Optimal income taxation with composition effects*

Laurence JACQUET† Etienne LEHMANN‡

February 14, 2017

Abstract

We study the optimal nonlinear income tax problem with unobservable multidimensional heterogeneity. The obtained optimal tax formula generalizes the standard one by taking means of sufficient statistics among the distinct individuals who earn the same income. The fact that these individuals differ along several characteristics brings a new source of endogeneity to the tax model that we call composition effects. We emphasize that composition effects may substantially affect optimal marginal tax rates, especially for very high income levels. We provide a sufficient condition for optimal marginal tax rates to be positive. Finally, we highlight the equivalence between the assumptions required to use the tax perturbation approach and the assumptions subjacent to the first-order mechanism design approach, both being used in the literature.

Keywords: Optimal taxation, multidimensional screening problems, tax perturbation, allocation perturbation, sufficient statistics.

I Introduction

Since the seminal paper of Mirrlees (1971), the optimal income tax literature provides formulas describing the optimal tax schedule. The most recent formulas express the optimal tax rates in terms of empirically meaningful "sufficient statistics"(Saez, 2001, 2002, Chetty, 2009). These sufficient statistics are behavioral responses to tax reforms, the income distribution and the social welfare weights which summarize the social preferences for redistribution (See e.g. Diamond (1998) and Saez and Stantcheva (2016)). Most of the papers in the literature restrict the unobserved heterogeneity to be one dimensional, typically skill (See e.g. Mirrlees (1971), Diamond (1998), Saez (2001)). This limitation is empirically counterfactual as individuals differ in many other dimensions. For instance, they differ also in their ability to change their income

*Part of the material we propose in this paper has been presented in working papers under the following titles "Optimal Nonlinear Income Taxation with Multidimensional Types: The Case with Heterogeneous Behavioral Responses" (2014), "Optimal Income Taxation when Skills and Behavioral Elasticities are Heterogeneous" (2015) and "Optimal Taxation with Heterogeneous Skills and Elasticities: Structural and Sufficient Statistics Approaches" (2016). The authors would like to deeply thank the Editor, Gita Gopinath and three referees for their extremely useful comments. We also acknowledge helpful comments from Pierre Boyer, Craig Brett, Bas Jacobs, Guy Laroque, Stéphane Robin, Emmanuel Saez, Florian Scheuer, Stefanie Stantcheva, Alain Trannoy as well as from participants at various seminars, conferences and workshops.

†THEMA - University of Cergy-Pontoise. Laurence Jacquet is also adjunct professor at Oslo Fiscal Studies, Economics Department, University of Oslo. Address: THEMA, Université de Cergy-Pontoise, 33 boulevard du Port, 95011 Cergy-Pontoise Cedex, France. Email: laurence.jacquet@u-cergy.fr.

‡CRED (TEPP) University Panthéon Assas Paris II. Address: 12 place du Panthéon 75 231, Paris Cedex 05, France. Email: etienne.lehmann@gmail.com. Etienne Lehmann is also research fellow at IZA, CESifo and CEPR and is junior member of IUF.

after a tax reform (their behavioral responses). In this paper, we explore the case where the unobserved heterogeneity is multidimensional.

For this purpose, we consider an economy where individuals are endowed with different characteristics, one of them being the skill level. Individuals that differ only in terms of skill belong to what we call a "group". For instance, a group may be characterized by the same behavioral elasticities. We impose the single-crossing assumption with respect to skill within each group. Thanks to this assumption, we derive an optimal tax formula in terms of sufficient statistics. We therefore confirm in Proposition 1 the conjecture of Saez (2001) that his tax formula is also valid in the multidimensional context if one averages sufficient statistics across individuals who belong to distinct groups but earn the same income.

One important limitation of sufficient statistics is their dependence to the tax schedule. While this is true in the one dimensional case, considering different groups implies a new source of endogeneity that we call "composition effects". For instance, if the economy is made of a low-elasticity group and a high-elasticity group and if the marginal tax rate is higher in the optimal economy than in the actual economy, the shift from the actual economy to the optimal one induces a lower reduction of earnings in the high-elasticity group than in the low-elasticity group. Therefore, the share of people, at a given income level, who belong to the low-elasticity group depends on the tax schedule. We argue that these composition effects may lead to substantial changes in the optimal tax schedule. We provide several examples of drastic changes that the optimal top marginal tax rate undergoes when one modifies the threshold from which one determines top incomes. Strikingly, our examples are consistent with the empirical finding of rather stable Pareto parameters for top income earners. Given the available empirical evidence, we cannot rule out that optimal top tax rate may lie between 46% and 77%.

Besides deriving the optimal tax formula, we also rewrite it in terms of policy-invariant functions instead of endogenous sufficient statistics. Under the additional assumption that individual preferences are additively separable, we obtain a "structural"formula in which optimal marginal tax rates are expressed as a function of the skill density and the derivatives of the individual and social utility functions (Proposition 2). This formula is not closed-form because it depends on the allocation where these functions are evaluated, but it is much more convenient to implement numerically. Moreover, we use it to extend the result of Mirrlees (1971) that optimal marginal tax rates are positive when social preferences are Utilitarian or Maximin (Proposition 3).

In the one-dimensional case, there exist two strategies to derive the optimal tax formula. On the one hand, Mirrlees (1971) considers the problem of finding the incentive-compatible allocation that maximizes the social objective subject to a resource constraint. In this "mechanism design"approach, the optimal allocation is obtained by verifying that any incentive-compatible perturbation of the optimal allocation does not lead to first-order improvements.[1] Describing

[1]Such a condition is generally obtained using an Hamiltonian or a Lagrangian technics.

how to decentralize the optimal allocation provides the optimal tax formula à la Diamond (1998). For technical convenience, it is frequently assumed that income is continuous and increasing in skill, which enables one to consider only the first-order incentive constraint.[2] This is the so-called "first-order approach". On the other hand, one can consider the tax schedule that maximizes the social objective given a budget constraint. In this case, the optimal tax schedule is obtained by verifying that any marginal tax reform to the optimal tax schedule does not lead to first-order improvements. This is the tax perturbation approach introduced in the optimal income tax literature by Piketty (1997) and Saez (2001). The tax perturbation approach requires that individuals behaviors responds smoothly to a tax reform (Golosov et al., 2014, Hendren, 2014), a rather ad-hoc assumption with nonlinear income tax schedules. We clarify the conditions under which the tax perturbation approach is valid. Intuitively, the income tax schedule should not be too concave so that indifference curves are more convex than the function relating pre-tax income to after-tax income. A tax schedule respecting this assumption and a regularity condition is said to be smooth. Saez (2001) shows the two approaches leads to the same formula only in the case where the heterogeneity is one-dimensional. We show in Proposition 4 that the assumptions required in the first-order approach and in the tax perturbation approaches are equivalent. More specifically, any smooth tax function induces an allocation where in each group income is continuous and increasing in skill; reciprocally, any allocation where in each group income is continuous and increasing in skill can be decentralized by a smooth income tax schedule. Therefore, the differences between the two approaches rely only in their exposition: while the first-order approach focuses on an incentive-compatible perturbation of the optimal allocation, the tax perturbation approach looks at the tax reform that decentralizes this perturbation. Therefore, both approaches are two faces of the same coin.

Related literature

There have been optimal income tax papers where heterogeneity is multidimensional. However, these models rely on the assumption that the action only depends on a one-dimensional aggregation of the multidimensional unobserved heterogeneity. This is the case, for instance, in models of optimal income taxation such as those proposed in Brett and Weymark (2003), Boadway et al. (2002), Choné and Laroque (2010), Lockwood and Weinzierl (2015), Rothschild and Scheuer (2013, 2016, 2014), Scheuer (2013) and Scheuer (2014a). By contrast, we relax this assumption.Since we do no rely on an aggregator, we are able to simultaneously consider, in an optimal income tax model, heterogeneity in income and heterogeneity in behavioral elasticities.

In Brett and Weymark (2003), Boadway et al. (2002), Choné and Laroque (2010), Lockwood and Weinzierl (2015), the additional source of heterogeneity is relevant only for the computation of the social welfare weights, but behaviors depends only on a single dimension. This additional source of heterogeneity rationalize the idea that some low-income individuals may not deserve redistribution, while some high-income individuals may deserve not to be taxed

[2]Noticeable exceptions are Lollivier and Rochet (1983) and Ebert (1992).

too heavily. Therefore, social welfare weights may be less decreasing in income and may even become increasing, open the possibility for optimal marginal tax rates to be negative.

Random participation models make up another strand of the literature where multidimensional heterogeneity is taken into account, although in a very specific way. In these models, individuals differ in skill and in a cost of participation (Rochet and Stole, 2002, Kleven et al., 2009, Jacquet et al., 2013) or of migration (Lehmann et al., 2014, Blumkin et al., 2014) and this latter dimension of heterogeneity matters only for the participation/migration margin. In these papers, the one-dimensional aggregation implies that individuals who earn the same income are characterized by the same level of skill and are therefore constrained to react identically to any tax reform. While departing from this restriction, we show (in an appendix available upon request) that our model can readily be extended to include a random participation constraint.

Finally, Scheuer (2013, 2014b), Rothschild (2013), Gomes et al. (2014b) consider optimal income tax models with different sectors and agents can migrate from one sector to the other.[3] This is a form of random participation across sectors. Again, once individuals choose in which sectors to work (or which combination), income depends only on single variable.

The remainder of the paper is organized as follows. In Section II, we present the model. Section III derives the optimal tax formula using the tax perturbation method. Section IV uses the latter formula for computing optimal top tax rates for top income earners and provides some plausible examples that illustrate that composition effects may have a serious effect on the optimal top tax rates. Section V provides the structural optimal tax formula and a sufficient condition for optimal marginal tax rates to be positive. Section VI discusses the connection between the first-order and the tax perturbation approaches and the last section concludes.

II Model

Every worker derives utility from consumption $c \in \mathbb{R}_+$ and disutility from effort.[4] More effort implies higher pre-tax income $y \in \mathbb{R}_+$ (for short, income hereafter). Following Mirrlees (1971), the government levies a non-linear tax $T(.)$ which depends on income y only. Consumption c is related to income y through the tax function $T(y)$ according to $c = y - T(y)$. Individuals differ along their skill level $w \in \mathbb{R}_+^*$ and along a vector of characteristics denoted $\theta \in \Theta$. We call a *group* a subset of individuals with the same θ. We assume that the set of groups Θ is measurable with a cumulative distribution function (CDF) denoted $\mu(\cdot)$. The set Θ can be finite or infinite and 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. Among individuals of the same group θ, skills are distributed according to the conditional skill density $f(\cdot|\theta)$ which is positive and differentiable over the support $\mathbb{R}_+^*$. The conditional CDF is denoted $F(w|\theta) \stackrel{\text{def}}{\equiv} \int_0^w f(x|\theta)dx$. We do not make any restriction on the correlation between w

[3]In Gomes et al. (2014a), types that derive the same utility across different sectors while supplying different labor are pooled together, whereas, in this paper, we will pool together individuals that generate the same income.

[4]Effort captures the quantity as well as the intensity of labor supply.

or θ. We normalize to unity the total size of the population.

Individuals of type (w, θ) have a twice continuously differentiable utility function with respect to c and y which is specified as: $\mathscr{U}(c, y; w, \theta)$ with $\mathscr{U}'_c > 0 > \mathscr{U}'_y$. We also assume that for each type (w, θ), indifference curves associated to $\mathscr{U}(\cdot, \cdot; w, \theta)$ are strictly convex in the income-consumption space. Earning a given income requires less effort to a more productive worker, so $\mathscr{U}'_w > 0$. A worker of type (w, θ), facing $y \mapsto T(y)$, solves:

$$U(w, \theta) \stackrel{\text{def}}{\equiv} \max_y \quad \mathscr{U}(y - T(y), y; w, \theta) \tag{1}$$

We call $Y(w, \theta)$ the solution to program (1),[5] $C(w, \theta) = Y(w, \theta) - T(Y(w, \theta))$ the consumption of a worker of type (w, θ) and $U(w, \theta)$ her utility. When the tax function is differentiable, the first-order condition associated to (1) implies that:

$$1 - T'(Y(w, \theta)) = \mathscr{M}(C(w, \theta), Y(w, \theta); w, \theta) \tag{2}$$

where:

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

denotes the marginal rate of substitution between (pre-tax) income and consumption (after-tax income). For a worker of a given type, the left-hand side of Equation (2) corresponds to the marginal gain of income after taxation while the right-hand side corresponds to the marginal cost of income in monetary terms.

We impose the single-crossing (Spence-Mirrlees) condition within each group of workers endowed with the same θ that the marginal rate of substitution function is a decreasing function of the skill level, i.e. more skilled workers find it less costly to earn y:

Assumption 1 (Within-group single-crossing condition)**.** *For each $\theta \in \Theta$, and each $(c, y) \in \mathbb{R}_+ \times \mathbb{R}_+$, function $w \mapsto \mathscr{M}(c, y; w, \theta)$ is differentiable with $\forall w \in \mathbb{R}_+^*$, $\mathscr{M}_w < 0$, $\lim_{w \mapsto 0} \mathscr{M}(c, y; w, \theta) = +\infty$ and $\lim_{w \mapsto \infty} \mathscr{M}(c, y; w, \theta) = 0$.*

Assuming that the marginal rate of substitution decreases from plus infinity to zero is a kind of Inada condition. It implies that for a given differentiable tax schedule $T(\cdot)$, a given income y and a given group θ, there is a unique skill level w such that within group θ, only individuals endowed with this skill level earn income y. Assumption 1 is automatically verified in the case where $\mathscr{U}(c, y; w, \theta)$ can be rewritten as:

$$\mathscr{U}(c, y; w, \theta) = u(c) - \frac{\theta}{1+\theta} \left(\frac{y}{w}\right)^{1+\frac{1}{\theta}} \qquad \text{with} \qquad \theta > 0 \quad \text{and} \quad u'(\cdot) > 0 \geq u''(\cdot). \tag{4}$$

We henceforth refer to this specification of preferences as the isoelastic ones. There θ stands for the individual Frisch labor supply elasticity, hereafter, "labor supply elasticity". The marginal

[5] If the maximization program (1) admits multiple solutions, we make the tie-breaking assumption that individuals choose among their best options the income level preferred by the government, i.e. the one with the largest tax liability.

rate of substitution equals $\mathscr{M}(c, y; w, \theta) = y^{\frac{1}{\theta}} / [u'(c)\ w^{1+\frac{1}{\theta}}]$ and is decreasing in w from infinity to zero.

The government's budget constraint takes the form:

$$\iint_{\theta \in \Theta, w \in \mathbb{R}_+^*} T\left(Y(w, \theta)\right)\ f(w|\theta) dw\ d\mu(\theta) = E \tag{5}$$

where $E \geq 0$ is an exogenous amount of public expenditures. The objective of the planner is to maximize a general social welfare function that sums over all types of individuals an increasing and weakly concave transformation $\Phi(U; w, \theta, \chi)$ of individuals' utility levels U:

$$\iint_{\theta \in \Theta, w \in \mathbb{R}_+^*} \Phi\left(U(w, \theta); w, \theta\right)\ f(w|\theta) dw\ d\mu(\theta) \tag{6}$$

Preferences for redistribution are induced by the concavity of Φ, $\Phi''_{UU} < 0$.[6]

This welfarist specification allows Φ to vary with type (w, θ) which makes it very general. *Weighted utilitarian* preferences are obtained with $\Phi(U; w, \theta) \equiv \varphi(w, \theta) \cdot U$ with weights $\varphi(w, \theta)$ depending on individual characteristics. The objective is *utilitarist* if $\varphi(w, \theta)$ is constant and $\Phi(U; w, \theta) \equiv U$ and it turns out to be *maximin (or Rawlsian)* if $\varphi(0, \theta) > 0$ while $\varphi(w, \theta) = 0$ $\forall w > 0$. When $\Phi(U; w, \theta)$ does not vary with its two last arguments and $\varphi(w, \theta) = 1$, we obtain a Bergson-Samuelson criterion which is a concave transformation of utility and does not depend on individual characteristics.

The government's problem consists in finding the tax schedule $T(\cdot)$ that maximizes social welfare function (6) subject to the budget constraint (5). Let $\lambda > 0$ be the Lagrange multiplier associated with the budget constraint (5). To solve the government maximization problem, we proceed in two steps. First, for a given value of the Lagrangian multiplier λ, one maximizes the Lagrangian:

$$\mathscr{L} \stackrel{\text{def}}{\equiv} \iint_{\theta \in \Theta, w \in \mathbb{R}_+^*} \left[T(Y(w, \theta) + \frac{\Phi\left(U(w, \theta); w, \theta\right)}{\lambda}\right]\ f(w|\theta) dw\ d\mu(\theta) \tag{7}$$

Second, one finds the value of λ which ensures that the solution found at the first step makes the budget constraint (5) binding. We define the social marginal welfare weights associated with workers of type (w, θ) expressed in terms of public funds by:

$$g\left(w, \theta\right) \quad \stackrel{\text{def}}{\equiv} \quad \frac{\Phi_U\left(U\left(w, \theta\right); w, \theta\right) \mathscr{U}_c'\left(C(w, \theta), Y(w, \theta); w, \theta\right)}{\lambda} \tag{8}$$

The government values giving one extra dollar to a worker (w, θ) as a gain of $g(w, \theta)$ dollar(s) of public funds.[7]

[6]Our specification also encompasses the case where function Φ equals a type-specific exogenous weight times the individuals' level of utility with $\Phi''_{UU} = 0$.

[7]We can easily extend our analysis to *non-welfarist* social criteria following the method of generalized marginal social welfare weights developed in Saez and Stantcheva (2016) to reflect non-welfarist views of justice which can be particularly relevant with heterogeneous preferences. Complementary to their approach, Fleurbaey and Maniquet (2011, 2015) connect the axioms of fair income tax theory and optimal income taxation and emphasize that it is not always straightforward to derive generalized marginal social welfare weights by income level.

III From tax reform responses to optimal tax formulas

A reform of a tax schedule $y \mapsto T(y)$ is defined thanks to its direction, which is a differentiable function $R(y)$ defined on $\mathbb{R}_+$, and its (algebric) magnitude $m \in \mathbb{R}$. After the reform, the tax schedule becomes $y \mapsto T(y) - mR(y)$ and the utility of an individuals of type (w, θ) is:

$$U^R(m; w, \theta) \stackrel{\text{def}}{\equiv} \max_y \quad \mathscr{U}\left(y - T(y) + m\, R(y), y; w, \theta\right) \tag{9}$$

We denote $Y^R(m; w, \theta)$ the income of workers of types (w, θ) after the reform and we have $C^R(m; w, \theta) = Y^R(m; w, \theta) - T(Y^R(m; w, \theta)) + m\, R(Y^R(m; w, \theta))$. When $m = 0$, we have $Y^R(0; w, \theta) = Y(w, \theta)$ and $C^R(0; w, \theta) = C(w, \theta)$. Applying the envelope theorem to (9), we get:

$$\frac{\partial U^R}{\partial m}(m; w, \theta) = \mathscr{U}_c\left(C^R(m; w, \theta), Y^R(m; w, \theta); w, \theta\right) \tag{10}$$

Using (3), we write the first-order condition associated to (9) which equalizes to zero the following expression:

$$\mathscr{Y}^R(y, m; w, \theta) \stackrel{\text{def}}{\equiv} 1 - T'(y) + m\, R'(y) - \mathscr{M}\left(y - T(y) + m\, R(y), y; w, \theta\right). \tag{11}$$

To compute the responses to a tax reform, one needs to apply the implicit function theorem to $\mathscr{Y}^R(Y^R(m; w, \theta), m; w, \theta) = 0$. At $m = 0$, $\mathscr{Y}^R_y$ does no longer depend on the direction R of the tax reform. To ease the notation, from now on, we then drop the superscript R and write $\mathscr{Y}_y(Y(w, \theta); w, \theta)$ for $\mathscr{Y}^R_y(Y(w, \theta), 0; w, \theta)$. To use the widespread tax perturbation method, one needs the following assumptions:

Assumption 2.

i) The tax function $T(\cdot)$ is twice differentiable.

ii) For all $(w, \theta) \in \mathbb{R}^*_+ \times \Theta$, the second-order condition holds strictly: $\mathscr{Y}_y\left(Y(w, \theta); w, \theta\right) < 0$.

iii) For all $(w, \theta) \in \mathbb{R}^*_+ \times \Theta$, the function $y \mapsto \mathscr{U}\left(y - T(y), y; w, \theta\right)$ admits a unique global maximum over $\mathbb{R}_+$.

Part *i)* of Assumption 2 ensures that first-order condition (11) is differentiable. Part *ii)* guarantees it is invertible in income y. Under *i)* and *ii)*, one can apply the implicit function theorem to (11) to describe how a local maximum of Program (9) changes after a tax reform. Part *iii)* is also crucial since it implies that the optimal allocation does smoothly respond to an infinitesimal tax perturbation. Without this, the tax perturbation approach would not be valid. Since the tax function is nonlinear, the function $y \mapsto \mathscr{U}\left(y - T(y) + mR(y), y; w, \theta\right)$ may in general admit several global maxima among which individuals of type (w, θ) are indifferent. Any small tax reform may then lead to a distinct unique global maximum. Moving from a global maximum to another one in the wake of a tax reform is associated with a jump in the chosen income (i.e. a jump in the supply of effort). Part *iii)* prevents this situation and ensures that the first-order condition corresponds to a unique global maximum.

We now describe circumstances under which the tax perturbation approach can be used because Assumption 2 is automatically satisfied. This is the case when the tax function $T(y)$ is restricted to be linear as the indifference curves associated to $\mathscr{U}(.,.;w,\theta)$ are strictly convex. Similarly, Assumption 2 is also satisfied when the tax function $T(y)$ is convex ($y \mapsto y - T(y)$ being concave, Parts *ii)* and *iii)* are then verified).[8] By continuity, Assumption 2 is also verified when $y \mapsto T(y)$ is "not too concave", more precisely when $y \mapsto y - T(y)$ is less convex than the indifference curve with which it has a tangency point in the (y,x)-plane (so that Part *ii)* of Assumption 2 is satisfied) and that this indifference curve is strictly above $y \mapsto y - T(y)$ for all other y (so that Part *iii)* of Assumption 2 is satisfied).

To study any tax reform, one simply applies the implicit function theorem to $\mathscr{Y}^R(y,m;w,\theta) = 0$ at $(y = Y(w,\theta), m = 0; w,\theta)$ which yields:

$$\frac{\partial Y^R}{\partial m} = -\frac{\mathscr{Y}^R_m}{\mathscr{Y}^R_y}$$

with:

$$\begin{aligned}
\mathscr{Y}^R_y(y,m;w,\theta) &= -T''(y) - \mathscr{M}_y(y - T(y) + m\,R(y), y; w,\theta) &\quad (12a)\\
&- \mathscr{M}(y - T(y) + m\,R(y), y; w,\theta)\ \mathscr{M}_c(y - T(y) + m\,R(y), y; w,\theta) &\\
\mathscr{Y}^R_m(y,m;w,\theta) &= R'(y) - R(y)\ \mathscr{M}_c(y - T(y) + m\,R(y), y; w,\theta). &\quad (12b)
\end{aligned}$$

To define the sufficient statistics that play a role in the tax formula, we study two specific reforms: one to capture substitution effects and another one to isolate income effects. To capture substitution effects, modify the marginal tax rate by a constant amount m around income $Y(w,\theta)$ and leave unchanged the level of tax at this income level so that we label this reform as *compensated*. Formally, this compensated reform has for direction $R(y) = y - Y(w,\theta)$; it does not modify the tax level in $y = Y(w,\theta)$ (i.e. $R(Y(w,\theta)) = 0$) and it uniformly modifies the marginal tax rate as can be seen from $R'(Y(w,\theta)) = 1$. The *compensated elasticity* of income with respect to the marginal retention rate $1 - T'(.)$ is:

$$\varepsilon(w;\theta) \overset{\text{def}}{\equiv} \frac{1 - T'(Y(w,\theta))}{Y(w,\theta)} \frac{\partial Y^c}{\partial m} = \frac{\mathscr{M}(C(w,\theta), Y(w,\theta); w,\theta)}{-Y(w,\theta)\ \mathscr{Y}_y(Y(w,\theta); w,\theta)} > 0 \quad (13a)$$

where the superscript "c" emphasizes that the change of $Y(w,\theta)$ is due to the compensated tax reform. The compensated elasticity is positive from Assumption 2.

To capture income effects, we consider a uniform transfer of money to all workers and call this reform a *lump-sum* one. This reform is obtained thanks to $R(y) \equiv 1$ and the *income effect* can then be written as:

$$\eta(w;\theta) \overset{\text{def}}{\equiv} \frac{\partial Y^L}{\partial m} = \frac{\mathscr{M}_c(C(w,\theta), Y(w,\theta); w,\theta)}{\mathscr{Y}_y(Y(w,\theta); w,\theta)} \quad (13b)$$

[8]From (3), we can write $-\mathscr{M}_y - \mathscr{M}\mathscr{M}_c = (\mathscr{U}_c)^{-3}\left(\mathscr{U}_{yy}\mathscr{U}_c^2 - 2\mathscr{U}_{cy}\mathscr{U}_c\mathscr{U}_y + \mathscr{U}_{cc}\mathscr{U}_y^2\right)$, which is negative due to the concavity of $\mathscr{U}(\cdot,\cdot;w,\theta)$. Therefore, from (1) and the associated first-order condition $\mathscr{Y}(y;w,\theta) \overset{\text{def}}{\equiv} 1 - T'(y) - \mathscr{M}(y - T(y), y; w,\theta) = 0$, we have the second-order condition $\mathscr{Y}_y = -T''(y) - \mathscr{M}_y - \mathscr{M}\ \mathscr{M}_c < 0$ whenever $T'' \geq 0$.

where the superscript "L" stresses that the change of $Y(w,\theta)$ is due to the lump-sum reform. If leisure is a normal good, one has $\mathscr{M}_c > 0$, in which case $\eta(w,\theta) < 0$.

Another relevant elasticity, the elasticity of earnings with respect to skill w, can be built up under Assumption 2. Apply the implicit function theorem to (11) with respect to skill w. Note that this ensures that income $Y(\cdot,\theta)$ is a continuously differentiable function in skill. Using $\mathscr{Y}_w^R = -\mathscr{M}_w$, the elasticity of earnings with respect to skill w is:

$$\alpha(w;\theta) \stackrel{\text{def}}{\equiv} \frac{w}{Y(w,\theta)}\,\dot{Y}(w,\theta) = \frac{w\,\mathscr{M}_w(C(w,\theta),Y(w,\theta);w,\theta)}{Y(w,\theta)\,\mathscr{Y}_y(Y(w,\theta);w,\theta)} > 0 \tag{13c}$$

which is positive from Assumption 1. Notice that Assumptions 1 and 2 rule out bunching.[9] This is because, under Assumption 1, bunching can only be decentralized by a kink in the tax function with increasing marginal tax rates (Saez, 2010). However, a tax function with such a kink violates Part *i*) of Assumption 2.

Combining (13a) (13b) with (12b), the way income of individuals (w,θ) reacts to any (general) tax reform $R(\cdot)$ is given by:

$$\frac{\partial Y^R}{\partial m}(0;w,\theta) = \varepsilon(w,\theta)\frac{Y(w,\theta)}{1-T'(Y(w,\theta))}R'(Y(w,\theta)) + \eta(w,\theta)\,R(Y(w,\theta)) \tag{14}$$

Our definitions of elasticities and income response (13a)-(13c) account for the nonlinearity of the income tax schedule. In the denominators of these definitions, the term $T''(Y(w,\theta))$ (which is incorporated in $\mathscr{Y}_Y$ (see Equation (12a))) emphasizes the role played by the local curvature of the tax schedule. We refer to these elasticities and income response as representing *total* responses i.e., including the circular process induced by the endogeneity of marginal tax rates. Recent papers in the literature e.g. Jacquet et al. (2013) and Scheuer and Werning (2017) use total elasticities and income responses which help streamline tax formulas. In contrast, previous papers (Saez, 2001, Golosov et al., 2014, Hendren, 2014) expressed optimal tax formulas in terms of responses that do not take into account the local curvature of the tax function, which we refer to *direct* responses. Let $\varepsilon^\star(w;\theta)$, $\eta^\star(w;\theta)$ and $\alpha^\star(w;\theta)$ denote these direct responses, i.e. the compensated elasticity of earnings with respect to the marginal retention rate, the income effect and the elasticity of earnings with respect to skill, when $T''=0$ in (13a)-(13c). These would be the relevant concepts if the tax function were linear. An exogenous change in w, or a tax reform induces a *direct* change in earnings $\Delta_1 y$ proportional to the direct response $\varepsilon^\star(w;\theta)$, $\eta^\star(w;\theta)$ and $\alpha^\star(w;\theta)$. However, when the tax schedule is nonlinear, the direct response in earnings Y modifies the marginal tax rate by $\Delta_1 T' = T''(Y(w,\theta)) \times \Delta_1 y$, thereby inducing a further change in earnings $\Delta_2 y = -Y(w,\theta)\frac{T''(Y(w,\theta))}{1-T'(Y(w,\theta))}\varepsilon^\star(w,\theta)\Delta_1 y$. This second change in earnings, in turn, induces a further modification in the marginal tax rate $T''(Y(w,\theta)) \times \Delta_2 y$ which induces an additional change in earnings. Therefore, a *circular process* takes place. The income level

[9] In our context of multidimensional characteristics, *bunching* refers to the specific situation where individuals who earn the same income belong to the same group θ but have distinct skills. In contrast, *pooling* refers to a situation where individuals who earn the same income belong to distinct groups. Since we address multidimensional problems, we can study pooling and neglect bunching without any loss in generality.

determines the marginal tax rate through the tax function, and the marginal tax rate affects the income level through the substitution effects. Using the identity $1 + x + x^2 + x^3 + ... = \frac{1}{1-x}$, the total effect is given by:

$$\begin{aligned}\Delta y = \sum_{i=1}^{\infty} \Delta_i y &= \Delta_1 y \sum_{i=1}^{\infty} \left(-Y(w,\theta) \frac{T''(Y(w,\theta))}{1 - T'(Y(w,\theta))} \varepsilon^\star(w,\theta) \right)^{i-1} \\ &= \Delta_1 y \, \frac{1 - T'(Y(w,\theta))}{1 - T'(Y(w,\theta)) + Y(w,\theta)\, T''(Y(w,\theta))\, \varepsilon^\star(w,\theta)}\end{aligned}$$

where the ratio in the last equality is positive whenever $-Y(w,\theta)\frac{T''(Y(w,\theta))}{1-T'(Y(w,\theta))}\varepsilon^\star(w,\theta)$ is lower than 1, i.e. whenever the second-order condition holds strictly. This ratio is the *corrective term* by which direct responses must be timed to obtain total responses as made explicit by the following equations:

$$\varepsilon(w,\theta) = \frac{1 - T'(Y(w,\theta))}{1 - T'(Y(w,\theta)) + Y(w,\theta)\, T''(Y(w,\theta))\, \varepsilon^\star(w,\theta)} \, \varepsilon^\star(w,\theta) \tag{15a}$$

$$\eta(w,\theta) = \frac{1 - T'(Y(w,\theta))}{1 - T'(Y(w,\theta)) + Y(w,\theta)\, T''(Y(w,\theta))\, \varepsilon^\star(w,\theta)} \, \eta^\star(w,\theta) \tag{15b}$$

$$\alpha(w,\theta) = \frac{1 - T'(Y(w,\theta))}{1 - T'(Y(w,\theta)) + Y(w,\theta)\, T''(Y(w,\theta))\, \varepsilon^\star(w,\theta)} \, \alpha^\star(w,\theta) \tag{15c}$$

Equalities (15a)-(15c) are obtained from the definitions of elasticities, income responses and from (2).[10] Note that the responses that are estimated by the empirical literature, using various methods, are the direct ones.[11]

Denote $h(y|\theta)$ the conditional income density within group θ at income y and $H(y|\theta) \stackrel{\text{def}}{\equiv} \int_{z=0}^{y} h(z|\theta)dz$ the corresponding conditional income CDF. According to (13c) and Assumption 1, income $Y(\cdot,\theta)$ is strictly increasing in skill within each group. We then have $H(Y(w,\theta)|\theta) \equiv F(w|\theta)$ for each skill level w. Differentiating both sides of this equality with respect to w and using (13c) leads to:

$$h(Y(w,\theta)|\theta) = \frac{f(w|\theta)}{\dot{Y}(w,\theta)} \qquad \Leftrightarrow \qquad Y(w,\theta)\, h(Y(w,\theta)|\theta) = \frac{w\, f(w|\theta)}{\alpha(w,\theta)} \tag{16}$$

We can now define the sufficient statistics that will play a role in the tax formula expressed using the income distribution and that are functions of income y. Let $W(\cdot,\theta)$ denote the reciprocal of $Y(\cdot,\theta)$ so that, within each group θ, individuals of type $(w = W(y,\theta),\theta)$ earn income

[10] From (11) and (13a) we can write:

$$\frac{\varepsilon(y,\theta)}{\varepsilon^\star(y,\theta)} = \frac{\mathscr{M}_y + \mathscr{M}\mathscr{M}_c}{T'' + \mathscr{M}_y + \mathscr{M}\mathscr{M}_c}$$

Substituting (3) into (2) and using the definition of $\varepsilon^\star(y,\theta)$ yields (15a). The same goes for Equations (15b) and (15c).

[11] In the real world, most of income tax schedules are piecewise linear in which case the distinction between direct and total responses are relevant only at the kinks of tax schedules. To estimate the behavioral responses to tax reforms, there are two main methodologies. The first one uses actual tax reforms as a quasi-experimental design. What identifies the causal effect of tax on behavior is the fact that, absent any behavioral responses, some tax payers face a change in marginal tax rate while some other do not (Feldstein, 1995, Auten and Carroll, 1999, Gruber and Saez, 2002), see Saez et al. (2012). Since this approach considers that marginal tax rates do locally not depend on taxable income, it identifies *direct* behavioral responses. The second approach identifies behavioral responses from discontinuities in the distribution of taxable income around kinks (Saez, 2010) or notches (Kleven and Waseem, 2013) observed in tax schedules. In particular, around a convex kink where $T''(\cdot) = +\infty$, the *total* skill elasticity $\alpha(w,\theta)$ is nil from (15c) which triggers bunching around the kink. One then uses the relation between the magnitude of this bunching and the *direct* compensated elasticity to identify the latter.

y. According to Assumption 1, $W(y,\theta)$ is the unique skill level w such that the individual first-order condition $1 - T'(y) = \mathscr{M}(y - T(y), y; w, \theta)$ is verified at income y. The *unconditional* income density is given by:

$$\hat{h}(y) \stackrel{\text{def}}{\equiv} \int_{\theta\in\Theta} h(y|\theta)\, d\mu(\theta) \tag{17a}$$

The mean *total* compensated elasticity at income level y is:

$$\hat{\varepsilon}(y) = \frac{\int_{\theta\in\Theta} \varepsilon\left(W(y,\theta),\theta\right)\; h(y|\theta)\, d\mu(\theta)}{\int_{\theta\in\Theta} h(y|\theta)\, d\mu(\theta)} \tag{17b}$$

The mean *total* income effect at income level y is:

$$\hat{\eta}(y) = \frac{\int_{\theta\in\Theta} \eta\left(W(y,\theta),\theta\right)\; h(y|\theta)\, d\mu(\theta)}{\int_{\theta\in\Theta} h(y|\theta)\, d\mu(\theta)} \tag{17c}$$

Finally, the mean social welfare weight at income level y is:

$$\hat{g}(y) = \frac{\int_{\theta\in\Theta} g\left(W(y,\theta),\theta\right)\; h(y|\theta)\, d\mu(\theta)}{\int_{\theta\in\Theta} h(y|\theta)\, d\mu(\theta)} \tag{17d}$$

Having defined the general tax reforms and described their impact on individual income, we now study when a given tax reform is desirable. To do so, we locally perturb the tax system (which can be optimal or suboptimal) in the direction $R(y)$ with magnitude m. If the initial tax schedule $T(\cdot)$ is optimal, such a perturbation should not yield any first-order effect on the Lagrangian (7).

Lemma 1. *Reforming the tax schedule in the direction $R(\cdot)$ triggers first-order effects on the Lagrangian (7) equal to:*

$$\begin{aligned}\frac{\partial \mathscr{L}^R}{\partial m} &= \int_{y=0}^{\infty} \left\{ \left[\hat{g}(y) - 1 + T'(y)\,\hat{\eta}(y)\right] \hat{h}(y) - \frac{d}{dy}\left[\frac{T'(y)}{1-T'(y)}\hat{\varepsilon}(y)\, y\, \hat{h}(y)\right]\right\} R(y)\, dy \\ &+ \lim_{y\mapsto\infty} \frac{T'(y)}{1-T'(y)}\,\hat{\varepsilon}(y)\, y\, \hat{h}(y)\, R(y) - \lim_{y\mapsto 0} \frac{T'(y)}{1-T'(y)}\,\hat{\varepsilon}(y)\, y\, \hat{h}(y)\, R(y)\end{aligned} \tag{18}$$

The proof which is relegated to Appendix A.1 is in the vein of Golosov et al. (2014). It is based on studying perturbations of a given non-linear tax system taking into account the partial (Gateaux) differential of government tax revenue and social welfare with respect to tax reforms in the direction $R(.)$. However, in Golosov et al. (2014), individuals make the same number of actions as they have individual characteristics so that their method is not applicable in our framework. With our method, one can solve models with one action and many types.[12]

[12] Golosov et al. (2014) assume that the mapping between the vector of types and the vector of income choices is injective. This assumption is necessary in their proof as they write the government's Lagrangian using the endogenous density of the vector of incomes. However, their assumption of injectivity is irrelevant in our context with one income and many types. In contrast, our proof makes explicit the way the unconditional income density depends on the skill density in each group through Equations (16) and (17a). When deriving a formula for inequality deflator, Hendren (2014) takes into account the consequences of the nonlinearity of the tax schedule but his main expression (his Equation (6)) holds only when the marginal rate is locally income-invariant. Gerritsen (2016) assumes that incomes are differentiable in tax reform but does not explicit the conditions under which this differentiability is ensured and instead quotes a previous version of our paper. However, he considers a tax reform that is not differentiable in income to derive his Equation (15) so that the tax perturbation he considers does not verify the conditions we present in Assumption 2 to validate the tax perturbation.

An important point to notice is that, in general, implementing a reform in the direction R implies a budget surplus or deficit. A first-order approximation of this budget surplus (or deficit) can be computed by putting social welfare weights $\hat{g}(\cdot)$ equal to zero in (18). One can then define a balanced-budget tax reform of magnitude m in the direction R by combining the tax reform of magnitude m and the direction R with the lump-sum rebate required to bind the budget constraint. Appendix A.1 shows that the first-order effect of this balanced-budget tax reform on the social objective is positively proportional to the first-order effect of the tax reform of magnitude m and the direction R on the Lagrangian. Therefore, if Expression (18) is positive, it is socially desirable to implement a tax reform in the direction $R(\cdot)$ (with a positive magnitude m) and to combine this tax reform with a lump-sum rebate to keep the government's budget balanced. Symmetrically, if Expression (18) is negative, it is socially desirable to implement a tax reform with a positive magnitude m in the direction $-R$ combined with a lump-sum transfer. Equation (18) is therefore useful to determine which tax reforms are desirable.

We now characterize the optimal tax schedule in the model with multidimensional types. The proof is in Appendix A.2.

Proposition 1. *Under assumptions 1 and 2, the optimal tax schedule satisfies:*

$$\frac{T'(y)}{1-T'(y)} = \frac{1}{\hat{\varepsilon}(y)} \frac{1-\hat{H}(y)}{y\,\hat{h}(y)} \left(1 - \frac{\int_y^\infty [\hat{g}(z) + \hat{\eta}(z)\,T'(z)]\,\hat{h}(z)dz}{1-\hat{H}(y)}\right) \tag{19a}$$

$$1 = \int_0^\infty [\hat{g}(z) + \hat{\eta}(z)\,T'(z)]\;\hat{h}(z)dz. \tag{19b}$$

If income effects were assumed away, Equation (19b) would imply that the weighted sum of social welfare weights is equal to 1. In the presence of income effects, a uniform increase in tax liability induces a change in tax revenue proportional to the marginal tax rate which explains the presence of the term $\hat{\eta}(z) \cdot T'(z)$.

The optimal tax rate given in Equation (19a) consists in three terms: *i*) the behavioral responses to taxes $\frac{1}{\hat{\varepsilon}(y)}$, which, in the vein of Ramsey (1927), is the inverse of the mean compensated elasticity; *ii*) the shape of the income distribution measured by the local Pareto parameter $\frac{1-\hat{H}(y)}{y\,\hat{h}(y)}$ of the income distribution and *iii*) the social preferences and income effects $1 - \frac{\int_y^\infty [\hat{g}(z)+\hat{\eta}(z)\cdot T'(z)]\hat{h}(z)dz}{1-\hat{H}(y)}$. Saez (2001) discusses how the optimal tax rate is affected by each of these three terms in the one-dimensional case. Shifting from the model with one dimension of heterogeneity to the model with multiple dimensions leads to replacing the marginal social welfare weight, the compensated elasticity and the income effect by their means calculated at a given income level. It is the mean of the *total* (rather than *direct*) compensated elasticity and income effect that must be computed.[13] Note that the averaging procedure is a far cry from the simple extension of the unidimensional case that would consist in computing the simple average of every estimated sufficient statistic and then multiplying each average by the same

[13]This is more intuitive than using the direct elasticity and income effect, which implies to encapsulate the circularity (described by (15a)-(15c)) in a so-called "virtual density" as in Saez (2001), Equation (13).

corrective term. Instead, every optimal sufficient statistic at any income level is a weighted average that requires as many corrective terms as there are groups in which individuals earn this income level and group-specific densities as weights (as described in Equations (15a)-(15b) and (17b)-(17d)). Remarkably, this procedure highlights the importance of *composition effects* which stem, at every income level, from the prevalence of distinct groups of individuals in the actual and optimal economies. This makes every weighted average of corrected sufficient statistic distinct in the actual and optimal economies.

IV Optimal tax rates at the top

In this Section, we illustrate that composition effects may be crucial in determining optimal marginal tax rates for top incomes. To do so, we adopt the usual simplifications. Income effects are assumed away, so that $\hat{\eta}(y) = 0$. The government does not value the welfare of top earners so that $\hat{g}(y) = 0$, i.e. we are interested in optimal top tax rates that maximize tax revenue. Equation (19a) can therefore be simplified to:

$$\frac{T'(y)}{1 - T'(y)} = \frac{1 - \hat{H}(y)}{\hat{\varepsilon}(y)\, y\, \hat{h}(y)} \tag{20}$$

For the purpose of our illustration, let us consider an example where the top 1 % of the population is composed of two groups indexed with $i = 1, 2$. For the sake of simplicity, we assume individual preferences are iso-elastic:

$$\mathscr{U}(c, y; w, \theta) = c - \frac{\theta}{1+\theta} \left(\frac{y}{w}\right)^{\frac{1+\theta}{\theta}}. \tag{21}$$

with $\theta = \theta_i$ in group $i = 1, 2$. Let τ_0 denote the marginal tax rate within the top bracket of the actual income tax schedule. We adopt the standard assumption that, in the actual economy, the upper part of the income density within group i is Pareto:

$$h_0(y|i) = k_i\, p_i\, y^{-(1+p_i)} \qquad \Leftrightarrow \qquad 1 - H_0(y|i) = k_i\, y^{-p_i} \tag{22}$$

where k_i is the scale parameter and p_i is the Pareto parameter, with $p_i = \frac{y\, h_0(y|i)}{1-H_0(y|i)}$ for all income levels. Appendix A.3 shows that the optimal tax rate at income y solves:

$$\begin{aligned} T'(y) &= \frac{1}{1 + \hat{\mu}_1(y, \tau_0, T'(y))\, \theta_1\, p_1 + \hat{\mu}_2(y, \tau_0, T'(y))\, \theta_2\, p_2} \qquad \text{where :} \\ \hat{\mu}_i(y, \tau_0, \tau_*) &\stackrel{\text{def}}{\equiv} \frac{\mu(\theta_i)\, k_i \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_i p_i} y^{-p_i}}{\mu(\theta_1)\, k_1 \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_1 p_1} y^{-p_1} + \mu(\theta_2)\, k_2 \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_2 p_2} y^{-p_2}} \end{aligned} \tag{23}$$

This Equation generalizes the optimal tax rate formula of Saez (2001) and Piketty et al. (2014) to the case with two groups.[14] In their formula, the top tax rate is a decreasing function of

[14]The generalization to $n > 2$ groups is straightforward. We consider two groups to ease the presentation.

the product of the compensated elasticity and the Pareto parameter.[15] In our formula (Equation (23)), the top tax rate depends on the products of the compensated elasticity and the Pareto parameter within each group. More precisely, it is a decreasing function of a weighted average of these products, the weights $\hat{\mu}$ being endogenous. They depend both on tax rates and on income levels.

On the one hand, they depend on the tax rates in the actual and in the optimal economies through the terms $\left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_i p_i}$. These terms capture the changes in income densities, within each group, caused by the responses of taxpayers to the variation of marginal tax rate from τ_0 in the actual economy to $T'(y)$ in the optimal economy. The magnitude of these changes is higher for the high-elasticity group because its members react more strongly to tax reforms.

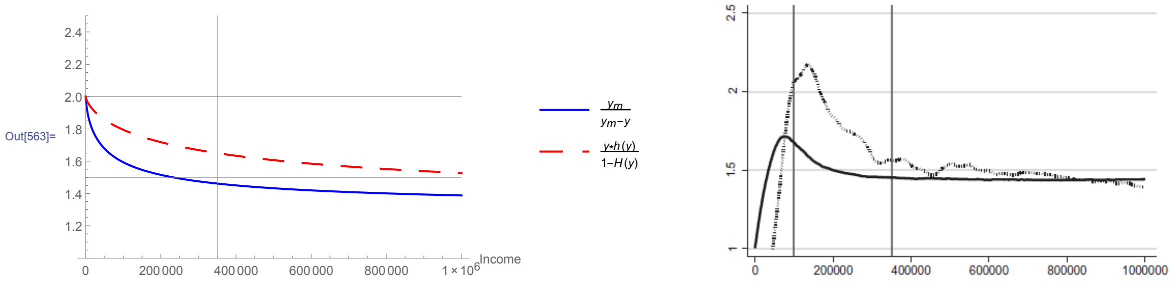
On the other hand, whenever the Pareto parameter p_i varies across groups, the weights $\hat{\mu}_i$ depend on income y. In particular, as income y tends to infinity, the very top earners come only from the group whose Pareto distribution has the fatter Pareto tail, i.e. the group with the lowest p_i. Therefore, in the optimal tax formula (Equation (23)), $\lim_{y\to\infty} \hat{\mu}_i \to 1$ for the group with the lowest p_i and $\lim_{y\to\infty} \hat{\mu}_i \to 0$ for the group with the highest p_i. The optimal asymptotic tax rate is then given by a formula identical to the one in Saez (2001), except that its elasticity and Pareto parameter have to be replaced by those of the group with the fatter Pareto tail (i.e. the group where the lowest p_i shows up).

Consider for instance the case where both groups have Pareto parameters close to 1.5 and where $\theta_1 = 0.2$ and $\theta_2 = 0.8$. One can think of the low-elasticity group as made of high salary workers who have few opportunities to adjust their income.[16] The high-elasticity group is made of capital owners and entrepeneurs. If the low-elasticity group has a Pareto parameter p_1 slightly above the one of the high-elasticity group then, the optimal asymptotic marginal tax rate is $1/(1 + 1.5 \times 0.8) \simeq 46\%$. In the opposite case, the optimal asymptotic marginal tax rate is $1/(1 + 1.5 \times 0.2) \simeq 77\%$. Now, if one instead implemented Equation (23) at an income level onwards from which both groups are of equal size, then one would take the mean of θ_1 and θ_2 to implement the optimal top tax rate. In our example, this yields $\theta = (\theta_1 + \theta_2)/2 = 0.5$. The asymptotic tax rate is then equal to $1/(1 + 1.5 \times 0.5) \simeq 57\%$.

When the Pareto parameters are heterogeneous across groups, the previous examples emphasize the role played by the income level from which one determines the optimal top tax rate. However, two out of our three examples are extreme as they require that income goes to infinity. We therefore provide a more general example. Even though it is well established in the literature that the top of the income distribution is extremely well approximated by a Pareto distribution, as far as we know, there are no empirical studies estimating Pareto parameters for distinct groups of population. Piketty and Saez (2013), who do *not* distinguish between several

[15] For comparison, note that Saez (2001) and Piketty et al. (2014) assume that the top tax rate above a fixed income level is constant and, neglecting multidimensional heterogeneity, they find $T'(\infty) = 1/(1 + p \cdot \theta)$ where p and θ are the Pareto parameter and elasticity.

[16] E.g., because their income is reported by their employers and their earnings are predetermined by their labor contracts.



(a) Our calibrated economy

(b) From Piketty and Saez (2013), p. 424

Figure 1: $y \mapsto \frac{y_m(y)}{y_m(y)-y}$ (solid lines) and $y \mapsto \frac{y\,\hat{h}_0(y)}{1-\hat{H}_0(y)}$ (dotted lines). The gray vertical line correspond to the 99^{th} percentile.

groups, estimate a single Pareto parameter using US tax return micro data. The right-hand side graph of Figure 1 displays their findings. The solid line is, for each income level y, the ratio $y_m(y)/(y_m(y)-y)$ where $y_m(z) = \mathbb{E}[z|z \geq y]$ stands for the mean income above y. The dotted line displays $y\,\hat{h}_0(y)/(1-\hat{H}_0(y))$ in the US. If the income distribution were Pareto, both ratios would be equal to the Pareto parameter. Piketty and Saez then argue that top incomes are well approximated by a Pareto distribution with parameter 1.5 since both ratios are constant around 1.5 above the 99^{th} percentile, that is above \$350, 500. Based on these empirical findings, we calibrate our economy so that the implied ratios $y_m(y)/(y_m(y)-y)$ and $y\,\hat{h}_0(y)/(1-\hat{H}_0(y))$ are also very close to 1.5, beyond \$350, 500, as can be seen on the left-hand side graph of Figure 1. To obtain this, we assume $\theta_1 = 0.2$ and $\theta_2 = 0.8$ as before and $p_1 = 1.3$ and $p_2 = 2$. We calibrate the scale parameters k_1 and k_2 so that each group represents half of the top percentile of the income distribution in the actual economy (see details in Appendix A.3). Being quite consistent with the available empirical findings, we consider our calibration as valuable. Of course, this is not the only possibility and we will propose another calibration exercise a bit further.

In Figure 2, we display the optimal top tax rates' formula (20) at different income levels. We represent the optimal tax rate as a function of income levels in the graph on the left and in terms of percentiles in the graph on the right. The thick (blue) curves represent the optimal top tax rates when one takes into account that the group composition of individuals who earn a given income level may be different in the actual and in the optimal economy (See Appendix A.3). Strikingly, optimal tax rates are far from being constant for very high income levels. The optimal marginal tax rate increases from 55.7% for the top 1% of the population to 60.2% for the top 0.5% and it reaches 70.2% for the top 0.1%. At the very top it rockets to 76%. The graph on the right shows a different pattern because the top 1, 0.5 and 0.1 percentiles correspond to substantially distinct income levels: \$350, 050 , \$537, 100 and \$1, 528, 500, respectively. This numerical example highlights the drastic change (up to 15 percentage points) that the optimal top marginal tax rate undergoes when one modifies the threshold from which one determines top incomes.

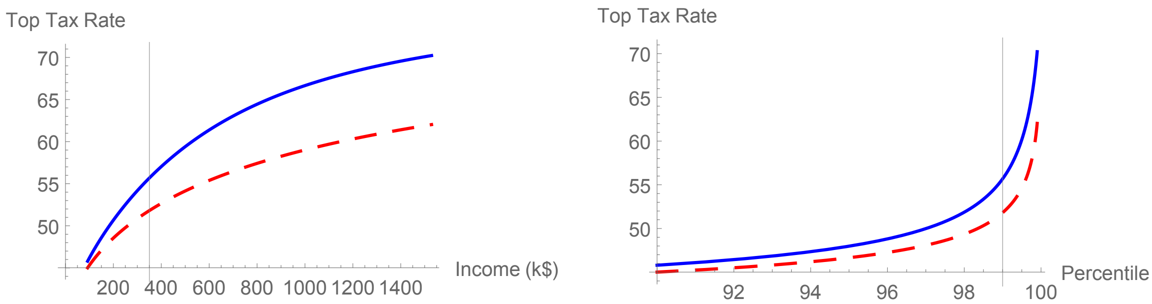


Figure 2: Optimal marginal tax rates for top incomes with (solid line) and without composition effects (dashed lines) when $p_1 = 1.3$ and $p_2 = 2$. The gray vertical line correspond to the 99^{th} percentile.

In Figure 2, the dashed (red) curves provides the optimal top tax rates from Equation (20) when one uses the actual densities (i.e. when one neglects composition effects) and when one approximates total elasticities by the direct ones. In other words, the terms $\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_i p_i}$ in the weights $\hat{\mu}_i$ of (23) are mystakenly replaced by 1. Therefore, the difference between the solid and the dashed curves shows the role played by composition effects, i.e. by the change in income densities between the current and the optimal economy. Taxpayers respond to the rise of marginal tax rates (between the actual and the optimal economy) by reducing their income. Importantly, taxpayers in the high-elasticity group respond more to this tax reform than taxpayers in the low-elasticity group. Since the income density is decreasing in income, at each very high income level, the proportion of taxpayers with a high elasticity is reduced. Optimal marginal tax rates that encapsulate composition effects are therefore larger than the ones derived when neglecting these effects. In our numerical example, these composition effects lead to a downward bias up to 8 percentage point around the 99.9 percentile.

Let us now consider another plausible scenario with Pareto parameters simply equal to 1.5 in both groups. From Equation (23), we know that in this case the weights $\hat{\mu}_i$ do not depend on income y anymore. Therefore, top tax rates are now independent of the income threshold. As illustrated in Figure 3, top tax rates are horizontal and, with our calibrations, $\tau_* = 59.9\%$. We now emphasize that composition effects also matter in this simple example. When increasing the tax rate from the actual $\tau_0 = 40\%$ to the optimal τ_*, tax payers in the high-elasticity group reduce their labor supply more than those in the low-elasticity group. As both within-group income densities are decreasing in income, this difference in behavioral responses leads to a relative increase of the proportion of tax payers in the low-elastic group at each top income level, thereby to a reduction of the mean income elasticity $\hat{\varepsilon}(y)$ according to (17b). The dashed line corresponds to the implementation of the optimal tax formula when one ignores the composition effect. It is a horizontal line lying at $1/(1 + 1.5 \times 0.5) = 57.1\%$. Neglecting composition effects leads to a 2,8 percentage downwards bias in the optimal tax rate.

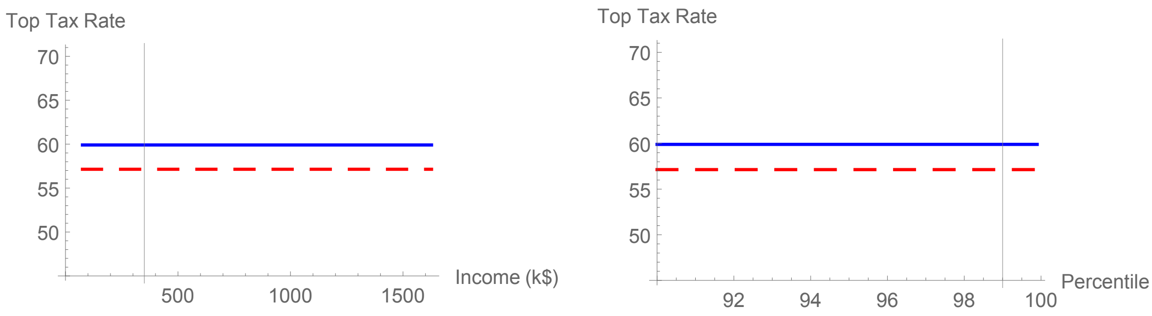


Figure 3: Optimal marginal tax rates for top incomes when composition effects are considered (solid lines) or not (dashed lines) and with $p_1 = p_2 = 1.5$. The gray vertical line correspond to the 99^{th} percentile.

These numerical results put the stress on the need for empirical studies on the description of the heterogeneity of elasticities and on the description of the income distributions across groups within the top percentile.

V Optimal structural tax formula

In this section, we derive additional results regarding optimal marginal tax rates by making the model more specific. Following Mirrlees (1971), we assume individuals preferences are additively separable:

Assumption 3. *The utility function is additively separable and takes the form:*

$$\mathscr{U}(c, y; w, \theta) = u(c) - v(y; w, \theta) \qquad \textit{with}: \quad u', v_y, v_{y,y} > 0 > v_w, v_{y,w} \quad , \quad u'' \leq 0$$

with $\lim_{w \mapsto 0} v_y(y; w) = +\infty$ *and* $\lim_{w \mapsto +\infty} v_y(y; w) = 0$

Assuming $u', v_y > 0 > v_w$ is necessary to retrieve $\mathscr{U}_c, \mathscr{U}_w > 0 > \mathscr{U}_y$ (our initial assumption). The convexity of the indifference curve associated to $\mathscr{U}$ under additively separable preferences is ensured by assuming that $v_{y,y} > 0 \geq u''$. Under Assumption 3, the marginal rate of substitution between pre-tax and after tax income is given by $\mathscr{M}(c, y; w, \theta) = v_y(y; w)/u'(c)$. Assumption 1 requires $v_{y,w} < 0$, $\lim_{w \mapsto 0} v_y(y; w) = +\infty$ and $\lim_{w \mapsto +\infty} v_y(y; w) = 0$. Therefore, Assumption 3 is more restrictive than Assumption 1.

Under Assumption 3, we first derive an optimal tax formula expressed in terms of the skill density and of direct (rather than total) behavioral elasticities. We also argue with this formula is more convenient for the numerical analysis than (19a). This new formula is also useful to show that optimal marginal tax rates are positive whenever the social objective is Utilitarian or Maximin.

Proposition 2. *Under Assumptions 2 and 3, the optimal structural tax formula verifies:*

$$\frac{T'(y)}{1-T'(y))} \int_{\theta\in\Theta} \frac{\varepsilon^*(W(y,\theta),\theta)}{\alpha^*(W(y,\theta),\theta)} W(y,\theta)\, f(W(y,\theta)|\theta)\, d\mu(\theta) \tag{24a}$$

$$= \quad u'(y-T(y)) \iint_{\theta\in\Theta, w\geq W(y,\theta)} \left(\frac{1}{u'(C(w,\theta))} - \frac{\Phi_U(U(w,\theta);w,\theta)}{\lambda}\right) f(w|\theta)dw\, d\mu(\theta)$$

for all income y and:

$$\iint_{\theta\in\Theta, w\in\mathbb{R}_+} \left(\frac{\Phi_U(U(w,\theta);w,\theta)}{\lambda} - \frac{1}{u'(C(w,\theta))}\right) f(w|\theta)dw\, d\mu(\theta) = 0. \tag{24b}$$

We prove Proposition 2 in Appendix A.4. For this purpose, we rewrite the optimal tax formula of Proposition 1 which is derived using the tax perturbation approach. In an appendix available upon request, we also proof Proposition 2 directly using an *allocation perturbation approach*. Note that the left-hand side of the sufficient statistics tax formula (19a) and the left-hand side of (24a) are equal even without assuming additive separability, as shown in Appendix A.4.

The optimal tax formula provided by Proposition 2 is much more convenient to implement numerically than the tax formula in terms of sufficient statistics provided by Proposition 1 (in Equation (19a)) and hence circumvents a significant limit of the latter. First, Equation (24a) is expressed in terms of the conditional skill density $f(\cdot|\theta)$ which is policy-invariant, while Equation (19a) is expressed in terms of the conditional income density $h(\cdot|\theta)$ which depends on the tax schedule. Second, Equation (24a) makes use of direct elasticities, which are the ones obtained in the empirical literature (see Footnote 11). Equation (24a) is a second-order differential equation that depends simply on the tax function and on its first-order derivative (i.e. the marginal tax rate). In contrast, the left-hand side of (19a) displays total elasticities (hence the second-order derivative of the tax function).[17] Therefore, to implement this equation numerically, one has to deal not only with the first order derivative of tax function (i.e. with the marginal tax rate) but also with the second-order derivative T'' that appears in Equations (17b) and (17c). Equation (19a) is thus a third-order differential equation. As such, it is much less convenient to implement numerically than the second-order differential equation (24a). Third, additive separability (Assumption 3) allows one to rewrite the income effects $\hat{\eta}(z)\, T'(z)$ in Equation (19a) in terms of the derivatives of the individual utility of consumption, $u'(\cdot)$, and of the derivative of welfare with respect to utility, Φ_u. Again, these two functions are policy-invariant which is extremely convenient for the numerical implementation.[18]

[17] More precisely, what appears in Equation (24a) is the ratio of the direct compensated elasticity ε to the direct skill density α. According to (16), the skill elasticity α shows up because the left-hand side is expressed in terms of the (policy-invariant) conditional skill density $f(\cdot|\theta)$ instead of the endogenous conditional income density $h(\cdot|\theta)$ found in Equation (19a). Note that, according to Equations (15a) and (15c), the total compensated ε and skill α elasticities differ from their direct counterparts ε^* and α^* by the same corrective term. Hence, the ratio of ε/α is equal to the ratio of the direct compensated elasticity ε^* to the direct skill elasticity α^*.

[18] Of course, (24a) and (24b) are not a closed-form solution. It depends on the arguments of functions ε^*/α^*, $u'(\cdot)$ and Φ_u, and are thus functions of the allocation $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$. Therefore one needs to iterate when implementing tax formula (24a). For given values of ε^*/α^*, $u'(\cdot)$ and Φ_u, Equations (24a) and (24b) provide an

Thanks to Proposition 2, one can now provide a sufficient condition for positive marginal tax rates.

Proposition 3. *Under Assumptions 1, 4 and 3, and utilitarian or maximin social preferences, optimal marginal tax rates are positive*

With multidimensional heterogeneity, the literature has highlighted that negative marginal tax rates can be optimal. In Boadway et al. (2002), Choné and Laroque (2010) and Lockwood and Weinzierl (2015), individuals differ along their skills and preferences for effort, and the social planner has weighted utilitarian preferences. In this context, individuals who pool at the same income level are characterized by different social marginal welfare weights and the mean social welfare weight may not be decreasing with income.[19] This *composition effect* may reduce marginal tax rates (Lockwood and Weinzierl, 2015) and may even induce them to become negative (Boadway et al., 2002, Choné and Laroque, 2010).[20] Proposition 3 shows that optimal marginal tax rates are positive as soon as all individuals who earn the same income y are characterized by the same marginal utility of consumption $\mathscr{U}_c$, which is ensured by the additive separability assumption, and by the same marginal social welfare Φ_u, which is ensured by utilitarian or maximin social objective. In such a case, all individuals who earn the same income are characterized by the same welfare weights. Therefore, the cause of negative marginal tax rates emphasized in Boadway et al. (2002), Choné and Laroque (2010) and Lockwood and Weinzierl (2015) does not apply. Proposition 3 generalizes, to the multidimensional case, Mirrlees (1971)'s result of positive optimal tax rates (which was obtained under additively separable preferences).

VI Tax perturbation method versus first-order mechanism design approach

Piketty (1997) and Saez (2001) were the first to derive the optimal income tax formula using the tax perturbation approach. However, the tax perturbation is valid only under some circumstances (see our Assumption 2). In particular, because of the nonlinearity of the tax schedule, a small tax reform may well induce a jump in the labor supply for taxpayers that were before the reform indifferent between two labor supply choices. That was the reason why Saez (2001) had to show that his formula was consistent with the one of Mirrlees (1971). In the latter, the government optimizes over incentive-compatible allocations and the optimal tax schedule is the one that decentralizes this schedule. However, Saez (2001) shows that both approaches

approximation for the tax schedule. This, in turn, provides an approximation of the optimal allocation by solving maximization program (1). This allows one to update the evaluations of ε^*/α^*, $u'(\cdot)$ and Φ_u. However, because the second-order derivatives of the tax function do not show up in (24a) and (24b), this iteration process is much easier than the one that hinges on the sufficient statistics formula (Equations (19a) and (19b)).

[19] However, in these papers, individuals who pool at the same income are characterized by the same compensated elasticity and income effects.

[20] For instance, this happens when some groups undervalued in the social objective are overrepresented at low income levels. In this case, individuals at the bottom of the income distribution receive lower social welfare weights than individuals with larger income levels. This yields negative marginal tax rates at the bottom of the distribution.

lead to the same formulas assuming one-dimensional heterogeneity. In this section, we discuss the connection between both approaches in the multidimensional context under Assumption 1. We show the equivalence between the tax perturbation approach under Assumption 2 and the well-established first-order mechanism design approach.

The mechanism design approach relies on the taxation principle (Hammond, 1979, Guesnerie, 1995) according to which it is equivalent for the government to select a nonlinear tax schedule taking into account the labor supply decisions as the ones described in (1), or to directly select an *incentive-compatible* allocation $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$ that verifies the self-selection constraints,

$$\forall w, \theta, w', \theta' \in (\mathbb{R}_+^* \times \Theta)^2 \qquad \mathscr{U}\left(C(w,\theta), Y(w,\theta); w, \theta\right) \geq \mathscr{U}\left(C(w',\theta'), Y(w',\theta'); w, \theta\right). \quad (25)$$

According to (25), individuals of type (w,θ) are better of with the bundle $(C(w,\theta), Y(w,\theta))$ designed for them than with bundles $(C(w',\theta'), Y(w',\theta'))$ designed for individuals of any other type (w',θ').

In the mechanism design approach, it is usual to assume that the government selects among incentive compatible allocations that are continuously differentiable. Then, incentive constraints (25) imply the first-order incentive constraints, i.e.

$$\forall (w,\theta) \in \mathbb{R}_+^* \times \Theta \qquad \dot{U}(w,\theta) = \mathscr{U}_w\left(C(w,\theta), Y(w,\theta); w, \theta\right) \quad (26)$$

The first-order incentive constraints (26) are necessary but not sufficient to verify the incentive constraints (25). A sufficient condition is that the allocation also verify a monotonicity constraint according to which in each group, $Y(\cdot,\theta)$ is nondecreasing in skill. We adopt a slightly more restrictive assumption.

Assumption 4. *The allocation $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$ is smooth if and only if it is continuously differentiable, it verifies (25) and $w \mapsto Y(w,\theta)$ admits a positive derivative for any group $\theta \in \Theta$ and at any skill level $w \in \mathbb{R}_+^*$.*

We get the following connection between Assumption 2 required for the tax perturbation approach and Assumption 4 in the first-order mechanism design approach (see Appendix A.6).

Proposition 4. *Under Assumption 1,*

i) Any tax schedule $y \mapsto T(\cdot)$ verifying Assumption 2 induces a smooth allocation that verifies Assumption 4.

ii) Any smooth allocation verifying Assumption 4 can be decentralized by a tax schedule that verifies Assumption 2.

Intuitively, under Assumption 1 (which states the single-crossing condition within group), elements of Assumptions 2 and 4 are equivalent. The fact that, for each group θ, the second-order condition of the individual program (1) holds strictly (Part *ii* of Assumption 2) is equivalent to $Y(\cdot,\theta)$ being strictly increasing in skill as required in Assumption 4. Moreover, the uniqueness of the global maximum from the individual maximization program (1) (Part *iii* of Assumption 2) is equivalent to $Y(\cdot, \theta)$ being continuous in skill as stated in Assumption 4.

Thanks to Proposition 4, the first-order mechanism design approach and the tax perturbation one are analog. The first-order mechanism design approach, pioneered by Mirrlees (1971), consists in choosing, among the allocations that verify Assumption 4, the one that maximizes the social objective (6) subject to the budget constraint (5). For this purpose, it computes the first-order effect, on the Lagrangian (7), of perturbing the optimal allocation. However, as the perturbed allocation has to verify Assumption 4, it is decentralized by a perturbed tax schedule that has to verify Assumption 2. Proposition 4 therefore assesses that it is equivalent to consider the effects of a perturbation of the allocation that preserves Assumption 4 (i.e. the first-order mechanism design approach), or to consider the responses of the allocation to a perturbation of the tax function that preserves Assumption 2. In other words, the first-order mechanism design approach focuses on the effects of an allocation perturbation while the tax perturbation approach focuses on the effects of the tax reform that decentralizes this perturbation. This is the reason why we claim that the first-order mechanism design approach and the tax perturbation one are actually the two faces of the same coin.

In the literature where the unobserved heterogeneity is one dimensional, the mechanism design approach can be developed under less restrictive assumptions than Assumption 4. In particular, Lollivier and Rochet (1983), Guesnerie and Laffont (1984), Ebert (1992), Boadway et al. (2000) study the case where individuals endowed with different skill levels choose the same consumption-income bundle. To decentralize such an allocation where *bunching* occurs, one would need a kink in the tax function. Therefore, bunching is never possible whenever a tax schedule is twice-continuously differentiable (i.e when the marginal tax rate is a continuously differentiable function of income) as required by Assumption 2. Note that the alternative "pathology" where individuals may be indifferent between two types of incomes appears much more plausible under twice continuously differentiable tax schedule. Surprisingly, this problem has attracted much less attention than bunching in the literature based on the mechanism design approach, a noticeable exception being Hellwig (2010).

VII Concluding Comments

In this paper, we have proposed a new structural method, based on an allocation perturbation, to derive optimal tax schedules and their optimal sufficient statistics, in the very general case where agents are heterogeneous in many dimensions. After contrasting this method with the usual tax perturbation approach, we have quantified the bias in marginal tax rates entailed by using observed sufficient statistics rather than the optimal ones. Using US data, we have shown that, even in a simple illustration, this bias can reach up to 10 percentage points. Our structural tax formula allows us to avoid such a bias and is necessary to correct the observed sufficient statistics.

To illustrate the generality of our framework, we have provided four possible interpretations of our tax formulae: income taxation with heterogeneous skills and heterogeneous labor supply elasticities, joint taxation of labor and non-labor income, joint income taxation of couples and income taxation with tax avoidance. More generally, our approach applies to any taxation problem in which the tax function depends on as many different sources of income as one wishes. It even extends beyond optimal taxation, e.g., to nonlinear pricing problems where consumers differ along several unobserved dimensions. We intend to implement these applications on real data in our future research.

A Theoretical Proofs

A.1 Proof of Lemma 1

Let $\mathscr{L}^R$ be the Lagrangian resulting from applying a reform of magnitude m in the direction R to the Lagrangian (7):

$$\mathscr{L}^R(m) \stackrel{\text{def}}{\equiv} \iint\limits_{\theta\in\Theta, w\in\mathbb{R}_+} \left[T(Y^R(m;w,\theta)) - m\, R(Y^R(m;w,\theta)) + \frac{\Phi\left(U^R(m;w,\theta);w,\theta\right)}{\lambda}\right] f(w|\theta)dw\, d\mu(\theta)$$

Computing the Gateaux differential of the Lagrangian with respect to m at $m=0$ yields:

$$\begin{aligned}
\frac{\partial \mathscr{L}^R}{\partial m} &= \iint\limits_{\theta\in\Theta, w\in\mathbb{R}_+} \left\{\frac{T'(Y(w,\theta))}{1-T'(Y(w,\theta))} Y(w,\theta)\, \varepsilon(w,\theta)\, R'(Y(w,\theta))\right.\\
&+ \left.\left[T'(Y(w,\theta))\,\eta(w,\theta) - 1 + g(w,\theta)\right] R(Y(w,\theta))\right\} f(w|\theta)dw\, d\mu(\theta)\\
&= \iint\limits_{\theta\in\Theta, y\in\mathbb{R}_+} \left\{\frac{T'(y)}{1-T'(y)} y\, \varepsilon\left(W(y,\theta),\theta\right)\, R'(y)\right.\\
&+ \left.\left[T'(y)\,\eta\left(W(y,\theta),\theta\right) - 1 + g\left(W(y,\theta),\theta\right)\right] R(y)\right\} h(y|\theta)dy\, d\mu(\theta)\\
&= \int\limits_{y\in\mathbb{R}_+} \left\{\frac{T'(y)}{1-T'(y)} y\, \hat{\varepsilon}(y)\, R'(y) + \left[T'(y)\,\hat{\eta}(y) - 1 + \hat{g}(y)\right] R(y)\right\} h(y)dy
\end{aligned}$$

We use (8), (10) and (14) to obtain the first equality. We use (16) for the change of variable from skill w to income y in the second equality. We use (17a)-(17d) for the last equality. Integrating by parts the integral of $\frac{T'(y)}{1-T'(y)} y\, \hat{\varepsilon}(y)\, \hat{h}(y)\, R'(y)$ leads to (18).

We now show that the first-order effect on the Lagragian (7) of a reform of magnitude m in the direction $R(\cdot)$ is positively proportional to the first-order effect on the social objective (6) of the reform denoted $\tilde{R}(m)$ which consists in implementing the tax reform in the direction $R(\cdot)$ with magnitude m and to rebate in a lump-sum way the induced net budget surplus. Let $\ell(m)$ denote this budget surplus. Under the balanced-budget tax reform $\tilde{R}(m)$ individuals solves:

$$U^{\tilde{R}}(m;w,\theta) \stackrel{\text{def}}{\equiv} \max_y \qquad \mathscr{U}\left(y - T(y) + m\, R(y) + \ell(m), y; w, \theta\right) \tag{27}$$

Applying the envelope theorem to (27) at $m=0$ leads to:

$$\frac{\partial U^{\tilde{R}}}{\partial m}(0;w,\theta) = \left(R(y) + \ell'(0)\right)\ \mathscr{U}_c\left(C(w,\theta), Y(w,\theta); w,\theta\right) \tag{28}$$

Applying the implicit function theorem on the first-order condition

$$1 - T'(y) + m\, R'(y) = \mathscr{M}\left(y - T(y) + m\, R(y) + \ell(m), y; w, \theta\right)$$

at $y = Y^{\tilde{R}}(m; w, \theta)$ and using (12b), (13b) and (14) leads to:

$$\frac{\partial Y^{\tilde{R}}}{\partial m}(0; w, \theta) = \frac{\partial Y^{R}}{\partial m}(0; w, \theta) + \eta(w, \theta)\, \ell'(m) \tag{29}$$

We now denote respectively $\mathscr{B}^R(m)$, $\mathscr{S}^R(m)$ and $\mathscr{L}^R(m)$ the budget surplus, the social objective and the Lagrangian when the tax function is perturbed in the direction R as a function of the magnitude m with $\mathscr{L}^R(m) = \mathscr{B}^R(m) + (1/\lambda)\mathscr{S}^R(m)$. We symmetrically denote $\mathscr{B}^{\tilde{R}}(m)$, $SWF^{\tilde{R}}(m)$ and $\mathscr{L}^{\tilde{R}}(m)$ the budget surplus, the social objective and the Lagrangian when the tax function is perturbed by the balanced-budget tax reform in the direction R and magnitude m. We get $\mathscr{L}^{\tilde{R}}(m) = \mathscr{B}^{\tilde{R}}(m) + (1/\lambda)\mathscr{S}^{\tilde{R}}(m)$. We also get that

$$0 = \mathscr{B}^{\tilde{R}}(m) = \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} \left\{T\left(Y^{\tilde{R}}(m; w, \theta)\right) - m\, R\left(Y^{\tilde{R}}(m; w, \theta)\right)\right\} f(w|\theta) dw\, d\mu(\theta) - \ell(m)$$

We thus get

$$\ell'(0) = \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} \left\{T'(Y(w,\theta))\, \frac{\partial Y^{\tilde{R}}}{\partial m}(0; w, \theta) - R(Y(w,\theta))\right\} f(w|\theta) dw\, d\mu(\theta)$$

Using (29), we thus get:

$$\ell'(0) = \frac{\partial \mathscr{B}^R}{\partial m}(0) + \ell'(0) \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta) dw\, d\mu(\theta)$$

so that:

$$\ell'(0) = \frac{1}{1 - \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta) dw\, d\mu(\theta)}\, \frac{\partial \mathscr{B}^R}{\partial m}(0) \tag{30}$$

Finally, using (28), we get:

$$\begin{aligned}
\frac{\partial \mathscr{S}^{\tilde{R}}}{\partial m}(0) &= \frac{\partial \mathscr{S}^R}{\partial m}(0) + \ell'(0) \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} \Phi'_u(U(w,\theta); w, \theta)\, \mathscr{U}_c(C(w,\theta), Y(w,\theta); w, \theta)\, f(w|\theta) dw\, d\mu(\theta) \\
&= \frac{\partial \mathscr{S}^R}{\partial m}(0) + \frac{\int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} \Phi'_u(U(w,\theta); w, \theta)\, \mathscr{U}_c(C(w,\theta), Y(w,\theta); w, \theta)\, f(w|\theta) dw\, d\mu(\theta)}{1 - \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta) dw\, d\mu(\theta)} \frac{\partial \mathscr{B}^R}{\partial m}(0) \\
&= \lambda\, \frac{\partial \mathscr{L}^R}{\partial m}(0)
\end{aligned} \tag{31}$$

where the last equality holds if and only if

$$\lambda = \frac{\int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} \Phi'_u(U(w,\theta); w, \theta)\, \mathscr{U}_c(C(w,\theta), Y(w,\theta); w, \theta)\, f(w|\theta) dw\, d\mu(\theta)}{1 - \int\limits_{(w,\theta)\in\mathbb{R}_+^*\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta) dw\, d\mu(\theta)} \tag{32}$$

A.2 Proof of Proposition 1

An optimal tax system implies that any tax reform $R(.)$ does yield zero first-order effect on the Lagrangian (7), i.e. (18) is nil at $m = 0$ for all directions $R(\cdot)$. This implies that $\lim_{y \mapsto 0} \frac{T'(y)}{1-T'(y)} \hat{\varepsilon}(y)\, y\, \hat{h}(y) = \lim_{y \mapsto \infty} \frac{T'(y)}{1-T'(y)} \hat{\varepsilon}(y)\, y\, \hat{h}(y) = 0$ and, for all income y:

$$\frac{d}{dy}\left[\frac{T'(y)}{1-T'(y)}\hat{\varepsilon}(y)\, y\, \hat{h}(y)\right] = \left[\hat{g}(y) - 1 + T'(y)\, \hat{\eta}(y)\right] \hat{h}(y)$$

Integrating the latter equality for all income z above y and using $\lim_{y \mapsto \infty} \frac{T'(y)}{1-T'(y)} \hat{\varepsilon}(y)\, y\, \hat{h}(y) = 0$ yields (19a). Making y tends to 0 in (19a) and using $\lim_{y \mapsto 0} \frac{T'(y)}{1-T'(y)} \hat{\varepsilon}(y)\, y\, \hat{h}(y) = 0$ leads to (19b).

A.3 Numerical implementation of the Top Tax Rates formula

We now describe how the income density of each group varies between the actual and optimal economy. Variables in the actual economy are labeled with a subscript zero and with an asterisk at the optimum. In the actual economy, the top tax rate is a constant τ_0.

From the first-order condition (2), an individual belonging to group θ earning income y in the optimal economy is endowed with a skill level equal to $W(y,\theta) = y^{\frac{1}{1+\theta}}\,(1-T'(y))^{-\frac{\theta}{1+\theta}}$ and earns income

$$\tilde{Y}_0(y,\theta) = \left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta} y.$$

in the actual economy. From this equation, we can write

$$H_*(y|i) = H_0(\tilde{Y}_0(y,\theta_i)|i) = H_0\left(\left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta_i} y \,\middle|\, i\right) \tag{33}$$

Differentiating both sides of (33) in y, we get:

$$y\, h(y|i) \quad = \quad \left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta_i} y\, h_0\left(\left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta_i} y \,\middle|\, i\right)\left[1 + \frac{\theta_i\, y\, T''(y)}{1-T'(y)}\right]$$

Using Equation (15a) and the fact that θ_i is the direct compensated elasticity within group i, we get:

$$\varepsilon(W(y,\theta_i),i)\, y\, h(y|i) = \theta_i \left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta_i} y\, h_0\left(\left(\frac{1-\tau_0}{1-T'(y)}\right)^{\theta_i} y \,\middle|\, i\right)$$

Using (22) we obtain:

$$\varepsilon(W(y,\theta),i)\, y\, h(y|i) = \theta_i\, p_i\, k_i \left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_i\, p_i} y^{-p_i}$$

The denominator of the Right-Hand side of (20) is thus equal to:

$$\hat{\varepsilon}(y)\, y\, \hat{h}(y) = \mu(\theta_1)\, k_1 \left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_1 p_1} y^{-p_1}\, \theta_1\, p_1 + \mu(\theta_2)\, k_2 \left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_2 p_2} y^{-p_2}\, \theta_2\, p_2 \tag{34}$$

Using (33), the numerator of the Right-Hand side of (20) is thus equal to:

$$1 - \hat{H}(y) = \mu(\theta_1)\, k_1 \left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_1 p_1} y^{-p_1} + \mu(\theta_2)\, k_2 \left(\frac{1-T'(y)}{1-\tau_0}\right)^{\theta_2 p_2} y^{-p_2} \tag{35}$$

Plugging (34) and (35) into (20) and solving for $T'(y)$ leads to (23).

In Figure 2, we plot the solution of this equation, for different values of income y and $\tau_0 = 40\%$. Parameters $\mu(\theta_1)$ and $\mu(\theta_2)$ are set such that at the income threshold $y = \$1,053,398$, which corresponds to the top 1% in the actual economy, both groups are equally large so that:

$$\mu(\theta_1)\, k_1\, y^{-p_1} = \mu(\theta_2)\, k_2\, y^{-p_2}.$$

A.4 Proof of Proposition 2

Let

$$X(y) \stackrel{\text{def}}{\equiv} \int_y^\infty \left\{1 - \hat{g}(z) - \hat{\eta}(z)\, T'(z)\right\} \hat{h}(z) dz \tag{36}$$

denote the right-hand side of (19a). Equation (19a) can be rewritten as:

$$\frac{T'(y)}{1 - T'(y)} \hat{\varepsilon}(y)\, y\, \hat{h}(y) = X(y) \tag{37}$$

Combine (17a) and (17b) to write:

$$\hat{\varepsilon}(y)\, y\, \hat{h}(y) = \int\limits_{\theta\in\Theta} \varepsilon\left(W(y,\theta),\theta\right)\, y\, h(y|\theta)\, d\mu(\theta) = \int\limits_{\theta\in\Theta} \frac{\varepsilon\left(W(y,\theta),\theta\right)}{\alpha\left(W(y,\theta),\theta\right)} W(y,\theta)\, f\left(W(y,\theta)|\theta\right)\, d\mu(\theta)$$

where (16) is used to obtain the second equality. Plug this equality in the left-hand side of (37) leads to:

$$\frac{T'(y)}{1 - T'(y)} \int\limits_{\theta\in\Theta} \frac{\varepsilon^*(W(y,\theta),\theta)}{\alpha^*(W(y,\theta),\theta)}\, W(y,\theta)\, f(W(y,\theta)|\theta)\, d\mu(\theta) = X(y) \tag{38}$$

From Assumption 3, one has $\mathscr{M}(c,y;w,\theta) = v_y(y;w,\theta)/u'(c)$, so we get that $\mathscr{M}_c(c,y;w,\theta)/\mathscr{M}(c,y;w,\theta)$ simplifies to $-u''(c)/u'(c)$. Therefore, according to Equations (13a) and (13b), we get:

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

Using (17b) and (17c) and the fact that individuals of different types (w,θ) who earn the same income y have to consume the same amount c, we get:

$$\hat{\eta}(y) = \hat{\varepsilon}(y)\, y\, \frac{u''(y - T(y))}{u'(y - T(y))} \tag{39}$$

We now define $J(y)$ by the equality $X(y) \stackrel{\text{def}}{\equiv} J(y)\, u'(y - T(y))$. We get:

$$\begin{aligned}
X'(y) &= J'(y)\, u'(y - T(y)) + X(y) \frac{u''(y - T(y))}{u'(y - T(y))}(1 - T'(y)) \\
&= J'(y)\, u'(y - T(y)) + T'(y)\, \hat{\varepsilon}(y)\, y\, \hat{h}(y) \frac{u''(y - T(y))}{u'(y - T(y))} \\
&= J'(y)\, u'(y - T(y)) + T'(y)\, \hat{\eta}(y)\, \hat{h}(y)
\end{aligned}$$

where the second equality uses (37) and the last uses (39). Differentiating in income both sides of (36) leads to:

$$\begin{aligned}
J'(y)\, u'(y - T(y)) + T'(y)\, \hat{\eta}(y)\, \hat{h}(y) &= \left\{-1 + \hat{g}(y) + T'(y)\hat{\eta}(y)\right\} \hat{h}(y) \\
J'(y)\, u'(y - T(y)) &= \left\{-1 + \hat{g}(y)\right\} \hat{h}(y) \\
J'(y) &= \int_{\theta\in\Theta} \left\{-\frac{1}{u'(y - T(y))} + \frac{\Phi_u\langle W(y,\theta),\theta\rangle}{\lambda}\right\} h(y|\theta)\, d\mu(\theta)
\end{aligned}$$

where the last equality uses (17a) and (17d). Integrating the last equality for all incomes z above y leads to:

$$\begin{aligned} J(y) &= \iint_{z\geq y,\theta\in\Theta}\left\{\frac{1}{u'(z-T(z))}-\frac{\Phi_u\langle W(z,\theta),\theta\rangle}{\lambda}\right\}h(z|\theta)dz\,d\mu(\theta) \\ &= \iint_{w\geq W(y,\theta),\theta\in\Theta}\left\{\frac{1}{u'(C(w,\theta))}-\frac{\Phi_u\langle w,\theta\rangle}{\lambda}\right\}f(w|\theta)dw\,d\mu(\theta) \end{aligned} \tag{40}$$

where the second equality uses (16) to change variables from income to skill. Plugging (40) in (38) leads to (24a). Using (36), Equation (19b) can be rewritten as $X(0)=0$, thereby $J(0)=0$. Using (40) leads to (24b).

A.5 Proof of Proposition 3

Under utilitarian preferences, $\Phi_u=1$ and we get:

$$J(y)\stackrel{\text{def}}{\equiv}\int_{z\geq y}\left(\frac{1}{u'(z-T(z))}-\frac{1}{\lambda}\right)\left(\int_\theta f\left(W(z,\theta)|\theta\right)d\mu\left(\theta\right)\right)\,dz$$

The derivative of $J(z)$ has the sign of $1/\lambda-1/u'(z-T(z))$, which is decreasing in w because of the concavity of $u(\cdot)$. Moreover, $\lim_{y\mapsto\infty}J(y)=0$ and Equation (24b) imply that $J(0)=0$. Therefore, $J(\cdot)$ first increases and then decreases. It is thus positive for all (interior) skill levels. Since $v_{yw}<0$ from (1), optimal marginal tax rates are positive.

Under Maximin, one has $U(x,\theta)>U(0,\theta)$ for all $x>0$ from (26). Therefore, within each group, the most deserving individuals are those whose skill $w=0$. The Maximin objective implies $\Phi_U\langle x,\theta\rangle=0$ for all $x>0$. Thereby,

$$J(y)\stackrel{\text{def}}{\equiv}\int_{z\geq y}\frac{1}{u'(z-T(z))}\left(\int_\theta f\left(W(z,\theta)|\theta\right)d\mu\left(\theta\right)\right)\,dz$$

for all $x>0$, which leads to positive marginal tax

A.6 Proof of Proposition 4

Part *i)* of Proposition 4.

Let $T(\cdot)$ be an income tax schedule satisfying Assumption 2. We already know that under Assumptions 1 and 2, one can apply the implicit function theorem to the first-order condition associated to (1). This implies that $Y(\cdot,\theta)$, thereby $C(\cdot,\theta)$ is continuously differentiable in w within each group θ. Moreover, $Y(\cdot,\theta)$ admits a positive derivative according to (13c). Finally, from (1) we get that:

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

Taking $y'=Y(w',\theta')$ leads to $C(w',\theta')=y'-T(y')$, so that the last inequality leads to (25). Therefore the allocation $w\mapsto(C(\cdot,\theta),Y(\cdot,\theta))$ induced by $T(\cdot)$ verifies (25), thereby Assumption 4.

Part *ii)* of Proposition 4

Let $(w,\theta)\mapsto(C(w,\theta),Y(w,\theta))$ be a mapping defined over $\mathbb{R}_+^*\times\Theta$ which verifies Assumption 4. Let $\mathbb{Y}$ denote the set of incomes that are assigned to some individuals along this allocation. To define the tax schedule that decentralizes this allocation, we first show that if two types (w,θ) and (w',θ') of individuals earn the same income $y=Y(w,\theta)=Y(w',\theta')$, then they have to be assigned the same consumption $C(w,\theta) = C(w',\theta')$. Otherwise, if by contradiction one has: $C(w,\theta) < C(w',\theta')$, then one would get that individuals of type (w,θ) would be better of with the bundle $(C(w'), Y(w'))$ designed for individuals of type (w',θ'), which would be in contradiction with (25). A symmetric argument applies if $C(w,\theta) < C(w',\theta')$ by inverting the role of (w,θ) and of (w',θ'). We can then unambiguously define the tax schedule denoted $T(\cdot)$ that decentralizes this allocation by:

$$\forall y \in \mathbb{Y} \qquad T(y) \stackrel{\text{def}}{\equiv} Y(w,\theta) - C(w,\theta) \qquad \text{where } (w,\theta) \text{ are such that: } y = Y(w,\theta) \tag{41}$$

Given this tax schedule, Program (1) of individuals of type (w,θ) is equivalent to solving : $\max_{(w',\theta') \in \mathbb{R}_+^* \times \Theta} \mathscr{U}\left(C(w',\theta'), Y(w',\theta'); w,\theta\right)$, whose solution is (w,θ) as $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$ verifies the incentive constraints (25). Therefore, the tax schedule $T(\cdot)$ defined by (41) decentralizes the given allocation.[21]

We then need to show a mathematical result. For each group $\theta \in \Theta$, as $Y(\cdot,\theta)$ is continuously differentiable, it admits a reciprocal denoted $Y^{-1}(\cdot,\theta)$ which is also continuously differentiable. Therefore the image of the (open) skill set $\mathbb{R}_+^*$ by $Y(\cdot,\theta)$ is an open set denoted $\mathbb{Y}(\theta) \subset \mathbb{R}_+$. Equation (41) can be rewritten on $\mathbb{Y}(\theta)$ by:

$$T(y) = y - C\left(Y^{-1}(y,\theta),\theta\right) \tag{42}$$

Moreover, we get that $\mathbb{Y} = \cup_{\theta\in\Theta}\mathbb{Y}(\theta)$ and is therefore an open set. Therefore, for each income $y \in \mathbb{Y}$, there exists a group θ such that $T(\cdot)$ verifies (42) in the neighborhood of y.

To show that $T(\cdot)$ verifies Part *i*) of Assumption 2, we first notice that from (42), $T(\cdot)$ is continuously differentiable as $Y^{-1}(\cdot,\theta)$ and $C(\cdot,\theta)$ are continuously differentiable. Moreover, we have from (2) that:

$$T'(y) = 1 - \mathscr{M}(y - T(y), y; Y^{-1}(w,\theta),\theta)$$

As $T(\cdot)$ and $Y^{-1}(\cdot,\theta)$ are continuously differentiable in y, and $\mathscr{M}(\cdot,\cdot;\cdot,\theta)$ is continuously differentiable in (c,y,w), $y \mapsto \mathscr{M}(y - T(y), y; Y^{-1}(w,\theta),\theta)$ is continuously differentiable, so $T'(\cdot)$ is continuously differentiable and $T(\cdot)$ verifies Part *i*) of Assumption 2..

To show that $T(\cdot)$ verifies Part *ii*) of Assumption 2,[22] we notice that the first-order condition (11) can be rewritten as $\mathscr{Y}\left(Y(w,\theta); w,\theta\right) \equiv 0$ for all skill levels. Differentiating this equality with respect to skill leads to: $\mathscr{Y}_y\left(Y(w,\theta); w,\theta\right)\ \dot{Y}(w,\theta) + \mathscr{Y}_w\left(Y(w,\theta); w,\theta\right) = 0$. As $\mathscr{Y}_w\left(Y(w,\theta); w,\theta\right) = -\mathscr{M}_w\left(C(w,\theta), Y(w,\theta); w,\theta\right)$ which is positive from Assumption 1 and $\dot{Y}(w,\theta) > 0$ from Assumption 4, then one must have $\mathscr{Y}_y\left(Y(w,\theta); w,\theta\right) < 0$, which is Part *ii*) of Assumption 2.

To show that $T(\cdot)$ verifies Part *iii*) of Assumption 2, we assume by contradiction that individuals of type (w^*,θ) are indifferent between earning income $Y(w^*,\theta)$ and earning an income level denoted $y' \in \mathbb{Y}$. We show that in such a case, some individuals with skill w close to w^* are better of with the bundle $(y' - T(y'), y')$ than with the bundle $(C(w,\theta), Y(w,\theta))$ designed for them, a contradiction. For this purpose we denote $\mathscr{C}(u,y;w,\theta)$ the after-tax income an individual of type (w,θ) should get to enjoy utility u while earning an pre-tax income y. Function $\mathscr{C}(\cdot,y;w,\theta)$ is the reciprocal of function $\mathscr{U}(\cdot,y;w,\theta)$ We get: $\mathscr{C}_u = 1/\mathscr{U}_c$, $\mathscr{C}_y = -\mathscr{U}_y/\mathscr{U}_c = \mathscr{M}$ and $\mathscr{C}_w = -\mathscr{U}_w/\mathscr{U}_c$. We then denote:

$$\mathscr{Q}(w) \stackrel{\text{def}}{\equiv} \mathscr{C}\left(U(w,\theta), y'; w,\theta\right) - y' + T(y')$$

To be indifferent between earning income $Y(w,\theta)$ and income y', individuals of type (w,θ) have to receive after-tax income $\mathscr{C}\left(U(w,\theta), y'; w,\theta\right)$ when earnings income y'. Therefore, $\mathscr{Q}(w)$ is a

[21] We have here followed Hammond (1979) very closely.

[22] We are grateful to Kevin Spiritus for encouraging us to emphasize this result.

measure in monetary units of the difference in well being for individuals of type (w,θ) between the bundle $(C(w,\theta),Y(w,\theta))$ designed for them (which leads them with utility $U(w,\theta)$) and the utility they would get by earning income y' and consuming $y'-T(y')$. We have by assumption $\mathscr{Q}(w^*)=0$. Moreover, we get:

$$\mathscr{Q}'(w)=\frac{\mathscr{V}\left(U(w,\theta),Y(w,\theta),w,\theta\right)-\mathscr{V}\left(U(w,\theta),y',w,\theta\right)}{\mathscr{U}_c\left(\mathscr{C}(U(w,\theta),Y(w,\theta);w,\theta),Y(w,\theta);w,\theta\right)}$$

where $\mathscr{V}(u,y;w,\theta)\stackrel{\text{def}}{\equiv}\mathscr{U}_w\left(\mathscr{C}(u,y;w,\theta),y;w,\theta\right)$ describes how $\mathscr{U}_w$ varies with income y along the indifference curve of individuals of type (w,θ) at utility u. We get that $\mathscr{V}_y=-\mathscr{U}_c\,\mathscr{M}_w$ which is strictly positive from Assumption 1. Therefore:

- If $y'>Y(w,\theta)$, then $\mathscr{Q}'(w^*)>0$, which implies that for some skills $w>w^*$ above w^* and sufficiently close to w^*, $\mathscr{Q}(w)>0$, i.e. $U(w,\theta)>\mathscr{U}(y'-T(y'),y';w,\theta)$. Therefore, individuals of type (w,θ) strictly prefers the bundle $(y'-T(y'),y')$ rather than the bundle $(C(w,\theta),Y(w,\theta)$ designed for them, a contradiction.
- If $y'<Y(w,\theta)$, then $\mathscr{Q}'(w^*)<0$, which implies that for some skills $w<w^*$ below w^* and sufficiently close to w^*, $\mathscr{Q}(w)>0$, i.e. $U(w,\theta)>\mathscr{U}(y'-T(y'),y';w,\theta)$. Therefore, individuals of type (w,θ) strictly prefers the bundle $(y'-T(y'),y')$ rather than the bundle $(C(w,\theta),Y(w,\theta)$ designed for them, a contradiction.

References

Auten, G. and R. Carroll (1999). The Effect Of Income Taxes On Household Income. *The Review of Economics and Statistics 81*(4), 681–693.

Blumkin, T., E. Sadka, and Y. Shem-Tov (2014). International tax competition: zero tax rate at the top re-established. *International Tax and Public Finance, Forthcoming*.

Boadway, R., K. Cuff, and M. Marchand (2000). Optimal income taxation with quasi-linear preferences revisited. *Journal of Public Economic Theory 2*(4), 435–60.

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.

Brett, C. and J. A. Weymark (2003). Financing education using optimal redistributive taxation. *Journal of Public Economics 87*(11), 2549–2569.

Chetty, R. (2009). Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods. *Annual Review of Economics 1*(1), 451–488.

Choné, P. and G. Laroque (2010). Negative marginal tax rates and heterogeneity. *American Economic Review 100*(5), 2532–47.

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.

Feldstein, M. (1995). Behavioral responses to tax rates: Evidence from tra86. *American Economic Review 85*, 170–174.

Fleurbaey, M. and F. Maniquet (2011). *Theory of Fairness and Social Welfare.* Cambridge: Cambride University Press.

Fleurbaey, M. and F. Maniquet (2015). Optimal taxation theory and principles of fairness. CORE Working Papers 2015/05.

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics 144*, 122 – 139.

Golosov, M., Tsyvinski, and N. Werquin (2014). Dynamic tax reforms. NBER Working Papers 20780.

Gomes, R., J.-M. Lozachmeur, and A. Pavan (2014a). Differential taxation and occupational choice. CESifo Working Paper 5054.

Gomes, R., J.-M. Lozachmeur, and A. Pavan (2014b). Differential taxation and occupational choice. CESifo Working Paper 5054.

Gruber, J. and E. Saez (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics 84*, 1–32.

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

Guesnerie, R. and J.-J. Laffont (1984). A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm. *Journal of Public Economics 25*, 329–369.

Hammond, P. J. (1979). Straightforward individual incentive compatibility in large economies. *The Review of Economic Studies 46*(2), 263–282.

Hellwig, M. F. (2010). Incentive Problems With Unidimensional Hidden Characteristics: A Unified Approach. *Econometrica 78*(4), 1201–1237.

Hendren, N. (2014). The Inequality Deflator: Interpersonal Comparisons without a Social Welfare Function. *NBER Working Papers* (20351).

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.

Jacquet, L., E. Lehmann, and B. Van der Linden (2013, September). Optimal redistributive taxation with both extensive and intensive responses. *148*(5), 1770–1805.

Kleven, H. J., C. T. Kreiner, and E. Saez (2009). The optimal income taxation of couples. *Econometrica 77*(2), 537–560.

Kleven, H. J. and M. Waseem (2013). Using Notches to Uncover Optimization Frictions and Structural Elasticities: Theory and Evidence from Pakistan. *The Quarterly Journal of Economics 128*(2), 669–723.

Lehmann, E., L. Simula, and A. Trannoy (2014). Tax me if you can! optimal nonlinear income tax between competing governments. *Quarterly Journal of Economics 129*(4), 1995–2030.

Lockwood, B. B. and M. Weinzierl (2015). De gustibus non est taxandum: Heterogeneity in preferences and optimal redistribution. *Journal of Public Economics* (124), 74–80.

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. A. (1971). An exploration in the theory of optimum income taxation. *Review of Economic Studies 38*, 175–208.

Piketty, T. (1997). La redistribution fiscale face au chômage. *Revue Française d'Economie 12*, 157–201.

Piketty, T. and E. Saez (2013). Optimal labor income taxation. In M. F. Alan J. Auerbach, Raj Chetty and E. Saez (Eds.), *Handbook of public economics, vol. 5*, Volume 5 of *Handbook of Public Economics*, Chapter 7, pp. 391 – 474. Elsevier.

Piketty, T., E. Saez, and S. Stantcheva (2014, February). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy 6*(1), 230–71.

Ramsey, F. P. (1927). A contribution to the theory of taxation. *Economic Journal 37*, 47–61.

Rochet, J.-C. and A. Stole, Lars (2002). Nonlinear pricing with random participation. *The Review of Economic Studies 69*(1), pp. 277–311.

Rothschild, C. (2013). Redistributive taxation in the roy model. *The Quarterly Journal of Economics 128*(2), 623–668.

Rothschild, C. and F. Scheuer (2013). Redistributive taxation in the Roy model. *The Quarterly Journal of Economics 128*(2), 623–668.

Rothschild, C. and F. Scheuer (2014). A theory of income taxation under multidimensional skill heterogeneity. NBER Working Papers 19822.

Rothschild, C. and F. Scheuer (2016). Optimal taxation with rent-seeking. *The Review of Economic Studies 83*(3), 1225–1262.

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *Review of Economic Studies 68*, 205–229.

Saez, E. (2002). Optimal income transfer programs: Intensive versus extensive labor supply responses. *Quarterly Journal of Economics 117*, 1039–1073.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy 2*(1), 180–212.

Saez, E., J. Slemrod, and S. H. Giertz (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature 50*(1), 3–50.

Saez, E. and S. Stantcheva (2016). Generalized social marginal welfare weights for optimal tax theory. *American Economic Review 106*(1), 24–45.

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

Scheuer, F. (2014a). Entrepreneurial taxation with endogenous entry. *American Economic Journal: Economic Policy 6*(2), 126–63.

Scheuer, F. (2014b). Entrepreneurial taxation with endogenous entry. *American Economic Journal: Economic Policy 6*(2), 126–63.

Scheuer, F. and I. Werning (2017). The taxation of superstars. *Quarterly Journal of Economics*, forth.