Thema
UMR 8184
THéorie Économique, Modélisation et Applications

Thema Working Paper n°2014-01
Université de Cergy Pontoise, France

Optimal Nonlinear Income Taxation with Multidimensional Types: The Case with Heterogeneous Behavioral Responses

Laurence JACQUET
Etienne LEHMANN

January, 2014

Optimal Nonlinear Income Taxation with Multidimensional Types: The Case with Heterogeneous Behavioral Responses*

Laurence JACQUET† Etienne LEHMANN‡

January 10, 2014

Abstract

This paper develops a general method to solve the optimal nonlinear income tax model with one action (individual pre-tax income) and multidimensional characteristics. Individuals differ in terms of skills and belong to different groups. A group is a subset of individuals with the same vector of characteristics but distinct skill levels. Assuming the Spence-Mirrlees single-crossing condition (with respect to the level of skill) in each group, we first derive the optimal second-best allocation. We then show how this optimality condition leads to a tax formula in terms of behavioral responses, social welfare weights and income density in the vein of Saez (2001). However, our multidimensional context implies that all these terms are averaged across individuals who earn the same income. We also show how our method can be used to solve a large set of policy relevant problems for which it is crucial to introduce multidimensional heterogeneity, e.g., joint taxation of households, nonlinear pricing of a monopoly.

*The authors acknowledge comments made by participants at seminars at CREST, THEMA,University of Oslo with a particular mention to Pierre Cahuc, Philippe Choné, Vidar Christiansen, Robert Gary-Bobo, Laurent Linnemer, David Martimort, Jean-Baptiste Michau, Julien Prat and Stéphane Robin. Any remaining errors are ours. Laurence Jacquet's research is supported by a Marie Curie FP7 Integration Grant within the 7th European Union Framework Program and by a labex MMEDII (ANR11-LBX-0023-01).

†THEMA - University of Cergy-Pontoise. Address: THEMA, Université de Cergy-Pontoise, 33 boulevard du Port, 95011 Cergy-Pontoise Cedex, France. Email: laurence.jacquet@u-cergy.fr. Laurence Jacquet is also research fellow at CESifo and IRES- Université Catholique de Louvain.

‡CRED (TEPP) University Panthéon Assas and CREST. Address: CREST-INSEE, Timbre J360, 15 boulevard Gabriel Péri, 92245, Malakoff Cedex, France. Email: etienne.lehmann@ensae.fr. A part of this research was undertaken while Etienne Lehmann was visiting UC Berkeley, which is gratefully acknowledged. Etienne Lehmann is also research fellow at IRES-Université Catholique de Louvain, IDEP, IZA and CESifo and is junior member of IUF.

I Introduction

The model of Mirrlees (1971), in which individual heterogeneity is uni-dimensional and individual choices are made along the intensive margin, has set a benchmark in several economic disciplines. Its applications range from public economics (to derive the income tax schedule on labor income) to industrial organization (to derive the nonlinear price schedule of a monopolistic firm). For technical reasons only, the Mirrlees model assumes the unobserved heterogeneity to be one-dimensional and imposes the Spence-Mirrlees single crossing condition. The assumption of heterogeneity along a single dimension is very restrictive. Realism calls for multidimensional heterogeneity. Workers differ along their skill levels but also along their labor supply elasticities for instance, and a consumer may have several uses for the same product, and her value of the product in each use may differ. However, it is well-known that multidimensional screening problems are typically challenging (Rochet and Choné, 1998).

We provide a method to solve a large class of optimal nonlinear income tax models when individuals differ along several dimensions and the observable action (the individual's amount of pre-tax income) is uni-dimensional. Individuals may both differ in terms of skills (that are continuously distributed) and belong to different groups. A group is a subset of individuals with distinct skill levels and with the same vector of other characteristics, e.g. gender, age, labor supply elasticity, level of non-labor income, etc. For instance, young, white women can be a single group with different skills. The set of groups may be finite or infinite and may be multidimensional. We show that keeping the assumption of single-crossing (with respect to skill) among individuals belonging to the same group makes the model tractable. We show that individuals of different groups pooled at the same income level are characterized by the same marginal rate of substitution between pre-tax and after-tax income. Intuitively, individuals of distinct groups who earn the same income level face the same marginal tax rate. From the individual maximization program, we know that identical marginal tax rates imply identical marginal rates of substitution. Using this equality in marginal rates of substitution together with the single-crossing condition within each group, we can fully characterize an incentive-compatible allocation from its restriction to a reference group Computing the first-order effects of a perturbation in this allocation yields a necessary condition for the second-best allocation (Proposition 1). The latter generalizes to multidimensional heterogeneity the standard necessary condition of the one-dimensional Mirrlees model.

We then rewrite this necessary condition in terms of behavioral responses, social welfare weights and income density in the spirit of Saez (2001). However, our multidimensional context implies that the behavioral responses, marginal social weights and income effects which appear in the tax formula are averaged across individuals pooled at the same income (Proposition 3). More precisely, the optimal marginal tax rate at a given income level is *ceteris paribus* inversely related (à la Ramsey (1927)) to the weighted average of the compensated labor supply elasticities of individuals earning the same income level. It also depends on an average of social welfare weights and income effects computed at larger income levels. Saez (2001) (p. 220) already conjectured that, by taking average behavioral responses and social welfare weights at any income level, his tax formula should also be valid in the case of multidimensional heterogeneity. We show that Saez's supposition is correct as soon as the single crossing condition (with respect to skill) is verified among individuals of the same group, and if the optimal allocation is smooth enough. Beside clarifying these conditions, we also establish the correct averaging procedure.

Another difference between Saez (2001)'s tax formula and ours lies in the way to integrate the following circular process: Consider a change in pre-tax income that occurs either due to substitution effects (in response to variations of the marginal tax rates), or due to income effects (in response to variation in tax liabilities). Because of the nonlinearity of the tax schedule, this change in pre-tax income affects in turn changes the marginal tax rate, which trigger a further change in pre-tax income due to a substitution effect. The tax formula in Saez (2001) integrates this circularity process by using, instead of the actual income density, the virtual income density defined as the density of income that would take place, at an income level, if the nonlinear tax schedule were replaced by a tangent linear tax at this income level. Using this non-intuitive virtual income density renders the tax formula not very transparent. Instead, we encapsulate the circularity process in our definitions of labor supply elasticities and income effects. This makes the circularity process part of the behavioral responses as we intuitively would expect it. We also make clear that this circularity process has to be taken into account to derive the optimal tax schedule.[1]

This paper gives a path to address a large scope of policy oriented problems for which it is crucial, but challenging, to include multidimensional heterogeneity. Under relative weak assumptions, our method can be used to determine the optimal nonlinear taxation of joint labor and non-labor income, of labour income in the presence of untaxable non-labor income, and of tax avoidance. These applications need to consider models where individuals earning the same income may have different labor supply elasticities. Our method also helps to characterize the optimal nonlinear monopoly pricing when consumers differ in the slope and the intercept of their demand curves.

Our method enables to treat in a simple way cases where individuals with the same income level have distinct labor supply elasticities. While this is an empirically relevant feature, previous screening models with multidimensional heterogeneity and one-dimensional action did not allow for it.[2] In random participation models applied to nonlinear pricing (Rochet and Stole, 2002) and to income taxation (Saez, 2002; Jacquet, Lehmann, and Van der Linden, 2013),

[1] Revecz (2003) criticizes Saez (2001) on the ground that the effects due to the circularity process does not need to be taken into account. Saez (2003) replies that including the circularity process is necessary to ensure the consistency between the optimal tax formula of Saez (2001) and the necessary condition provided in Mirrlees (1971). We show that this is also the case with multidimensional heterogeneity.

[2] An exception is Laffont, Maskin, and Rochet (1987) who study the nonlinear pricing model when consumers differ both in the slope and the intercept of their demand curves. However, they need much more restrictive assumptions on preferences and the distributions of characteristics than we do.

individuals differ along two unobserved characteristics, one being the participation cost which drives the decision to participate. When an individual participates, her choice (along the intensive margin) only depends on her other characteristic. Take individuals who participate and take the same action. They must have the same other characteristic. In the optimal tax context, this means that workers who earn the same income, but have distinct participation costs, are endowed with the same skill. Therefore, these workers cannot have distinct labor supply elasticities.

Choné and Laroque (2010) consider an income tax problem with a bi-dimensional unobserved heterogeneity. The tractability of their model relies on an aggregator as used in Brett and Weymark (2003), i.e. an exogenous unidimensional combination of the agents' characteristics. The individual intensive choices then only depend on this unidimensional combination. Hence, two individuals who earn the same income must be characterized by the same level of the aggregator and therefore cannot have distinct labor supply elasticities. Choné and Laroque (2010) show moreover that optimal marginal tax rates may become negative thanks to the additional source of heterogeneity. In their model, individuals who earn the same income can have distinct social welfare weights so that (average) social welfare weights may not be decreasing with income (see also Boadway, Marchand, Pestieau, and del Mar Racionero (2002)). We show that, once the social preferences are restricted such that individuals who earn the same income, have the same social welfare weights, optimal marginal tax rates are positive (Proposition 2). Therefore, our general framework highlights that allowing for heterogeneous labor supply elasticities is not sufficient to obtain negative marginal tax rates.

Rothschild and Scheuer (2013a,b,c) and Scheuer (2013a,b) consider optimal income tax models with several sectors. Individuals are characterized by their productivities in these sectors. They choose their total labor effort and how to split it across each sector. The latter decision depends on the relative price of labor across sectors, which is determined at equilibrium. In their model, the disutility of effort of each individual turns out to be a function of the ratio of their income over an individual wage rate, which is itself a function of her productivities in all sectors and of the prices in all sectors. Therefore, two individuals who earn the same income need to have the same wage rate, thereby the same labor supply elasticity.

The paper is organized as follows. The next section introduces the model. Section III explains how it applies to many relevant problems. Section IV characterizes incentive-compatible allocation, with a particular emphasis on the pooling mechanism. Section V derives the necessary condition for the optimal allocation. Section VI reinterprets this condition in terms of behavioral responses, social welfare weights and income density.

II The model

II.1 Individuals

We assume that individuals differ along their skill level $w \in \mathbb{R}_+$ and along a vector of characteristics denoted $\theta \in \Theta$. These characteristics can be labor supply elasticity, gender, levels of non-labor income, etc. We call a *group* a subset of individuals with the same θ. We assume that the set of groups Θ is measurable with a distribution $\mu(\cdot)$. Θ may be finite or infinite. Moreover, Θ may be of any dimension. The distribution $\mu(.)$ of the population across the different groups may be continuous, but it may also exhibit mass points. Within individuals of the same group θ, skills are continuously distributed according to the conditional density $f(\cdot|\theta)$. This density is assumed positive over the support $\mathbb{R}_+$ and continuous. The conditional cumulative distribution function is denoted $F(w|\theta) \overset{\text{def}}{\equiv} \int_0^w f(x|\theta)dx$. The size of the total population is normalized to 1, so that:

$$\int_{\theta\in\Theta} \left\{ \int_0^{+\infty} f(w|\theta)\, dw \right\} d\mu(\theta) = 1$$

An individual of type (w,θ) admits preferences over after-tax income (or consumption) c and pre-tax income (income for short) y that are described by the twice differentiable utility function:

$$\mathscr{U}(c,y;w,\theta) = u(c) - v(y;w,\theta) \qquad \text{with} \qquad u'(\cdot), v'_y(\cdot), v''_{yy}(\cdot) > 0 > v'_w(\cdot) \tag{1}$$

Utility increases with consumption while it decreases with pre-tax income as a higher pre-tax income is obtained thanks to a higher supply of effort. Moreover, utility increases with productivity because earning a given income requires less effort to a more productive agents. The latter assumption is standard. For instance, when preferences depend on effort ℓ and when pre-tax income is equal to the product of effort and skill, $y = w \times \ell$, we get:

$$\mathscr{U}(c,y;w,\theta) \equiv u(c) - \mathcal{V}\left(\frac{y}{w};w,\theta\right) \qquad \text{with} \qquad \mathcal{V}'_\ell(\cdot) > 0 \quad \mathcal{V}''_{\ell\ell}(\cdot) > 0 \tag{2}$$

assuming $\mathcal{V}'_\ell > 0$ implies $v'_y > 0 > v'_w$ thereby $\mathscr{U}'_y < 0 < \mathscr{U}'_w$.

It is very common to assume that preferences are quasilinear in consumption (i.e $u(\cdot)$ is linear). Such assumption rules out income effects on the labor supply. Our additively separable specification of preferences in (1) generalizes the quasilinear assumption by allowing for income effects.

The marginal rate of substitution between (pre-tax) income and consumption is:

$$\mathscr{M}(c,y;w,\theta) \overset{\text{def}}{\equiv} -\frac{\mathscr{U}'_y(c,y;w,\theta)}{\mathscr{U}'_c(c,y;w,\theta)} = \frac{v'_y(y;w,\theta)}{u'(c)} \tag{3}$$

We impose a strict single-crossing (Spence-Mirrlees) condition within each group of individuals endowed with the same θ. For each θ, starting from any positive level of consumption and pre-tax income, more skilled agents need to be compensated with a smaller increase in their consumption to accept increasing their pre-tax income by one unit. For each $\theta \in \Theta$, the marginal rate of substitution $\mathscr{M}(c, y; w, \theta)$ thus decreases in the skill level, as imposed by the following assumption.

Assumption 1 (Withing group single-crossing condition)**.** *For each $\theta \in \Theta$, and each (c, y), function $w \mapsto \mathscr{M}(c, y; w, \theta)$ maps $\mathbb{R}_+$ into $\mathbb{R}_+$ with a strictly negative derivative everywhere, so:*

$$\mathscr{M}'_w(c, y; w, \theta) < 0 \quad \Leftrightarrow \quad v''_{yw}(y; w, \theta) < 0 \tag{4}$$

That $v''_{yw}(y; w, \theta) < 0$ is not a restrictive assumption as it holds for instance when preferences verify (2). Assumption 1 imposes in addition that the marginal rate of substitution decreases from plus infinity to zero. This is a kind of INADA condition.

II.2 The Government

Let $Y(w, \theta)$ and $C(w, \theta) = Y(w, \theta) - T(Y(w, \theta))$ respectively denote the pre-tax and after-tax income of an individual of type (w, θ). The government's budget constraint is:

$$\int_{\theta \in \Theta} \left\{ \int_0^{+\infty} [Y(w, \theta) - C(w, \theta)] \, f(w|\theta) \, dw \right\} d\mu(\theta) \geq E \tag{5}$$

where E is an exogenous amount of public expenditures.

Turning now to the government's objective function, we assume that the government uses a type-specific cardinal representation of individuals' utility. Let $U(w, \theta) = u(C(w, \theta)) - v(Y(w, \theta); w, \theta)$ denote the utility level enjoyed by an individual of type (w, θ) and $\Phi(\cdot; w, \theta)$ denote the type-specific increasing and weakly concave social transformation of individuals' utility. The government maximizes:

$$\int_{\theta \in \Theta} \left\{ \int_0^{+\infty} \Phi \left(U(w, \theta); w, \theta \right) \; f(w|\theta) \, dw \right\} d\mu(\theta) \tag{6}$$

This general specification admits the *Benthamite* social preferences where $\Phi(U; w, \theta)) \equiv U$ as a particular case. The social objective is then:

$$\int_{\theta \in \Theta} \left\{ \int_0^{+\infty} U(w, \theta) \, f(w|\theta) \, dw \right\} d\mu(\theta)$$

Another particular case is weighted utilitarianism with type-specific weights $\varphi(w, \theta)$ and $\Phi(U; w, \theta)) \equiv \varphi(w, \theta) \cdot U$. The social objective is then:

$$\int_{\theta \in \Theta} \left\{ \int_0^{+\infty} \varphi \left(w, \theta \right) \; U(w, \theta) \, f(w|\theta) \, dw \right\} d\mu(\theta)$$

Lastly, our social objective includes the Bergson-Samuelson case where the social objective is a type-independent transformation of individuals utility so $\Phi(U; w, \theta)$ does not vary with its two last arguments. The social objective takes then the form:

$$\int_{\theta \in \Theta} \left\{ \int_0^{+\infty} \Phi \left(U(w, \theta) \right) \; f(w|\theta) \, dw \right\} d\mu(\theta)$$

III Applications

We now provide examples of different policy problems that can be solved with our framework, where the Mirrlees (1971) model is extended for heterogeneity in θ. All the additional sources of heterogeneity that are introduced in each example are not mutually exclusive such that each argument of the vector θ can be a distinct source of heterogeneity.

III.1 Heterogeneous labor supply elasticities

Our leading example is the case where heterogeneity in θ represent distinct labor supply elasticities. For instance, the empirical literature suggests that labor supply elasticity is higher for women than for men of similar earnings. We can rewrite the individual preferences (1) as isoelastic ones with:

$$\mathscr{U}(c,y;w,\theta) = u(c) - \frac{\theta}{1+\theta}\left(\frac{y}{w}\right)^{1+\frac{1}{\theta}} \qquad \text{with} \qquad \theta > 0 \quad \text{and} \quad u'(\cdot) > 0 \geq u''(\cdot) \qquad (7)$$

where θ is the elasticity of the labor supply.[3] Typically, women have a larger labor supply elasticity θ than men. In this case, the distribution $\mu(.)$ is two mass points with two values for labor supply elasticity θ depending on the gender. Alternatively, we can also consider that $\Theta = \mathbb{R}_+$ if labor supply elasticities are continuously distributed among individuals.

III.2 Joint taxation of non-labor income

We can also consider cases where θ stands for the ability to earn some non-labor income z that the government observe but cannot distinguish from labor income. This can for instance be entrepreneurial income, or the rents from housing capital. When the income tax schedule treats identically labor and non-labor income, as in France, labor income is equal to taxable income y minus non-labor income z. Let $V(y-z,z;w,\theta)$ denote the joint disutility of earning labor income $y-z$ and non-labor income z for an individual of skill w who belongs to the group θ, with $V'_{y-z}, V'_z > 0$. Parameter θ may here stand for the ability of earning non-labor income. Individuals of type (w,θ) solves:

$$\max_{y,z} \quad u(y - T(y)) - V(y-z,z;w,\theta)$$

This program can be solved sequentially, the last step being the choice of non-labor income z for a given taxable income y. Our model can then be retrieved by defining:

$$v(y;w,\theta) \stackrel{\text{def}}{\equiv} \min_z \quad V(y-z,z;w,\theta)$$

Our framework applies when the second-order derivatives of $V(\cdot)$ are such that $v''_{yw} < 0$.[4] Such a restriction arises naturally when non-labor income is exogenous, as it is the case when z stands for rents perceive by landlords who have inherited the property they rent.

[3] θ is the labor supply elasticity if $u(\cdot)$ is linear (no income effect) and is otherwise the Frish labor supply elasticity.

[4] The envelope theorem induces that $v'_y = V'_{y-z}$ and $v'_w = V'_w$. Hence, one obtains $v'_y > 0 > v'_w$, whenever $V'_{y-z} > 0 > V'_w$, which is a natural assumption. In contrast, the within-group single-crossing property $v''_{yw} < 0$ needs further restrictions on the second-order derivatives of V.

III.3 Nonlinear household joint income taxation

The preceding case applies in particular to the optimal joint income taxation of households when the members of a household take their decisions cooperatively. Then, w and θ denote the skill levels of the members of the couple and their labor incomes are respectively $y - z$ and z.

III.4 Tax avoidance

Our framework is also relevant to solve an income tax model with fiscal avoidance. Assume that θ is the individual ability to avoid taxation. We denote z the amount of income that is not taxed (because the individual has found means that allows her to reduce her taxable income) and y the taxable income, so that labor income is equal to $y + z$. Consumption is the sum of after-tax taxable income $c = y - T(y)$ plus untaxed income z. We assume here that preferences are quasi-linear in consumption:

$$c + z - V\left(y + z, z; w, \theta\right)$$

where $V'_{y+z}, V''_{y+z} > 0$. Moreover, $V'_z, V''_z > 0$; for a given income, avoiding taxation is more and more costly (i.e., requires more and more effort). To retrieve our model, we simply define:

$$v\left(y; w, \theta\right) \stackrel{def}{\equiv} \min_z V\left(y + z, z; w, \theta\right) - z$$

and assume that the second-order derivatives of $V(\cdot)$ are such that $v''_{yw}\left(y; w, \theta\right) < 0$ to ensure the within-group single-crossing property.

III.5 Untaxable non-labor income

We can also consider the case where θ is some exogenous untaxable non-labor income. For instance, we may think of θ as the imputed rent of owner-occupied housing. In this context, consumption is $c + \theta$. Let us assume that the social objective is $\widetilde{\Phi}(U; w, \theta)$ and that individual preferences over consumption exhibit constant absolute risk aversion (CARA), the individual utility can be stated as $\mathscr{U}\left(c, y; w, \theta\right) = -e^{-\gamma(c+\theta)} - \tilde{v}(y; w)$.[5] To solve this model, we simply divide this utility function by $a(\theta) \stackrel{\text{def}}{\equiv} e^{-\gamma\theta}$ which yields individual preferences (1) where:

$$v(y; w, \theta) \stackrel{\text{def}}{\equiv} \frac{\tilde{v}(y; w)}{a(\theta)}$$

and $u(c) = -e^{-\gamma c}$ and, we multiply by $a(\theta)$ the individual utility in the objective function so that the social objective is

$$\Phi(U; w, \theta) \stackrel{\text{def}}{\equiv} \widetilde{\Phi}\left(a\left(\theta\right) \cdot U; w, \theta\right)$$

[5]We here obviously assume that $\tilde{v}'_y > 0 > \tilde{v}'_w$, that $\tilde{v}''_{yw} < 0$ and that $\widetilde{\Phi}'_u > 0 \geq \widetilde{\Phi}''_{uu}$.

III.6 Nonlinear pricing

Our framework is an adverse selection model where the *principal* is the government and individuals are the *agents*. The principal observes a one-dimensional *action* (individuals' pre-tax income) but individuals' unobserved characteristics are multidimensional ($(w,\theta) \in \mathbb{R}_+ \times \Theta$). Although we focus in this paper on the optimal taxation applications, this setup can also be applied to other economic contexts. This simply requires a reinterpretation of the variables. For instance, Laffont, Maskin, and Rochet (1987) study the problem of a monopolist who sells a single product and needs to determine a nonlinear price schedule observing only a one-dimensional action (how much consumers are demanding), while consumers (the agents) differ both in the slope and in the intercept of their demand curves. They derive the optimal quantity assignment function when these two characteristics are independently and uniformly distributed and under restrictive assumptions on preferences (they are assumed linear in income and quadratic in consumption). Our model can then be used to solve this nonlinear pricing model under less restrictive assumptions on the distributions of characteristics and on preferences.[6] We believe our framework can also be reinterpreted to other economic applications of adverse selection problems.

IV Incentive-compatible allocations

This section characterizes incentive-compatible allocations. An individual of type (w,θ), facing the nonlinear income tax $y \mapsto T(y)$ solves:

$$\max_{y} \quad \mathscr{U}\left(y - T(y), y; w, \theta\right) \tag{8}$$

The[7] solution to the maximization program (8) is $Y(w,\theta)$ and we get $C(w,\theta) = Y(w,\theta) - T(Y(w,\theta))$ and $U(w,\theta) = u(C(w,\theta)) - v\left(Y(w,\theta); w,\theta\right)$. When the tax function is differentiable, the first-order condition associated to (8) implies with (3):

$$1 - T'\left(Y\left(w,\theta\right)\right) = \mathscr{M}\left(C\left(w,\theta\right), Y\left(w,\theta\right); w,\theta\right) \tag{9}$$

where the right-hand side is the marginal rate of substitution between pre-tax income and consumption. As the individual's objective is maximized for $y = Y(w,\theta)$, we must have:

$$\forall (w,\theta,y') \in \mathbb{R}_+ \times \Theta \times \mathbb{R}_+ \quad \mathscr{U}\left(C(w,\theta), Y(w,\theta); w,\theta\right) \geq \mathscr{U}\left(y' - T(y'), y'; w,\theta\right)$$

Taking $y' = Y(w',\theta')$ leads to the following set of incentive constraints:

$$\forall (w,w',\theta,\theta') \in \mathbb{R}_+^2 \times \Theta^2 \quad \mathscr{U}\left(C(w,\theta), Y(w,\theta); w,\theta\right) \geq \mathscr{U}\left(C(w',\theta'), Y(w',\theta'); w,\theta\right) \tag{10}$$

[6] It should be noticed that we restrict in this paper to deterministic mechanism. This is restriction is natural in the context of optimal income taxation but may be less natural in other contexts.

[7] If the maximization program (8) admits multiple solutions, we make the tie-breaking assumption that individuals choose among its best options the income level preferred by the government, i.e. the one with the largest tax liability.

i.e. the individuals characterized by (w, θ) prefer the bundle $(C(w, \theta), Y(w, \theta))$ they have chosen to any other bundle intended to any other type of workers. According to the *taxation principle* (Hammond, 1979; Rochet, 1985; Guesnerie, 1995), for any incentive-compatible allocation, there exists a nonlinear income tax schedule $T(.)$ that decentralizes it. It is therefore equivalent for the government to find an income tax schedule and to take into account the reactions of individuals to this tax schedule incorporating the individual's maximization problem (8), or to choose *directly* an incentive-compatible allocation (i.e. an allocation that verifies the set of incentive constraints (10)).

We characterize the set of incentive-compatible allocations in two steps. We first characterize incentive-compatible allocations within each group. In this step, the within-group single crossing assumption 1 enables to retrieve the properties that are usual when unobserved heterogeneity is one-dimensional as in Mirrlees (1971). The novelty lies in the second step where we characterize how these within-group allocations need to be set to ensure incentive-compatibility.

IV.1 Withing-group incentive constraints

An incentive-compatible allocation has to satisfy (10). It thus has to verify the set of incentive-compatible constraints within each θ-group, that we call "within-group incentive compatible constraints", i.e.:

$$\forall (w, w', \theta) \in \mathbb{R}_+^2 \times \Theta \quad \mathscr{U}\left(C(w, \theta), Y(w, \theta); w, \theta\right) \geq \mathscr{U}\left(C(w', \theta), Y(w', \theta); w, \theta\right) \tag{11}$$

For each θ, characterizing the within-group allocations $w \mapsto (C(w, \theta), Y(w, \theta))$ that verify the within-group incentive constraints (11) is the same problem as characterizing incentive compatible allocations when unobserved heterogeneity is one-dimensional, due to the within-group single crossing assumption 1. We can therefore retrieve the following standard result:

Lemma 1. *Under Assumption 1, The function $w \mapsto Y(w, \theta)$ is nondecreasing for each $\theta \in \Theta$.*

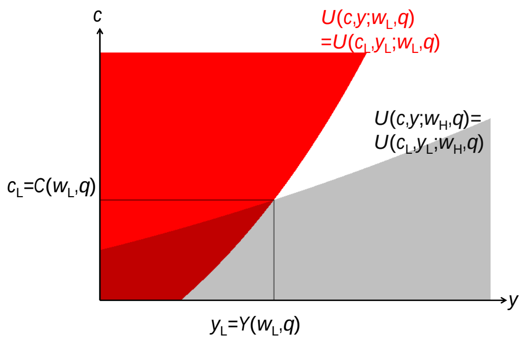

Figure 1: Proof of Lemma 1

The usual graphical proof applies. Figure 1 displays the indifference curves of individuals belonging to the same group θ but endowed with two distinct skill levels $w_L < w_H$. These indifference curves are labelled $\mathscr{U}(c, y; w_L, \theta)$ and $\mathscr{U}(c, y; w_H, \theta)$. The within-group single-crossing assumption implies that the indifference curve of the low-skilled workers is steeper than the one of the high-skilled worker. The indifference curves intersect at the bundle $(C(w_L, \theta), Y(w_L, \theta))$ that the government designs for individuals of type (w_L, θ). To respect the incentive constraints (11), the government needs to assign a bundle $(C(w_H, \theta), Y(w_H, \theta))$ to the high-skilled workers that is above the indifference curve of the high-skilled workers $\mathscr{U}(c, y; w_H, \theta)$ (otherwise, the individuals of type (w_H, θ) would prefer the bundle $(C(w_L, \theta), Y(w_L, \theta))$ to the bundle $(C(w_H, \theta), Y(w_H, \theta))$ designed for them) and below the indifference curve of the low-skilled workers (otherwise, individuals of type (w_L, θ) would prefer the bundle $(C(w_H, \theta), Y(w_L, \theta))$ to the bundle $(C(w_H, \theta), Y(w_H, \theta))$ designed for them). Consequently, the bundle $(C(w_H, \theta), Y(w_H, \theta))$ designed for the high-skilled workers should be located in the non-shaded area in Figure 1, which implies that $Y(w_L, \theta) \leq Y(w_H, \theta)$.

$Y(\cdot; \theta)$ being nondecreasing, it may exhibit discontinuities over a countable set and it may also exhibit bunching (i.e. portion where it is constant). It is however standard to consider only smooth allocations where these two "pathologies" do not arise. We therefore make the following smoothness assumption.

Assumption 2 (Smooth allocations). *For each θ, $w \mapsto Y(w, \theta)$ is differentiable and maps $\mathbb{R}_+$ into $\mathbb{R}_+$ with a strictly positive derivative.*

We henceforth use a dot to denote the derivatives with respect to w of functions $Y(\cdot, \theta)$, $C(\cdot, \theta)$ and $U(\cdot, \theta)$. The following lemma shows the equality between the marginal rate of substitution between pre-tax income and consumption and the ratio $\dot{C}(w, \theta)/\dot{Y}(w, \theta)$.

Lemma 2. *Under Assumptions 1 and 2, for each θ, the mapping $w \mapsto U(w, \theta)$ is differentiable with a derivative*

$$\dot{U}(w, \theta) = \mathscr{U}'_w(C(w, \theta), Y(w, \theta); w, \theta) = -v'_w\left(Y(w, \theta); w, \theta\right) \tag{12}$$

Moreover, Equation (12) *is equivalent to:*

$$\frac{\dot{C}(w, \theta)}{\dot{Y}(w, \theta)} = \mathscr{M}\left(C(w, \theta), Y(w, \theta); w, \theta\right) \tag{13}$$

The proof is in Appendix A. This lemma will be useful to characterize the set of incentive compatible allocations. Integrating Equation (12) leads to:

$$U(w, \theta) = U(0, \theta) - \int_0^w v'_w\left(Y(t, \theta); t, \theta\right)\, dt \tag{14}$$

If the government was able to observe the group θ to which each taxpayer belongs to, the government would propose group-specific income tax schedules $T(\cdot; \theta)$. Only the within-group incentive constraints (11) (and not (10)) would then be taken into account.[8] The observation of θ then reduces the set of incentive constraints and increases the possibility for the government to redistribute income as highlighted in the so-called tagging literature (Akerlof, 1978; Cremer, Gahvari, and Lozachmeur, 2010). In contrast, this paper considers the no-tagging case where θ is unobserved. We thus need to describe how the various within-group allocations coexist to verify the full set of incentive constraints (10).

IV.2 Pooling types across θ-groups at each income level

In our context of multidimensional heterogeneity, each level of labor income Y is obtained by individuals having distinct couples of individual characteristics. In other words, individuals belonging to distinct θ-groups and endowed with distinct w pool at the same income level $Y(w, \theta)$.[9]

Let us choose a reference group $\theta_0 \in \Theta$. A key of our analysis relies in the construction and characterization of the subjacent pooling function. At each income level, this function pools together (*i*) individuals of skill w belonging to the reference group θ_0 that allow them to earn $Y(w, \theta_0)$, and (*ii*) individuals in another group θ and whose level of skill allows them to earn the same level of labor income $Y(w, \theta_0)$ as individuals of skill w belonging to the group θ_0. More precisely, the pooling function, that we denote $W(w, \theta; \theta_0)$ is the mapping that associates to a skill level w the skill level that is needed to individuals in an another group θ to obtain the same income level $Y(w, \theta_0)$ as individuals of skills θ in the reference group. For each $\theta \in \Theta$, we have $w \overset{Y(\cdot,\theta_0)}{\longmapsto} Y(w, \theta_0) \overset{Y^{-1}(\cdot,\theta)}{\longmapsto} W(w, \theta; \theta_0)$ which, according to Assumption 2, is a one-to-one differentiable mapping, with a strictly positive derivative everywhere.

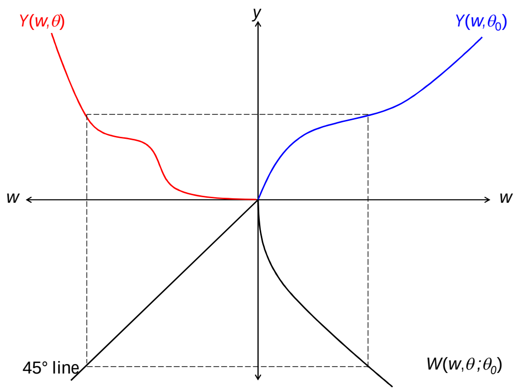

Figure 2: The pooling function

[8]To be more precise, this remark holds only if the government was furthermore allowed to condition taxation on θ. For instance, despite the government can *observe* whether a taxpayer is a woman or a man, gender-based taxation is in practice ruled out for horizontal equity reasons, preventing the government for using an information that would be otherwise relevant (Alesina, Ichino, and Karabarbounis, 2011). A similar issue arises to condition income taxation on height (Mankiw and Weinzierl, 2010).

[9] We call "pooling" the case where individuals belonging to different groups θ obtain exactly the same pre-tax income while "bunching" occurs when individuals characterized by distinct w but belonging to the same group θ obtain exactly the same pre-tax income. Assumption 2 rules out bunching and makes pooling unavoidable.

Figure 2 illustrates the pooling mechanism. Let us choose a reference group denoted θ_0. For any skill level w, the resulting income level, $Y(w,\theta_0)$, is represented in the upper-right quarter of Figure 2. Now, let us choose another element denoted θ of Θ. From the upper-left quarter of Figure 2, we can find the skill level which gives the same level of income as $Y(w,\theta_0)$, for this alternative θ. Using the 45°-line in the lower-left quarter allows to draw the latter skill level as a function of the initial skill level w in the lower-right quarter. Repeating this exercise for any w determines the pooling function $W(\cdot,\theta;\theta_0)$.

By definition of the pooling function, we have $Y(W(w,\theta;\theta_0),\theta) \equiv Y(w,\theta_0)$. Provided that the allocation is incentive-compatible, it is not possible that individuals endowed with $W(w,\theta;\theta_0)$ and θ and agents endowed with w and θ_0 obtain the same labor income $Y(w,\theta_0)$ but distinct consumption levels. Therefore, we know that for each (w,θ), we simultaneously have

$$Y(W(w,\theta;\theta_0),\theta) \equiv Y(w,\theta_0) \qquad \text{and} \qquad C(W(w,\theta;\theta_0),\theta) \equiv C(w,\theta_0). \tag{15}$$

This simultaneous equality can then be used to derive a very useful property of incentive-compatible allocations: Individuals of different θ-groups that pool at the same income level needs to have the same marginal rate of substitution between pre-tax income and consumption. This property is formally presented in the following lemma.

Lemma 3. *Under Assumptions 1 and 2, along an incentive-compatible allocation, the bundle designed for individuals of type $(W(w,\theta;\theta_0),\theta)$ coincides with the bundle $(C(w,\theta_0),Y(w,\theta_0))$ designed for individuals of type (w,θ_0), where $W(w,\theta;\theta_0)$ is the unique solution in ω to*

$$\mathscr{M}\left(C(w,\theta_0),Y(w,\theta_0);w,\theta_0\right) = \mathscr{M}\left(C(w,\theta_0),Y(w,\theta_0);\omega,\theta\right) \tag{16}$$

Proof According to Assumption 1, Equation (16) admits exactly one solution in ω. Differentiating in w the two equalities in (15) leads to:

$$\dot{Y}(W(w,\theta;\theta_0),\theta)\ \dot{W}(w,\theta) = \dot{Y}(w,\theta_0) \qquad \text{and} \qquad \dot{C}(W(w,\theta;\theta_0),\theta)\ \dot{W}(w,\theta) = \dot{C}(w,\theta_0)$$

where $\dot{W}(w,\theta)$ denotes the partial derivative of W in its first argument (skill). Hence,

$$\frac{\dot{C}(W(w,\theta;\theta_0),\theta)}{\dot{Y}(W(w,\theta;\theta_0),\theta)} = \frac{\dot{C}(w,\theta_0)}{\dot{Y}(w,\theta_0)}$$

If the allocation is incentive-compatible, then, according to Lemma 2, (13) holds, which implies:

$$\mathscr{M}\left(C(w,\theta_0),Y(w,\theta_0);w,\theta_0\right) = \mathscr{M}\left(C(w,\theta_0),Y(w,\theta_0);W(w,\theta;\theta_0),\theta\right) \tag{17}$$

□

Intuitively, if individuals of type (w,θ_0) and of type $(W(w,\theta;\theta_0),\theta)$ choose the same income Y, they face the same marginal tax rate $T'(.)$. Hence, from the first-order condition (9) they must face the same marginal rate of substitution. The key point to note is that, because of the within-group single crossing Assumption 1, Equation (16) admits exactly one solution in ω. So, (16) fully characterizes the pooling function $W(\cdot,\theta;\theta_0)$. The following lemma, which is proved in Appendix B, shows that once an incentive-compatible allocation is set for the reference group θ_0, the allocation for another group θ is determined by the equality between their marginal rates of substitution (17).

Lemma 4. *Let $w \mapsto (C(w,\theta_0), Y(w,\theta_0))$ be a within-group allocation that verifies Assumption 2 and the within-group incentive-compatible Equation* (13). *For each $w \in \mathbb{R}_+$ and each group $\theta \in \Theta$, let $\underline{W}(w,\theta)$ be the unique skill level ω that solves* (16). *Then, there exists a unique incentive-compatible allocation $(w,\theta) \mapsto (\underline{C}(w,\theta), \underline{Y}(w,\theta))$ whose restriction to group θ_0 is $w \mapsto (C(w,\theta_0), Y(w,\theta_0))$ and that verifies Assumption 2 if and only if, for each θ, $\underline{W}(\cdot,\theta)$ maps $\mathbb{R}_+$ into $\mathbb{R}_+$ and admits a positive derivative $\underline{\dot{W}}(w,\theta) > 0$ everywhere.*

The assumption that $\underline{W}(\cdot,\theta)$ admits a positive derivative everywhere plays in our analysis a role similar to the "first-order approach" in the Mirrleesian literature with a one-dimensional unobserved heterogeneity. In the sequel, we therefore select the allocation only for the reference group θ_0 and assume that the allocation triggered for the other groups verifies Assumption 2.

Using (3), the equality of the marginal rates of substitution in Equation (17) can be rewritten as:

$$\frac{v'_y\left(Y(w,\theta_0);w,\theta_0\right)}{u'(C(w,\theta_0))} = \frac{v'_y\left(Y(w,\theta_0);W(w,\theta;\theta_0),\theta\right)}{u'(C(w,\theta_0))}$$

which simplifies to:

$$v'_y\left(Y(w,\theta_0);w,\theta_0\right) = v'_y\left(Y(w,\theta_0);W(w,\theta;\theta_0),\theta\right) \tag{18}$$

Therefore, the pooling function $W(\cdot,\theta;\theta_0)$ that enables to retrieve $(C(\cdot,\theta), Y(\cdot,\theta))$ from $(C(\cdot,\theta_0), Y(\cdot,\theta_0))$ depends on $Y(\cdot,\theta)$. However the pooling function does not depend on $C(\cdot,\theta)$.[10] Consider as an illustration the case of isoelastic preferences as in (7). Equation (18) then implies that the pooling function is:

$$W(w,\theta,\theta_0) = \left(w^{\frac{1+\theta_0}{\theta_0}} \cdot \left(Y(w,\theta_0)\right)^{\frac{1}{\theta}-\frac{1}{\theta_0}}\right)^{\frac{\theta}{1+\theta}}$$

The pooling function thus depend on the choice of $Y(\cdot,\theta_0)$ whenever the different groups are endowed with different labor supply elasticities (i.e. when $\theta \neq \theta_0$).

To the best of our knowledge, the literature has only considered models where what we call the pooling function depends neither on $Y(.,.)$ nor on $C(.,.)$. In our words, the pooling function thus depends only on exogenous characteristics and not on chosen variables. This is in particular the case when income decisions depends on a one-dimensional aggregator of the multi-dimensional unobserved types (Choné and Laroque, 2010). Individuals that earn the same income level are then characterized by the same level of the aggregator and are therefore characterized by the same labor supply elasticity. In Rothschild and Scheuer (2013b,c,a) and Scheuer (2013a,b), as detailed in our introduction, individuals who earn the same pre-tax income have the same labor supply elasticity. In Random participation models (Rochet and Stole,

[10]This simplification relies on the commonly assumed separability in the utility function, see (1).

2002; Jacquet, Lehmann, and Van der Linden, 2013) or in optimal income tax models with migration (Blumkin, Sadka, and Shem-Tov, 2012; Lehmann, Simula, and Trannoy, 2013), people with identical skill levels earn the same pre-tax income, whatever their participation costs. The pooling function is thus an identity hence, is exogenous.

V The optimal allocation

In this section, we derive necessary conditions for the optimal tax problem. Applying variational calculus to a small perturbation in the optimal allocation, we derive the necessary conditions of the optimal allocation. The importance of the pooling function we characterized in the previous section reveals itself here.

Let $\mathscr{C}(u, y; w, \theta)$ denote the consumption level the governments need to provide to a worker of type (w, θ) that earns y to ensure her with a utility level u. Function $\mathscr{C}(\cdot, y; w, \theta)$ is the reciprocal of $\mathscr{U}(\cdot, y; w, \theta)$ and we have:

$$\mathscr{C}'_u(u, y; w, \theta) = \frac{1}{u'(c)} \qquad \text{and} \qquad \mathscr{C}'_y(u, y; w, \theta) = \frac{v'_y(y; w, \theta)}{u'_c(c)} \tag{19}$$

where the various derivatives are evaluated at $c = \mathscr{C}(u, y; w, \theta)$. Let λ denote the Lagrange multiplier associated to the government's budget constraint. We define the Lagrangian $\mathscr{L}$ of the government's problem as:

$$\mathscr{L} \stackrel{\text{def}}{\equiv} \iint_{(w,\theta)} \left[Y(w,\theta) - \mathscr{C}\left(U(w,\theta), Y(w,\theta); w, \theta\right) + \frac{\Phi\left(U(w,\theta); w, \theta\right)}{\lambda} \right] f(w|\theta) dw \, d\mu(\theta) \tag{20}$$

The government's problem consists in maximizing the Lagrangian $\mathscr{L}$ within the subset of incentive-compatible allocations (i.e. of allocations that verify (10)).

V.1 The optimal allocation when θ is observable

Before deriving the necessary conditions of our problem with multidimensional heterogeneity, we first remind the necessary conditions of the optimal allocation when the unobserved heterogeneity has one dimension only. This case arises either when individuals differ in skills only or in an economy where tagging can be used hence the government observes θ and can condition its tax schedule on both the income Y and the type of group θ (see, e.g., Akerlof (1978); Boadway and Pestieau (2006); Cremer, Gahvari, and Lozachmeur (2010)). In that case, only the within-group incentive constraints (11) have to be taken into account. The optimal tax schedule can then be found by taking income $Y(\cdot, \theta)$ as the control variable, utility $U(\cdot, \theta)$ as the state variable and maximizing (20) subject to (12). The necessary conditions are:

$$\frac{1 - \dfrac{v'_y\left(Y(w,\theta); w, \theta\right)}{u'\left(C(w,\theta)\right)}}{-v''_{yw}\left(Y(w,\theta); w, \theta\right)} f(w|\theta) = \int_w^\infty \left(\frac{1}{u'(C(\omega,\theta))} - \frac{\Phi'_u(U(\omega,\theta); \omega, \theta)}{\lambda} \right) f(\omega|\theta) d\omega \tag{21a}$$

$$0 = \int_0^\infty \left(\frac{1}{u'(C(\omega,\theta))} - \frac{\Phi'_u(U(\omega,\theta); \omega, \theta)}{\lambda} \right) f(\omega|\theta) d\omega \tag{21b}$$

These necessary conditions can be using the usual Hamiltonian approach,[11] which is equivalent to considering the effects of an infinitesimal variation ΔY of the control variable $Y(.,\theta)$ over a small interval $[w-\delta_w, w]$ of the skill distribution. Because of incentive-compatibility, from (14), the levels of utility remain unchanged below the skill level $w-\delta_w$ and are changed by a uniform amount of $\Delta_U = -v''_{yw}(Y(w); w, \theta) \cdot \Delta_Y \cdot \delta_w$ for all skill levels above w (where the perturbation takes place). This small perturbation of the optimal allocation, which is illustrated in Figure 3, has two first-order effects on the Lagrangian (20). First, the direct change in

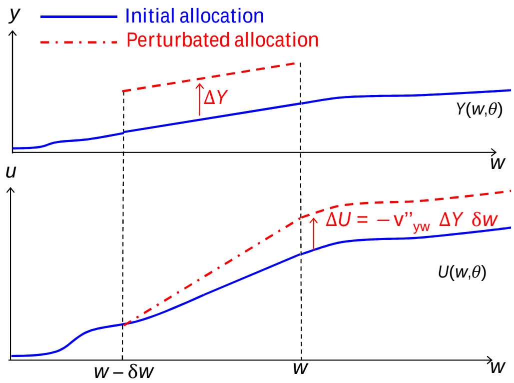

Figure 3: A perturbation of the optimal allocation

the control $Y(w,\theta)$ in $[w-\delta_w, w]$ affects (20) by:

$$\left(1 - \frac{v'_y\left(Y(w,\theta); w, \theta\right)}{u'\left(C(w,\theta)\right)}\right) \Delta_Y\, f(w|\theta)\delta_w = \frac{1 - \dfrac{v'_y\left(Y(w,\theta); w, \theta\right)}{u'\left(C(w,\theta)\right)}}{-v''_{yw}(Y(w); w, \theta)}\, f(w|\theta)\, \Delta_U$$

where the second equation in (19) has been used and where the right-hand side is obtained thanks to $\Delta_U = -v''_{yw}(Y(w); w, \theta) \cdot \Delta_Y \cdot \delta_w$. Second, the uniform change in utility levels $U(w,\theta)$ above w affects (20) by:

$$\int_w^\infty \left(\frac{\Phi'_u(U(\omega,\theta))}{\lambda} - \frac{1}{u'(C(\omega,\theta))}\right) f(\omega|\theta) d\omega\, \Delta_U$$

where (19) has again been used. The sum of these two terms must be equal zero if the initial allocation was optimal. Making Δ_Y and δ_w tending to 0 leads to (21a). Finally, (21b) is obtained by a uniform change in the utility levels. The latter affects (20) by the latest expression and is equal to zero since this kind of perturbation of the optimal allocation does not change the state variable according to (14).

V.2 The optimal allocation when θ is not observable

We can now extend the reasoning of the previous subsection to the case with unobserved heterogeneity in θ. The following proposition gives the necessary conditions in this context.

[11] See Appendix C.

Proposition 1. *Under Assumptions 1 and 2, the optimal allocation must verify:*

$$\int_{\theta\in\Theta}\left\{-\frac{1-\frac{v'_y\left(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}}{v''_{yw}\left(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)}f(W(w,\theta;\theta_0)|\theta)\right\}d\mu(\theta) \tag{22a}$$
$$= \int_{\theta\in\Theta}\left\{\int_{\omega\geq W(w,\theta;\theta_0)}\left(\frac{1}{u'\left(C(\omega,\theta)\right)}-\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda}\right)f(\omega|\theta)d\omega\right\}d\mu(\theta)$$

for all $w\in\mathbb{R}_+$ and

$$\int_{\theta\in\Theta}\left\{\int_{\omega\in\mathbb{R}_+}\left(\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda}-\frac{1}{u'\left(C(\omega,\theta)\right)}\right)f(\omega|\theta)d\omega\right\}d\mu(\theta)=0 \tag{22b}$$

Proof We derive Equation (22a) by considering the first-order effects of a small perturbation in the optimal allocation. Firstly, we study the impact this perturbation has on each group. Secondly, we aggregate these effects to compute the total impact on the Lagrangian (20). The perturbation we consider is triggered by an infinitesimal change $\Delta_Y(\theta_0)$ of income $Y(\cdot,\theta_0)$ among individuals in the reference group θ_0 endowed with a skill level in a small interval $[w-\delta_w(\theta_0),w]$.

First, for $\theta\neq\theta_0$, Equation (18) implies that this perturbation modifies the pooling function $W(.,\theta;\theta_0)$ only within the small interval $[w-\delta_w(\theta_0),w]$. Therefore, $Y(.,\theta)$ is modified only in the skill interval:

$$[W(w,\theta;\theta_0)-\delta_w(\theta),W(w,\theta;\theta_0)]\overset{\text{def}}{\equiv}[W(w-\delta_w(\theta_0),\theta),W(w,\theta;\theta_0)].$$

where we denote $\delta_w(\theta)$ the size of this interval. We moreover denote $\Delta_Y(\theta)$ the average (unweighted) change in Y within this interval.

Second, according to (14), because of this change in $Y(\cdot,\theta)$ within-group incentive compatibility implies that utility levels of individuals whose skills are above this interval must be modified uniformly by an amount of:

$$\Delta_U(\theta)=-v''_{yw}(Y(w,\theta);w,\theta)\cdot\Delta_Y(\theta)\cdot\delta_w(\theta)$$

Figure 3 illustrates the perturbation.

Third, as $U(w,\theta)=u(C(w,\theta))-v(Y(w,\theta);w,\theta)$ (see Equation (1)) and that $Y(w,\theta)$ is fixed for $\omega>W(w,\theta;\theta_0)$, the uniform change in $U(\omega,\theta)$ must be entirely due to a uniform change in $u(C(\omega,\theta))$ hence in $C(\omega,\theta)$ for all $\omega>W(w,\theta;\theta_0)$. Thanks to incentive compatibility and (15), we know that $C(\omega,\theta_0)$ and $C(W(\omega,\theta),\theta)$ are modified in a similar way for all $\omega>W(w,\theta;\theta_0)$. Consequently, the uniform change $\Delta_U(\theta)$ in utility that occurs in each group for skills levels ω above $W(w,\theta)$ must be equal and we have $\Delta_U(\theta_0)=\Delta_U(\theta)\equiv\Delta_U$. From the previous equation, we get for all θ that:

$$\Delta_U=-v''_{yw}(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta)\cdot\Delta_Y(\theta)\cdot\delta_w(\theta) \tag{23}$$

which determines $\Delta_Y(\theta)$ as a function of Δ_U.

Using infinitesimal calculus, we now determine the impact of our perturbation on the Lagrangian (20). For each θ, the change $\Delta_Y(\theta)$ of income over skills in $[W(w,\theta;\theta_0) - \delta_w(\theta), W(w,\theta;\theta_0)]$ has two effects on the Lagragian (20).

1. From the first equality in (19) and (20), the change $\Delta_Y(\theta)$ of income for individuals of group θ and skills in $[W(w,\theta;\theta_0) - \delta_w(\theta), W(w,\theta;\theta_0)]$ modifies tax revenue by:

$$\left(1 - \frac{v'_y\left(Y(W(w,\theta;\theta_0),\theta); W(w,\theta;\theta_0),\theta\right)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}\right) \cdot \Delta_Y(\theta) \cdot \delta_w(\theta) \cdot f(W(w,\theta;\theta_0)|\theta) \cdot \mu(\theta)$$
$$= -\frac{1 - \dfrac{v'_y\left(Y(W(w,\theta;\theta_0),\theta); W(w,\theta;\theta_0),\theta\right)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}}{v''_{yw}(Y(W(w,\theta;\theta_0),\theta); W(w,\theta;\theta_0),\theta)} \cdot \Delta_U \cdot f(W(w,\theta;\theta_0)|\theta) \cdot \mu(\theta)$$

   where the density $f(.|\theta)$ is evaluated at the skill level where the change of pre-tax income takes place (i.e., at $\omega = W(w,\theta;\theta_0)$) and where the right-hand side is obtained using (23).

2. The uniform increase Δ_U of utility for all $\omega > W(w,\theta;\theta_0)$ affects for each θ, the Lagragian by:

$$\int_{\omega \geq W(w,\theta;\theta_0)} \left(\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda} - \frac{1}{u'\left(C(\omega,\theta)\right)}\right) f(\omega|\theta) d\omega \cdot \Delta_U \cdot \mu(\theta)$$

   where the first equation in (19) has been used to obtain the second term in the brackets.

Integrating these effects for each group θ, the total effect on the Lagrangian is equal to:

$$\begin{aligned}
\Delta\mathscr{L} &= \int_{\theta\in\Theta} \left\{ -\frac{1 - \dfrac{v'_y\left(Y(W(w,\theta;\theta_0),\theta); W(w,\theta;\theta_0),\theta\right)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}}{v''_{yw}\left(Y(W(w,\theta;\theta_0),\theta); W(w,\theta;\theta_0),\theta\right)} f(W(w,\theta;\theta_0)|\theta) \right\} d\mu(\theta) \cdot \Delta_U \\
&+ \int_{\theta\in\Theta} \left\{ \int_{\omega \geq W(w,\theta;\theta_0)} \left(\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda} - \frac{1}{u'\left(C(\omega,\theta)\right)}\right) f(\omega|\theta) d\omega \right\} d\mu(\theta) \cdot \Delta_U
\end{aligned}$$

Such a modification should not have any first-order effect if the original allocation is optimal. This term must then be equal to zero at an optimum, which leads to (22a).

To derive (22b), we consider a perturbation of the optimal allocation that increases uniformly the utility by Δ_U for all agents. As the change in utility levels is uniform, $\dot{U}(w,\theta)$ are unchanged. Hence, according to (12), incomes $Y(w,\theta)$ are unchanged,[12] so the impact on the Lagrangian is simply:

$$\int_{\theta\in\Theta} \left\{ \int_{\omega\in\mathbb{R}_+} \left(\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda} - \frac{1}{u'\left(C(\omega,\theta)\right)}\right) f(\omega|\theta) d\omega \right\} d\mu(\theta) \cdot \Delta_U$$

that is the aggregation of the welfare variations across θ-groups. This impact has to be equal to zero at the optimum, which is (22b). $\square$

[12]Therefore, the perturbation only modifies $C(\cdot,\cdot)$ in such way that $u(C(w,\theta))$ is modified by the same amount of all individuals. Hence the perturbation preservers incentive-compatibility.

V.3 A case where optimal marginal tax rates are positive

This subsection emphasizes that allowing for heterogeneous labor supply elasticities at each income level, in the Mirrlees' model, is not sufficient to obtain negative marginal tax rates. Boadway, Marchand, Pestieau, and del Mar Racionero (2002) and Choné and Laroque (2010) introduce a second source of heterogeneity and show that this may lead to negative optimal marginal income tax rates. The additional heterogeneity induces that individuals who pool at the same income level $Y(w, \theta_0)$ are characterized by different social marginal utilities of consumption:

$$\Phi'_u(U(W(w,\theta)); w, \theta) \cdot u'(C(W(w,\theta;\theta_0)))$$

Therefore, although the concavity of the social welfare function implies that the social marginal utility of consumption is decreasing in skill *within each group* θ, the average social marginal utility of consumption may not be decreasing in income because of aggregation of these social marginal utilities across groups. This may happen when some groups undervalued in the social objective are overrepresented at low income levels. In this case, the heterogeneity reduces the welfare weights put on the low income types, relatively to those of the larger income types. This yields negative marginal tax rates at the bottom of the income distribution.

As previously stated, their paper however assumes an exogenous pooling function. One can therefore ask whether an endogenous pooling function may also invalidate the result of positive marginal tax rates. Consider a Benthamite social objective, so that the argument of Boadway, Marchand, Pestieau, and del Mar Racionero (2002) and Choné and Laroque (2010), (i.e. welfare weights increasing with income) for obtaining negative marginal tax rates does not apply. Proposition 2 states that endogenous pooling *alone* does not generate negative optimal marginal tax rate. In our general framework, only the argument of Boadway, Marchand, Pestieau, and del Mar Racionero (2002) and Choné and Laroque (2010) may thus explain why optimal marginal tax rates may become negative.

Proposition 2. *Under Benthamite social preferences,* $\Phi(U; w, \theta) \equiv U$ *and if* $u(\cdot)$ *is concave in consumption, optimal marginal tax rates are positive.*

Proof Let us define:

$$I(w) \stackrel{\text{def}}{\equiv} \int_{\omega \geq w} \left(\frac{1}{u'\left(C\left(\omega,\theta\right)\right)} - \frac{1}{\lambda} \right) \cdot \left(\int_{\theta \in \Theta} f\left(W(\omega,\theta;\theta_0) \middle| \theta\right) d\mu\left(\theta\right) \right) \cdot d\omega$$

Using the individual first-order condition (9) and (1), we can rewrite $1 - T'(Y(w, \theta)) = v'_y / u'$. Under Benthamite preferences, $\Phi'_u = 1$. So, Equations (22a) and (22b) can be respectively rewritten as:

$$\begin{aligned} T'(Y(w,\theta_0)) \cdot \int_{\theta \in \Theta} \frac{f\left(W\left(w,\theta\right) \middle| \theta\right)}{-v''_{yw}\left(Y\left(W\left(w,\theta\right),\theta\right); W\left(w,\theta\right),\theta\right)} d\mu(\theta) &= I(w) \\ 0 &= I(0) \end{aligned}$$

The derivative of $I(w)$ has the sign of $1/\lambda - 1/u'(C(\omega,\theta))$, which is decreasing in w because of the concavity of u. $I(w)$ needs thus to be first increasing and then decreasing. It is thus positive for all skill levels. As $v''_{yw} < 0$, optimal marginal tax rates must be positive, by the same argument as Mirrlees (1971). $\square$

VI An elasticity-based optimal tax formula

This section describes how the conditions describing the second-best optimum may lead to an optimal tax formula expressed in terms of behavioral responses. This is particularly important to grasp the economic intuition behind the second-best optimum, to give tax policy recommendations. We first define a set of useful elasticities and welfare weights. Then we reinterpret Proposition 1 in these terms.

VI.1 Defining elasticities and welfare weights

We define a compensated tax reform around earnings $Y(w,\theta)$ as a reform that changes the marginal tax by a constant amount τ around $Y(w,\theta)$, while leaving unchanged the level of tax at $Y(w,\theta)$. The tax function is changed to $T(Y) - \tau(Y - Y(w,\theta))$. The income response effect is defined as the response of the labor supply to a small lump-sum change ρ in tax liability, so the tax function becomes $T(Y) - \tau(Y - Y(w,\theta)) - \rho$. The (w,θ)-individual then solves:

$$\max_{y} \quad u(y - T(y) + \tau(y - Y(w,\theta)) + \rho)) - v(y;w,\theta) \tag{24}$$

We compute various elasticities of the solution y to this program when τ, ρ or w are marginally changed.[13] The compensated elasticity[14] $\varepsilon(w,\theta)$ of labor supply at (w,θ) is the elasticity of income y to a uniform increase τ of one minus the marginal tax $1 - T'(Y(w,\theta))$. The income response $\eta(w,\theta)$ is the derivative of income y to a lump-sum transfer ρ. We denote $\alpha(w,\theta)$ the elasticity of earnings to skill levels w. Formally, we have:

$$\varepsilon(w,\theta) \stackrel{\text{def}}{\equiv} \frac{1 - T'\left(Y(w,\theta)\right)}{Y(w,\theta)} \frac{\partial y}{\partial \tau} \qquad \eta(w,\theta) \stackrel{\text{def}}{\equiv} \frac{\partial y}{\partial \rho} \qquad \alpha(w,\theta) \stackrel{\text{def}}{\equiv} \frac{w}{Y(w,\theta)} \dot{Y}(w,\theta) \tag{25}$$

These derivatives are computed by applying the implicit function theorem to the first-order condition associated to the individual's program (24), which is $\mathscr{Y}(Y(w,\theta),0,0;w,\theta) = 0$ where:

$$\begin{aligned}\mathscr{Y}(y,\tau,\rho;w,\theta) \quad \stackrel{\text{def}}{\equiv} \quad & \left(1 - T'\left(Y(w,\theta)\right) + \tau\right) \cdot u'_c\left(y - T(y) + \tau\left(y - Y(w,\theta)\right) + \rho, y; w,\theta\right) \\ & - v'_y\left(y - T(y) + \tau\left(y - Y(w,\theta)\right) + \rho, y; w,\theta\right)\end{aligned}$$

[13] As the income tax is nonlinear, such marginal changes may induce discrete jumps in the income y chosen. This may occur when individuals of type (w,θ) are initially indifferent between two global maxima. However, because of Assumption 1, this type of indifference is associated with a discontinuity of the mapping $w \mapsto Y(w,\theta)$. As Assumption 2 rules out such discontinuities, we can safely ignore them.

[14] This elasticity is compensated in the sense that the tax level is unchanged at earnings level $Y(w,\theta)$. This is an Hicksian elasticity.

We get:

$$\begin{aligned}
\mathscr{Y}'_y(y,0,0;w,\theta) &= -T''\, u'_c + (1-T')^2 u''_{cc} + 2(1-T')u''_{cy} - v''_{yy} && \text{(26a)}\\
\mathscr{Y}'_\tau(y,0,0;w,\theta) &= u'_c + (y - Y(w,\theta)) \cdot \Big((1-T')u''_{cc} + u''_{cy}\Big) = u'_c && \text{(26b)}\\
\mathscr{Y}'_\rho(y,0,0;w,\theta) &= (1-T')u''_{cc} + u''_{cy} = \frac{u''_{cy}u'_c + u''_{cc}v'_y}{u'_c} && \text{(26c)}\\
\mathscr{Y}'_w(y,0,0;w,\theta) &= (1-T')u''_{cw} + u''_{yw} = \frac{v''_{yw}u'_c + u''_{cw}v'_y}{u'_c} && \text{(26d)}
\end{aligned}$$

where we use the first-order condition (9). The second-order condition (of the individual maximization program) is $\mathscr{Y}'_y \leq 0$. As soon as the second-order condition strictly holds for all types of individuals, the implicit function theorem applies and income is a differentiable function of skill within each group. Hence, bunching does not take place and discontinuities in the function are ruled out. Throughout this section, we thus assume the following regularity conditions:

Assumption 3. *The tax function $T(\cdot)$ is twice differentiable and, for all $(w,\theta) \in \mathbb{R}_+ \times \Theta$, the second-order condition holds strictly: $\mathscr{Y}'_y\,(Y(w,\theta),0,0;w,\theta) < 0$.*

Assumption 2 is automatically satisfied whenever Assumptions 1 and 3 are imposed.[15] Assumption 3 is thus a further restriction on the regularity of the optimal allocation compared to Assumption 2. Applying the implicit function theorem at $y = Y(w,\theta), \tau = 0, \rho = 0; w, \theta$, we get $\partial y/\partial x = -\mathscr{Y}'_x/\mathscr{Y}'_y$ for $x = \tau, \rho, w$. Hence:

$$\varepsilon(w,\theta) = -\frac{v'_y}{Y(w,\theta)\cdot \mathscr{Y}'_y} \qquad \eta(w,\theta) = -\frac{u''\cdot v'_y}{u'\cdot \mathscr{Y}'_y} \qquad \alpha(w,\theta) = \frac{w\, v''_{yw}}{Y(w,\theta)\, \mathscr{Y}'_y} \tag{27a}$$

As $v'_y > 0$ and the second-order condition is $\mathscr{Y}'_y < 0$, one gets $\varepsilon(w,\theta) > 0$. The sign of $\eta(w,\theta)$ is negative as the additively separable specification of preference in (1) implies that leisure is a normal good. The single-crossing assumption 1 moreover ensures that $\alpha(w,\theta) > 0$. We in particular get:

$$\frac{\varepsilon(w,\theta)}{\alpha(w,\theta)} = -\frac{v'_y\,(Y(w,\theta);w,\theta)}{w\cdot v''_{yw}\,(Y(w,\theta);w,\theta)} \tag{27b}$$

Hence, the ratio between the compensated and the skill elasticities does not depend on the curvature of the tax function. Moreover, we get:

$$\eta(w,\theta) = Y(w,\theta)\cdot \frac{u''(C(w,\theta))}{u'(C(w,\theta))}\cdot \varepsilon(w,\theta) \tag{27c}$$

so the ratio between income effects and the compensated elasticities is the same among individuals earning the same income level.

These elasticities differ from those in the tax literature (see e.g.Saez (2001)) by the presence of T'' in $\mathscr{Y}'_y < 0$ (see Equation (26a)) which appears in the denominators of $\varepsilon(w,\theta)$, $\eta(w,\theta)$

[15]Under Assumption 3, the implicit function theorem applies. It in particular induces that within each group, income is a differentiable function of skill with an elasticity $\alpha(w,\theta)$ given in Equation (27a). Assumption 1 implies the latter to be strictly positive.

and $\alpha(w,\theta)$, and which accounts for the nonlinearity of the income tax schedule. An exogenous change in either w, τ, or ρ induces a direct change in earnings $\Delta_1 Y(w,\theta)$. However, this change in turn modifies the marginal tax rate by $\Delta_1 T' = T''(Y(w,\theta)) \times \Delta_1 Y(w,\theta)$, thereby inducing a further change in earnings $\Delta_2 Y(w,\theta)$. Therefore, a "circular process" (Saez (2001), Saez (2003)) takes place: The income level determines the marginal tax rate through the tax function, and the marginal tax rate affects the income level through the substitution effect. The term $T''(Y(w,\theta)) \cdot u'_c$ appears in the denominators of our elasticities and income responses $\varepsilon(w,\theta)$, $\alpha(w,\theta)$, $\eta(w,\theta)$ (through $\mathscr{Y}'_y$ to take this circular process into account.

Let $h(\cdot|\theta)$ denote the conditional density of income y within the group θ and $H(\cdot|\theta)$ be the associated conditional cumulative distribution function. One has from Lemma 1 that $H(Y(w,\theta)|\theta) = F(w|\theta)$. Differentiating both sides in skill w and using (25), we get:

$$h(Y(w,\theta)|\theta) = \frac{f(w|\theta)}{\dot{Y}(w,\theta)} \qquad \Leftrightarrow \qquad Y(w,\theta)\, h(Y(w,\theta)|\theta) = \frac{w\, f(w|\theta)}{\alpha(w,\theta)} \tag{28}$$

The unconditional income density at income $y = Y(w,\theta_0)$, that is the mass of individuals (endowed with distinct w and θ) who earn the same income $Y(w,\theta_0)$, is:

$$\hat{h}(Y(w,\theta_0)) \stackrel{\text{def}}{\equiv} \int_{\theta\in\Theta} h\left(W(w,\theta;\theta_0)|\theta\right)\, d\mu(\theta) \tag{29}$$

We can now define the mean compensated elasticity at income level $Y(w,\theta_0)$ as:

$$\hat{\varepsilon}(Y(w,\theta_0)) \stackrel{\text{def}}{\equiv} \frac{\int_{\theta\in\Theta} \varepsilon\left(W(w,\theta;\theta_0)|\theta\right)\; h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)}{\int_{\theta\in\Theta} h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)} \tag{30}$$

We similarly define the mean income effect at income level $Y(w,\theta_0)$ as:

$$\hat{\eta}(Y(w,\theta_0)) \stackrel{\text{def}}{\equiv} \frac{\int_{\theta\in\Theta} \eta\left(W(w,\theta;\theta_0)|\theta\right)\; h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)}{\int_{\theta\in\Theta} h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)} \tag{31}$$

From Equation (27c), we get:

$$\hat{\eta}(Y(w,\theta_0)) = Y(w,\theta_0) \cdot \frac{u''\left(C(w,\theta_0)\right)}{u'\left(C(w,\theta_0)\right)} \cdot \hat{\varepsilon}(Y(w,\theta_0)) \tag{32}$$

We define $g(w,\theta)$ the endogenous marginal social weight associated with workers of type (w,θ), expressed in terms of public funds by:

$$g(w,\theta) \stackrel{\text{def}}{\equiv} \frac{\Phi'_u\left(U(w,\theta);w,\theta\right) \cdot \mathscr{U}'_c\left(C(w,\theta),Y(w,\theta);w,\theta\right)}{\lambda} \tag{33}$$

The government values giving one extra dollar to a worker (w,θ) as a gain of $g(w,\theta)$ in terms of public funds. The mean marginal social weight at income $Y(w,\theta_0)$ is thus:

$$\hat{g}(Y(w,\theta_0)) \stackrel{\text{def}}{\equiv} \frac{\int_{\theta\in\Theta} g\left(W(w,\theta;\theta_0)|\theta\right)\; h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)}{\int_{\theta\in\Theta} h\left(W(w,\theta;\theta_0)|\theta\right)\; d\mu(\theta)} \tag{34}$$

It gives the average of the marginal social weights of people endowed with distinct w and θ but earning the same income level $Y(w,\theta)$.

VI.2 Obtaining an elasticity-based optimal tax formula

Using the definitions of the above elasticities, income effects, income densities and welfare weights, we now derive the optimal tax formula in terms of empirically meaningful variables.

Proposition 3. *Under assumptions 1 and 2, the optimal tax schedule has to satisfy:*

$$\frac{T'(y)}{1-T'(y)} = \frac{1}{\hat{\varepsilon}(y)} \cdot \frac{1-\hat{H}(y)}{y\,\hat{h}(y)} \cdot \left(1 - \frac{\int_y^\infty \left[\hat{g}(z) + \hat{\eta}(z) \cdot T'(y(z))\right] \cdot \hat{h}(z)dz}{1-\hat{H}(y)}\right) \quad \forall y \in \mathbb{R}_+ \tag{35a}$$

$$1 = \int_0^\infty \left[\hat{g}(z) + \hat{\eta}(z) \cdot T'(Y(z))\right] \cdot \hat{h}(z)dz \tag{35b}$$

The proof is provided in Appendix D and consists in reinterpreting the first-order conditions that were derived in Proposition 1. Equation (35a) generalizes the elasticity-based optimal tax formula of Saez (2001) to the case where θ is not unique. It differs from the latter in three important aspects. First, our tax formula depends on the actual income density and not on the virtual density. Second, the elasticities $\hat{\varepsilon}(\cdot)$ and $\hat{\eta}(\cdot)$ are weighted averages of elasticities $\varepsilon(\cdot,\cdot)$ and $\eta(\cdot,\cdot)$ that take into account the nonlinearity of the optimal tax schedule through the term $T''(\cdot)$ in $\mathscr{Y}'_y$ and in Equation (27a). However our compensated labor elasticity times the actual density is equal to the traditional compensated elasticity times the virtual income density of Saez (2001). Last, we show that our tax formula applies in a context with multidimensional unobserved heterogeneity.

According to Equation (35a), there are three different arguments to depart from a linear income tax schedule. The first is that the mean compensated elasticity of labor supply $\hat{\varepsilon}(y)$ may vary with income. In line with the Ramsey (1927) inverted elasticity rule, optimal marginal tax rates are *ceteris paribus* larger at income levels where the compensated elasticity $\hat{\varepsilon}(y)$ is lower. In our context, this elasticity is endogenous for different reasons. First, it depends on the curvature of the income tax function, as it becomes apparent with the term $-T''\ u'_c$ in $\mathscr{Y}'_y$ (see (26a)), hence in the expression of the elasticity $\varepsilon(\cdot,\cdot)$ (see (27a)). This endogeneity already prevails in the Mirrlees framework where θ is homogeneous, see Jacquet, Lehmann, and Van der Linden (2013). Moreover, the pool of types (w,θ) that bunch at any income level is endogenous. Therefore, the mean elasticity $\hat{\varepsilon}(y)$ increases when among the types (w,θ) such that $Y(w,\theta) = y$, the proportion of individual with an elasticity $\varepsilon(w,\theta)$ larger than the average $\hat{\varepsilon}(y)$ increases. This *composition effect* is a key novelty of the present setting, compared to the Mirrleesian model.

The second reason to depart from a flat tax is that, at a given level of income, the optimal marginal tax rate increases (in absolute terms) with the *distribution term* $(1-\hat{H}(y))/(y\hat{h}(y))$ of the income distribution at income level y, the second term in the right hand-side of (35a). This is because substitution effects at income level y are larger the bigger is the income density $\hat{h}(y)$ and the larger is the income y, while the tax level effects (i.e., the change in the level of tax collected) are larger the larger is the mass $1-\hat{H}(y)$ of individuals with income above y. The distribution term is very likely to vary with income, unless incomes are Pareto distributed everywhere. Diamond (1998) and Saez (2001) argue that the distribution term is empirically constant in the upper part of the income distribution but is not everywhere constant. Moreover, the income distribution is typically unimodal. So the distribution term is decreasing beyond this mode, implying lower marginal tax rates in absolute values. Finally, it is worth stressing that the distribution term of the income distribution is endogenous, for two main reasons. The first is the endogeneity of the elasticity $\alpha(w,\theta)$ of income with respect to the level of skill.[16] Even when θ does not vary, the nonlinearity of the income tax schedule implies that this elasticity depends on the curvature of the tax schedule (see (27a)), and is thus endogenous. Second, the pool of individuals (w,θ) who pool at any income level is endogenous, so the distribution term may also be affected by a composition effect, as emphasized by Cremer, Gahvari, and Lozachmeur (2010).

Third, as emphasized by the third term of the right hand-side of (35a), optimal marginal tax rates varies with the mean of social welfare weights $\hat{g}(z)$ and the mean of income effects $\hat{\eta}(z) \cdot T'(z)$ for income levels z above y. The larger $\hat{g}(z)$, the more the government values the well-being of agents at this level of income z hence, the lower should be the marginal tax rate below this level. The larger $(\hat{\eta}(z)(< 0))$ in absolute value, the higher should be the marginal tax rate below income z, because an increase in the level of tax paid by workers with income z induces them to work more through the income effect. Our formulation for the contribution of the income effect differs from Saez (2001) who defines the income effect response along a linear tax schedule. Saez therefore needs to use the virtual earnings density instead of the actual one to correct the income effect term for the circularity process. Conversely, we define η taking this circular process into account in (27a), which simplifies our formula of the optimal marginal tax rates. Again, the means of social welfare weights $\hat{g}(z)$ and of income effects $\hat{\eta}(z) \cdot T'(z)$ typically vary with earnings. An exception is when the government is Maximin (Rawlsian) and the utility function is quasilinear (i.e. $u(\cdot)$ is linear) so that $\hat{g}(z) = 0$ and $\hat{\eta}(z) \cdot T'(z) = 0$. Again, composition effects are new sources of endogeneity for $\hat{g}(\cdot)$ and $\hat{\eta}(\cdot)$, compared to the framework of Mirrlees. The composition effects on social welfare weights are central in the analyses of Boadway, Marchand, Pestieau, and del Mar Racionero (2002) and Choné and Laroque (2010).

A Proof of Lemma 2

Following, e.g. Salanié (2005), from the taxation principle, individuals choose the type w', θ' that they want to mimic, i.e. they solve

$$\max_{w',\theta'} \quad \mathscr{U}\left(C(w',\theta'), Y(w',\theta'); w, \theta\right)$$

[16] From (28), the distribution terms of the conditional skill (exogenous) and of the income (endogenous) distribution are linked through:

$$\frac{1 - H(Y(w,\theta)|\theta)}{Y(w,\theta) \cdot h(Y(w,\theta)|\theta)} = \alpha(w,\theta) \cdot \frac{1 - F(w|\theta)}{w \cdot f(w|\theta)}$$

Function $(w',\theta') \mapsto \mathscr{U}\left(C(w',\theta'),Y(w',\theta');w,\theta\right)$ admits a partial derivative with respect to w' that is equal to:

$$\dot{C}(w',\theta')\ \mathscr{U}'_c\left(C(w',\theta'),Y(w',\theta');w,\theta\right) + \dot{Y}(w',\theta')\ \mathscr{U}'_y\left(C(w',\theta'),Y(w',\theta');w,\theta\right)$$

The first-order condition implies that this expression must be nil at $(w',\theta') = (w,\theta)$. Using (3) leads to (13). Differentiating in w both sides of $U(w,\theta) = \mathscr{U}\left(C(w,\theta),Y(w,\theta);w,\theta\right)$ leads to:

$$\begin{aligned}
\dot{U}(w,\theta) &= \mathscr{U}'_c\left(C(w,\theta),Y(w,\theta);w,\theta\right)\ \dot{C}(w,\theta) + \mathscr{U}'_y\left(C(w,\theta),Y(w,\theta);w,\theta\right)\ \dot{Y}(w,\theta) \\
&+ \mathscr{U}'_w\left(C(w,\theta),Y(w,\theta);w,\theta\right) \\
&= \left(\frac{\dot{C}(w,\theta)}{\dot{Y}(w,\theta)} - \mathscr{M}\left(C(w,\theta),Y(w,\theta);w,\theta\right)\right) \mathscr{U}'_c\left(C(w,\theta),Y(w,\theta);w,\theta\right)\ \dot{Y}(w,\theta) \\
&+ \mathscr{U}'_w\left(C(w,\theta),Y(w,\theta);w,\theta\right)
\end{aligned}$$

Hence, (12) holds if and only if (13) holds.

B Proof of Lemma 4

We first show that there exists at most one allocation $(w,\theta) \mapsto (\underline{C}(w,\theta),\underline{Y}(w,\theta))$ that verifies Assumption 2 and such that $(\underline{C}(w,\theta_0),\underline{Y}(w,\theta_0)) = C(w,\theta_0),Y(w,\theta_0))$. We then show that this allocation verifies the whole set of incentive constraints (10).

To build the entire incentive compatible allocation $(w,\theta) \mapsto (\underline{C}(w,\theta),\underline{Y}(w,\theta))$, we must obviously choose $(\underline{C}(w,\theta_0),\underline{Y}(w,\theta_0)) = C(w,\theta_0),Y(w,\theta_0))$ for any skill level in the reference group.

For each group θ, $\underline{Y}(\cdot;\theta)$ verifies Assumption 2 if and only if its reciprocal $\underline{Y}^{-1}(\cdot;\theta)$ is differentiable with a strictly positive derivative and maps $\mathbb{R}_+$ into $\mathbb{R}_+$. Let then $y \in \mathbb{R}_+$ be an income level. As $Y(\cdot,\theta_0)$ satisfies Assumption 2, there exists a unique skill level w such that $y = Y(w,\theta_0)$. Then according to Lemma 3, among individuals of group θ, only thos of skill $\underline{W}(w,\theta)$ are assigned to the income level $y = Y(w,\theta_0)$.[17] Therefore, $\underline{Y}^{-1}(\cdot;\theta)$ must be defined by:

$$\underline{Y}^{-1}(\cdot;\theta): \qquad y \overset{Y^{-1}(\cdot,\theta_0)}{\longmapsto} w = Y^{-1}(y,\theta_0) \overset{\underline{W}(\cdot,\theta)}{\longmapsto} Y^{-1}(y,\theta)$$

Hence, $\underline{Y}^{-1}(\cdot,\theta)$ is differentiable and is defined over $\mathbb{R}_+$. It admits a positive derivative everywhere and takes value on the whole $\mathbb{R}_+$ if and only if $\underline{W}(\cdot,\theta)$ does. Hence, $\underline{Y}(\cdot,\theta)$ is a differentiable increasing function with positive derivatives that maps $\mathbb{R}_+$ onto $\mathbb{R}_+$.

We now show that $\underline{C}(w,\theta)$ is also uniquely determined for any skill level ω and group θ. This is because we now from above that for each type (ω,θ), there exists a single skill level such that $\underline{Y}(\omega,\theta) = Y(w,\theta_0)$. Incentive compatibility then requires that $\underline{C}(\omega,\theta)$ needs also to be equal to $\underline{C}(w,\theta_0)$.

This ends the proof that given an incentive-compatible allocation $w \mapsto (C(w,\theta_0),Y(w,\theta_0))$ defined with the reference group that verifies Assumption 2, there exists *at most* a unique allocation $(w,\theta) \mapsto (\underline{Y}(w,\theta),\underline{C}(w,\theta))$ that can be incentive-compatible. We now verify that this allocation does verify the incentive constraint (10).

We first notice that the allocation $(w,\theta) \mapsto (\underline{Y}(w,\theta),\underline{C}(w,\theta))$ is built in such a way that one has:

$$\underline{Y}(\omega,\theta) = Y(w,\theta_0) \qquad \text{and} \qquad \underline{C}(\omega,\theta) = C(w,\theta_0)$$

if and only if $\omega = \underline{W}(w,\theta)$ and (17) holds. Differentiating in w both sides of the two equations:

$$\underline{Y}(\underline{W}(w,\theta),\theta) = Y(w,\theta_0) \qquad \text{and} \qquad \underline{C}(\underline{W}(w,\theta),\theta) = C(w,\theta_0)$$

[17] Hence function $\underline{W}(\cdot,\theta)$ coincides with the pooling function $W(\cdot,\theta;\theta_0)$.

and rearranging terms leads to:

$$\frac{\dot{\underline{C}}\left(\underline{W}(w,\theta),\theta_0\right)}{\dot{\underline{Y}}\left(w,\theta_0\right)} = \frac{\dot{\underline{C}}\left(\underline{W}(w,\theta),\theta_0\right)}{\dot{\underline{Y}}\left(w,\theta_0\right)}$$

As $w \mapsto (C(w,\theta_0), Y(w,\theta_0))$ is assumed to verify the within-group incentive-compatible Equation (13), the left-hand side of the latter equation is equal to $\mathscr{M}(C(w,\theta_0), Y(w,\theta_0); w, \theta_0)$. Given the definition of function $\underline{W}(\cdot,\theta)$, we have that $w \mapsto (\underline{C}(w,\theta), \underline{Y}(w,\theta))$ also verifies Equation (13). From Lemma 2, it thus verifies within-group incentive constraints (11).

We finally verify whether the inequality (10) is verified for any $(w, w', \theta, \theta') \in \mathbb{R}_+^2 \times \Theta^2$. We now there exists $\omega \in \mathbb{R}_+$ such that

$$\underline{Y}(\omega,\theta) = \underline{Y}(w',\theta') \qquad \text{and} \qquad \underline{C}(\omega,\theta) = \underline{C}(w',\theta')$$

Hence (10) is equivalent to:

$$\mathscr{U}\left(C(w,\theta), Y(w,\theta); w, \theta\right) \geq \mathscr{U}\left(C(\omega,\theta), Y(\omega,\theta); w, \theta\right)$$

and is verified as $w \mapsto (\underline{C}(w,\theta), \underline{Y}(w,\theta))$ also verifies Equation (13), and thus from Lemma 2, it verifies the within-group incentive constraints (11).

C Equations (21a) and (21b)

Define the Hamiltonian for each θ as

$$\left(Y(w,\theta) - \mathscr{C}\left(Y(w,\theta), U(w,\theta); w, \theta\right) + \frac{\Phi\left(U(w,\theta); w, \theta\right)}{\lambda}\right) \cdot f(w|\theta) - q(w|\theta) \cdot v'_w\left(Y(w,\theta); w, \theta\right)$$

The necessary conditions are, using (19):

$$\begin{aligned}
0 &= \left(1 - \frac{v'_y(Y(w,\theta); w,\theta)}{u'(C(w,\theta))}\right) \cdot f(w|\theta) - q(w|\theta) \cdot v''_{yw}\left(Y(w,\theta); w, \theta\right) && \text{(36a)}\\
-\dot{q}\left(w|\theta\right) &= \left(\frac{\Phi'_U\left(U(w,\theta); w,\theta\right)}{\lambda} - \frac{1}{u'(C(w,\theta))}\right) \cdot f(w|\theta) && \text{(36b)}\\
0 &= q(0|\theta) && \text{(36c)}\\
0 &= \lim_{w \mapsto \infty} q(w|\theta) && \text{(36d)}
\end{aligned}$$

Combining (36b) with (36d) leads to

$$q(w|\theta) = \int_w^\infty \left(\frac{\Phi'_U\left(U(\omega,\theta); \omega,\theta\right)}{\lambda} - \frac{1}{u'(C(\omega,\theta))}\right) \cdot f(\omega|\theta) d\omega \tag{36e}$$

Combining (36a) with (36e) leads to (21a). Combining (36c) with (36e) leads to (21b).

D Proof of Proposition 3

Let $y \in \mathbb{R}_+$. According to Assumption 2, there exists a single skill level w such that $y = Y(w,\theta_0)$. From (9), we know that the first-order condition of the individual maximization program can be written as:

$$1 - T'(Y(w,\theta)) = \frac{v'_y(Y(w,\theta); w,\theta)}{u'(C(w,\theta))} \tag{37}$$

The term between brackets in the right-hand side of (22a) can then be rewritten as:

$$
\begin{aligned}
& -\frac{1-\frac{v'_y\left(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}}{v''_{yw}\left(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)}\cdot f(W(w,\theta;\theta_0)|\theta) \\
= \quad & T'\left(Y(w,\theta_0)\right)\cdot\frac{\varepsilon\left(W(w,\theta;\theta_0),\theta\right)}{\alpha\left(W(w,\theta;\theta_0),\theta\right)}\cdot\frac{W(w,\theta;\theta_0)}{v'_y\left(Y(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)}\cdot f(W(w,\theta;\theta_0)|\theta) \\
= \quad & \frac{T'\left(Y(w,\theta_0)\right)}{1-T'\left(Y(w,\theta_0)\right)}\cdot\frac{\varepsilon\left(W(w,\theta;\theta_0),\theta\right)}{\alpha\left(W(w,\theta;\theta_0),\theta\right)}\cdot\frac{W(w,\theta;\theta_0)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}\cdot f(W(w,\theta;\theta_0)|\theta) \\
= \quad & \frac{T'\left(Y(w,\theta_0)\right)}{1-T'\left(Y(w,\theta_0)\right)}\cdot\frac{Y(w,\theta_0)}{u'\left(C(W(w,\theta;\theta_0),\theta)\right)}\cdot\varepsilon\left(W(w,\theta;\theta_0),\theta\right)\cdot h(Y(w,\theta_0)|\theta)
\end{aligned}
$$

The first equality is obtained using Equations (27b), (37) and the fact that $Y(W(w,\theta;\theta_0),\theta) = Y(w,\theta_0)$ from the definition (15) of $W(\cdot,\theta;\theta_0)$. The second equality uses (37) again. The third equality follows (28). It implies with (30) that Equation (22a) can be rewritten as:

$$
\frac{T'\left(Y(w,\theta_0)\right)}{1-T'\left(Y(w,\theta_0)\right)}\cdot\hat{\varepsilon}\left(Y(w,\theta_0)\right)\cdot Y(w,\theta_0)\cdot\hat{h}(Y(w,\theta_0)) = J(w) \tag{38}
$$

where

$$
J(w) \stackrel{\text{def}}{\equiv} u'\left(C(w,\theta_0)\right)\cdot\int_{\theta\in\Theta}\left\{\int_{\omega\geq W(w,\theta;\theta_0)}\left(\frac{1}{u'\left(C(\omega,\theta)\right)}-\frac{\Phi'_u(U(\omega,\theta);\omega,\theta)}{\lambda}\right)f(\omega|\theta)d\omega\right\}d\mu(\theta) \tag{39}
$$

$J(\cdot)$ admits for derivative $\dot{J}(w)$ where:

$$
\begin{aligned}
& \dot{J}(w) = \dot{C}(w,\theta_0)\cdot\frac{u''\left(C(w,\theta_0)\right)}{u'\left(C(w,\theta_0)\right)}\cdot J(w) \\
& \int_{\theta\in\Theta}\left\{\frac{\Phi'_u\left(U(W(w,\theta;\theta_0),\theta);W(w,\theta;\theta_0),\theta\right)\ u'\left(C(W(w,\theta;\theta_0),\theta)\right)}{\lambda}-1\right\}\dot{W}(w,\theta)\ f\left(W(w,\theta;\theta_0)|\theta\right)d\mu(\theta) \\
& = \int_{\theta\in\Theta}\left\{g\left(W(w,\theta;\theta_0),\theta\right)-1\right\}\cdot\dot{W}(w,\theta)\cdot f\left(W(w,\theta;\theta_0)|\theta\right)\cdot d\mu(\theta)+\dot{C}(w,\theta_0)\cdot\frac{u''\left(C(w,\theta_0)\right)}{u'\left(C(w,\theta_0)\right)}\cdot J(w)
\end{aligned}
$$

where (33) has been used. Deriving with respect to the skill w both sides of (15) and of $C(w,\theta) = Y(w,\theta) - T\left(Y(w,\theta)\right)$, we get that:

$$
\dot{W}(w,\theta) = \frac{\dot{Y}\left(w,\theta_0\right)}{\dot{Y}\left(W(w,\theta;\theta_0),\theta\right)} \qquad \text{and} \qquad \dot{C}(w,\theta) = \left(1-T'\left(Y(w,\theta)\right)\right)\cdot\dot{Y}(w,\theta)
$$

We thus obtain:

$$
\begin{aligned}
\dot{J}(w) \quad = \quad & \left(\int_{\theta\in\Theta}\left\{g\left(W(w,\theta;\theta_0),\theta\right)-1\right\}\cdot\frac{f\left(W(w,\theta;\theta_0)|\theta\right)}{\dot{Y}(W(w,\theta;\theta_0),\theta)}\cdot d\mu(\theta)\right. \\
+ \quad & \left.\left(1-T'\left(Y(w,\theta_0)\right)\right)\cdot\frac{u''\left(C(w,\theta_0)\right)}{u'\left(C(w,\theta_0)\right)}\cdot J(w)\right)\cdot\dot{Y}(w,\theta_0)
\end{aligned}
$$

Using (28) and (38), $\dot{J}(w)$ can be rewritten as:

$$
\begin{aligned}
\dot{J}(w) \quad = \quad & \left(\int_{\theta\in\Theta}\left\{g\left(W(w,\theta;\theta_0),\theta\right)-1\right\}\cdot h\left(Y(w,\theta_0)|\theta\right)\ d\mu(\theta)\right. \\
+ \quad & \left. T'\left(Y(w,\theta_0)\right)\cdot Y(w,\theta_0)\cdot\frac{u''\left(C(w,\theta_0)\right)}{u'\left(C(w,\theta_0)\right)}\cdot\hat{\varepsilon}(Y(w,\theta_0))\cdot\hat{h}(Y(w,\theta_0))\right)\cdot\dot{Y}(w,\theta_0)
\end{aligned}
$$

Using (32) and (34), we get:

$$-\dot{J}(w) = \left\{1 - \hat{g}(Y(w,\theta_0)) - \hat{\eta}(Y(w,\theta_0)) \cdot T'\left(Y(w,\theta_0)\right)\right\} \cdot \hat{h}\left(Y(w,\theta)\right) \cdot \dot{Y}(w,\theta_0)$$

As $J(w) = \int_{x \geq w} (-\dot{J}(x))dx$, we get

$$J(w) = \int_{x \geq w} \left\{1 - \hat{g}(Y(x,\theta_0)) - \hat{\eta}(Y(x,\theta_0)) \cdot T'\left(Y(x,\theta_0)\right)\right\} \cdot \hat{h}\left(Y(x,\theta)\right) \cdot \dot{Y}(x,\theta_0) \cdot dx$$

Changing variables by posing $z = Y(x,\theta_0)$, we get

$$J(w) = \int_{z \geq Y(w,\theta_0)} \left\{1 - \hat{g}(z) - \hat{\eta}(z) \cdot T'\left(Y(z)\right)\right\} \cdot \hat{h}\left(Y(x,\theta)\right) \cdot dz \tag{40}$$

Plugging (40) into (38) and rearranging terms gives (35a). Combining (22b), (39) and (40) leads to (35b).

References

AKERLOF, G. A. (1978): "The Economics of "Tagging" as Applied to the Optimal Income Tax, Welfare Programs, and Manpower Planning," *The American Economic Review*, 68(1), 8–19.

ALESINA, A., A. ICHINO, AND L. KARABARBOUNIS (2011): "Gender-Based Taxation and the Division of Family Chores," *American Economic Journal: Economic Policy*, 3(2), 1–40.

BLUMKIN, T., E. SADKA, AND Y. SHEM-TOV (2012): "International Tax Competition: Zero Tax Rate at the Top Re-established," Discussion Paper 3820, CESifo WP.

BOADWAY, R., M. MARCHAND, P. PESTIEAU, AND M. DEL MAR RACIONERO (2002): "Optimal Redistribution with Heterogeneous Preferences for Leisure," *Journal of Public Economic Theory*, 4(4), 475–498.

BOADWAY, R., AND P. PESTIEAU (2006): "Tagging and redistributive taxation," *Annals of Economics and Statistics*, (83-84), 123–147.

BRETT, C., AND J. WEYMARK (2003): "Financing Education Using Optimal Redistributive Taxation," 87, 2549–2569.

CHONÉ, P., AND G. LAROQUE (2010): "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 100(5), 2532–47.

CREMER, H., F. GAHVARI, AND J.-M. LOZACHMEUR (2010): "Tagging and Income Taxation: Theory and an Application," *American Economic Journal: Economic Policy*, 2(1), 31–50.

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

GUESNERIE, R. (1995): *A Contribution to the Pure Theory of Taxation*. Cambridge University Press, Cambridge.

HAMMOND, P. J. (1979): "Straightforward Individual Incentive Compatibility in Large Economies," *The Review of Economic Studies*, 46(2), 263–282.

JACQUET, L., E. LEHMANN, AND B. VAN DER LINDEN (2013): "Optimal redistributive taxation with both extensive and intensive responses," *Journal of Economic Theory*, 148(5), 1770–1805.

LAFFONT, J.-J., E. MASKIN, AND J.-C. ROCHET (1987): "Optimal Nonlinear Pricing with Two-Dimensional Characteristics," in *Information, Incentives and Economic Mechanism*, ed. by T. Grove, R. Radner, and S. Reiter, pp. 256–266. Univ. of Minnesota Press.

LEHMANN, E., L. SIMULA, AND A. TRANNOY (2013): "Tax Me If You Can! Optimal Nonlinear Income Tax between Competing Governments," IZA Discussion Papers 7646, Institute for the Study of Labor (IZA).

MANKIW, N. G., AND M. WEINZIERL (2010): "The Optimal Taxation of Height: A Case Study of Utilitarian Income Redistribution," *American Economic Journal: Economic Policy*, 2(1), 155–76.

MIRRLEES, J. A. (1971): "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 38, 175–208.

RAMSEY, F. P. (1927): "A Contribution to the Theory of Taxation," *Economic Journal*, 37, 47–61.

REVECZ, J. (2003): "Comparing elasticities-based optimal income tax formulas," *Public Finance*, 53(3-4), 470–479.

ROCHET, J. C. (1985): "The taxation principle and multi-time Hamilton-Jacobi equations," *Journal of Mathematical Economics*, 14(2), 113 – 128.

ROCHET, J.-C., AND P. CHONÉ (1998): "Ironing, Sweeping, and Multidimensional Screening," 66(4), 783–826.

ROCHET, J.-C., AND A. STOLE, L (2002): "Nonlinear Pricing with Random Participation," *The Review of Economic Studies*, 69(1), pp. 277–311.

ROTHSCHILD, C., AND F. SCHEUER (2013a): "Optimal Taxation with Rent-Seeking," NBER Working Paper No. 17035.

——— (2013b): "Redistributive Taxation in the Roy Model," *The Quarterly Journal of Economics*, 128(2), 623–668.

——— (2013c): "A Theory of Income Taxation under Multidimensional Skill Heterogeneity," mimeo.

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.

——— (2003): "Reply on Comparing Elasticity-based Optimal Income Tax Formulas by John T. Revecz," *Public Finance*, 53(3-4), 480–485.

SALANIÉ, B. (2005): *The Economics of Contracts*. MIT Press, Cambridge, MA.

SCHEUER, F. (2013a): "Adverse selection in credit markets and regressive profit taxation," *Journal of Economic Theory*, 148(4), 1333–1360.

——— (2013b): "Entrepreneurial Taxation with Endogenous Entry," NBER Working Papers 19235, *American Economic Journal: Economic Policy* , forthcoming.