



PARIS SCHOOL OF ECONOMICS
ÉCOLE D'ÉCONOMIE DE PARIS

WORKING PAPER N° 2007 - 15

**Optimality conditions and comparative static properties of
non-linear income taxes revisited**

Laurent Simula

JEL Codes: H21, H31

Keywords: Optimal tax, income tax, comparative statics



PARIS-JOURDAN SCIENCES ÉCONOMIQUES
LABORATOIRE D'ÉCONOMIE APPLIQUÉE - INRA



48, Bd JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES – ÉCOLE NORMALE SUPÉRIEURE

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

Laurent SIMULA*
EHESS, PSE, GREQAM and IDEP†

May 20, 2007

Abstract

Optimality conditions and comparative static properties of the optimal Mirrleesian non-linear 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 gross income, which is 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.

JEL classification: H21; H31.

*I am grateful to Guy Gilbert, Emmanuelle Taugourdeau, Gwenola Troitin, Alain Trannoy and John Weymark for very helpful comments and suggestions. The usual caveat applies.

†Affiliations: Ecole des Hautes Etudes en Sciences Sociales, Paris School of Economics, Research Group in Quantitative Economics of Aix-Marseilles 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 Mirrlees (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 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. 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 and the population is discrete. Their derivation uses a reduced form of the optimal tax problem, which does only involve the choice of the consumption good (Weymark, 1986b) and concentrates on the fully separating social allocation distinguished from those involving bunching in Weymark (1986a). 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). The assumption that individual preferences are quasilinear in leisure is maintained in this paper. 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) and Boone and Bovenberg (2007). 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 offer 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. This allows the reduced-form optimal income tax problem to have an explicit solution (Weymark, 1986b). When quasilinear-in-consumption preferences are considered, the linearity with respect to the observable variable is lost, the social objective of the reduced-form depends on productivity through the disutility of labour 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 efficiency/rent-extraction trade-off reflected by the income tax schedule is thus very specific. On the contrary, when 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), Boadway and Pestieau (2007), d'Autume (2000), Diamond (1998), Piketty (1997), Salanié (1998) or Saez (2002) 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. This reduced form can be seen as a special case of Chambers (1989) "concentrated" objective function derived for separable preferences. However, thanks to our quasilinearity assumption, it clearly reflects the trade-off between equity and efficiency. On the one hand, it incorporates the fact that the only incentive-compatibility constraints which matter are the downward adjacent ones. Accordingly, in the optimum, each individual is indifferent

between his own bundle and that of his nearest less productive neighbour. This is in accordance with the optimality features derived in the principal/agent model and reflects efficiency considerations. On the other hand, at any given gross-income/consumption bundle, the angle between the indifference curve of the individual for whom the bundle is designed and the indifference curve of the nearest more productive individual is entirely determined by the social weights. Altogether, these features allow a very simple geometric characterization of the optimal allocation and of the bunching pattern. In particular, it is sufficient to know the indifference curves and the individual social weights to construct the complete optimal allocation geometrically. 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. It appears that varying the skill level of an individual only alters the optimal allocation locally, through three channels: it involves a local substitution effect, an incentive effect and an informational externality which modifies the behaviour of the nearest less productive individual.

This analysis can be regarded as a special case but also as a generalization of Stiglitz (1982)'s model. First, it investigates quasilinear preferences instead of well-behaved generic preferences and focuses on the maximization of social weighted functions rather than characterizing Pareto efficient tax schedules. This latter restriction amounts to investigating the case Stiglitz (1982) refers to as "normal" as soon as the individual social weights are declining with individual productivity. Second, it casts light on a population consisting of more than two classes of individuals, which is of interest for the applicability of the model to real situations.

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 $i \in \mathcal{I} := \{1, \dots, I\}$. There are two goods, consumption and leisure. Person i 's consumption and labour supply are denoted x_i and ℓ_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 θ_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 $\theta := (\theta_1, \dots, \theta_I)$ is taken to be

monotonically increasing,

$$0 < \theta_1 < \dots < \theta_I. \quad (1)$$

An individual with productivity θ_i working ℓ_i units of time has gross income

$$z_i := \theta_i \ell_i, \quad i \in \mathcal{I}. \quad (2)$$

All individuals have the same preferences over consumption and leisure, represented by the utility function $U : \mathbb{R}_+^2 \rightarrow \mathbb{R}$,

$$U(x_i, \ell_i) := \gamma x_i - v(\ell_i), \quad i \in \mathcal{I}, \quad (3)$$

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}^3 -function which satisfies $v' > 0$, $v'' > 0$, $v''' > 0$, $v(0) = 0$, $v'(0) = 0$, and $v'(\bar{\ell}) \rightarrow \infty$ where $\bar{\ell}$ 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 $u : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ in the gross-income/consumption space,

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

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

$$s(z_i; \theta_i) := -\frac{u'_{z_i}(x_i, z_i; \theta_i)}{u'_{x_i}(x_i, z_i; \theta_i)} = \frac{v'(z_i/\theta_i)}{\gamma \theta_i}, \quad i \in \mathcal{I}. \quad (5)$$

In particular, the higher is γ , 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 $x = (x_1, \dots, x_I)$ and $z = (z_1, \dots, z_I)$. 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

$$\begin{cases} \theta & \longrightarrow & a & \mathbb{R}_+ \times \mathbb{R}_+ \\ \theta_i & & & (x_i, z_i), \end{cases} \quad (6)$$

which satisfies the *incentive compatibility constraints*

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

and the *tax revenue constraint*

$$\sum_{i=1}^I z_i \geq \sum_{i=1}^I x_i. \quad (8)$$

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

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

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

in which $\lambda := (\lambda_1, \dots, \lambda_I)$ 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.

$$0 < \lambda_I < \dots < \lambda_1. \quad (10)$$

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. When $I > 2$, one can therefore expect that the only binding incentive compatibility constraints will be the downward adjacent ones stating that the θ_{i+1} -individual must be indifferent between his own bundle and that of the θ_i -individual, for $i = 1, \dots, I - 1$.

As $W(a)$ is homogeneous of degree one in λ , 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

$$\Lambda(\theta_I) = I. \quad (11)$$

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

$$\mathcal{P} := \mathbb{R}_{++}^I \times \mathbb{R}_{++} \times \{\lambda \mid (10) \text{ and } (11) \text{ are satisfied}\}. \quad (12)$$

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

Problem 1 (Optimal Non-linear Income Tax Problem). *For $(\theta, \gamma, \lambda) \in \mathcal{P}$, 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 $(\theta, \gamma, \lambda) \in \mathcal{P}$, there is a unique solution to Problem 1. I denote by $g^x : \mathcal{P} \rightarrow \mathbb{R}_+^I$ and $g^z : \mathcal{P} \rightarrow \mathbb{R}^I$ the functions which relate $(\theta, \gamma, \lambda)$ to the optimal consumption and gross income vectors for Problem 1 respectively, with

$$g^x(\theta, \gamma, \lambda) := (g_1^x(\theta, \gamma, \lambda), \dots, g_I^x(\theta, \gamma, \lambda)), \quad (13)$$

$$g^z(\theta, \gamma, \lambda) := (g_1^z(\theta, \gamma, \lambda), \dots, g_I^z(\theta, \gamma, \lambda)). \quad (14)$$

The *indirect utilities* $V_i : \mathcal{P} \rightarrow \mathbb{R}$ are thus obtained as

$$V_i(\theta, \gamma, \lambda) := u(g_i^x(\theta, \gamma, \lambda), g_i^z(\theta, \gamma, \lambda); \theta_i), \quad i \in \mathcal{I}. \quad (15)$$

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, to interpret 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 \mathcal{Z} .

Subproblem 1. *Given a gross income vector $z \in \mathcal{Z}$ and the parameters $(\theta, \gamma, \lambda) \in \mathcal{P}$, choose the consumption vector $x \in \mathbb{R}_+^I$ to maximize the social welfare function $W(x)$ 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 x to Problem 1 is obtained as

$$\arg \max_{x \in \mathcal{X}^*} W(x^*(z; \theta, \gamma, \lambda), z). \quad (16)$$

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 $W(x^*(z; \theta, \gamma, \lambda), z)$.*

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 in productivity. These restrictions can be used to derive sufficient conditions under which an allocation a satisfies the incentive-compatibility constraints (7). We proceed in two steps.

First, if an allocation a satisfies (7), then gross income and net income must be non-decreasing in productivity, i.e.

$$(x_1, z_1) \leq \dots \leq (x_I, z_I), \quad (17)$$

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

$$\mathcal{Z} := \{z \in \mathbb{R}^I \mid 0 \leq z_1 \leq \dots \leq z_I\}. \quad (18)$$

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

Second, given $z \in \mathcal{Z}$, a sufficient condition for an allocation to satisfy the incentive compatibility constraints (7) employs the concept of simple monotonic chain to the left. Following Guesnerie and Seade (1982), a *simple monotonic chain to the left* is an allocation a such that

$$u(x_{i+1}, z_{i+1}; \theta_{i+1}) = u(x_i, z_i; \theta_{i+1}), \quad i = 1, \dots, I-1. \quad (19)$$

Given quasilinear-in-consumption preferences, (19) is equivalent to

$$u(x_{i+1}, z_{i+1}; \theta_{i+1}) - u(x_i, z_i; \theta_i) = v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right). \quad (20)$$

In words, the adjacent downward incentive compatibility constraints are active for all $i = 2, \dots, I$.

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 the incentive compatibility constraints (7).*

Proof. Guesnerie and Seade (1982). □

This pattern expresses a specific efficiency/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 (θ_i, θ_{i+1}) , this rate basically depends on

$$R_{i+1} := v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right), \quad i = 1, \dots, I-1, \quad (21)$$

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 $(\theta, \gamma, \lambda) \in \mathcal{P}$. 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 $(\theta, \gamma, \lambda) \in \mathcal{P}$, 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. □

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),

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

and so

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

As any solution to Subproblem 1 is production efficient, by Lemma 1, the binding tax revenue constraint (8) can be substituted in $\sum_{i=1}^I x_i$, obtained from (23), to get

$$\sum_{i=1}^I z_i = Ix_1 + \frac{1}{\gamma} \sum_{i=2}^I \sum_{j=2}^i \left[v\left(\frac{z_j}{\theta_j}\right) - v\left(\frac{z_{j-1}}{\theta_j}\right) \right] = Ix_1 + \frac{1}{\gamma} \sum_{i=2}^I (I+1-i) \left[v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_{i-1}}{\theta_i}\right) \right]. \quad (24)$$

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 $\mathcal{X}^*(z; \theta, \gamma, \lambda)$, which is independent of the social weights λ and inherits the differentiability properties of v .

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

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

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

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 $\beta = (\beta_1, \dots, \beta_I)$ with

$$\beta_i := \Lambda(\theta_i) - i, \quad i \in \mathcal{I}. \quad (27)$$

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 $i = 1, \dots, I-1$.

The parameters β_i summarize in a transparent way the redistributive taste of the government. First, all λ_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 β_i express the policymaker's *strict* aversion to income inequality. Second, to get further insight into β_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. Since γ is the marginal utility of money, the utility of each of them is increased by γ . Accordingly, the gross social benefit amounts to $\gamma \Lambda_i(\theta_i)$. However, the tax revenue is decreased by i euros, which corresponds to a social cost γi . 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. So, β_i is this net social benefit expressed in monetary units. Third, the parameters β_i can alternatively be defined as

$$\beta_i := I - i - \sum_{j=i+1}^I \lambda_j, \quad i = 1, \dots, I-1, \quad (28)$$

and $\beta_I = 0$, because $\Lambda(\theta_I) = I$. 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*. They allow us to rewrite the reduced-form optimal non-linear income tax problem (i.e. Subproblem 2) as follows.

Problem 2 (Reduced Form). For $(\theta, \gamma, \lambda) \in \mathcal{P}$, choose z in \mathcal{Z} so as to maximize the social objective function $\mathcal{W}^*(z; \theta, \gamma, \lambda) : \mathcal{Z} \times \mathcal{P} \rightarrow \mathbb{R}$, with¹

$$\mathcal{W}^*(z; \theta, \gamma, \lambda) := \sum_{i=1}^I \left[\gamma z_i - v\left(\frac{z_i}{\theta_i}\right) \right] - \sum_{i=1}^I \beta_i R_{i+1}. \quad (29)$$

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 $(\theta, \gamma, \lambda) \in \mathcal{P}$, the optimal solution to Problem 2 is $g^z(\theta, \gamma, \lambda)$, the gross income vector solution to Problem 1, and the optimal consumption vector for Problem 1 is

$$g^x(\theta, \gamma, \lambda) = x^*(g^z(\theta, \gamma, \lambda); \theta, \gamma). \quad (30)$$

Proof. See the Appendix. □

An important implication of Proposition 3 is that the social allocation solution to the optimal non-linear income tax problem 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 θ_i -individual for whom the $(x_i^*(z_i; \theta, \gamma), z_i)$ -bundle is designed, on the one hand, and the implicit marginal tax rate $T'(z_i; \theta_{i+1})$ the nearest more productive θ_{i+1} -individual would face if he were mimicking the θ_i -individual. They are formally defined as

$$T'(z_i; \theta_j) := 1 - s(z_i; \theta_j) = 1 - \frac{v'(z_i/\theta_j)}{\gamma\theta_j}, \quad (i, j) \in \mathcal{I}^2. \quad (31)$$

The implicit marginal tax rates allows us to get further understanding of the social objective function of the reduced-form optimal income tax problem $\mathcal{W}^*(z; \theta, \gamma, \lambda)$. Indeed, let z be a fixed gross income vector and consider that the gross income z_i of the θ_i -individual is increased at the

¹Since $\beta_I = 0$, R_{I+1} is an arbitrary number.

margin. As

$$\frac{\partial \mathcal{W}^*(z; \theta, \gamma, \lambda)}{\partial z_i} = \gamma T'(z_i; \theta_i) - \beta_i \frac{\partial R_{i+1}}{\partial z_i}, \quad i \in \mathcal{I}, \quad (32)$$

by (29) and (31), the impact on social welfare may be thought of as proceeding in two steps. In the first step, the θ_i -individual pays $T'(z_i; \theta_i)$ additional euros in taxes, which relaxes the tax revenue constraint (8). As γ 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. The θ_i -individual 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 the θ_i -individual. So, cheating becomes more attractive to them. In order to restore individual truth-telling and obtain a new monotonic chain to the left, the policymaker has to increase the marginal information rent left to each of them by $\frac{1}{\gamma} \frac{\partial R_{i+1}}{\partial z_i}$ euros. Because $\gamma \beta_i$ is the net social cost of marginally increasing the consumption of the $I - i$ most productive individuals, social welfare is reduced by $\beta_i \frac{\partial R_{i+1}}{\partial z_i}$. 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.

$$\frac{\partial \mathcal{W}^*(z; \theta, \gamma, \lambda)}{\partial z_i} = 0 \Leftrightarrow T'(z_i; \theta_i) = \frac{\beta_i}{\gamma} \frac{\partial R_{i+1}}{\partial z_i}, \quad i \in \mathcal{I}, \quad (33)$$

whose sign is given by β_i because $\gamma > 0$ and $\partial R_{i+1} / \partial z_i > 0$. Consequently, as $\beta_i > 0$ for $i < I$ and $\beta_I = 0$, the optimal marginal tax rates $T'(z_i; \theta_i)$ faced by all but the most productive individuals is strictly positive. The labour supply is thus distorted except at the top.

3.4. Characterization of the Social Optimum

Characterizing the social optimum requires to know whether it involves bunching or is fully separating. If the solution $z = g^z(\theta, \gamma, \lambda)$ to Problem 2 is such that $0 < z_1 < \dots < z_I$, then it must satisfy (33). Using (31), (32) and (21), this is equivalent to

$$z_i = \alpha_i^{-1}(1/\beta_i; \theta, \gamma), \quad i = 1, \dots, I-1, \quad (34)$$

where

$$\alpha_i(z_i; \theta, \gamma) = \frac{T'(z_i; \theta_{i+1}) - T'(z_i; \theta_i)}{T'(z_i; \theta_i)}, \quad i = 1, \dots, I-1, \quad (35)$$

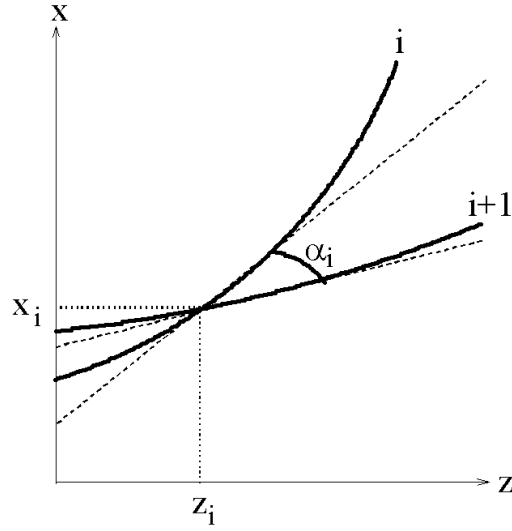


FIGURE 1: Geometric interpretation of the first-order conditions: at the optimum and at the (x_i, z_i) -bundle, the angle α_i between the indifference curves i and $i+1$ of the θ_i and θ_{i+1} -individual respectively is determined by the cumulative social weights β_i .

and $z_I = \theta_I (v')^{-1}(\gamma \theta_I)$. As $\alpha_i(z_i; \theta, \gamma)$ is a strictly increasing function of z_i^2 , z_i is a strictly decreasing function of β_i , with $z_i = \alpha_i^{-1}(1/\beta_i; \theta, \gamma)$. Hence, $0 < z_1 < \dots < z_I$ is possible if and only if $0 < \beta_I < \dots < \beta_1$.

Proposition 4. *The optimal allocation is fully separating if and only*

$$(\theta, \gamma, \lambda) \in \mathcal{P}^0 := \{(\theta, \gamma, \lambda) \in \mathcal{P} \mid 0 < \beta_I < \dots < \beta_1\}. \quad (36)$$

This necessary and sufficient condition has a very clear interpretation. Bunching does not occur when giving one euro to each of the i less productive individuals is socially more effective than giving one euro to each of the $i+1$ less productive ones, for $i = 1, \dots, I-1$. This corresponds to an additional restriction on the individual social weights λ_i . From now on, attention is restricted to the case where this restriction is met. The optimality condition can then be written in a strikingly simple form.

Proposition 5. *For $(\theta, \gamma, \lambda) \in \mathcal{P}^0$, z is socially optimal if and only if*

$$\alpha_i(z_i; \theta, \gamma) = \frac{1}{\beta_i}, \quad i = 1, \dots, I-1, \quad (37)$$

²As $d[T'(z_i; \theta_{i+1}) - T'(z_i; \theta_i)]/dz_i = \frac{1}{\gamma} \left[\frac{v''(z_i/\theta_i)}{\theta_i^2} - \frac{v''(z_i/\theta_{i+1})}{\theta_{i+1}^2} \right] > 0$ while $dT'(z_i; \theta_i)/dz_i < 0$, (35) implies $d\alpha_i(z_i; \theta, \gamma)/dz_i > 0$ for $i = 1, \dots, I-1$.

and $T'(z_I; \theta_I) = 0$.

For a gross income z_i , $\alpha_i(z_i; \theta, \gamma)$ tells us to which extent the indifference curves of the θ_{i+1} -individual must be flatter than those of the θ_i -individual. Geometrically, it thus corresponds to the angle between the tangents to the indifference curves of the θ_i and θ_{i+1} -individuals depicted in Figure 1. This angle is closely related to the single-crossing condition and thus henceforth referred to as the *Spence-Mirrlees angle*. The single-crossing condition is a restriction on the *sign* of $\alpha_i(z_i; \theta, \gamma)$, which must be strictly positive. Here, this condition is automatically satisfied because individual preferences are quasilinear in consumption. The conditions for social optimality (37) introduce an additional restriction on $\alpha_i(z_i; \theta, \gamma)$: an allocation is socially optimal only if, at each observed gross income level z_i , the Spence-Mirrlees angle is entirely determined by the exogenously given cumulative social weight β_i . In addition, the labour supply of the more productive individuals is not distorted since the marginal tax rate at the top is equal to zero.

Thanks to Proposition 5, the optimal allocation can be constructed geometrically in two steps illustrated in Figure 2. *In the first step*, the tax revenue constraint is ignored. Starting from zero, gross income is gradually increased until $\alpha_1 = 1/\beta_1$ and the bundle (\bar{x}_1, z_1) is determined. Then, gross income is increased along the indifference curve of the θ_2 -individual through the (\bar{x}_1, z_1) -bundle until the angle with the indifference curve of the θ_3 -individual is equal to $1/\beta_2$. Proceeding recursively, a monotonic chain to the left (\bar{x}, z) is obtained. This allocation is incentive compatible, but not necessarily budget-balanced. That is why, *in the second step*, each \bar{x}_i is varied by a same amount ε so as to get a binding tax revenue constraint. The resulting allocation $(\bar{x} + \varepsilon, z)$, which is both incentive compatible and production efficient, 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 sum $\tilde{u}(x_i, z_i; \theta_i)$ over all i in \mathcal{I} , get

$$\sum_{i=1}^I \tilde{u}(x_i, z_i; \theta_i) = \sum_{i=1}^I \theta_i h(x_i) - \gamma \sum_{i=1}^I z_i, \quad (38)$$

and replace $\sum_{i=1}^I z_i$ by $\sum_{i=1}^I x_i$. This step is required to obtain a reduced-form optimal income tax problem. Consequently, 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 Λ_i . So, the impact of the policy-maker's taste for redistribution is less transparent because social weights and productivity levels are mixed together.

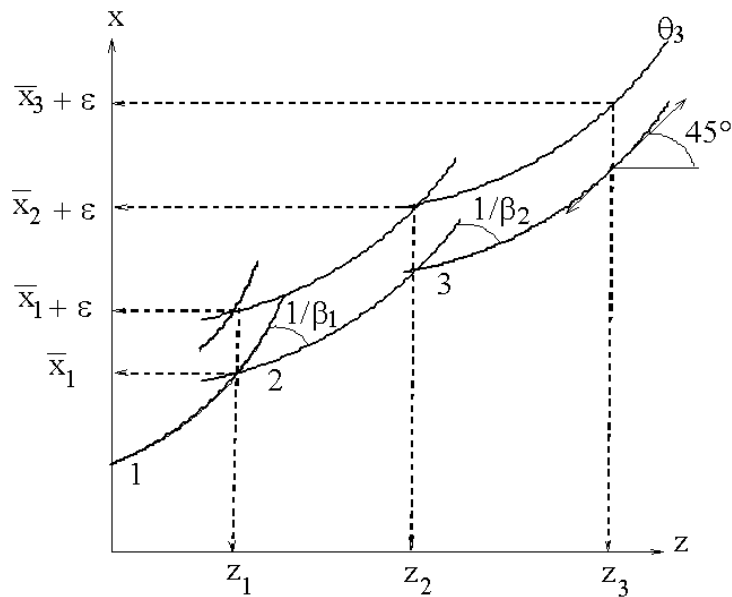


FIGURE 2: Two-step geometric construction of the optimal allocation. Step 1: the simple monotonic chain to the left (\bar{x}, z) is obtained by setting $\alpha_i = 1/\beta_i$ for $i = 1, 2$ and choosing z_3 such that the θ_3 -individual is not taxed at the margin. Step 2 : (\bar{x}, z) is translated by $(0, \epsilon)$ so as to obtain a production efficient allocation.

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 variables. Since the disutility of labour v is \mathcal{C}^2 , the implicit function theorem implies that $g_i^z(\theta, \gamma, \lambda)$ is \mathcal{C}^1 . It thus follows from Proposition 2 that $g_i^x(\theta, \gamma, \lambda) = x^*(g_i^z(\theta, \gamma, \lambda); \theta, \gamma, \lambda)$ is also \mathcal{C}^1 . These results can be summarized as follows.

Proposition 6. *The functions g^x, g^z, T' and V are \mathcal{C}^1 at every $(\theta, \gamma, \lambda)$ in \mathcal{P}^0 .*

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

4.1. Comparative Statics for the Marginal Utility of Money

The marginal utility of money corresponds to the unit of count in welfare of our economy. Namely, one additional euro of consumption for a given individual increases his well-being by γ utils. Varying γ at the margin is thus likely to modify the individual labour supply.

Proposition 7. For $(\theta, \gamma, \lambda) \in \mathcal{P}^0$ and $i \in \mathcal{I}$,

$$\frac{\partial g_i^z(\theta, \gamma, \lambda)}{\partial \gamma} > 0. \quad (39)$$

Proof. See the Appendix. □

A small increase in the marginal utility of money γ raises the gross income of every individual. Indeed, when γ goes up, an extra unit of consumption contributes more to individual well-being as previously. So, the indifference curves become flatter in the (z, x) -space. Every θ_i -individual is thus willing to work more in order to increase his consumption by a given amount.

Unfortunately, the comparative statics of the optimal implicit marginal tax rates cannot be obtained in the general case. In fact, as $T'(z_i; \theta_j) = 1 - v'(z_i/\theta_j) / (\gamma\theta_j)$, 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 has 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 θ 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 θ_{i+1} can be summarized as follows.

Proposition 8. For $(\theta, \gamma, \lambda) \in \mathcal{P}^0$ and $(i, j) \in \{1, \dots, I-1\} \times \mathcal{I}$,

$$\frac{\partial T'(z_i; \theta_i)}{\partial \theta_{i+1}} > 0, \quad \frac{\partial T'(z_i; \theta_{i+1})}{\partial \theta_{i+1}} > 0, \quad (40)$$

$$\frac{\partial T'(z_{i+1}; \theta_{i+1})}{\partial \theta_{i+1}} < 0, \quad \frac{\partial T'(z_{i+1}; \theta_{i+2})}{\partial \theta_{i+1}} < 0, \quad (41)$$

$$\frac{\partial T'(z_j; \theta_j)}{\partial \theta_{i+1}} = \frac{\partial T'(z_j; \theta_{j+1})}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i+1\}, \quad (42)$$

and

$$\frac{\partial z_i}{\partial \theta_{i+1}} < 0, \quad (43)$$

$$\frac{\partial z_{i+1}}{\partial \theta_{i+1}} > 0, \quad (44)$$

$$\frac{\partial z_j}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i+1\}, \quad (45)$$

where $z \equiv \bar{g}(\theta, \gamma, \lambda)$.

Proof. See the Appendix. □

Increasing the productivity of the θ_{i+1} -individual does only alter his gross income and that of his nearest less productive neighbour. Indeed, by Proposition 5, only α_i and α_{i+1} depend on θ_{i+1} . So, the optimality condition $\alpha_j = 1/\beta_j$, which implicitly defines z_j as a function of θ_{i+1} , is unaffected except for the θ_i and θ_{i+1} -individuals. Accordingly, the gross income levels of all other individuals remain unaltered. As regards the θ_i and θ_{i+1} -individuals, the adjustment process combines *three effects*.

First, the variation in θ_{i+1} gives rise to a *local substitution effect*. The increase in the productivity of the θ_{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 θ_{i+1} has an *incentive effect*. As he becomes more efficient, the θ_{i+1} -individual has to provide less effort if he wants to imitate the θ_i -individual. Consequently, his indifference curve through the gross-income/consumption bundle of the θ_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 θ_i -individual incurs an *informational externality* induced by the incentive effect. Since the cumulative social weight β_i is unaltered, the angle α_i between the indifference curves of the θ_i and θ_{i+1} -individuals through the (x_i, z_i) -bundle must stay constant (Proposition 5). 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 θ_i -individual. Finally, the substitution effect leads the θ_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 production-efficient allocation. 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 5 basically involves the cumulative social weights β_i , and thus the individual social weights λ_i . As emphasized previously, these social weights express the government's strict aversion to income inequality. Changing them marginally is thus likely to alter the progressivity of the tax schedule.

The impact of a change in the cumulative social weight β_i , with $i < I$, is examined first because it will be useful when deriving the comparative statics with respect to the individual social weights. As $z_i = \alpha_i^{-1}(1/\beta_i; \gamma, \lambda)$ where $\alpha_i' > 0$, an increase in β_i is associated with a reduction in z_i . Indeed, given $T'(z_i; \theta_{i+1})$, an increase in β_i requires the implicit marginal tax rate $T'(z_i; \theta_i)$ faced by the θ_i -individual to be raised. The induced substitution effect leads the θ_i -individual to work less.

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

$$\frac{\partial g_i^z(\theta, \gamma, \lambda)}{\partial \beta_i} < 0 \text{ and } \frac{\partial g_j^z(\theta, \gamma, \lambda)}{\partial \beta_i} = 0 \quad (46)$$

Proof. See the Appendix. □

This result can now be used to consider the *impact of an increase in the individual social weight of the θ_i -individual to the detriment of a more productive θ_j -individual*. By definition of $\Lambda(\theta_k)$, every β_k is increased for $k \in \{i, \dots, j-1\}$ while all other β_k remain unaltered. By Proposition 9, it is thus optimal to decrease the gross income z_k of each θ_k -individual, with $k \in \{i, \dots, j-1\}$, and to hold that of the others constant. The impact on the consumption levels and indirect utilities can also be signed for all $k \notin \{i, \dots, j-1\}$.

Proposition 10. Let $(\bar{\theta}, \bar{\gamma}, \bar{\lambda}) \in \mathcal{P}^0$, $i \in \{1, \dots, I-1\}$ and $j \in \{i+1, \dots, I\}$. Let $\lambda : S \rightarrow \mathbb{R}^I$, where $S = (-1, 1)$, be \mathcal{C}^1 with

$$\begin{cases} \lambda_k(0) = \bar{\lambda}_k, & k = i, j, \\ \lambda_k(s) \equiv \bar{\lambda}_k, & \forall s \in S, \forall k \neq i, j, \\ d\lambda_i(s)/ds = -d\lambda_j(s)/ds, & \forall s \in S. \end{cases} \quad (47)$$

³A change in β_I is impossible since, by definition, $\beta_I \equiv 0$.

Then, if λ_i is increased to the detriment of λ_j^4 ,

$$\begin{cases} dz_k/ds < 0, & \forall k \in \{i, \dots, j-1\}, \\ dz_k/ds = 0, & \forall k \notin \{i, \dots, j-1\}, \end{cases} \quad (48)$$

$$\begin{cases} dx_k^*/ds > 0, & \forall k < i, \\ dx_k^*/ds < 0, & \forall k \geq j, \end{cases} \quad (49)$$

$$\begin{cases} dV_k/ds > 0, & \forall k < i, \\ dV_k/ds < 0, & \forall k \geq j, \end{cases} \quad (50)$$

where $z \equiv g^z(\bar{\theta}, \bar{\gamma}, \bar{\lambda})$ and $x^* \equiv x^*(z; \bar{\theta}, \bar{\gamma})$.

Proof. See the Appendix. □

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 (31), they only depend on λ through gross income z . It thus follows from (48) that $T'(x_k^*, z_k, \theta_k)$ and $T'(x_k^*, z_k, \theta_{k+1})$ are unaltered for $k \notin \{i, \dots, j-1\}$ whilst $T'(x_k^*, z_k, \theta_k)$ and $T'(x_k^*, z_k, \theta_{k+1})$ are increased for $k \in \{i, \dots, j-1\}$.

For concreteness, let us consider that the population consists of three individuals and that the social weight of the θ_2 -individual is increased at the expense of the θ_3 -individual. Let (\bar{x}, \bar{z}) be the initial allocation and denote by (x, z) the new one. The changes in gross income have been explained previously. 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 (48), the gross income levels of the θ_1 and θ_3 -individuals are held fixed, i.e. $z_1 = \bar{z}_1$ and $z_3 = \bar{z}_3$, while \bar{z}_2 is reduced (by $d\bar{z}_2$). In consequence, the requirement that the θ_3 -individual is indifferent between his own bundle and that of the θ_2 -individual induces a decrease in the consumption levels \bar{x}_2 and \bar{x}_3 of both more productive individuals (by $-d\bar{x}_2$ and $-d\bar{x}_3$ respectively) as well as in the indirect utility of the θ_3 -individual. A new monotonic chain to the left is obtained. As the θ_2 -individual faces a strictly positive marginal tax rate, he reduces his gross income by a smaller amount than his consumption, i.e. $-d\bar{x}_2 > -d\bar{z}_2$. So, $d\bar{x}_3$ can be sufficiently small for

$$\sum_{i=1}^3 \bar{x}_i - d\bar{x}_2 - d\bar{x}_3 < \sum_{i=1}^3 \bar{z}_i - d\bar{z}_2, \quad (51)$$

which means that the new monotonic chain to the left is not production efficient. As $\sum_{i=1}^3 \bar{x}_i =$

⁴The inequalities in (48)–(50) are reversed if λ_i is decreased to the benefit of λ_j .

$\sum_{i=1}^3 \bar{z}$, the second step consists therefore in giving

$$\varepsilon = \frac{1}{3} (d\bar{x}_2 + d\bar{x}_3 - dz_2) > 0 \quad (52)$$

euros of consumption to each individual. The new consumption levels are thus the following:

$$x_1 = \bar{x}_1 + \frac{1}{3} (d\bar{x}_2 + d\bar{x}_3 - dz_2) > \bar{x}_1, \quad (53)$$

$$x_2 = \bar{x}_2 - \frac{2}{3} d\bar{x}_2 + \frac{1}{3} (d\bar{x}_3 - dz_2), \quad (54)$$

$$x_3 = \bar{x}_3 - \frac{2}{3} d\bar{x}_3 + \frac{1}{3} (d\bar{x}_2 - dz_2) < \bar{x}_3. \quad (55)$$

Hence, the θ_1 -individual enjoys greater consumption, contrary to the θ_3 -individual. The change in the consumption of the θ_2 -individual is ambiguous. It is positive if and only if $dz_2 < d\bar{x}_3 - 2d\bar{x}_2$. The variations (50) 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. This offers two advantages. First, the link between the social weights and the conditions for social optimality is very transparent. Second, clear-cut comparative statics properties can easily be derived as regards changes in the productivity levels which are the key parameters of the optimal income tax model.

APPENDIX

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

(i) *a is a simple monotonic chain to the left.*

The proof is ad absurdum. Assume a is not a simple monotonic chain to the left, i.e. $u(x_j^*, z_j; \theta_j) \neq u(x_{j-1}^*, z_{j-1}; \theta_j)$. As $u(x_j^*, z_j; \theta_j) < u(x_{j-1}^*, z_{j-1}; \theta_j)$ implies $u(x_j^*, z_j; \theta_j) < u(x_{j-1}^*, z_{j-1}; \theta_{j-1})$ which violates (7), our assumption is equivalent to considering that there exists $j \geq 2$ for which

$$\gamma x_j^* - v\left(\frac{z_j}{\theta_j}\right) > \gamma x_{j-1}^* - v\left(\frac{z_{j-1}}{\theta_j}\right). \quad (56)$$

By (17), it must be $x_j^* > x_{j-1}^*$ and $z_j > z_{j-1}$ for (56) to be satisfied. If z does not satisfy the second inequality, then (i) is established.

Let $\bar{x}_i = x_i^* + \varepsilon_1$ for $i = 1, \dots, j-1$ and $\bar{x}_i = x_i^* - \varepsilon_2$ for $i = j, \dots, I$, where $\varepsilon_1 > 0$ and $\varepsilon_2 > 0$. The incentive compatibility constraints remain satisfied for sufficiently small ε_1 and ε_2 . ε_1 and ε_2 are chosen such that $\bar{x} = x^*$, i.e.

$$(j-1)\varepsilon_1 = (I-j+1)\varepsilon_2. \quad (57)$$

Therefore, using (57), one gets:

$$W(\bar{x}, z) - W(x^*, z) = \gamma \sum_{i=1}^I \lambda_i (\bar{x}_i - x_i^*) = \gamma \left[\sum_{i=1}^{j-1} \lambda_i \varepsilon_1 - \sum_{i=j}^I \lambda_i \varepsilon_2 \right], \quad (58)$$

which can be minored thanks to (10):

$$\begin{aligned} W(\bar{x}, z) - W(x^*, z) &\geq \gamma [\lambda_{j-1} (j-1) \varepsilon_1 - \lambda_j (I-j+1) \varepsilon_2] \\ &= \gamma (I-j+1) \varepsilon_2 [\lambda_{j-1} - \lambda_j] > 0. \end{aligned} \quad (59)$$

This contradicts the fact that x^* is in $\mathcal{X}^*(z; \theta, \gamma, \lambda)$.

(ii) *a is production efficient.*

Fix \bar{z} in \mathcal{Z} . The self-selection constraints (7) are satisfied, with $z_i = \bar{z}_i$ and $x_i = x_i^*$ 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 $\varepsilon > 0$. But the self-selection constraints (7) are equivalent to

$$\gamma x_i^* + \gamma \varepsilon - v\left(\frac{\bar{z}_i}{\theta_i}\right) \geq \gamma x_j^* + \gamma \varepsilon - v\left(\frac{\bar{z}_j}{\theta_i}\right), \forall (i, j) \in \mathcal{I}^2. \quad (60)$$

Increasing every x_i^* by ε is thus incentive compatible. Since it is also Pareto improving, x^* cannot be in $\mathcal{X}^*(\bar{z}; \theta, \gamma, \lambda)$. A contradiction.

(iii) *Existence of an optimal allocation.*

Fix \bar{z} in \mathcal{Z} and pick an arbitrary value \bar{x}_1 for x_1 . Then, proceed sequentially to select \bar{x} which

solves

$$x_{i+1} = \bar{x}_i + \frac{1}{\gamma} \left[v \left(\frac{\bar{z}_{i+1}}{\theta_{i+1}} \right) - v \left(\frac{\bar{z}_i}{\theta_{i+1}} \right) \right], \quad i = 1, \dots, I-1. \quad (61)$$

By construction, \bar{a} is a monotonic chain to the left. Since $\bar{z} \in \mathcal{Z}$, it follows from Proposition 1 that \bar{a} satisfies (7). If \bar{a} 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 $\varepsilon > 0$. Consequently, the constraint set is not empty.

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

$$\gamma x_i - v \left(\frac{\bar{z}_i}{\theta_i} \right) \geq \gamma x_1 - v \left(\frac{\bar{z}_1}{\theta_i} \right) \Leftrightarrow \gamma [x_i - x_1] \geq v \left(\frac{\bar{z}_i}{\theta_i} \right) - v \left(\frac{\bar{z}_1}{\theta_i} \right) \geq 0, \quad \forall i \in \mathcal{I}. \quad (62)$$

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.

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

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

$$u(x_i^*, z_i; \theta_i) = u(x_{i-1}^*, z_{i-1}; \theta_{i-1}) + v \left(\frac{z_{i-1}}{\theta_{i-1}} \right) - v \left(\frac{z_{i-1}}{\theta_i} \right), \quad i = 2, \dots, I, \quad (63)$$

from which

$$u(x_i^*, z_i; \theta_i) = u(x_1^*, z_1; \theta_1) + \sum_{j=1}^{i-1} \left[v \left(\frac{z_j}{\theta_j} \right) - v \left(\frac{z_j}{\theta_{j+1}} \right) \right], \quad i = 2, \dots, I. \quad (64)$$

Consequently,

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

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

$$\sum_{i=1}^I u(x_i^*, z_i; \theta_i) = \gamma \sum_{i=1}^I z_i - \sum_{i=1}^I v \left(\frac{z_i}{\theta_i} \right). \quad (66)$$

Plugging (66) in (65) and solving for $u(x_1^*, z_1; \theta_1)$,

$$u(x_1^*, z_1; \theta_1) = \frac{1}{I} \left\{ \gamma \sum_{i=1}^I z_i - \sum_{i=1}^I v\left(\frac{z_i}{\theta_i}\right) - \sum_{i=1}^{I-1} (I-i) \left[v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right] \right\}. \quad (67)$$

By (64) and (67), the social value function reads

$$\begin{aligned} W &= u(x_1^*, z_1; \theta_1) \sum_{i=1}^I \lambda_i + \sum_{i=2}^I \sum_{j=1}^{i-1} \lambda_i \left[v\left(\frac{z_j}{\theta_j}\right) - v\left(\frac{z_j}{\theta_{j+1}}\right) \right], \\ &= u(x_1^*, z_1; \theta_1) I + \sum_{i=1}^{I-1} \left(\sum_{j=i+1}^I \lambda_j \right) \left[v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right], \\ &= \sum_{i=1}^I \left\{ \gamma z_i - v\left(\frac{z_i}{\theta_i}\right) + \left(i - \sum_{j=1}^i \lambda_j \right) \left[v\left(\frac{z_i}{\theta_i}\right) - v\left(\frac{z_i}{\theta_{i+1}}\right) \right] \right\}, \end{aligned} \quad (68)$$

in which $v(z_I/\theta_{I+1})$ is an arbitrary number. \square

Proof of Proposition 7. Let $z_i \equiv g_i^z(\theta, \gamma, \lambda)$ and write the first-order conditions given in Proposition 5 as

$$\gamma - \frac{1 + \beta_i}{\theta_i} v' \left(\frac{z_i}{\theta_i} \right) + \frac{\beta_i}{\theta_{i+1}} v' \left(\frac{z_i}{\theta_{i+1}} \right) = 0, \quad i = 1, \dots, I-1, \quad (69)$$

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

$$\frac{\partial \phi_i^\gamma(z_i, \gamma)}{\partial z_i} = \frac{\beta_i}{\theta_{i+1}^2} v'' \left(\frac{z_i}{\theta_{i+1}} \right) - \frac{1 + \beta_i}{\theta_i^2} v'' \left(\frac{z_i}{\theta_i} \right) < 0, \quad i = 1, \dots, I-1. \quad (70)$$

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

$$\frac{\partial g_i^z(\theta, \gamma, \lambda)}{\partial \gamma} \equiv \frac{d\varphi_i^\gamma(\gamma)}{d\gamma} = - \frac{\partial \phi_i^\gamma(z_i, \gamma) / \partial \gamma}{\partial \phi_i^\gamma(z_i, \gamma) / \partial z_i} = - \left[\frac{\partial \phi_i^\gamma(z_i, \gamma)}{\partial z_i} \right]^{-1} > 0. \quad \square \quad (71)$$

Proof of Proposition 8. Let $z_i \equiv g_z^i(\theta, \gamma, \lambda)$ and use (69) to define $\phi_j^\theta(z_j, \theta_{i+1}) = 0$ for $j =$

1, ..., I - 1. Hence,

$$\frac{\partial \phi_i^\theta(z_i, \theta_{i+1})}{\partial \theta_{i+1}} = -\frac{\beta_i}{\theta_{i+1}^2} \left[v' \left(\frac{z_i}{\theta_{i+1}} \right) + \frac{z_i}{\theta_{i+1}} v'' \left(\frac{z_i}{\theta_{i+1}} \right) \right] < 0, \quad (72)$$

$$\frac{\partial \phi_{i+1}^\theta(z_{i+1}, \theta_{i+1})}{\partial \theta_{i+1}} = \left(\frac{1 + \beta_{i+1}}{\theta_{i+1}^2} \right) \left[v' \left(\frac{z_{i+1}}{\theta_{i+1}} \right) + \frac{z_{i+1}}{\theta_{i+1}} v'' \left(\frac{z_{i+1}}{\theta_{i+1}} \right) \right] > 0, \quad (73)$$

$$\frac{\partial \phi_j^\theta(z_j, \theta_{i+1})}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i+1\}. \quad (74)$$

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

$$\frac{d\varphi_j^\theta(\theta_{i+1})}{d\theta_{i+1}} = -\frac{\partial \phi_j^\theta(z_j, \theta_{i+1}) / \partial \theta_{i+1}}{\partial \phi_j^\theta(z_j, \theta_{i+1}) / \partial z_j}. \quad (75)$$

As $\partial \phi_j^\theta(z_j, \theta_{i+1}) / \partial z_j \equiv \partial \phi_j^\gamma(z_j, \gamma) / \partial z_j < 0$ by (70), it follows from (72)–(75) that

$$\frac{\partial g_j^z(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} \equiv \frac{d\varphi_j^\theta(\theta_{i+1})}{d\theta_{i+1}} = \begin{cases} < 0 & \text{if } j = i, \\ > 0 & \text{if } j = i + 1, \\ = 0 & \text{if } j \notin \{i, i + 1\}. \end{cases} \quad (76)$$

In consequence,

$$\frac{\partial T'(z_i; \theta_i)}{\partial \theta_{i+1}} = -\frac{v''(z_i/\theta_i)}{\gamma \theta_i^2} \frac{\partial g_i^z(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} > 0, \quad (77)$$

$$\frac{\partial T'(z_{i+1}; \theta_{i+2})}{\partial \theta_{i+1}} = -\frac{1}{\gamma \theta_{i+2}^2} v'' \left(\frac{z_{i+1}}{\theta_{i+2}} \right) \frac{\partial g_{i+1}^z(\theta, \gamma, \lambda)}{\partial \theta_{i+1}} < 0, \quad (78)$$

$$\frac{\partial T'(z_j; \theta_{j+1})}{\partial \theta_{i+1}} = 0 \text{ for } j \notin \{i, i + 1\}. \quad (79)$$

By Proposition 5,

$$T'(z_i; \theta_{i+1}) \equiv \left(1 + \frac{1}{\beta_i} \right) T'(z_i; \theta_i), \quad i = 1, \dots, I - 1, \quad (80)$$

where $\beta_i > 0$. Therefore, (77) and (78) imply

$$\frac{\partial T'(z_i; \theta_{i+1})}{\partial \theta_{i+1}} > 0 \text{ and } \frac{\partial T'(z_{i+1}; \theta_{i+1})}{\partial \theta_{i+1}} < 0. \quad \square \quad (81)$$

Proof of Proposition 9. Let $z_i \equiv g_z^i(\theta, \gamma, \lambda)$ and use (69) to define 5 to define $\phi_j^\beta(z_j, \beta_i) = 0$ for

$j = 1, \dots, I-1$. Since $\theta_i < \theta_{i+1}$ and $v'' > 0$,

$$\frac{\partial \phi_j^\beta(z_j, \beta_i)}{\partial \beta_i} = \begin{cases} -\frac{1}{\theta_i} v' \left(\frac{z_i}{\theta_i} \right) + \frac{1}{\theta_{i+1}} v' \left(\frac{z_i}{\theta_{i+1}} \right) < 0 & \text{if } j = i, \\ 0 & \text{otherwise.} \end{cases} \quad (82)$$

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

$$\frac{d\phi_j^\beta(\beta_i)}{d\beta_i} = -\frac{\partial \phi_j^\beta(z_j, \beta_i) / \partial \beta_i}{\partial \phi_j^\beta(z_j, \beta_i) / \partial z_j}. \quad (83)$$

As $\partial \phi_j^\beta(z_j, \beta_i) / \partial z_j \equiv \partial \phi_j^\gamma(z_j, \gamma) / \partial z_j < 0$ by (70), it follows from (82)–(83) that

$$\frac{\partial g_j^z(\theta, \gamma, \lambda)}{\partial \beta_i} \equiv \frac{d\phi_j^\beta(\beta_i)}{d\beta_i} \begin{cases} < 0 & \text{if } j = i, \\ = 0 & \text{otherwise.} \end{cases} \quad \square \quad (84)$$

Proof of Proposition 10. Since β_k is increased for all $k \in \{i, \dots, j-1\}$ and unaltered otherwise, Proposition 9 implies (48). We then prove (49). By Proposition 2,

$$x_k^*(z; \theta, \gamma) = \frac{1}{I} \left[\sum_{h=1}^I z_h - \frac{1}{\gamma} \sum_{h=2}^I (I+1-h) \left(v \left(\frac{z_h}{\theta_h} \right) - v \left(\frac{z_{h-1}}{\theta_h} \right) \right) \right] + \frac{1}{\gamma} \sum_{h=2}^k \left[v \left(\frac{z_h}{\theta_h} \right) - v \left(\frac{z_{h-1}}{\theta_h} \right) \right]. \quad (85)$$

For $k < i$, differentiating (85) and using (48),

$$\begin{aligned} \frac{dx_k^*(z; \theta, \gamma)}{ds} &= \frac{1}{I} \sum_{h=i}^{j-1} \left[1 - \frac{1}{\gamma \theta_h} v' \left(\frac{z_h}{\theta_h} \right) + \frac{I-h}{\gamma} \left(\frac{1}{\theta_{h+1}} v' \left(\frac{z_h}{\theta_{h+1}} \right) - \frac{1}{\theta_h} v' \left(\frac{z_h}{\theta_h} \right) \right) \right] \frac{dz_h}{ds} \\ &= \frac{1}{I} \sum_{h=i}^{j-1} \frac{T'(z_h, \theta_h)}{\beta_h} [\beta_h - I + h] \frac{dz_h}{ds}. \end{aligned} \quad (86)$$

As, for $h = 1, \dots, j-1$, (i) $\beta_h > 0$, (ii) $\beta_h - I + h < 0$ by (28) and (1), (iii) $T'(z_h, \theta_h) > 0$ [cf. (33)] and (iv) $dz_h/ds < 0$, one gets $dx_k^*(z; \theta, \gamma)/ds > 0$ for $k < i$.

For $k \geq j$,

$$\frac{d}{ds} \left\{ \frac{1}{\gamma} \sum_{h=2}^k \left[v \left(\frac{z_h}{\theta_h} \right) - v \left(\frac{z_{h-1}}{\theta_h} \right) \right] \right\} = \frac{1}{\gamma} \sum_{h=i}^{j-1} \left[\frac{1}{\theta_h} v' \left(\frac{z_h}{\theta_h} \right) - \frac{1}{\theta_{h+1}} v' \left(\frac{z_h}{\theta_{h+1}} \right) \right] \frac{dz_h}{ds} \quad (87)$$

$$= \sum_{h=i}^{j-1} T'(z_h, \theta_h) \alpha_h \frac{dz_h}{ds} = \sum_{h=i}^{j-1} \frac{T'(z_h, \theta_h)}{\beta_h} \frac{dz_h}{ds}. \quad (88)$$

This additional term is added to (86) to get

$$\frac{dx_k^*(z; \theta, \gamma)}{ds} = \frac{1}{I} \sum_{h=i}^{j-1} \frac{T'(z_h, \theta_h)}{\beta_h} [\beta_h + h] \frac{dz_h}{ds}. \quad (89)$$

As, for $h = i, \dots, j-1$, (i) $\beta_h + h = \Lambda(\theta_h) > 0$, (ii) $T'(z_h, \theta_h) > 0$ [cf. (33)] and (iii) $dz_h/ds < 0$, one obtains $dx_k^*(z; \theta, \gamma)/ds < 0$ for $k \geq j$.

By (15), (50) is a direct implication of (49) and (48). \square

References

- ATKINSON, A. (1990): "Public Economics and the Economic Public," *European Economic Review*, 34(2-3), 225–248.
- BLUNDELL, R. (1992): "Labour Supply and Taxation: A Survey," *Fiscal Studies*, 13, 15–40.
- BLUNDELL, R., AND T. MACURDY (1999): "Labor Supply: A Review of Alternative Approaches," in *Handbook of Labor Economics*, ed. by O. Ashenfelter, and D. Card, vol. 3A, pp. 1559–1695. North-Holland, Amsterdam.
- BOADWAY, R., K. CUFF, AND M. MARCHAND (2000): "Optimal Income Taxation With Quasi-Linear Preferences Revisited," *Journal of Public Economic Theory*, 2(4), 435–460.
- BOADWAY, R., AND P. PESTIEAU (2007): "Optimal Income Taxation with Tagging," *Annales d'Economie et de Statistiques*.
- BOONE, J., AND L. BOVENBERG (2007): "The Simple Economics of Bunching: Optimal Taxation with Quasi-Linear Preferences," *Journal of Public Economic Theory*, 9, 89–105.
- BRETT, C., AND J. WEYMARK (2004): "Public Good Provision and the Comparative Statics of Optimal Nonlinear Income Taxation," *International Economic Review* (forthcoming).
- CHAMBERS, R. (1989): "Concentrated Objective Functions for Nonlinear Taxation Models," *Journal of Public Economics*, 39, 365–375.

- D'AUTUME, A. (2000): "L'imposition optimale du revenu: une application au cas français," *Revue Française d'Economie*, 3, 3–63.
- DIAMOND, P. (1998): "Optimal Income Taxation: An Example with U-Shaped Pattern of Optimal Marginal Tax Rates," *American Economic Review*, 88(1), 83–95.
- EBERT, U. (1992): "A Reexamination of the Optimal Nonlinear Income Tax," *Journal of Public Economics*, 49(1), 47–73.
- GUESNERIE, R., AND J. SEADE (1982): "Nonlinear Pricing in a Finite Economy," *Journal of Public Economics*, 57, 157–179.
- HAMILTON, J., AND P. PESTIEAU (2005): "Optimal Income Taxation and the Ability Distribution : Implications for Migration Equilibria," *International Tax and Public Finance*, 12, 29–45.
- LOLLIVIER, S., AND J.-C. ROCHET (1983): "Bunching and Second-Order Conditions: A Note on Optimal Tax Theory," *Journal of Economic Theory*, 31, 392–400.
- MIRRLEES, J. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38(2), 175–208.
- PIKETTY, T. (1997): "La redistribution fiscale face au chômage," *Revue Française d'Economie*, 12, 157–201.
- SAEZ, E. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–229.
- (2002): "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 117, 1039–1073.
- SALANIÉ, B. (1998): "Un Exemple de Taxation Optimale," in *Fiscalité et Redistribution*, pp. 87–90. La Documentation Française.
- STIGLITZ, J. (1982): "Self-Selection and Pareto Efficient Taxation," *Journal of Public Economics*, 17(2), 213–240.
- TUOMALA, M. (1990): *Optimal Income Tax and Redistribution*. Clarendon Press, Oxford.
- WEYMARK, J. (1986a): "Bunching Properties of Optimal Nonlinear Income Taxes," *Social Choice and Welfare*, 3, 213–232.

——— (1986b): “A Reduced-Form Optimal Nonlinear Income Tax Problem,” *Journal of Public Economics*, 30, 199–217.

——— (1987): “Comparative Static Properties of Optimal Nonlinear Income Taxes,” *Econometrica*, 55, 1165–1185.