g_{\tilde{f}}(\gamma, p, \lambda)\) is \(\mathcal{C}^1\). It thus follows from Proposition 2 that \(g_{\tilde{f}}(\theta, \gamma, \lambda) = x^* (g_{\tilde{f}}(\theta, \gamma, \lambda); \theta, \gamma, p)\) is also \(\mathcal{C}^1\). If \(\mathcal{P}\) is defined as

\[ \mathcal{P} := \left\{ (\theta, \gamma, \lambda) \in \mathbb{R}^2_{+} \times \mathcal{D} \mid z \in \mathcal{D}^0 \right\}, \]  

(35)

these results can be summarized as follows:
Proposition 5. Let \((\theta, \gamma, \lambda) \in \mathcal{P}\). Then, the functions \(g^x, g^z, T^t\) and \(V\) are \(C^1\) at every \((\theta, \gamma, \lambda)\) in \(\mathcal{P}\).

The effects of changing an underlying parameter can now be investigated.

4.1. Comparative Statics for the Marginal Utility of Money

A small increase in the marginal utility for money \(\gamma\) raises the gross income of every individual.

Proposition 6. For every \((\theta, \gamma, \lambda) \in \mathcal{P}\) and every \(i \in \mathcal{I}\),

\[
\frac{\partial g^z_i(\theta, \gamma, \lambda)}{\partial \gamma} > 0.
\]  

(36)

Proof. See the Appendix.

Indeed, when \(\gamma\) goes up, an extra unit of consumption contributes more to individual well-being as previously. Therefore, the indifference curves become flatter in the \((z, x)\)-space. Every \(\theta_i\)-individual is thus willing to work more in order to increase his consumption by a given amount and decides to work more. A small increase in the disutility of labour has exactly the reverse effect. Unfortunately, the comparative statics of the optimal implicit marginal tax rates cannot be obtained in the general case. In fact, as \(T'(z_i; \theta_j) = 1 - v'(z_i/\theta_i)/(\gamma \theta_i)\), the reduction in \(1/\gamma\) goes in the opposite direction to the associated increase in gross income.

4.2. Comparative Statics for Individual Productivities

Varying the skill levels have more subtle effects on the optimal allocation. This is of particular interest since productivities are probably the most basic ingredients of the Mirrleesian optimal income tax model. They are indeed the sole source of heterogeneity within the population and give rise to the adverse selection problem which is the key of Mirrleesian income taxation. The fact that the productivity vector \(\theta\) is strictly monotonically increasing ensures that (1) remains satisfied once a given individual productivity is changed at the margin. The effects of a variation in \(\theta_{i+1}\) can be summarized as follows.

Proposition 7. For every \((\theta, \gamma, \lambda) \in \mathcal{P}\) and every \((i, j) \in \{1, ..., I - 1\} \times \mathcal{I}\), the impact of a
small variation in $\theta_{i+1}$ on the implicit marginal tax rates is

$$\frac{\partial T'}{(g^*_i(\theta, \gamma, \lambda); \theta_i)} > 0,$$

$$\frac{\partial T'}{(g^*_i(\theta, \gamma, \lambda); \theta_{i+1})} > 0,$$

$$\frac{\partial T'}{(g^*_{i+1}(\theta, \gamma, \lambda); \theta_{i+1})} < 0,$$

$$\frac{\partial T'}{(g^*_{i+1}(\theta, \gamma, \lambda); \theta_{i+2})} < 0,$$

$$\frac{\partial T'}{(g^*_j(\theta, \gamma, \lambda); \theta_j)} = \frac{\partial T'}{(g^*_j(\theta, \gamma, \lambda); \theta_{j+1})} = 0 \text{ for } j \notin \{i, i+1\},$$

while the gross income are changed as follows:

$$\frac{\partial g^*_i(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} < 0,$$

$$\frac{\partial g^*_{i+1}(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} > 0,$$

$$\frac{\partial g^*_j(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i+1\}.$$

**Proof.** See the Appendix.

Increasing the productivity of the $\theta_i$-individual does only alter his gross income and that of his nearest less productive neighbour. Indeed, by Proposition 4, only $\alpha_i$ and $\alpha_{i+1}$ depend on $\theta_{i+1}$. So, the optimality condition $\alpha_j = 1/\beta_j$, which implicitly defines $z_j$ as a function of $\theta_{i+1}$, is unchanged except for the $\theta_i$ and $\theta_{i+1}$-individuals. Accordingly, the gross income levels of all other individuals remain unaltered. The local adjustment process combines three effects.

First, the variation in $\theta_{i+1}$ gives rise to a local substitution effect. Indeed, the increase in the productivity of the $\theta_{i+1}$-individual results in a rise in his net-of-tax wage rate, which leads him to increase his labour supply in efficiency units, $z_{i+1}$.

Second, changing $\theta_{i+1}$ has an incentive effect. As he becomes more efficient, the $\theta_{i+1}$-individual has to provide less effort if he wants to imitate the $\theta_i$-individual. Consequently, his indifference curve through the gross-income/consumption bundle of the $\theta_i$-individual flattens. This corresponds to an increase in the implicit marginal tax rate $T'(x_i, z_i; \theta_{i+1})$ he would face if he were cheating.

Third, the $\theta_i$-individual incurs an informational externality induced by the incentive effect. Since the cumulative social weight $\beta_i$ is unaltered, the angle $\alpha_i$ between the indifference curves
of the $\theta_i$ and $\theta_{i+1}$-individuals through the $(x_i,z_i)$-bundle must stay constant (Proposition 3). Consequently, the increase in $T'(x_i,z_i;\theta_{i+1})$ must be associated with an increase in the implicit marginal tax rate $T'(x_i,z_i;\theta_i)$ and thus with a reduction in the net-of-tax wage rate of the $\theta_i$-individual. Finally, the substitution effect leads the $\theta_i$-individual to work less.

The changes in gross income ensure that a new monotonic chain to the left is obtained. However, this incentive-compatible allocation is not necessarily budget balanced. Therefore, in a second step, the consumption levels are adjusted in order to obtain a binding tax revenue constraint. However, the comparative static results as regards consumption cannot be derived in the general case.

4.3. Comparative Statics for the Social Weights

The geometric characterization of the solution to the optimal income tax problem found in Proposition 4 basically involves the cumulative social weights $\beta_i$, and thus the individual social weights $\lambda_i$. That is why it is instructive to examine the impact of changing these weights at the margin.

The impact of a change in the cumulative social weight of a $\theta_i$-individual, with $i < I^1$, is examined first. Since all cumulative weights sum to $I$, the increase in $\beta_i$ must be accompanied by a decrease in some other cumulative weights. Different scenarios can be contemplated, but we prefer to concentrate on the case in which the increase in $\beta_i$ is totally compensated by a reduction in some other social weight $\beta_j$. This situation is of particular interest because the changes in all endogenous variables can be signed when $j = i + 1$, i.e. when $\lambda_{i+1}$ is decreased to place greater individual social weight on the $\theta_i$-individual. Given Proposition 4 and for a given $T'(g_i^T(\theta,\gamma,\lambda);\theta_{i+1})$, an increase in $\beta_i$ requires the implicit marginal tax rate $T'(g_i^T(\theta,\gamma,\lambda);\theta_i)$ faced by the $\theta_i$-individual to be raised. The induced substitution effect leads the $\theta_i$-individual to reduce his labour supply.

Proposition 8. For all $(\theta,\gamma,\lambda) \in \mathcal{D}$ and all $(i,j) \in \{1,...,I-1\} \times \mathcal{I}$ with $i \neq j$,

$$\frac{\partial g_i^T(\theta,\gamma,\lambda)}{\partial \beta_i} < 0 \quad \text{and} \quad \frac{\partial g_j^T(\theta,\gamma,\lambda)}{\partial \beta_i} = 0$$

(45)

Proof: See the Appendix.

This result can now be used to consider the variation in which the individual social weight of the $\theta_i$-individual is increased to the detriment of a more productive $\theta_j$-individual (i.e. $i < j$).

\footnote{A change in $\beta_i$ is impossible since, by definition, $\beta_j \equiv 0$.}
By definition of $\Lambda(\theta_k)$, every $\beta_k$ is increased for $k \in \{i, \ldots, j-1\}$ while all other $\beta_k$ remain unaltered. By Proposition 8, it is thus optimal to decrease the gross income $z_k$ of each $\theta_k$-individual, with $k \in \{i, \ldots, j-1\}$, and to hold that of the others constant. The impact on consumption can also be signed for all $k \notin \{i+1, \ldots, j-1\}$.

**Proposition 9.** Let $(\overline{s}, \overline{p}, \overline{p}, \overline{\lambda}) \in \mathcal{S}$, $i \in \{1, \ldots, I-1\}$ and $j \in \{i+1, \ldots, I\}$. Let $\lambda : S \rightarrow \mathbb{R}^I$, where $S = (-1, 1)$ be $C^1$ with

\begin{align}
\lambda_k(s) &\equiv \overline{\lambda}_k, \quad \forall s \in S, \quad \forall k \neq i, j \\
\lambda_k(0) &\equiv \overline{\lambda}_k, \quad k = i, j, \\
d\lambda_i(s)/ds &= -d\lambda_j(s)/ds, \quad \forall s \in S.
\end{align}

Then, if $d\lambda_i(0)/ds > 0$,

\begin{align}
dg_k^i(\overline{s}, \overline{p}, \overline{p}, \overline{\lambda})/ds &= 0, \quad \forall k < i \text{ and } \forall k \geq j, \quad (49) \\
dg_k^i(\overline{s}, \overline{p}, \overline{p}, \overline{\lambda})/ds &< 0, \quad \forall k \in \{i, \ldots, j-1\}, \quad (50) \\
dx_i^k(z; \overline{s}, \overline{p}, \overline{p})/ds &> 0, \quad \forall k \leq i, \quad (51) \\
dx_i^k(z; \overline{s}, \overline{p}, \overline{p})/ds &< 0, \quad \forall k \geq j. \quad (52)
\end{align}

The inequalities in (50)–(52) are reversed if $d\lambda_i(0)/ds < 0$.

**Proof.** See the Appendix.

For all $k \notin \{i, \ldots, j-1\}$, the variation in $\lambda$ has no effect on gross income. In addition, the implicit optimal marginal tax rates do only depend on $\lambda$ through gross income $z$. In consequence, the impact on the implicit marginal tax rates and indirect utility levels directly follows from Proposition 9.

**Corollary 1.** Consider the same situation as in Proposition 9. Then, if $d\lambda_i(0)/ds > 0$,

\begin{align}
dT'(x_k^i, \overline{g}_k^i; \theta_k) /ds &= dT'(x_k^i, \overline{g}_k^i; \theta_{k+1}) /ds = 0, \quad \forall k < i \text{ and } \forall k \geq j, \quad (53) \\
dT'(x_k^i, \overline{g}_k^i; \theta_k) /ds &> 0, \quad \forall k \in \{i, \ldots, j-1\}, \quad (54) \\
dT'(x_k^i, \overline{g}_k^i; \theta_{k+1}) /ds &> 0, \quad \forall k \in \{i, \ldots, j-1\}, \quad (55) \\
dV_k(\overline{s}, \overline{p}, \overline{p}, \overline{\lambda}) /ds &> 0, \quad \forall k \leq i, \quad (56) \\
dV_k(\overline{s}, \overline{p}, \overline{p}, \overline{\lambda}) /ds &< 0, \quad \forall k \geq j. \quad (57)
\end{align}

The inequalities in (53)–(57) are reversed if $d\lambda_i(0)/ds < 0$.

17
When the $\theta_i$-individual is the nearest less productive neighbour of the $\theta_j$-individual, the comparative statics of all endogenous variables is obtained. For the sake of clarity, the interpretation of these results concentrate on this special case.

The social weight of the $\theta_i$-individual is increased at the expense of the $\theta_{i+1}$-individual. This naturally places the former in a more favourable position and reinforces the incentive to mimic the latter. In order to dissuade the latter from cheating, the policymaker increases $T'(x^*_i, g^*_i; \theta_{i+1})$. By Proposition 4, this gives rise to an externality incurred by the $\theta_i$-individual: the implicit marginal tax rate $T'(x^*_i, g^*_i; \theta_i)$ is increased, so the $\theta_i$-individual chooses to work less. The gross income levels of the other individuals are unaffected.

The adjustments in consumption parallel those in gross income investigated by Weymark (1987, pp. 1177-1178). They can be thought of as proceeding in two steps. To be concrete, suppose $\lambda_i$ is decreased to the benefit of the $\theta_{i+1}$-individual. In the first step, the budget constraint (8) is left aside and the incomes of all $\theta_k$-individuals, with $k \neq i$, is held fixed. Since $g^*_i$ is increased, the requirement that the $\theta_{i+1}$-individual is indifferent between the $(x^*_i, g^*_i)$-bundle and his own bundle can only be satisfied if he is made better off. By induction, all more productive individuals are better off when the new monotonic chain to the left is obtained. In consequence, the new allocation Pareto dominates the initial one. However, as the initial allocation is optimal given $\left(\bar{\theta}, \bar{\gamma}, \bar{p}, \bar{\lambda}\right)$, changing $\lambda_i$ while all other parameters are held constant must have no first-order effect. Hence, the new allocation must not be budget balanced. In the second step, each individual’s consumption is decreased in order to satisfy the tax revenue constraint (8). The slope of his indifference curve at the $(x^*_i, g^*_i)$-bundle is less than 1 because additional income are taxed above 100%. Accordingly, the increase in the gross income of the $\theta_i$-individual which took place in the first-step is more than sufficient for him to buy his extra consumption. In the second step, the decrease in each individual’s consumption can thus be sufficiently small for each individual with productivity above $\theta_i$ to benefit from a net increase in his consumption.

5. CONCLUSION

Thanks to the absence of income effects on labour supply, the trade-off between equity and efficiency is very pure when individual preferences are quasilinear in consumption. This case has been investigated in depth in the continuous population version of Mirrlees model (Atkinson (1990), Diamond (1998), Piketty (1997), Salanié (1998) or d’Autume (2000)), but the analysis carried out for a finite population has concentrated on the situation where preferences are quasilinear in leisure. In this extent, the present paper contributes to filling this gap.

When preferences are quasilinear in consumption, it is not necessary to work with skilled-normalized social weights as in Weymark (1987). Therefore, the respective influences of individ-
ual productivities and social weights are easier to separate in the social objective function of the
reduced-form optimal income tax problem. This offers two advantages. First, the link between
the social weights and the conditions for social optimality is very transparent. For each observed
gross-income level \( z_i \), the angle between the indifference curve of the \( \theta_i \) and \( \theta_{i+1} \)-individual
must be equal to the inverse of the cumulative social weight \( \beta_i \): Second, clear-cut comparative
statics properties can easily be derived as regards changes in the productivity levels.

**APPENDIX**

**Proof of Lemma 1.** The proof proceeds in three steps.

(i) *An Optimal Allocation is Production Efficient.*

Fix \( z \) in \( \mathcal{Z} \) and suppose \( x^* \in \mathcal{X}^*(z; \gamma, p, \lambda) \). The self-selection constraints (7) are satisfied, with
\( z_i = \bar{x}_i \) and \( x_j = x^*_j \) for every \( i \in \mathcal{I} \). The proof proceeds by contradiction. Assume (8) is not
binding. Hence, (8) remains satisfied if all \( x^*_i \) are increased by a sufficiently small \( \epsilon > 0 \). But
then, the self-selection constraints (7) become

\[
\gamma x_i^* + \epsilon - v \left( \frac{z_i}{\theta_i} \right) \geq \epsilon - v \left( \frac{z_j}{\theta_{j+1}} \right), \forall (i, j) \in \mathcal{I}^2,
\]

which reduces to (7). Increasing every \( x^*_i \) by \( \epsilon \) is thus incentive compatible. Since it is also
Pareto improving, \( x^* \) cannot belong to \( \mathcal{X}^*(z; \theta, \gamma, \lambda) \). A contradiction. Therefore, (8) holds
with equality.

(ii) *Existence of an Optimal Allocation.*

Fix \( z \) in \( \mathcal{Z} \) and pick an arbitrary value \( x_1 \) for \( x_1 \). Then, proceed sequentially to select \( x \) which
solves

\[
\gamma x_{i+1} - v \left( \frac{z_{i+1}}{\theta_{i+1}} \right) = \gamma x_i - v \left( \frac{z_i}{\theta_i} \right), \forall i = 2, ..., I - 1. \tag{59}
\]

By construction, \( \bar{x} \) is a monotonic chain to the left. Since \( z \in \mathcal{Z} \), it follows from Proposition 1
that \( \bar{x} \) satisfies (7). If \( \bar{x} \) does not satisfy (8), it is sufficient to change each \( x_i \) by a sufficient large
amount. If (8) is not binding, the argument used in (i) applies: it is sufficient to increase each \( \bar{x}_i \)
by a well-chosen \( \epsilon > 0 \). Consequently, the constraint set is not empty.

Now, write (7) for \( j = 1 \) to obtain

\[
\gamma x_i - v \left( \frac{z_i}{\theta_i} \right) \geq \gamma x_1 - v \left( \frac{z_1}{\theta_1} \right) \iff \gamma [x_i - x_1] \geq v \left( \frac{z_i}{\theta_i} \right) - v \left( \frac{z_1}{\theta_1} \right) \geq 0, \forall i \in \mathcal{I}. \tag{60}
\]

Since \( x_1 \geq 0 \), all \( x_i \) are bounded from below. In addition, all \( x_i \) must be bounded from above for
(8) to be binding. Consequently, if the set \( \mathcal{X}^*(z; \theta, \gamma, \lambda) \) is non-empty, it is a bounded subset of
the feasible set.

19
Finally, $W$ is continuous while the constraint set is compact (because the inequalities are weak) and non-empty. Hence, by Weierstrass theorem, $\mathcal{F}^*(z; \theta, \gamma, \lambda) \neq \emptyset$.

(iii) Any Optimal Allocation is a Monotonic Chain to the Left.

\[ \square \]

Proof of Proposition 3. It is sufficient to establish that substitution of $x^* (z; \theta, \gamma)$ into $W$ yields $\mathcal{W}^*$ for all $z \in \mathcal{F}$. By (20),

\[ u(x_i^*, z_i; \theta_i) = u(x_{i-1}^*, z_{i-1}; \theta_{i-1}) + v\left(\frac{z_{i-1}}{\theta_{i-1}}\right) - v\left(\frac{z_i}{\theta_i}\right), \quad i = 2, ..., I, \tag{61} \]

from which

\[ u(x_i^*, z_i; \theta_i) = u(x_1^*, z_1; \theta_1) + \sum_{j=1}^{i-1} v\left(\frac{z_j}{\theta_j}\right) - v\left(\frac{z_i}{\theta_i}\right), \quad i = 2, ..., I. \tag{62} \]

Consequently,

\[ \sum_{i=1}^{I} u(x_i^*, z_i; \theta_i) = I u(x_1^*, z_1; \theta_1) + \sum_{i=1}^{I-1} (I - i) \left[ v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right]. \tag{63} \]

In addition, summing (4) over $i$ on $\mathcal{F}$ and employing (8),

\[ \sum_{i=1}^{I} u(x_i^*, z_i; \theta_i) = \gamma \sum_{i=1}^{I} z_i - \sum_{i=1}^{I} v\left(\frac{z_i}{\theta_i}\right). \tag{64} \]

Plugging (64) in (63) and solving for $u(x_1^*, z_1; \theta_1)$,

\[ u(x_1^*, z_1; \theta_1) = \frac{1}{I} \left\{ \gamma \sum_{i=1}^{I} z_i - \sum_{i=1}^{I} v\left(\frac{z_i}{\theta_i}\right) - \sum_{i=1}^{I-1} (I - i) \left[ v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right] \right\}. \tag{65} \]

By (62) and (65), the social value function reads

\[ W = u(x_1^*, z_1; \theta_1) I + \sum_{i=1}^{I-1} \sum_{j=1}^{I} \lambda_i \left[ v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_j}{\theta_j}\right) \right], \]

\[ = u(x_1^*, z_1; \theta_1) I + \sum_{i=1}^{I-1} \left( \sum_{j=1}^{i} \lambda_j \right) \left[ v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right], \]

\[ = \sum_{i=1}^{I} \left\{ \gamma z_i - \delta v\left(\frac{z_i}{\theta_i}\right) + \left( i - \sum_{j=1}^{i} \lambda_j \right) \left[ v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right] \right\}. \tag{66} \]
Proof of Proposition 6. Let $z_i \equiv g^j_i(\theta, \gamma, \lambda)$ and write the first-order conditions given in Proposition 4 as

$$
\gamma - \frac{1 + \beta_i}{\theta_i} v'(z_i) + \frac{\beta_i}{\theta_{i+1}} v'(\frac{z_i}{\theta_{i+1}}) = 0, \quad i = 1, \ldots, I - 1,
$$

(67)

to define $\phi_i(z_i, \gamma) = 0$. Since $v'' > 0$ and $0 < \theta_i < \theta_{i+1},$

$$
\frac{\partial \phi_i(z_i, \gamma)}{\partial z_i} = \frac{\beta_i}{\theta_i^2} v''(z_i) - \frac{1 + \beta_i}{\theta_{i+1}^2} v'' \left( \frac{z_i}{\theta_{i+1}} \right) < 0, \quad i = 1, \ldots, I - 1.
$$

(68)

By the implicit function theorem, for every $\gamma \in \mathbb{R}_{++}, \phi_i(z_i, \gamma) = 0$ has a unique solution which defines $z_i$ as a $C^1$-function $z_i = \phi_i(\gamma)$, with derivative

$$
\frac{\partial g^j_i(\theta, \gamma, \lambda)}{\partial \gamma} \equiv \phi'_i(\gamma) = - \frac{\partial \phi_i(z_i, \gamma)}{\partial \gamma} \frac{\partial \phi_i(z_i, \gamma)}{\partial z_i} = - \left[ \frac{\partial \phi_i(z_i, \gamma)}{\partial z_i} \right]^{-1} > 0.
$$

(69)

Proof of Proposition 7. Let $z_i \equiv g^j_i(\theta, \gamma, \lambda)$ and use (67) to define $\phi_j(z_j, \theta_{i+1}) = 0$ for $j = 1, \ldots, I - 1$. Hence,

$$
\frac{\partial \phi_i(z_i, \theta_{i+1})}{\partial \theta_{i+1}} = - \frac{\beta_i}{\theta_{i+1}^2} \left[ v'(z_i) + \frac{z_i}{\theta_{i+1}} v''(\frac{z_i}{\theta_{i+1}}) \right] < 0, \quad j \neq i, i + 1.
$$

(70)

$$
\frac{\partial \phi_{i+1}(z_{i+1}, \theta_{i+1})}{\partial \theta_{i+1}} = \left( 1 + \beta_{i+1} \right) \left[ v'(\frac{z_{i+1}}{\theta_{i+1}}) + \frac{z_{i+1}}{\theta_{i+1}} v'' \left( \frac{z_{i+1}}{\theta_{i+1}} \right) \right] > 0, \quad j \neq i, i + 1.
$$

(71)

$$
\frac{\partial \phi_j(z_j, \theta_{i+1})}{\partial \theta_{i+1}} = 0 \text{ for } j \neq \{i, i + 1\}.
$$

(72)

By the implicit function theorem, for every $\theta_{i+1}, i = 1, \ldots, I - 1, \phi_j(z_j, \theta_{i+1}) = 0$ has a unique solution which defines $z_j$ as a $C^1$-function $z_j = \phi_j(\theta_{i+1})$, with derivative

$$
\phi'_j(\theta_{i+1}) = - \frac{\partial \phi_j(z_j, \theta_{i+1})}{\partial \theta_{i+1}} \frac{\partial \phi_j(z_j, \theta_{i+1})}{\partial z_j}.
$$

(73)

As $\partial \phi_j(z_j, \theta_{i+1})/\partial z_j \equiv \partial \phi_j(z_j, \gamma) / \partial z_j < 0$ by (68), it follows from (70)–(73) that

$$
\frac{\partial g^j_i(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} \equiv \phi'_j(\theta_{i+1}) = \begin{cases} < 0 & \text{if } j = i, \\ > 0 & \text{if } j = i + 1, \\ = 0 & \text{if } j \notin \{i, i + 1\}. \end{cases}
$$

(74)
In consequence,
\[ \frac{\partial T'(z_i; \theta_i)}{\partial \theta_i} = \nu' \left( \frac{z_i}{\theta_i} \right) \frac{\partial \phi_i}{\partial \theta_i} (\theta, \gamma, \lambda) > 0, \quad (75) \]
\[ \frac{\partial T'(z_{i+1}; \theta_{i+1})}{\partial \theta_{i+1}} = - \frac{1}{\gamma \theta_i^2} \nu'' \left( \frac{z_{i+1}}{\theta_{i+1}} \right) \frac{\partial g_{\gamma \lambda}}{\partial \theta_i} (\theta, \gamma, \lambda) < 0, \quad (76) \]
\[ \frac{\partial T'(z_j; \theta_{j+1})}{\partial \theta_{j+1}} = 0 \text{ for } j \notin \{i, i+1\}. \quad (77) \]

By Proposition 4,
\[ T'(z_i; \theta_{i+1}) \equiv \left( 1 + \frac{1}{\beta_i} \right) T'(z_i; \theta_i), \text{ } i = 1, \ldots, I - 1, \quad (78) \]
where \( \beta_i > 0 \). Therefore, (75) and (76) imply
\[ \frac{\partial T'(z_i; \theta_{i+1})}{\partial \theta_i} > 0 \text{ and } \frac{\partial T'(z_{i+1}; \theta_{i+1})}{\partial \theta_{i+1}} < 0. \quad (79) \]

**Proof of Proposition 8.** Let \( z_j = g_j' (\theta, \gamma, \lambda) \) and use (67) to define 4 to define \( \phi_j (z_j, \beta_i) = 0 \) for \( j = 1, \ldots, I - 1 \). Since \( \theta_i < \theta_{i+1} \) and \( \nu'' > 0 \),
\[ \frac{\partial \phi_j (z_j, \beta_i)}{\partial \beta_i} = \begin{cases} - \frac{1}{\beta_i} \nu' \left( \frac{z_j}{\theta_i} \right) + \frac{1}{\theta_i} \nu' \left( \frac{z_i}{\theta_i} \right) & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (80) \]

By the implicit function theorem, for every \( \theta_{i+1}, i = 1, \ldots, I - 1 \), \( \phi_j (z_j, \beta_i) = 0 \) has a unique solution which defines \( z_j \) as a \( C^1 \)-function \( z_j = \phi_j (\beta_i) \), with derivative
\[ \phi_j' (\beta_i) = - \frac{\partial \phi_j (z_j, \beta_i)}{\partial \beta_j} / \frac{\partial \phi_j (z_j, \beta_i)}{\partial z_j}. \quad (81) \]

As \( \frac{\partial \phi_j (z_j, \beta_i)}{\partial z_j} \equiv \frac{\partial \phi_j (z_j, \gamma)}{\partial z_j} < 0 \) by (68), it follows from (80)–(81) that
\[ \frac{\partial g_j' (\theta, \gamma, \lambda)}{\partial \beta_i} \equiv \phi_j' (\beta_i) \begin{cases} < 0 & \text{if } j = 1, \\ 0 & \text{otherwise.} \end{cases} \quad (82) \]

**Proof of Proposition 9.** Since \( \beta_k \) is increased for all \( k \in \{i, \ldots, j-1\} \) and unaltered otherwise, Proposition 8 implies (49) and (50).

I now establish (51) and (52). By Proposition 2,
\[ x_k^* (z; \theta, \gamma, p) = \frac{1}{I} \left\{ \sum_{h=1}^{I} z_j - \frac{1}{\gamma_{h+1}} \sum_{k=2}^{I} (I + 1 - h) \left[ v \left( \frac{z_h}{\theta_h} \right) - v \left( \frac{z_{h-1}}{\theta_h} \right) \right] \right\} + \frac{1}{\gamma_{h+1}} \sum_{h=2}^{I} \left[ v \left( \frac{z_h}{\theta_h} \right) - v \left( \frac{z_{h-1}}{\theta_h} \right) \right]. \]
Hence, for $k \geq j$,

$$
\frac{d\Delta^k(z; \theta, \gamma, p)}{ds} = \frac{d\Delta^j(z; \theta, \gamma, p)}{ds} = \frac{1}{\gamma!} \sum_{h=i}^{j-1} \left[ \gamma + (h - 1) \frac{\{z_h\}}{\theta_h} \right] \frac{dz_h}{ds} > 0, \quad (84)
$$

where the inequality follows from (50). For $k \leq i$,

$$
\frac{d\Delta^k(z; \theta, \gamma, p)}{ds} = \frac{d\Delta^i(z; \theta, \gamma, p)}{ds} = \frac{1}{\gamma!} \sum_{h=i}^{i-1} \left[ \gamma + (h - 1 - 1) \frac{\{z_h\}}{\theta_h} \right] \frac{dz_h}{ds}. \quad (85)
$$

Using (67) and the definition of $\beta$,

$$
\gamma + (h - 1 - 1) \frac{\{z_h\}}{\theta_h} = \beta_h + h - I \frac{\{z_h\}}{\theta_h} - \frac{\beta_h}{\theta_{h+1}} \frac{\{z_h\}}{\theta_{h+1}}, \quad (86)
$$

$$
= \sum_{j=1}^{h} \lambda_j \frac{\{z_h\}}{\theta_h} - \frac{\beta_h}{\theta_{h+1}} \frac{\{z_h\}}{\theta_{h+1}} < 0. \quad (87)
$$

Consequently, by (50) and (86), (85) is positive. \qed

References


