Optimality Conditions and Comparative Static Properties of Non-Linear Income Taxes Revisited

Laurent SIMULA*

PSE, Paris, and LMU, Munich†

October 11, 2007

Abstract

Optimality conditions and comparative static properties of the Mirrleesian optimal nonlinear income tax are obtained for a finite population and quasilinear-in-consumption preferences. The first implication of the linearity in consumption is that the welfare weights in the reduced form are not a function of the individual skill levels, contrary to Weymark (1987). This permits skills to be separated from underlying welfare considerations, which gives greater insight into the characterization of the optimum and its comparative properties. Interior optima are characterized by the equality between a wedge involving the ratio between the marginal tax rate of the individual for whom the bundle is designed and cumulative social weights. The second significant result concerns the effect of a change in the skill level of an individual. A slight alteration in the skill level of any individual has the surprising effect of only affecting the implicit marginal tax rates and pre-tax incomes of himself and his nearest less productive neighbour; the implicit marginal tax rate of no other consumer is modified. The third set of results deals with the comparative statics with respect to pure changes in the social weights.

Keywords: Optimal Tax, Income Tax, Comparative Statics.

JEL classification: D82; H21; H31.

*I am particularly indebted to Guy Laroque, Etienne Lehmann, Alain Trannoy and John Weymark for very detailed and insightful comments and suggestions. I am also very grateful to Peter Diamond, Marc Fleurbaey, Guy Gilbert, Nicolas Gravel, Laurence Jacquet, Emmanuelle Taugourdeau and Gwenola Trotin. The usual caveat applies.

1. INTRODUCTION

This article adds to literature that develops qualitative properties of solutions to screening problems. It analyses a finite population version of Mirrlees (1971)’s model, as in Guesnerie and Seade (1982), but concentrates on the special case where individual utility is separable in consumption and leisure, and linear in consumption. It provides a geometric characterization of the optimum allocation and derives comparative static results. The informational framework corresponds to Roberts’s (1984) assignment uncertainty. The policymaker 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 policymaker faces an adverse-selection problem and is restricted to setting a tax scheme based on gross earnings.

Merely assuming that individual preferences are concave and increasing both in consumption and leisure does not yield many general results. Indeed, the optimal tax structure is the product of different sorts 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. Moreover, 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, further 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 computational approaches, which are very useful for they allows the optimal tax rates to be quantified, but are not ideally suited for shedding light on the economic intuition behind the results.

That is why more restrictive functional forms of individual preferences have been considered in order to get clear-cut general results. Lollivier and Rochet’s (1983) observation that quasilinear-in-income preferences yield closed-form solution for the optimal income tax problem has proved particularly fruitful. Applying their basic insights to the finite population framework, Weymark (1986b) has derived a reduced-form optimal income tax problem, characterized bunching (Weymark, 1987) and provided comparative static results (Weymark, 1987). The obtained reduced form has consumption as only variable. The income distribution is then implied by the incentive compatibility constraints. Moreover, the reduced-form problem has the form of an unconstrained maximization of a weighted utilitarian social welfare function. This methodology has then 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). In the same vein, Hamilton and Pestieau (2005) have examined the impact of changing individual productivity in an economy with two classes of agents where the government adopts a maximin or maximax objective function. With a continuous population, the tractability allowed by quasilinear-in-income preferences has also 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, and by Boadway, Cuff, and Marchand (2000) and Boone and Bovenberg (2007) to characterize the optimum allocation.
Assuming quasilinear-in-income preferences offers technical advantages because they are linear with respect to the variable observed by the government and used as the tax base. This notably allows the reduced-form optimal income tax problem to have an explicit solution. There are however some limitations in the linear-in-income model. The first is that the disutility of labour is constant. Hence, when a price is varied, the change in individual consumption does only depend on the substitution effect while all income effects are absorbed by labour supply. The second is that the reduced-form problem is not as informative as it first seems. Indeed, the weights in the reduced form are a combination of the weights in the underlying welfare function and the levels of skill describing the population. This complicates the interpretation of the optimality conditions and of some comparative static results.

At first sight, working with quasilinear-in-consumption preferences is technically less tractable: the linearity with respect to the observable variable is lost, the role of the labour supply elasticity is more complex in determining the optimal tax and the reduced-form problem has no explicit solution. However, this kind of preferences are worth examining for at least three reasons. First, most of the empirical studies, though not all, give credence to small income effects relative to substitution effects as regards labour supply (Blundell, 1992, Blundell and MacCurdy, 1999). Accordingly, the case with no income effect on labour supply provides a relevant benchmark, which has been extensively used in the continuous population model since the work by Diamond (1998) (cf. Atkinson (1990), Boadway and Pestieau (2007), d’Autume (2000), Piketty (1997), Saez (2001, 2002), Salanié (1998)). Second, from the theoretical viewpoint, assuming that all income effects are absorbed by consumption is a more satisfying assumption. Otherwise, the optimal tax schedule is independent of the labour response (Boadway, Cuff, and Marchand, 2000). Third, the comparative static properties of the optimal non-linear income tax problem could differ significantly from those obtained under quasilinear-in-income preferences.

This paper exploits the fact that, when some redistribution from the more to the less productive individuals is desirable, all conditions for incentive compatibility reduce to pairwise comparisons of utility levels between adjacent individuals (Guesnerie and Seade, 1982, Hellwig, 2007, Röell, 1985). This places structure on the optimal allocation and allows the derivation of a reduced-form income tax problem involving only the allocation of gross incomes within the population. This reduced form can be seen as a special case of Chambers (1989) "concentrated" objective function derived for separable preferences. The first implication of the linearity in consumption is that the welfare weights in the reduced form are not a function of the individual skill levels. This permits skills to be separated from underlying welfare considerations, which gives greater insight into the characterization of the optimum and its comparative properties. In particular, the trade-off between equity and efficiency is very transparent when the optimum is interior: at each observed gross income level, a wedge involving the ratio between the marginal tax rate of the individual for whom the bundle is designed and that of his nearest more productive neighbour, which reflects efficiency, must be equal to a cumulative social weight. This wedge allows a very simple geometric characterization of the optimal allocation in the absence of bunching and plays an important part in the comparative static analysis. It is closely related to the single-crossing property which corresponds to a restriction on its sign, whilst the optimality conditions restrict its magnitude. This observation is novel in the literature and is potentially useful for other kinds of screening problems employing
the single-crossing property. The second significant result concerns the effect of a change in the skill level of an individual, which had not been addressed in the literature. In Mirrlees model, productivity is the sole parameter of heterogeneity within the population and the fundamental source of the self-selection problem since it basically conditions the effort a given individual has to undertake when he chooses to mimic everyone else. It is shown that a slight alteration in the skill level of any individual has the surprising effect of only affecting the implicit marginal tax rates and pre-tax incomes of himself and his nearest less productive neighbour; the implicit marginal tax rate of no other consumer is modified. The last set of results examines in which ways the optimal allocation is affected by changes in the social weights. Since our reduced form has not to use skill-normalized social weights as in Weymark (1987), our comparative static results isolate the impact of varying the redistributive tastes of the policy-maker.

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. Units of the consumption good are chosen so that one unit costs one euro. 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. This simplification is not particularly restrictive as the distance between two productivity levels is free to vary. 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 \( U : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \),

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

where \( \gamma \in \mathbb{R}_{++} \) is the marginal utility of money.\(^1\) It is assumed that the disutility of labour \( v : \mathbb{R}_+ \to \mathbb{R}_+ \) is a \( C^3 \)-function which satisfies \( v' > 0, \ v'' > 0, \ v''' > 0, \ v(0) = 0 \) and \( \lim_{\ell_i \to \infty} v'(\ell_i) = \infty \). The assumption \( v''' > 0 \) means that the marginal disutility of labour is convex. Hence, providing a marginal

\(^1\) Similarly, it is typically assumed that \( \ell_i \in \mathbb{R} \) in the quasilinear-in-income version of Mirrlees’s model (Lollivier and Rochet, 1983, Weymark, 1986b, 1987).
unit of labour time becomes increasingly unattractive when \( t_i \) goes up, which seems quite reasonable.\(^2\)

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

\[
u (x_i, z_i; \theta_i) := \gamma x_i - v (z_i / \theta_i), \quad i \in I.
\]

The marginal rate of substitution \( s (z_i; \theta_i) \) of person \( i \) at the \((x_i, z_i)\)-bundle is independent of his consumption level, with

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

Three points are worth noting in the gross-income/consumption space. First, \( i \)'s indifference curves are parallel vertical displacements of each other. Second, the Spence-Mirrlees condition is met: for a given gross income level, the higher the productivity of an individual, the flatter his indifference curves. Third, when the marginal utility of money \( \gamma \) increases, indifference curves become flatter; so a lower increase in consumption is required to compensate for an increase in gross income while keeping utility constant.

A social allocation specifies a consumption and gross income level 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

\[
\begin{aligned}
\theta & \to a : \mathbb{R} \times \mathbb{R}^+ \to \mathbb{R}^I \times \mathbb{R}^I_+ \\
\theta_i & \to (x_i, z_i),
\end{aligned}
\]

which satisfies the incentive compatibility constraints

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

and the tax revenue constraint

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

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,

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

---

\(^2\)This assumption plays an important role in the analysis. It ensures that the reduced-form optimal income tax problem has a unique solution and also appears in some comparative static results. The role of third derivatives in comparative static exercises has been emphasized by the literature devoted to risk and uncertainty. For instance, see Kimball (1990), Laffont (1989), Menegatti (2001).
in which \( \lambda := (\lambda_1, \ldots, \lambda_I) \) are individual social weights. The 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.

\[
0 < \lambda_I < \ldots < \lambda_1.
\]

If \( I = 2 \), this assumption amounts to considering the “normal” case studied by Stiglitz (1982) in which only the incentive compatibility constraint of the high type is binding. By extension, when \( I > 2 \), one can therefore expect that the only binding incentive compatibility constraints will be the downward adjacent ones stating that \( i + 1 \) must be indifferent between his own bundle and \( i \)’s one, for \( i = 1, \ldots, I - 1 \).

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

\[
\Lambda (\theta_I) = I.
\]

Consequently, admissible parameters \( (\theta, \gamma, \lambda) \) belong to the set

\[
P := [\theta] (1) \text{ is satisfied} \times \mathbb{R}_{++} \times [\lambda](10) \text{ and (11) are satisfied}.
\]

The optimal non-linear income tax problem can thus be formulated as follows.

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

### 3. THE OPTIMAL ALLOCATION

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 sets. The reduced-form problem makes it easier to interpret the social value function as well as the optimality conditions 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 \( Z \) (which will be formally defined below).

**Subproblem 1.** Given a gross income vector \( z \in Z \) and the parameters \( (\theta, \gamma, \lambda) \in P \), choose the consumption vector \( x \in \mathbb{R}^I \) to maximize the social welfare function \( W (a) \) subject to the incentive-compatibility 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

\[
\arg \max_{z \in Z} W \left( x^* (z; \theta, \gamma, \lambda), z \right).
\]
So, the reduced-form problem can be stated as follows.

Subproblem 2. Given the parameters \((\theta, \gamma, \lambda) \in \mathcal{P}\), choose \(z \in \mathcal{Z}\) to maximize the social welfare function \(W(x^* (z; \theta, \gamma, \lambda), z)\).

3.1. Implications of the Incentive-Compatibility Constraints

The incentive-compatibility constraints (7) place structure on the solution to Problem 1. These restrictions can be used to derive sufficient conditions under which an allocation \(a\) is incentive-compatible. We proceed in two steps.

First, if an allocation \(a\) satisfies (7), then the two adjacent incentive-compatibility constraints \(IC_{i+1,j}\) and \(IC_{i,j+1}\) are satisfied for \(i = 1, \ldots, I - 1\). Adding both of them yields

\[
R(z_i; \theta_i, \theta_{i+1}) \leq R(z_{i+1}; \theta_i, \theta_{i+1}), \quad i = 1, \ldots, I - 1.
\]

where the function \(R : \mathbb{R}_+ \to \mathbb{R}\) is defined by

\[
R(z_k; \theta_i, \theta_j) := v(z_k/\theta_i) - v(z_k/\theta_j), \quad (k, i, j) \in \mathcal{T}^3.
\]

Lemma 1. Let \((\theta_i, \theta_j) \in \mathbb{R}^2_+\). Then, \(R(z_k; \theta_i, \theta_j)\) is strictly increasing and strictly convex in \(z_k\) if and only if \(\theta_i > \theta_j\).

Proof. See the Appendix.

For \(i = 1, \ldots, I - 1\), applying Lemma 1 to (14) yields \(z_{i+1} \geq z_i\). So, since \(v' > 0\), \(IC_{i+1,j} \Leftrightarrow x_{i+1} - x_i \geq v(z_{i+1}/\theta_{i+1}) - v(z_i/\theta_{i+1})\) implies \(x_{i+1} \geq x_i\). Consequently, the gross income and net income vectors of an incentive-compatible allocation \(a\) must be non-decreasing in productivity, i.e. such that

\[
(x_1, z_1) \leq \cdots \leq (x_I, z_I),
\]

with \((x_{i-1}, z_{i-1}) < (x_i, z_i)\) if \((x_{i-1}, z_{i-1}) \neq (x_i, z_i), i = 2, \ldots, I\). Accordingly, the set \(\mathcal{Z}\) in which the solution in \(z\) to Problem 1 must lie is defined as

\[
\mathcal{Z} := \left\{ z \in \mathbb{R}^{|\mathcal{I}|} | 0 \leq z_1 \leq \cdots \leq z_I \right\}.
\]

The condition \(z \in \mathcal{Z}\) corresponds to the second-order condition for incentive compatibility in the continuum model.

Second, given \(z \in \mathcal{Z}\), it proves sufficient to check the downward and upward adjacent incentive compatibility constraints to get an incentive-compatible allocation provided \(z \in \mathcal{Z}\) (Cooper, 1984).

Lemma 2. Given \(z\) in \(\mathcal{Z}\),

\[
\begin{align*}
\quad u(x_i, z_i; \theta_i) & \geq u(x_{i-1}, z_{i-1}; \theta_i), \quad i = 2, \ldots, I, \quad (18a) \\
\quad u(x_i, z_i; \theta_i) & \geq u(x_{i+1}, z_{i+1}; \theta_i), \quad i = 1, \ldots, I - 1, \quad (18b)
\end{align*}
\]
propose (7).

Proof. See the Appendix. \qed

There are many ways of satisfying (18). This is notably the case if all adjacent downward incentive-
compatibility constraints are binding, i.e. if
\[ u(x_i, z_i - 1) \leq x_i \leq z_i - 1, \]
\[ u(x_i, z_i - 1) \leq z_i - 1 \leq x_i \leq z_i - 1, \]
(19)

An allocation satisfying (19) is called a simple monotonic chain to the left by Guesnerie and Seade (1982).

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

This result is established geometrically in Figure 1. Consider three successive gross income/consumption
bundles \( C_{i-1} \leq C_i \leq C_{i+1} \) and suppose person \( k \) is indifferent between \( C_k \) and \( C_{k-1} \) for \( k = i, i + 1, \)
i.e. \( u(C_i; \theta_i) = u(C_{i-1}; \theta_i) \) and \( u(C_i; \theta_{i+1}) = u(C_{i+1}; \theta_{i+1}) \). Since the Spence-Mirrlees condition is met, \( i \) (weakly) prefers \( C_i \) to \( C_{i+1} \). As a consequence, the local upward incentive-compatibility constraint is necessarily satisfied. By Lemma 2, all non-adjacent incentive-compatibility constraints for \( i \) are also verified. Repeating the argument for \( i = 2, \ldots, I - 1 \), shows that any simple monotonic chain to the left is incentive-compatible.

A simple-monotonic chain to the left reflects a specific efficiency/rent-extraction trade-off. Indeed,
given quasilinear-in-consumption preferences, (19) is equivalent to
\[ u(x_{i+1}, z_{i+1}; \theta_{i+1}) - u(x_i, z_i; \theta_i) = R(z_i; \theta_i, \theta_{i+1}), \quad i = 1, \ldots, I - 1. \]  

By Lemma 1, (20) indicates at which rate utility must be convexly increased for the tax schedule to induce individual truth telling. So, for each pair of adjacent productivity levels \( \theta_i, \theta_{i+1} \), \( R(z_i; \theta_i, \theta_{i+1}) \) may be regarded as the "marginal rent" the policymaker has to leave to the more productive \( i + 1 \)-individual 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

Since \( Z \) is closed and bounded whilst \( W(\alpha) \) is continuous, there exist solutions to Subproblem 1 for all gross income vector \( z \in Z \) and all \( (\theta, \gamma, \lambda) \in \mathcal{P} \). They all share the following remarkable structure.

**Lemma 3.** Given \( z \in Z \) and \( (\theta, \gamma, \lambda) \in \mathcal{P} \), any allocation \( a = (x^*, z) \) where \( x^* \in X^*(z; \theta, \gamma, \lambda) \), is a simple monotonic chain to the left which is production efficient.

**Proof.** See the Appendix. \( \boxdot \)

Combined with Proposition 1, Lemma 3 ensures that all implications of the incentive-compatibility constraints (7) are embedded in any solution to Subproblem 1, provided \( z \in Z \). Moreover, the fact that \( a \) is a simple monotonic chain to the left gives rise to a specific consumption pattern. Indeed, by (19),
\[ x_i = x_{i-1} + \frac{1}{\gamma} \left[ v(z_i/\theta_i) - v(z_{i-1}/\theta_i) \right], \quad i = 2, \ldots, I, \]  
and so
\[ x_i = x_1 + \frac{1}{\gamma} \sum_{j=2}^{I} \left[ v(z_j/\theta_j) - v(z_{j-1}/\theta_j) \right], \quad i = 2, \ldots, I. \]  

As any solution to Subproblem 1 is production efficient, by Lemma 3, the binding tax revenue constraint (8) can be substituted in \( \sum_{i=1}^{I} x_i \), obtained from (22), to get
\[ \sum_{i=1}^{I} z_i = Ix_1 + \frac{1}{\gamma} \sum_{i=2}^{I} \sum_{j=2}^{I} \left[ v \left( \frac{z_i}{\theta_i} \right) - v \left( \frac{z_{i-1}}{\theta_i} \right) \right] = Ix_1 + \frac{1}{\gamma} \sum_{i=2}^{I} (I + 1 - i) \left[ v \left( \frac{z_i}{\theta_i} \right) - v \left( \frac{z_{i-1}}{\theta_i} \right) \right]. \]  

This equation admits a unique solution in \( x_1 \). Substituting the latter in (22) and proceeding sequentially show that there is a unique consumption vector in \( X^*(z; \theta, \gamma, \lambda) \), which is independent of the social weights \( \lambda \) and inherits the differentiability properties of \( v \).

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

\[
x^i_1(z; \theta, \gamma) = \frac{1}{I} \left[ \sum_{j=1}^I z_j - \frac{1}{\gamma} \sum_{j=2}^I (I + 1 - j) \left( v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_{j-1}} \right) \right) \right],
\]

(24)

\[
x^i_2(z; \theta, \gamma) = x^i_1(z; \theta, \gamma) + \frac{1}{\gamma} \sum_{j=2}^i \left[ v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_{j-1}}{\theta_{j-1}} \right) \right], \quad i = 2, \ldots, I.
\]

(25)

### 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 \( \beta = (\beta_1, \ldots, \beta_I) \) with

\[
\beta_i := \Lambda(\theta_i) - i, \quad i \in \mathcal{I}.
\]

(26)

Because of (10) and (11), the graph of \( i \to \Lambda(\theta_i) \) is hump-shaped and above the 45\(^{-}\)-line. Hence, \( \beta_i > 0 \) for \( i = 1, \ldots, I - 1 \), whilst \( \beta_I = 0 \). Skill levels do not appear directly in (26), unlike in the analogous expression in Weymark (1986b).

The parameters \( \beta_i \) summarize in a transparent way the redistributive taste of the government. First, they would all be equal if the government adopted pure utilitarianism as a social objective, in which case \( \beta_i = 0 \) for every \( i \). Therefore, the social parameters \( \beta_i \) express the policymaker’s strict aversion to income inequality. Second, to get further insight into \( \beta_i \), it is instructive to consider the effects of the government’s decision to give each of the \( i \) less productive individuals one extra euro of consumption in the absence of incentive-compatibility constraints. On the one hand, the utility of each of them is increased by \( \gamma \), the marginal utility of money. Hence, the gross social benefit amounts to \( \gamma \Lambda(\theta_i) \). On the other hand, the consumption of every individual \( i = 1, \ldots, I \) must be decreased by \( i/I \) in order to satisfy the tax revenue constraint (8). This reduces individual welfare by \( \gamma i/I \) and thus social welfare by \( \Lambda(\theta_i) \times \gamma i/I = \gamma i \). Summing both effects, it appears that \( \gamma \beta_i \) is the net social benefit of marginally increasing the consumption of the \( i \) less skilled individuals whilst ignoring informational externalities. So, \( \beta_i \) is this net social benefit expressed in monetary units. Third, the parameters \( \beta_i \) can alternatively be defined as

\[
\beta_i := I - i - \sum_{j=i+1}^I \lambda_j, \quad i = 1, \ldots, I - 1.
\]

(27)

They thus also corresponds to the net social cost, expressed in euros, of marginally increasing the consumption of the \( I - i \) most productive individuals. That is why they are henceforth referred to as net cumulative social weights. Thanks to them Subproblem 2 can be rewritten as follows.

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

$$
\mathcal{W}^* (z; \theta, \gamma, \lambda) := \sum_{i=1}^{I} \left[ y z_i - v (z_i / \theta_i) \right] - \sum_{i=1}^{I} \beta_i R (z_i; \theta_i, \theta_{i+1}).
$$

(28)

$\mathcal{W}^*$ is strictly concave over the convex set $\mathcal{Z}$ because $v''$ is positive and increasing.\(^4\) Hence, there is a unique gross income vector which maximizes $\mathcal{W}^*$ for every $(\theta, \gamma, \lambda) \in \mathcal{P}$. It remains to substitute it in $x^*_i (g^z (\theta, \gamma, \lambda); \theta, \gamma)$ to get the optimal allocation.

**Proposition 3.** For $(\theta, \gamma, \lambda) \in \mathcal{P}$, there is a unique allocation $a = (g^z (\theta, \gamma, \lambda), g^z (\theta, \gamma, \lambda))$ solution to Problem 2, with

$$
\begin{align*}
\left\{ \begin{array}{l}
g^z_i (\theta, \gamma, \lambda) = \arg \max_{z_i \in \mathcal{Z}} \mathcal{W}^* (z; \theta, \gamma, \lambda) \\
g^z_i (\theta, \gamma, \lambda) = x^*_i (g^z (\theta, \gamma, \lambda); \theta, \gamma)
\end{array} \right., \quad i \in \mathcal{I}.
\end{align*}
$$

(29)

**Proof.** See the Appendix.

A first observation is that both $g^z (\theta, \gamma, \lambda)$ and $g^z (\theta, \gamma, \lambda)$ are functions. Moreover, Proposition 3 implies that the social allocation solution to the optimal non-linear income tax problem is a monotonic chain to the left. As a 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 as

$$
T' (z_i; \theta_j) := 1 - s (z_i; \theta_j) = 1 - \frac{v' (z_i / \theta_j)}{\gamma \theta_j}, \quad (i, j) \in \mathcal{I}^2.
$$

(30)

Every $T' (z_i; \theta_j)$ is less than one since $v' > 0$ and decreases in $z_i$ since $v'' > 0$. Since at the optimum only the adjacent downward incentive-compatibility 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 $i$ for whom the $(x^*_i (z_i; \theta_i), z_i)$-bundle is designed and the implicit marginal tax rate $T' (z_i; \theta_{i+1})$ the nearest more productive $i + 1$-individual would face if he were mimicking the $i$-individual.

The implicit marginal tax rates allow us to get further understanding of the reduced-form objective function $\mathcal{W}^* (z; \theta, \gamma, \lambda)$. 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

$$
\frac{d \mathcal{W}^* (z; \theta, \gamma, \lambda)}{dz_i} = \gamma T' (z_i; \theta_i) - \beta_i R' (z_i; \theta_i, \theta_{i+1}), \quad i \in \mathcal{I},
$$

(31)

by (28) and (30), the impact on social welfare may be thought of as proceeding in two steps. In the first step, the $\theta_i$-individual pays $T' (z_i; \theta_i)$ additional euros in taxes, which relaxes the tax revenue constraint

\(^{3}\)Since $\beta_i = 0$, $R (z_i; \theta_i, \theta_{i+1})$ is defined arbitrarily.

\(^{4}\)Since $v'' > 0$, $\beta_i \geq 0$ and Lemma 1 holds, $d^2 \mathcal{W}^* (z; \theta, \gamma, \lambda) / dz_i^2 = -v'' \left( \frac{z_i}{\theta_i} \right) / \theta_i^2 - \beta_i R'' (z_i; \theta_i, \theta_{i+1})$ is strictly negative.
(8). As \(\gamma\) is the marginal utility of money, the positive effect on social welfare amounts to \(\gamma \, T' (z_i; \theta_i)\). In the second step, the effect on incentives is taken into account. Person \(i\) receives \(1 - T' (z_i; \theta_i)\) extra euro of consumption. As a result, the \(I - i\) more productive individuals have to sacrifice less consumption when they decide to mimic \(i\). So, cheating becomes more attractive to them. For \(i + 1\), the marginal rent is \(R' (z_i; \theta_i, \theta_{i+1})\). This person’s binding incentive-compatibility constraint \(IC_{i+1, i}\) is restored by adjusting consumption by \(R' (z_i; \theta_i, \theta_{i+1}) / \gamma\). Increasing the consumption of everybody of higher ability by the same amount preserves the monotonic chain to the left. Because \(\gamma \beta_i\) is the net social cost of marginally increasing consumption of the \(I - i\) most productive individuals, social welfare is reduced by \(\beta_i R' (z_i; \theta_i, \theta_{i+1})\). If the social optimum is interior, it is therefore obtained when the positive effect on social welfare due to the relaxation of the tax revenue constraint offsets the negative one stemming from private information, i.e.

\[
\frac{dW^* (z; \theta, \gamma, \lambda)}{dz_i} = 0 \Rightarrow T' (z_i; \theta_i) = \frac{\beta_i}{\gamma} R' (z_i; \theta_i, \theta_{i+1}), \quad i \in I.
\]

(32)

### 3.4. Characterization of the Optimal Allocation

In order to characterize the optimal allocation, it is useful to consider the relaxed form of Problem 2, obtained when all monotonicity conditions on \(z\) are removed (but not the non-negativity constraint \(z \geq 0\)). Its unique solution is denoted \(\tilde{z} (\theta, \gamma, p) = (\tilde{z}_1 (\theta, \gamma, p), \ldots, \tilde{z}_I (\theta, \gamma, p))\). If \(\tilde{z} (\theta, \gamma, p)\) is non-decreasing, then it is equal to the socially optimal gross income vector \(g^* (\theta, \gamma, \lambda)\). Otherwise, the optimal allocation involves bunching.

**Proposition 4.** For \((\theta, \gamma, \lambda) \in \mathcal{P}\), the optimal allocation is such that:

(i) \(g^*_i (\theta, \gamma, \lambda) = \tilde{z}_i (\theta, \gamma, p)\), except on a finite number \(K\) of disjoint compact sets \(B^k := \{i^k, \ldots, j^k\}\), \(k = 1, \ldots, K\), \(i^k\) increasing with \(k\), where \(g^*_n (\theta, \gamma, \lambda) = \tilde{z}^k\) for every \(n \in B^k\);

(ii) \(0 < T' (z_i; \theta_i) < 1\) for \(i = 1, \ldots, I - 1\), \(T' (z_1; \theta_1) = 0\) and bunching at the top is ruled out;

(iii) if \(i^k = 1\),

\[
\frac{1}{\#B^k} \sum_{n \in B^k} \left. \frac{\partial W^* (z; \theta, \gamma, \lambda)}{\partial z_n} \right|_{z_n = \tilde{z}^k} \leq 0 \quad \left(= 0 \text{ if } \tilde{z}^k > 0\right);
\]

(33)

(iv) for each interior \(B^k\),

\[
\frac{1}{\#B^k} \sum_{n \in B^k} \left. \frac{\partial W^* (z; \theta, \gamma, \lambda)}{\partial z_n} \right|_{z_n = \tilde{z}^k} = 0.
\]

(34)

**Proof.** See the Appendix. \(\square\)

The optimal gross income vector is thus solution to the relaxed Problem 2, except on a finite number of bunching sets. Two types of bunching must be distinguished. Bunching due to the violation of the non-negativity constraints, called \(z = 0\) bunching by Boone and Bovenberg (2007), can only occur at the bottom, in which case \(\tilde{z}^1 = 0\). The other kind of bunching stems from the violation of the monotonicity constraints and can happen either at the bottom or in the interior of the skill distribution. In this situation, the gross income level \(\tilde{z}^k\) of the individuals who are bunch together is implicitly determined by the average of their first-order conditions for the relaxed Problem 2. The fact that there is no bunching
and non distortion at the top follows from an efficiency argument, illustrated in Figure 2. To see why, assume $T'(z_I; \theta_I) \neq 0$ and consider the 45° line through $(x_I, z_I)$. Moving along this line above $(x_I, z_I)$ increases $I$’s utility without hurting neither the tax revenue constraint (8) nor the incentive-compatibility constraints (7) since a new simple monotonic chain to the left can easily be constructed. Therefore, the initial situation is strongly Pareto-dominated and $T_0 = \frac{z_I}{D_0}$ at the optimum. Now, assume that $I = 1$ and $I$ are bunched together (possibly with other individuals). They both face a marginal tax rate equal to zero. Hence,

$$T'(z_I; \theta_I) = 0 \iff z_I = \theta_I v'^{-1}(\gamma \theta_I),$$

$$T'(z_{I-1}; \theta_{I-1}) = 0 \iff z_{I-1} = \theta_{I-1} v'^{-1}(\gamma \theta_{I-1}).$$

Since they are bunched together, it must be $z_I = z_{I-1}$. However, this is impossible because $v'' > 0$ implies that $\theta v'^{-1}(\gamma \theta)$ is strictly increasing in $\theta$.

To gain further insights into the optimal allocation, it is useful to define $a_i : [0, \theta_i v'^{-1}(\gamma \theta_I)] \rightarrow \mathbb{R}_+$ as

$$a_i(z_i; \theta_i, \theta_{i+1}, \gamma) := \frac{R'(z_i; \theta_i, \theta_{i+1})}{\gamma T'(z_i; \theta_i)} = \frac{T'(z_i; \theta_{i+1}) - T'(z_i; \theta_i)}{T'(z_i; \theta_i)}, \quad i = 1, \ldots, I - 1. \quad (37)$$

For $i < I$, the domain of $a_i$ corresponds to the gross incomes for which $T'(z_i; \theta_i)$ is strictly positive; hence, by Proposition 4, $g^*_I(\theta, \gamma, \lambda) \in [0, \theta_i v'^{-1}(\gamma \theta_i)]$. Moreover, $a_i$ is continuous and strictly increasing over its domain.\(^5\) Using (32) and (37), \(z_i\) must satisfy

$$a_i(\tilde{z}_i; \theta_i, \theta_{i+1}, \gamma) \geq \frac{1}{\beta_i} \left(= \text{if } \tilde{z}_i > 0\right), \quad i = 1, \ldots, I - 1. \quad (38)$$

\(^5\)For $i < I$, $a_i$ is strictly increasing since $dT''(z_i; \theta_i)/dz_i < 0$ and $R''(z_i; \theta_i, \theta_{i+1}) > 0$ by Lemma 1.
So, if \( z_1 > 0, z_i = a_i^{-1}(1/\beta_i; \theta_i, \theta_{i+1}, \gamma) \). Since the optimal gross income vector \( g^* (\theta, \gamma, \lambda) \) lies in the interior of \( Z \) if and only if \( 0 < z_1 < \ldots < z_I \), the following characterization is obtained.

**Proposition 5.** For \( (\theta, \gamma, \lambda) \in \mathcal{P} \), the optimum is fully separating if and only if \( \beta_i \in a_i ([0, \theta_i \upsilon_i^{-1}(\gamma \theta_i)] \) for \( i = 1, \ldots, I - 1 \) with

\[
0 < a_i^{-1}(1/\beta_1; \theta_1, \theta_2, \gamma) < \ldots < a_i^{-1}(1/\beta_{I-1}; \theta_{I-1}, \theta_I, \gamma) < \theta_i \upsilon_i^{-1}(\gamma \theta_I). \tag{39}
\]

This proposition implicitly characterizes the set of parameters for which there is no bunching, which is denoted \( \mathcal{P}^0 \) for easy reference.\(^6\) When \( (\theta, \gamma, \lambda) \in \mathcal{P}^0 \), it follows from (38) that the optimality conditions can be written in a strikingly simple form.

**Proposition 6.** For \( (\theta, \gamma, \lambda) \in \mathcal{P}^0 \), \( z \) is socially optimal if and only if

\[
a_i (z_i; \theta_i, \theta_{i+1}, \gamma) = 1/\beta_i, \quad i = 1, \ldots, I - 1, \tag{40}
\]

and \( T^i (z_i; \theta_i) = 0. \)

For a gross income \( z_i, a_i (z_i; \theta_i, \theta_{i+1}, \gamma) \) tells us to to which extent \( i + 1 \) must face a higher implicit marginal tax rate than \( i \). Geometrically, it thus corresponds to the tangent of the angle between the indifference curves of \( i \) and \( i + 1 \) divided by \( T^i_i (z_i; \theta_i) \), as shown in Figure 3. The wedge \( a_i (z_i; \theta_i, \theta_{i+1}, \gamma) \) is closely related to the single-crossing condition and thus henceforth referred to as the *Spence-Mirrlees wedge*. Indeed, the single-crossing condition corresponds to a restriction on its sign, which must be strictly positive. This condition is herein automatically satisfied because individual preferences are quasi-linear. The conditions for social optimality (40) introduce an additional restriction on the magnitude

\(^6\) \( a_i \) is type-dependent. Therefore, contrary to Weymark (1986a), condition (39) does not only depend on \( \beta \), but also on \( \theta \) and \( \gamma \).
FIGURE 4: Construction of the Optimal Allocation

An allocation is socially optimal only if, at each observed gross income level $z_i$, the Spence-Mirrlees wedge is entirely determined by the exogenously given cumulative social weight $\beta_i$. In addition, efficiency requires the labour supply of the more productive individuals not to be distorted. Hence, by (5), one obtains $z_i = (\gamma \theta_i) - \sum_{j=1}^{i-1} z_j$, which is independent of the social weights.

In the absence of bunching, Proposition 6 allows a simple two-step geometric construction of the optimal allocation, as illustrated in Figure 4. In the first step, the tax revenue constraint is ignored. Start at $(0, \bar{x}_i)$, where $\bar{x}_i$ is arbitrary. Move along $\theta_i$'s indifference curve until $\alpha_i (z_i; \theta_i, \theta_{i+1}, \gamma) = 1/\beta_i$. This point is $(\bar{x}_i, z_i)$. Then, move along $\theta_2$'s indifference curve through $(\bar{x}_i, z_i)$ until $\alpha_2 (z_i; \theta_2, \theta_3, \gamma) = 1/\beta_2$. This point is $(\bar{x}_i, z_i)$. And so on until $i = I - 1$. The determination of $z_I$ exploits the no-distortion-at-the-top result. Starting from $(\bar{x}_{I-1}, z_{I-1})$, gross income is increased along the indifference curve of the most productive individual until the tangent to this line has slope one. By construction, the allocation $(\bar{x}, z)$ is a simple monotonic chain to the left and is thus incentive compatible. However, it is not necessarily budget-balanced. That is why, in the second step, each $\bar{x}_i$ is varied by a same amount $\epsilon = \frac{1}{I} \sum_{i=1}^{I} (z_i - \bar{x}_i)$ so as to get a binding tax revenue constraint. The resulting incentive-compatible and production-efficient allocation $(x, z) = (\bar{x} + \epsilon, z)$ is socially optimal.

Before going further and derive comparative static properties, it is instructive to examine one main source of differences between our results and those derived in Weymark (1986a,b, 1987). In the latter papers, the quasilinear-in-leisure utility function $u (x_i, z_i; \theta_i) := h (x_i) - \gamma z_i / \theta_i$ 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 replace $\sum_{i=1}^l z_i$ by $\sum_{i=1}^l x_i$ and get
\[
\sum_{i=1}^l \tilde{u}(x_i, z_i; \theta_i) = \sum_{i=1}^l \theta_i h(x_i) - \gamma \sum_{i=1}^l z_i = \sum_{i=1}^l \theta_i h(x_i) - \gamma \sum_{i=1}^l x_i.
\]

This step is required to obtain a reduced-form optimal income tax problem. This is why skill-normalized social weights $e_i$ are used in the social objective $P_i d_1 e_i u(x_i, z_i; \theta_i)$. The first-order conditions of the reduced-form problem involve therefore skill-normalized cumulative social weights $\sum_{j=1}^l \tilde{e}_j$ instead of $\Lambda_i$. So, the impact of the policymaker’s taste for redistribution is less transparent because social weights and productivity levels are mixed together.

In Weymark’s reduced-form problem, the social objective is maximized with respect to consumption levels. The parameters $\sum_{j=1}^l \tilde{\lambda}_j, \theta_i$ and $\theta_{i+1}$ appear in the first-order condition for $x_i$. As a consequence, $x_i$ is a function of all social weights $\tilde{\lambda}_j$ and all productivity levels. A contrario, in the present setting, the variable with respect to which the reduced-form objective is maximized, $z_i$, does not depend on $\theta_k$ for $k \neq i, i + 1$. Therefore, one expects the comparative statics to differ significantly, from one kind of quasilinearity to the other, when skill levels are directly involved in the analysis.

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. For this purpose, it is first necessary to examine the differentiability properties of the main functions $g^e(\theta, \gamma, \lambda)$ and $g^z(\theta, \gamma, \lambda)$. Using Property 4 is is possible to establish that both functions are continuously differentiable provided a change in $(\theta, \gamma, \lambda)$ does not modify the sets $B^k$ of individuals bunched together. In this case, calculus techniques makes it possible to obtain comparative static properties. However, for expositional reasons, the analysis will henceforth consider that the initial allocation at stake is fully separating.

**Proposition 7.** The functions $g^e$ and $g^z$ are $C^1$ at every $(\theta, \gamma, \lambda)$ in $P^0$.

**Proof.** See the Appendix.

4.1. Comparative Statics for the Marginal Utility of Money

The marginal utility of money $\gamma$ measures the intensity of the individual preference for private consumption. When it goes up, the marginal utility of leisure expressed in consumption good is reduced. So, the indifference curves of every individual become flatter in the $(x, z)$-space. Everyone is thus willing to sacrifice more leisure to obtain a certain amount of additional consumption. From a more general viewpoint, changing $\gamma$ at the margin casts light on how taxes should be adjusted in a country where all individuals would like to work more to consume more.

**Proposition 8.** For $(\theta, \gamma, \lambda) \in P^0$ and $i \in I$, \( \partial g^z(\theta, \gamma, \lambda) / \partial \gamma > 0 \).

\footnote{Cf. (A.13) in Weymark (1986b).}
Proof. See the Appendix.

The changes in the tax system should not discourage more hard-working individuals to earn more. National income is increased and thus, since production efficiency is preserved by Lemma 3, total consumption goes up. However, as shown by Proposition 2, the change in person $i$’s consumption depends on how the gaps $v(z_j/\theta_i) - v(z_{j-1}/\theta_i)$, $j = 1, \ldots, i$, react to a change in the marginal utility of money. Because there is no reason why these gaps should be affected in a systematic manner, the effect of varying consumption cannot be signed without introducing further restrictions. In the same vein, the changes in the implicit marginal tax rates $T(z_i/\theta_i)$ cannot be obtained since $z_i$ and $\gamma$ are simultaneously increased.

4.2. Comparative Statics for Individual Productivities

In Mirrlees’s model, individuals are born with different abilities to turn effort into output. These fixed skill levels are the sole source of heterogeneity within the population. They probably constitute the most basic ingredients of the model as they give rise to the adverse selection problem. Indeed, the effort level $z_j/\theta_i$ a given $i$-individual must provide to earn the gross income of everyone else depends on his productivity. In practice, skill levels can be subject to changes. For instance, a high skilled individual can catch an illness which impairs his productivity while a low skilled can benefit from on-the-job training. Another argument rests on the technological side of the economy. Since a person’s productivity depends on the capitalistic intensity in his branch of activity, a new investment can make him more productive. It seems therefore worthwhile to examine what is the impact of changing a person’s skill level on his own choices, but also on the policy-maker and the whole population.\(^8\)

**Proposition 9.** For $(\theta, \gamma, \lambda) \in \mathcal{P}^0$ and $(i, j) \in \{1, \ldots, I - 1\} \times \mathcal{I}$,

\[
\frac{\partial z_i}{\partial \theta_{i+1}} < 0, \quad \frac{\partial z_{i+1}}{\partial \theta_{i+1}} > 0, \quad \frac{\partial z_j}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i + 1\},
\]

and

\[
d T'(z_i; \theta_i)/d\theta_{i+1} > 0, \quad d T'(z_{i+1}; \theta_{i+1})/d\theta_{i+1} > 0, \quad d T'(z_{i+1}; \theta_{i+2})/d\theta_{i+1} < 0 \text{ when } i \neq I - 1, \quad d T'(z_{i+1}; \theta_{i+1})/d\theta_{i+1} = 0 \text { when } i = I - 1, \quad d T'(z_j; \theta_j)/d\theta_{i+1} = d T'(z_{j+1}; \theta_{j+1})/d\theta_{i+1} = 0 \text{ for } j \notin \{i, i + 1\},
\]

where $z = g^z(\theta, \gamma, \lambda)$.

\(^8\)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.
Increasing the productivity of the \( i+1 \)-individual does only alter his gross income and that of his nearest less productive neighbour. Indeed, by Proposition 6, only \( a_i (z_i; \theta_i, \theta_i+1, \gamma) \) and \( a_{i+1} (z_{i+1}; \theta_i+1, \theta_{i+2}, \gamma) \) depend on \( \theta_{i+1} \). Accordingly, the gross income levels of all other individuals, \( j \neq i, i+1 \), remain unaltered. As regards persons \( i \) and \( i+1 \), the adjustment process combines three effects, which are illustrated in Figure 5.

First, the variation in \( \theta_{i+1} \) gives rise to a local substitution effect. The increase in person \( i+1 \)’s productivity 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, \( i+1 \) has to provide less effort if he wants to imitate \( i \). Consequently, his indifference curve through \( i \)’s gross-income/consumption bundle 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, \( i \) incurs an informational externality induced by the incentive effect. Since the cumulative social weight \( \beta_i \) is unaltered, the wedge \( a_i (z_i; \theta_i, \theta_{i+1}, \gamma) \) must stay constant (Proposition 6). 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 \( i \)’s net-of-tax wage rate. Finally, the substitution effect leads \( i \) to work less.

The changes in gross income ensure that a new monotonic chain to the left \( a' \) is obtained. However,
this incentive-compatible allocation is not necessarily budget-balanced. Therefore, in a second step, all consumption levels are adjusted by a same amount in order to obtain a production-efficient allocation. This corresponds to a vertical displacement of all indifference curves through the bundles of $a'$. However, the direction of this displacement cannot be signed in the general case, which explains why comparative static as regards consumption is ambiguous.

### 4.3. Comparative Statics for the Social Weights

Since welfare weights in the reduced form are not a function of the skill levels, it is possible to examine how pure changes in the policy-maker tastes for redistribution alter the optimal allocation. Different kinds of changes could be considered. We concentrate herein on the scenario where the policy-maker simultaneously decides to give more weight to an individual and less to another one. In practice, social and political changes often promote greater social concern for some groups of the population while the social weight of others diminishes.

It is useful to start with the impact of changing $\beta_i$, with $i < I$.\(^9\) It directly follows from Proposition 6. Indeed, by (40), $g^*_k (\theta, \gamma, \lambda)$ is independent of $\beta_i$ for $i \neq j$. Moreover, since $a'_i (\cdot; \theta_i, \theta_{i+1}, \gamma)$ is strictly increasing around the optimum, (40) also implies that a rise in $\beta_i$ reduces $g^*_j (\theta, \gamma, \lambda)$.

**Proposition 10.** For $(\theta, \gamma, \lambda) \in \mathcal{P}^0$ and $(i, j) \in \{1, \ldots, I-1\} \times \mathcal{I}$ with $i \neq j$, $\partial g^*_i (\theta, \gamma, \lambda) / \partial \beta_i < 0$ and $\partial g^*_j (\theta, \gamma, \lambda) / \partial \beta_i = 0$.

Given these preliminary results, the impact of an increase in the individual social weight of the $\theta_i$-individual to the detriment of a more productive $\theta_j$-individual can now be considered. Formally, the decrease in $\lambda_j$ is fully compensated by an increase in $\lambda_i$, i.e. $d\lambda_i = -d\lambda_j$, while all other social weights are kept constant. 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 10, 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 the consumption levels and indirect utilities $V_k (\theta, \gamma, \lambda) := u \left( g^*_k (\theta, \gamma, \lambda), g^*_k (\theta, \gamma, \lambda); \theta_k \right)$ can also be signed for all $k \notin \{i, \ldots, j-1\}$.

**Proposition 11.** Let $(\theta, \gamma, \lambda) \in \mathcal{P}^0$, $i \in \{1, \ldots, I-1\}$ and $j \in \{i+1, \ldots, I\}$. Let $\lambda : S \to \mathbb{R}^I$, where $S = (-1, 1)$, be $C^1$ with

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

\(^9\)A change in $\beta_I$ is impossible since, by definition, $\beta_I = 0$. 

19
Then, if \( i \) is increased to the detriment of \( j \),

\[
\begin{align*}
\frac{dz_k}{ds} &< 0, & \forall k \in [i, \ldots, j - 1], \\
\frac{dz_k}{ds} &> 0, & \forall k \not\in [i, \ldots, j - 1], \\
\frac{dx_k^*}{ds} &< 0, & \forall k < i, \\
\frac{dx_k^*}{ds} &> 0, & \forall k \geq j, \\
\frac{dV_k}{ds} &< 0, & \forall k < i, \\
\frac{dV_k}{ds} &> 0, & \forall k \geq j,
\end{align*}
\]

(50)

(51)

(52)

where \( z \equiv g^2(\overline{z}, \overline{\tau}, \overline{\xi}) \) and \( x^* \equiv x^*(z; \overline{u}, \overline{\tau}) \).

**Proof.** See the Appendix. \( \square \)

Before interpreting these results, it is worth examining the impact of this change in the social weights on the implicit optimal marginal tax rates. By (30), they only depend on \( z \) through gross income \( z \).

It thus follows from (50) that \( T_0(x_k^*, z_k, \theta_k) \) and \( T_0(x_k^*, z_k, \theta_{k+1}) \) are unaltered for \( k \not\in [i, \ldots, j - 1] \) whilst \( T'(x_k^*, z_k, \theta_k) \) and \( T'(x_k^*, z_k, \theta_{k+1}) \) are increased for \( k \in [i, \ldots, j - 1] \).

For concreteness, let us consider that the population consists of three individuals and that person 2’s social weight is increased at the expense of person 3. Let \((\overline{x}, \overline{z})\) be the initial allocation and denote by \((x, z)\) the new one. The changes in gross income have already been explained. The adjustments in consumption can be thought of as proceeding in two steps. In the first step, the budget constraint (8) is left aside. By (50), the gross income levels of persons 1 and 3 are held fixed, i.e. \( z_1 = \overline{z}_1 \) and \( z_3 = \overline{z}_3 \), while \( z_2 \) is reduced (by \( dz_2 \)). As a consequence, the requirement that person 3 is indifferent between his own bundle and person 2’s one induces a decrease in the consumption levels \( \overline{x}_2 \) and \( \overline{x}_3 \) of both more productive individuals (by \( dx_2 \) and \( dx_3 \) respectively) as well as in person 3’s indirect utility. Person 2’s is restored to his initial indifference curve and a new monotonic chain to the left is obtained. As person 2 faces a strictly positive marginal tax rate, he reduces his consumption by a smaller amount than his gross income, i.e. \(-d\overline{x}_2 > -d\overline{z}_2\). So, \( d\overline{x}_3 \) can be sufficiently small for

\[
\sum_{i=1}^{3} \overline{x}_i - d\overline{x}_2 - d\overline{x}_3 < \sum_{i=1}^{3} \overline{z}_i - d\overline{z}_2.
\]

(53)

As \( \sum_{i=1}^{3} \overline{x}_i = \sum_{i=1}^{3} \overline{z}_i \), (53) means that the new monotonic chain to the left is not production efficient: total labour is in excess relative to total consumption. The second step consists therefore in giving

\[
\epsilon = \frac{1}{3} (d\overline{x}_2 + d\overline{x}_3 - d\overline{z}_2) > 0
\]

(54)

\[10\]The inequalities in (50)–(52) are reversed if \( i \) is decreased to the benefit of \( j \).
additional euros of consumption to each individual. The new consumption levels are thus the following:

\[
x_1 = \bar{x}_1 + \frac{1}{3} (d\bar{x}_2 + d\bar{x}_3 - d\bar{x}_2) > \bar{x}_1, \tag{55}
\]

\[
x_2 = \bar{x}_2 - \frac{2}{3} d\bar{x}_2 + \frac{1}{3} (d\bar{x}_3 - d\bar{x}_2), \tag{56}
\]

\[
x_3 = \bar{x}_3 - \frac{2}{3} d\bar{x}_3 + \frac{1}{3} (d\bar{x}_2 - d\bar{x}_2) < \bar{x}_3. \tag{57}
\]

Hence, person 1’s enjoys greater consumption, contrary to person 3. The change in person 2’s consumption is ambiguous. It is positive if and only if \(d\bar{x}_2 < d\bar{x}_3 - 2d\bar{x}_2\). The variations in the indirect utilities directly follow from those in gross income and 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. Therefore, the respective influences of individual productivities and social weights are easier to separate in the social objective function of the reduced-form optimal income tax problem. The Spence-Mirrlees wedge is a key determinant of the properties of the optimal solution and plays an important role in signing the comparative statics. This observation is novel and is potentially useful in deriving comparative statics for any self-selection problem, not just the optimal tax problem, in which adjacent incentive constraints bind.

APPENDIX

Proof of Lemma 1. Differentiating \( R \left( z_k; \theta_i, \theta_j \right) \) yields \( R' \left( z_k; \theta_i, \theta_j \right) = v' \left( z_k / \theta_j \right) / \theta_i - v' \left( z_k / \theta_j \right) / \theta_j \) and \( R'' \left( z_k; \theta_i, \theta_j \right) = v'' \left( z_k / \theta_i \right) / \theta_i^2 - v'' \left( z_k / \theta_j \right) / \theta_j^2 \). The results follow from \( v'' > 0 \) and \( v'' > 0 \) respectively.

Proof of Lemma 2. Let \((i, j, k) \in T^3 \) with \(i < j < k\). Adding \( IC_{ij} \) and \( IC_{jk} \), one obtains

\[
x_i - v \left( \frac{z_i}{\theta_j} \right) \geq v \left( \frac{z_j}{\theta_j} \right) - v \left( \frac{z_k}{\theta_j} \right) - v \left( \frac{z_j}{\theta_i} \right) + x_k. \tag{A.1}
\]

(A.1) implies that \( IC_{ik} \) holds if and only if

\[
v \left( \frac{z_i}{\theta_j} \right) - v \left( \frac{z_j}{\theta_i} \right) - v \left( \frac{z_j}{\theta_j} \right) \geq -v \left( \frac{z_k}{\theta_j} \right) \Leftrightarrow R \left( z_i; \theta_i, \theta_j \right) \leq R \left( z_k; \theta_i, \theta_j \right), \tag{A.2}
\]
which is satisfied since $R(A; \theta_i, \theta_j)$ is strictly increasing and $z_j \leq z_k$ by (16). Hence, $IC_{ij}$ and $IC_{jk}$ imply $IC_{ik}$. As a consequence, (i) $IC_{1,2}$ implies $IC_{1,k}$ for $k > 2$, (ii) for $i = 2, \ldots, I - 1$, $IC_{i,i-1}$ and $IC_{i,i+1}$ imply $IC_{i,k}$ for $k \neq i - 1, i, i + 1$ and (iii) $IC_{1,I-1}$ implies $IC_{i,k}$ for $k < I - 1$. □

Proof of Lemma 3. (i) $a$ is a simple monotonic chain to the left. The proof proceeds by contradiction. Assume $a$ is not a simple monotonic chain to the left, i.e. $u \left( x^*_j, z_j; \theta_j \right) \neq u \left( x^*_{j-1}, z_{j-1}; \theta_j \right)$ for some $j \geq 2$. This is equivalent to considering that there exists $j \geq 2$ for which

$$\gamma x^*_j - v \left( z_j/\theta_j \right) > \gamma x^*_{j-1} - v \left( z_{j-1}/\theta_j \right).$$  
(A.3)

For (A.3) to be satisfied, $(x_j, z_j) > (x_{j-1}, z_{j-1})$. Hence, (i) is established if $z_j \leq z_{j-1}$. If $z_j > z_{j-1}$, let $x_i = x^*_i + \epsilon$ for $i = 1, \ldots, j - 1$, and $x_j = x^*_j + \frac{j-1}{j+1} \epsilon$ for $i = j, \ldots, I$, where $\epsilon > 0$ is arbitrarily chosen. The new allocation $(\bar{x}, z)$ satisfies all incentive compatibility constraints for sufficiently small $\epsilon$ (Lemma 2) and is feasible because $\sum x_i = \sum x^*_i$. In addition,

$$W(\bar{x}, z) - W(x^*, z) = \gamma \sum_{i=1}^{j-1} \lambda_i (x^*_i - x^*_i) = \gamma \left[ \sum_{i=1}^{j-1} \lambda_i \epsilon - \sum_{i=j}^{l-1} \lambda_i \frac{j-1}{j+1} \epsilon \right],$$  
(A.4)

which can be minoried thanks to (10) to get

$$W(\bar{x}, z) - W(x^*, z) \geq \gamma \left( j - 1 \right) \epsilon \left[ \hat{\lambda}_{j-1} - \hat{\lambda}_j \right] > 0,$$

(A.5)

contradicting $x^* \in X^*(\bar{x}, z, \gamma, \lambda)$.

(ii) $a$ is production efficient. Fix $\bar{x}$ in $\mathcal{Z}$. The constraints (7) are satisfied, with $z_i = \bar{x}_i$ and $x_i = x^*_i$. The proof proceeds by contradiction. Assume (8) is not binding. As a consequence, (8) is still satisfied if every $x^*_i$ is increased by a sufficiently small $\epsilon > 0$. This increase is incentive compatible since a same amount $\gamma \epsilon$ is added to both sides of $IC_{i,j}$ for every $(i, j) \in \mathcal{I}^2$. A Pareto-improving change is thus feasible, contradicting $x^* \in X^*(\bar{x}, z, \gamma, \lambda)$.

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

$$u \left( x^*_j, z_i; \theta_i \right) = u \left( x^*_1, z_i; \theta_i \right) + \sum_{j=1}^{I-1} R \left( z_j; \theta_j, \theta_{j+1} \right), \quad i = 2, \ldots, I,$$  
(A.6)

from which

$$\sum_{i=1}^{I} u \left( x^*_i, z_i; \theta_i \right) = I u \left( x^*_1, z_1; \theta_1 \right) + \sum_{i=1}^{I-1} (I - i) R \left( z_i; \theta_i, \theta_{i+1} \right).$$  
(A.7)

In addition, summing (4) over $i$ on $\mathcal{I}$ and employing the equality form of (8),

$$\sum_{i=1}^{I} u \left( x^*_i, z_i; \theta_i \right) = \gamma \sum_{i=1}^{I} z_i - \sum_{i=1}^{I} v \left( z_i/\theta_i \right).$$  
(A.8)
Using (A.6) and (A.9),
\[
u (x^*_i, z_1; \theta_1) = \frac{1}{T} \left[ \gamma \sum_{i=1}^{I} z_i - \sum_{i=1}^{I} v \left( \frac{z_i}{\theta_i} \right) - \sum_{i=1}^{l-1} (I-i) R (z_i; \theta_i, \theta_{i+1}) \right],
\]
(A.9)

Using (A.6) and (A.9),
\[
W = u (x^*_i, z_1; \theta_1) \sum_{i=1}^{I} \lambda_i + \sum_{i=2}^{I} \sum_{j=1}^{l-1} \lambda_i R (z_j; \theta_j, \theta_{j+1}) = u (x^*_i, z_1; \theta_1) I + \sum_{i=1}^{l-1} \left( \sum_{j=1}^{l-1} \lambda_j \right) R (z_i; \theta_i, \theta_{i+1})
\]
\[
= \sum_{i=1}^{I} \left[ \gamma z_i - v \left( \frac{z_i}{\theta_i} \right) + \left( i - \sum_{j=1}^{i} \lambda_j \right) R (z_i; \theta_i, \theta_{i+1}) \right],
\]
(A.10)
in which \( R (z_i; \theta_i, \theta_{i+1}) \) is an arbitrary number. \( \square \)

Proof of Proposition 4. The Lagrangian for Problem 2 writes
\[
\mathcal{L} = W^* (z; \theta, \gamma, \lambda) + \mu_1 z_1 + \mu_2 (z_2 - z_1) + \cdots + \mu_I (z_I - z_{I-1}),
\]
yielding the following first-order and complementarity-slackness conditions:
\[
\frac{\partial \mathcal{L}}{\partial z_i} = \frac{\partial W^* (z; \theta, \gamma, \lambda)}{\partial z_i} / \partial z_i + \mu_i - \mu_{i+1} = 0, \quad i = 1, \ldots, I - 1, \quad (A.12)
\]
\[
\frac{\partial \mathcal{L}}{\partial z_I} = \frac{\partial W^* (z; \theta, \gamma, \lambda)}{\partial z_I} / \partial z_I + \mu_I = 0, \quad (A.13)
\]
\[
\mu_1 \geq 0 \quad (= \text{if } z_1 > 0), \quad (A.14)
\]
\[
\mu_i \geq 0 \quad (= \text{if } z_i > z_{i-1}), \quad i = 2, \ldots, I. \quad (A.15)
\]

Point (i). If \( i \notin B^k \forall k, z_{i-1} < z_i < z_{i+1} \). Then, by (A.15), \( \mu_i = \mu_{i+1} = 0 \). Therefore, by (A.12), \( z_I \) satisfies \( \partial W^* (z; \theta, \gamma, \lambda) / \partial z_i = 0 \), whose unique solution is \( \tilde{z}_i \). Hence, \( z_i = \tilde{z}_i \).

Point (ii). \( T' (z_i; \theta_i) > 0 \) for \( i < I \) is established by Guesnerie and Seade (1982, Proposition 7). The remainder is shown in the text.

Points (iii)--(iv). Two kinds of bunching are to consider: (a) bunching at the bottom due to the constraint \( z_i \geq 0 \) or (b) bunching at the bottom or in the interior due to the violation of the monotonicity constraints. Assume \( \{k, \ldots, j^k \} \) are bunched together at the optimum, with gross income \( \bar{z}^k > 0 \). Case (a): \( B^k = \{1, \ldots, j^k \} \), \( \mu_1 > 0 \) and \( \mu_{j^k+1} = 0 \). Hence, summing (A.12) for \( n = 1, \ldots, j^k \) yields
\[
\sum_{n \in B^k} \partial W^* (z; \theta, \gamma, \lambda) / \partial z_n \bigg|_{z_1=0} = -\mu_1 \leq 0, \quad (A.16)
\]
with equality if \( \bar{z}^k > 0 \), which can be divided by \( \#B^k > 0 \). Case (b): \( \mu_{ik} = \mu_{j^k+1} = 0 \). Summing (A.12) for \( n = i^k, \ldots, j^k \) yields the equality form of (A.16). \( \square \)

Proof of Proposition 7. By Proposition 6, \( g^*_i (\theta, \gamma, \lambda) = \alpha_i^{-1} (1/\beta_i; \theta_i, \theta_{i+1}, \gamma) \) for \( i < I \) and \( g^*_I (\theta, \gamma, \lambda) = 23 \)
By the implicit function theorem, (A.17) defines \( z \). Since
\[
\frac{\partial \phi_i}{\partial z_i} = \frac{1 + \beta_i}{\theta_i} v''(z_i) + \frac{\beta_i}{\theta_{i+1}} v''(z_i) = 0.
\]
(A.17)

Since \( v'' > 0 \) and \( 0 < \theta_i < \theta_{i+1} \),
\[
\frac{\partial \phi_i}{\partial z_i} = \frac{\beta_i}{\theta_{i+1}} v''(z_i) + \frac{1 + \beta_i}{\theta_i} v''(z_i) < 0.
\]
(A.18)

By the implicit function theorem, (A.17) defines \( z \) as a \( C^1 \)-function of \( \gamma \), \( z_i = \phi_i^\gamma(\gamma) \), with derivative
\[
\frac{\partial z_i}{\partial \gamma} = \frac{d\phi_i^\gamma(\gamma)}{d\gamma} = -\frac{\partial \phi_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial z_i} > 0.
\]
(A.19)

Proof of Proposition 8. Let \( z = g^\gamma(\theta, \gamma, \lambda) \). For \( i = I, z_I = \theta_I v''^{-1}(\gamma \theta_I) \) by (35) and, since \( v'' > 0 \), \( \partial z_I / \partial \gamma > 0 \). For \( 1, \ldots, l = I - 1 \), the first-order conditions in Proposition 6 write
\[
\phi_i := \gamma - \frac{1 + \beta_i}{\theta_i} v''(z_i) + \frac{\beta_i}{\theta_{i+1}} v''(z_i) = 0.
\]
(A.17)

Proof of Proposition 9. Let \( z = g^\gamma(\theta, \gamma, \lambda) \). It is clear from Proposition 6 that \( z_I \) does not depend on \( \theta_{i+1} \) except for \( j = i, i + 1 \). Hence, \( \partial z_J / \partial \theta_{i+1} = 0 \) for \( j \neq [i, i + 1] \). Consequently, (44) follows from (30).

It remains to examine the effect of a change in \( \theta_{i+1} \) on \( z_i \) and \( z_{i+1} \). If \( \theta_{i+1} < \theta_I \), (A.17) implicitly defines \( z_i \) and \( z_{i+1} \) as \( C^1 \)-functions of \( \theta_{i+1} \), \( z_i = \phi_i^\theta(\theta_{i+1}) \) and \( z_{i+1} = \phi_i^\theta(\theta_{i+1}) \) respectively, with derivatives
\[
\frac{d\phi_i^\theta(\theta_{i+1})}{d\theta_{i+1}} = -\frac{\partial \phi_i}{\partial \phi_i} \frac{\partial \phi_i}{\partial z_i} \quad \text{and} \quad \frac{d\phi_i^\theta(\theta_{i+1})}{d\theta_{i+1}} = -\frac{\partial \phi_i^\theta}{\partial \phi_i} \frac{\partial \phi_i^\theta}{\partial z_i^\theta}.
\]
(A.20)

(42) and (43) hold because \( \partial \phi_i / \partial z_i < 0 \) and \( \partial \phi_i / \partial z_i < 0 \) by (A.18) while
\[
\frac{\partial \phi_i}{\partial \theta_{i+1}} = -\frac{\beta_i}{\theta_{i+1}^2} \left[ v'(z_i) \left( \theta_{i+1} \right) + z_i \frac{\beta_i}{\theta_{i+1}} v''(z_i) \right] < 0,
\]
(A.21)
\[
\frac{\partial \phi_i}{\partial \theta_{i+1}} = \left( 1 + \frac{\beta_i}{\theta_{i+1}^2} \right) \left[ v'(z_i) \left( \theta_{i+1} \right) + z_i + \frac{\beta_i}{\theta_{i+1}} v''(z_i) \right] > 0.
\]
(A.22)

If \( \theta_{i+1} = \theta_i \), the change in \( z_{i-1} \) is obtained as previously and that in \( z_i \) comes directly from the observation that \( z_I = \theta_I v''^{-1}(\gamma \theta_I) \) (Proposition 6) and \( v'' > 0 \).

Two cases must be distinguished as regards marginal tax rates. Case (a): \( i < I - 1 \). As \( \partial z_i / \partial \theta_{i+1} < 0 \), (30) and \( v'' > 0 \) imply an increase in \( T'(z_i; \theta_I) \). Similarly, \( \partial z_{i+1} / \partial \theta_{i+1} > 0 \) implies a reduction in \( T'(z_{i+1}; \theta_{i+2}) \). In addition, for \( i = 1, \ldots, I - 1 \), (18) writes \( T'(z_i; \theta_{i+1}) \equiv (1 + 1/\beta_i) T'(z_i; \theta_i) \) where \( \beta_i > 0 \). So, \( T'(z_i; \theta_{i+1}) \) increases and \( T'(z_{i+1}; \theta_{i+2}) \) decreases. Case (b): \( i = I - 1 \). \( T'(z_{i+1}; \theta_{i+1}) \) is unaltered, equal to zero, by Proposition 4. The changes in \( T'(z_i; \theta_i) \) and \( T'(z_i; \theta_{i+1}) \) are the same as in (a).

Proof of Proposition 11. Since \( \beta_k \) is increased for all \( k \in \{i, \ldots, j - 1\} \) and unaltered otherwise, Propo-
sition 10 implies (50). We then prove (51). By Proposition 2,
\begin{equation}
\begin{aligned}
x_k^* (z; \theta, \gamma) &= \frac{1}{I} \left[ \sum_{h=1}^I z_h - \frac{1}{\gamma} \sum_{h=2}^I (I + 1 - k) \left( v \left( \frac{z_h}{\theta_h} \right) - v \left( \frac{z_{h-1}}{\theta_{h-1}} \right) \right) \right] \\
&\quad + \frac{1}{\gamma} \sum_{h=2}^k \left[ v \left( \frac{z_h}{\theta_h} \right) - v \left( \frac{z_{h-1}}{\theta_{h-1}} \right) \right].
\end{aligned}
\end{equation}
(A.23)

the last term of which is omitted when \( k = 1 \).

For \( k < i \), differentiating (A.23), using (50), and rearranging thanks to (A.15), (40) and (30),
\begin{equation}
\begin{aligned}
\frac{dx_k^* (z; \theta, \gamma)}{ds} &= \frac{1}{I} \sum_{h=i}^{i-1} \left[ 1 - \frac{v' (z_h/\theta_h)}{\gamma \theta_h} + \frac{1}{\gamma} \left( \frac{1}{\theta_h} v' \left( \frac{z_h}{\theta_h} \right) - \frac{1}{\theta_{h+1}} v' \left( \frac{z_h}{\theta_{h+1}} \right) \right) \right] \frac{dz_h}{ds} \\
&= \frac{1}{I} \sum_{h=i}^{i-1} T' (z_h; \theta_h) \left[ 1 - (I - h) \alpha_h (z_h; \theta_h, \theta_{h+1}) \right] \frac{dz_h}{ds} \\
&= \frac{1}{I} \sum_{h=i}^{i-1} T' (z_h; \theta_h) \beta_h \frac{dz_h}{ds}.
\end{aligned}
\end{equation}
(A.24)

For \( h = 1, \ldots, j - 1 \), \( \beta_h - I + h < 0 \) by (27) and (10), \( T' (z_h; \theta_h) > 0 \) by Proposition 4 and \( dz_h/ds < 0 \). Hence, \( dx_k^* (z; \theta, \gamma) /ds > 0 \) for \( k < i \).

For \( k \geq j \),
\begin{equation}
\begin{aligned}
\frac{d}{ds} \left[ \frac{1}{\gamma} \sum_{h=2}^k \left( v \left( \frac{z_h}{\theta_h} \right) - v \left( \frac{z_{h-1}}{\theta_{h-1}} \right) \right) \right] &= \frac{1}{I} \sum_{h=i}^{i-1} \left[ \frac{1}{\theta_h} v' \left( \frac{z_h}{\theta_h} \right) - \frac{1}{\theta_{h+1}} v' \left( \frac{z_h}{\theta_{h+1}} \right) \right] \frac{dz_h}{ds} \\
&= \sum_{h=i}^{i-1} T' (z_h; \theta_h) \alpha_h \frac{dz_h}{ds} = \sum_{h=i}^{i-1} \frac{T' (z_h; \theta_h) \alpha_h}{\beta_h} \frac{dz_h}{ds}.
\end{aligned}
\end{equation}
(A.25)

This additional term must be added to (A.24) to get
\begin{equation}
\begin{aligned}
\frac{dx_k^* (z; \theta, \gamma)}{ds} &= \frac{1}{I} \sum_{h=i}^{i-1} T' (z_h; \theta_h) \beta_h \frac{dz_h}{ds}.
\end{aligned}
\end{equation}
(A.27)

For \( h = i, \ldots, j - 1 \), \( \beta_h + h = \Lambda (\theta_h) > 0 \), \( T' (z_h; \theta_h) > 0 \) by Proposition 4 and \( dz_h/ds < 0 \). Hence, \( dx_k^* (z; \theta, \gamma) /ds < 0 \) for \( k \geq j \). (52) is a direct implication of (51) and (50).

\textbf{References}


