CRED WORKING PAPER N° 2019-1

Optimal Income Taxation with Composition Effects

January 9, 2019

LAURENCE JACQUET* AND ETIENNE LEHMANN†

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

†CRED(TEPP) - Université Paris II. Email : elehmann@u-paris2.fr. Etienne Lehmann is also research fellow at IZA, CESifo and CEPR. Address CRED Université Paris II Panthéon-Assas, 12 Place du Panthéon, 75231 Paris Cedex 05, France.

Abstract

We study the optimal nonlinear income tax problem with multidimensional individual characteristics on which taxes cannot be conditioned. We obtain an optimal tax formula that generalizes the standard one by averaging the sufficient statistics of individuals who earn the same income. Dependence on the tax schedule is a well-known limitation of sufficient statistics: they have distinct values in the actual and optimal economies, which can bias the recommended tax schedule. This bias is generally considered to be negligible, but, as a first main contribution, we show that multidimensional unobserved individual heterogeneity actually makes it quite substantial. Multidimensional heterogeneity brings a new source of endogeneity to the sufficient statistics (due to changes in average behavioral responses) that we call composition effects. Using U.S. data, we highlight that composition effects substantially affect optimal marginal tax rates. Our results put the stress on the need for more empirical studies on the distribution of sufficient statistics and not only on their means conditional on income. As a second main contribution, we show the equivalence between the tax perturbation and mechanism design approaches which bridges the gap between both methods that have, so far, been used separately in the literature.

I   Introduction

The recent years have seen an increase in empirical analyses that provide so-called "sufficient statistics"(e.g., reduced-form elasticities) to give policy prescriptions that are easily implementable and relatively easy to explain to the general public. As a compromise between reduced-form and structural analyses, the approach based on sufficient statistics has applications in macroeconomics, labor economics, development economics, industrial organization, political economy and in international trade (e.g., Chetty (2009), Hornstein et al. (2011), Arkolakis et al. (2012), Bierbrauer and Boyer (2018)). It is however in the optimal tax literature that researchers have come to rely extensively on empirically meaningful sufficient statistics to express tax formulas (e.g., Saez (2001, 2002), Saez and Stantcheva (2018) and references in Chetty (2009) and in Kleven (2018)). For this reason, in the present paper, we select optimal tax policy as the field of choice to illustrate a more general point regarding the limitations of sufficient statistics in the presence of multidimensional heterogeneity.

In the optimal tax literature, the key sufficient statistics are *total*[1] 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 (2001)). One important and well-known limitation of sufficient statistics is their dependence on the tax schedule. Their values are distinct in the actual and optimal economies, which can bias the recommended tax schedule. This bias is generally considered to be negligible, but we show in the present paper that multidimensional unobserved individual heterogeneity actually makes it quite substantial. Going beyond one dimension of unobserved heterogeneity is a requisite in order to obtain more empirically-meaningful tax formulas. In the case of the income tax, for instance, individuals who differ in many dimensions (such as health, gender, marital status and ethnicity) are more likely to respond differently to any tax reform, even if they initially earn the same income.

The main contribution of this paper is thus to show how allowing for multidimensional heterogeneity in a model *à la* Mirrlees (1971) exacerbates the aforementioned bias in the tax schedule. In the presence of multidimensional individual heterogeneity, we identify a new source of endogeneity in the sufficient statistics that we call *composition effects*. Composition effects are changes in the composition of the population that may occur, at each income level, between the actual and optimal economies, when individuals who earn the same income have distinct multidimensional characteristics and therefore distinct behavioral responses.

More precisely, as conjectured by Saez (2001), optimal marginal tax rates depend on the averages of sufficient statistics taken among taxpayers who earn the same income. Composition effects arise because, at a given income level, the share of taxpayers for whom the sufficient statistics are relatively large may vary between the actual economy and the optimum. For instance, composition effects in the elasticities of earnings with respect to the marginal retention

---

[1]Total sufficient statistics encapsulate the circularity process of optimal tax systems, i.e. they are evaluated along the nonlinear income tax schedule (as in Jacquet et al. (2013) and Scheuer and Werning (2016)) and not along a linearized one.

rate (i.e. one minus the marginal tax rate) arise as follows. The optimal marginal tax formula states that, ceteris paribus, the optimal marginal tax rate at a given income level decreases with the average of elasticities at this income level. Assume now that marginal tax rates are larger in the optimal economy than in the actual one. When going from the actual to the optimal economy, behavioral responses (due to larger marginal tax rates) tend to decrease all taxpayers' incomes. Behavioral responses are however larger for taxpayers with higher elasticities. Therefore, their income distribution is much more shifted to the left than the one of taxpayers with lower elasticities. The average elasticity tends to decrease at high income levels while it tends to increase at low income levels. In turn, this tends to increase optimal marginal tax rates that apply on high incomes and to reduce them on low incomes. We provide several numerical examples of drastic changes optimal marginal tax schedules undergo when one incorporates composition effects. Using U.S. data, we isolate pure composition effects by comparing elasticities with and without composition effects. We find that the elasticities which incorporate composition effects are higher for lower income levels and lower for higher income levels. The discrepancy can reach up to 25 percentage points. This implies that, when taking composition effects into account, optimal marginal tax rates increase by up to 20 percentage points on high incomes and decrease by more than 6 percentage points on low incomes.

These findings have a crucial implication for the empirical literature that provides sufficient statistics. One cannot simply rely on estimates of the conditional means of sufficient statistics estimated in the actual economy. One needs to know the distribution of sufficient statistics conditional on income to implement optimal tax schedules. Consequently, multidimensional heterogeneity and the implied composition effects in behavioral responses appear as an important issue for policy design.

Another important contribution of our paper consists in bridging the gap between the mechanism design method, which is widely used in various fields in economics, and the tax perturbation method, which is more specific to the optimal taxation literature (although it may find applications in industrial organization as well, e.g. in nonlinear monopoly pricing problems). On the one hand, since Mirrlees (1971), the mechanism design method allows one to derive optimal tax profiles by finding the incentive-compatible allocation that maximizes the social objective subject to a resources constraint. The optimal allocation is obtained by verifying (usually using a Hamiltonian or a Lagrangian) that no incentive-compatible perturbation leads to any first-order improvement. On the other hand, the tax perturbation approach seeks the tax reform that decentralizes such a perturbation. The optimal tax literature uses either method to determine optimal tax policies, but the link between both methods is missing when heterogeneity is multidimensional. In the present paper, we formally show the equivalence between these methods, which had not been done until now.

To do so, we first provide conditions under which the tax perturbation method of Piketty (1997) and Saez (2001) is valid with multidimensional heterogeneity. In a nutshell, these conditions imply that the marginal tax rate does not decrease too fast with income which ensures

that individual behaviors respond smoothly to a tax reform. We then proceed to show that, under multidimensional heterogeneity, the assumptions required by the tax perturbation and first-order mechanism design[2] approaches are equivalent. Therefore, these two approaches are two faces of the same coin.

The paper is organized as follows. After a survey of the related literature in Section II, we introduce the framework in Section III. We begin our analysis in Section IV with the simple linear tax model to explain what composition effects are and to illustrate the empirical bias they impose. In Section V, we provide the conditions for using the tax perturbation method when individual heterogeneity is multidimensional and give the relevant sufficient statistics. In Section VI, we then derive the optimal nonlinear tax formula in terms of these sufficient statistics and we shed the light on composition effects. Section VII numerically investigates the sensitivity of the optimal tax function to composition effects and the magnitude of these effects. Section VIII shows the equivalence between the mechanism design and tax perturbation approaches. Section IX concludes.

II Related literature

The use of tax perturbation arguments in order to derive optimal nonlinear tax formulas goes back to Christiansen (1984), Piketty (1997), Roberts (2000) and Saez (2001).[3] Saez (2001) rigorously shows the consistency of these tax formulas with the ones obtained thanks to the first-order mechanism design approach of Mirrlees (1971), when heterogeneity is one-dimensional. Saez (2001) conjectures that his tax formula expressed in terms of sufficient statistics would also be valid with multidimensional heterogeneity. In this paper, we confirm that it is indeed the case by providing properly-constructed average measures of the usual sufficient statistics. We do so by adapting the general tax perturbation approach used by Golosov et al. (2014), Sachs et al. (2016), Gerritsen (2016) and Spiritus (2017) to our framework. Indeed Golosov et al. (2014) have a model where the number of individual characteristics is lower than the number of actions (e.g. the choice of labor and capital income levels) while in our framework the number of characteristics is larger. Scheuer and Werning (2016), who show that the optimal linear commodity tax model of Diamond and Mirrlees (1971) encapsulates the model of Mirrlees (1971), point out that including multidimensional heterogeneity in the Mirrlees (1971) model leads to an optimal tax formula with simple averages of the sufficient statistics. Werning (2007) derives a condition for Pareto efficient tax schedules when heterogeneity is one-dimensional and also writes that his condition can simply be extended to multidimensional heterogeneity by averaging the sufficient statistics by group of individuals. Hendren (2017) develops inequality

---

[2]With multidimensional heterogeneity, we say that the mechanism design approach is first-order when income admits a strictly positive derivative with respect to skill, in each group of workers. This assumption implies (and plays a role similar to) the satisfaction of the second-order incentive compatibility condition when heterogeneity is one-dimensional.

[3]Kleven and Kreiner (2006) also use a tax perturbation to derive the marginal cost of public funds.

deflators which assess Pareto improvements when making second-best transfers through the income tax schedule. The method he uses to derive these inequality deflators is close to the tax perturbation approach and also allows for multidimensional heterogeneity. Beyond deriving the optimal tax formula under multidimensional heterogeneity, we contribute to this literature by studying the role of composition effects in the determination of optimal tax schedules and by connecting the first-order mechanism design and tax perturbation approaches.

Cuff (2000), Boadway et al. (2002), Brett and Weymark (2003), Choné and Laroque (2010), Lockwood and Weinzierl (2015) introduce in the Mirrlees (1971) model an additional source of heterogeneity, typically preferences for leisure or work opportunity cost, that matters only for the computation of social welfare weights. The heterogeneity of preferences raises ethical questions which challenge the design of redistributive schemes. Individuals that have the same skill but distinct preferences for leisure, will earn, at the laissez-faire, different levels of income, with individuals having higher preferences for leisure earning less. The fraction of individuals with a high preference for leisure can be relatively higher among low income earners than among high income earners. The government may then be reluctant to redistribute towards the former, since some of them are also enjoying a higher quantity of leisure. In this context, the mean social welfare weight (the mean computed across the social welfare weights of individuals who earn the same level of income) may become less decreasing than in the tax model with a single dimension of heterogeneity (Lockwood and Weinzierl, 2015) and may even become non monotonic with income, which opens the possibility for optimal marginal tax rates to be negative (e.g., Choné and Laroque (2010)). In these papers, although agents differ in productivity and preference for leisure, their behavior is assumed to depend only on a unidimensional combination of the two underlying parameters. This aggregation implies that all individuals with a given income are constrained to respond identically to any tax reform despite the heterogeneity in social welfare weights. Therefore, contrary to our paper, this literature assumes away composition effects on behavioral responses.

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., 2015) and this latter dimension of heterogeneity matters only for the participation/migration margin. Scheuer (2013, 2014), Rothschild and Scheuer (2013), Gomes et al. (2017) consider optimal income tax models with different sectors[4] where agents can migrate from one sector to the other.[5] This is also a form of random participation across sectors. Again, once individuals choose in which

---

[4]Differing from our paper and from Scheuer (2013) and Rothschild and Scheuer (2013), nonlinear income taxation is sector-specific in Gomes et al. (2017) and in an extension of Scheuer (2014).

[5]In Gomes et al. (2017), 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. In Rothschild and Scheuer (2014), at each income level the composition of individuals across sectors changes. This modifies terms related to general equilibrium effects (resulting from effort choice among different intensive margins) in the optimal tax formula but, again, average behavioral responses are identical in the actual and optimal economies.

sectors to work (or which combination in Rothschild and Scheuer (2016)), income depends only on a single variable. While departing from this restriction, our model can readily be extended to include a participation margin simply following Jacquet et al. (2013). The tax formulas then simply incorporate new terms with the behavioral elasticities implied by the participation margin (see Jacquet et al. (2013)).

III Model

Every worker derives utility from consumption $c \in \mathbb{R}_+$ and disutility from effort. Effort captures the quantity as well as the intensity of labor supply. 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.[6] 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 or θ. We normalize to unity the total size of the population.

III.1 Individual choice

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),[7] $C(w, \theta) = Y(w, \theta) - T(Y(w, \theta))$ 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:

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

---

[6]Our definition of "group" is identical to the one in Werning (2007, p.15).

[7]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.

where:

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

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 that, within each group of workers endowed with the same θ, the marginal rate of substitution is a decreasing function of the skill level, i.e. that more skilled workers find it less costly to increase their income 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 $\forall w \in \mathbb{R}_+^*$, $\mathscr{M}_w < 0$.*

Assumption 1 is for instance verified in the case where $\mathscr{U}(c, y; w, \theta)$ is specified 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 Frisch labor supply elasticity. The marginal 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.

III.2   Government

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 $\Phi(U; w, \theta)$ of individuals' utility levels U:

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

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 $\varphi(w, \theta) = 0 \ \forall w > 0$. When $\Phi(U; w, \theta)$ does not vary with its two last arguments and $\Phi_{UU} \leq 0$, 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 the social welfare function subject to the budget constraint, i.e. that maximizes (6) subject to (5).

Let $\lambda > 0$ be the Lagrange multiplier associated with the budget constraint (5). The Lagrangian (expressed in monetary terms) is:

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

We define the social marginal welfare weights associated with workers of type (w, θ) expressed in terms of public funds by:

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

The government values giving one extra dollar to a worker (w, θ) as a gain of $g(w, \theta)$ dollar(s) of public funds.[8]

IV   Optimal linear tax and composition effects

In this section, we illustrate, using the very simple case of linear taxation, how composition effects bias the empirical implementation of the optimal tax rate. We thus constrain the tax schedule to be linear with a tax rate denoted τ and a demogrant D so that the tax schedule is: $T(y) = \tau y - D$. We also assume income effects away with quasilinear preferences of the form:

$$\mathscr{U}(c, y; w, \theta) = c - \frac{\theta}{1 + \theta} y^{1 + \frac{1}{\theta}} w^{-\frac{1}{\theta}} \qquad w \in \mathbb{R}_+^*, \theta \in \Theta \tag{9}$$

Under this utility and tax functions, the solution to program (1) for individuals of type (w, θ) is an income of

$$Y(w, \theta) = (1 - \tau)^\theta w. \tag{10}$$

The average income in group θ is thus given by:

$$\overline{y}(\theta, \tau) \overset{\text{def}}{\equiv} (1 - \tau)^\theta \int_{w \in \mathbb{R}_+^*} w f(w|\theta) \, dw. \tag{11}$$

Note that due to the iso-elasticity of preferences in (9), parameter θ is equal to the elasticity of average income $\overline{y}(\theta, \tau)$ in group θ with respect to the net-of-tax rate $1 - \tau$. Aggregate earnings (or the average income across all individuals) are equal to the sum of all individual incomes and are denoted by:

$$\overline{Y}(\tau) \overset{\text{def}}{\equiv} \int_{\theta \in \Theta} \overline{y}(\theta, \tau) \, d\mu(\theta) = \int_{\theta \in \Theta} (1 - \tau)^\theta \int_{w \in \mathbb{R}_+} w f(w|\theta) dw \, d\mu(\theta). \tag{12}$$

The elasticity of aggregate earnings with respect to the net-of-tax rate $(1 - \tau)$ is equal to the income-weighted average elasticity and is given by:

$$e \overset{\text{def}}{\equiv} \frac{1 - \tau}{\overline{Y}(\tau)} \frac{\partial \overline{Y}}{\partial (1 - \tau)} = \int_{\theta \in \Theta} \theta \, \frac{\overline{y}(\theta, \tau)}{\overline{Y}(\tau)} \, d\mu(\theta). \tag{13}$$

---

[8]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, 2017) 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.

The government is maximin which is equivalent to maximizing tax revenue $\tau \cdot Y - E$ or to maximizing D (Boadway and Jacquet, 2008). Following Piketty and Saez (2013), the tax rate τ_L that maximizes tax revenue is such that $\overline{Y} \cdot (1 - \tau) - \tau \cdot d\overline{Y}/d(1 - \tau) = 0$, i.e.:

$$\tau_L = \frac{1}{1 + e}.$$

The higher the elasticity e of aggregate earnings, the more distortive is taxation and so the lower the revenue maximizing linear tax rate τ_L.

From (10), we directly see that individuals do not earn the same income under the actual tax rate and the optimal (revenue maximizing) one as soon as these two rates are distinct. The average income within each group θ, $\overline{y}(\theta, \tau)$, is then also distinct under the actual and optimal tax rates. In the calculation of the elasticity of aggregate earnings e, this average income for each group multiplies the elasticity θ (see (13)). Therefore the elasticity e is distinct under the actual and the optimal tax rates. To put it differently, because the composition of the population behind a given income (and a given average income) differs under the actual and the optimal tax rates, the elasticity of aggregate earnings takes distinct values under each tax schedule. Implementing the optimal linear tax formula with the elasticity of aggregate earnings as a function of the *actual* tax rate instead of the *optimal* one, i.e. neglecting composition effects, may strongly bias the outcome.

We now give a mathematical rationale for this. To implement the tax rate that maximizes the tax revenue, one can combine Equations (10)-(13) and rewrite elasticity e in terms of statistics observable in the current economy, making the optimal (revenue maximizing) linear tax rate τ_L and the actual linear tax rate τ_0 appear. [9] This yields:

$$\tau_L = \frac{1}{1 + \int_{\theta \in \Theta} \theta \, \frac{\overline{y}(\theta, \tau_L)}{\overline{Y}(\tau_L)} \, d\mu(\theta)} = \frac{1}{1 + \int_{\theta \in \Theta} \theta \, \frac{\left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{\theta} \overline{y}(\theta, \tau_0)}{\int_{t \in \Theta} \left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{t} \overline{y}(t, \tau_0) \, d\mu(t)} \, d\mu(\theta)}. \tag{14}$$

In the denominator, the elasticity of aggregate earnings e is written as the sum of all elasticities θ, each of them being weighted (from (11) and (12)) by

$$\overline{y}(\theta, \tau_L)/\overline{Y_L}(\tau_L) = \left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{\theta} \overline{y}(\theta, \tau_0) / \int_{t \in \Theta} \left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{t} \overline{y}(t, \tau_0) \, d\mu(t),$$

i.e. by the ratio of average income within group θ to the sum of all individual incomes. In this expression, the terms $\left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{\theta}$ and $\left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{t}$ show up due to composition effects. Indeed actual and optimal linear tax rates are typically different so that these terms are not equal to 1. Therefore, the composition of the population at each income level is different which makes the elasticity of aggregate earnings distinct under the actual and the optimal linear tax rates.

---

[9] Indeed, $\overline{y}_L(\theta, \tau) = \left(\frac{1 - \tau_L}{1 - \tau_0}\right)^{\theta} \overline{y}_0(\theta, \tau)$ from (11).

Assume $\tau_L > \tau_0$. We have $\left(\frac{1-\tau_L}{1-\tau_0}\right)^\theta < 1$ and $Y_L(w,\theta) < Y_0(w,\theta)$ since, from (10), $Y_L(w,\theta) = \left(\frac{1-\tau_L}{1-\tau_0}\right)^\theta Y_0(w,\theta)$ where $Y_i(w,\theta) \overset{\text{def}}{\equiv} (1-\tau_i)w$ with $i = L,0$. From (11), the average income in group θ is then lower under τ_L than under τ_0, $\overline{y}(\theta,\tau_L) < \overline{y}(\theta,\tau_0)$ and the same then applies for aggregate earnings, $\overline{(\tau_L)} < \overline{(\tau_0)}$. We also know that, due to behavioral responses, the reduction of average income $\overline{y}(\theta,\tau)$ in groups with a larger θ is stronger than in groups with a lower θ. Therefore, we have $\overline{y}(\theta,\tau_L)/\overline{y}(\theta,\tau_0) < \overline{Y}(\tau_L)/\overline{Y}(\tau_0)$, i.e. $\overline{y}(\theta,\tau_L)/\overline{Y}(\tau_L) < \overline{y}(\theta,\tau_0)/\overline{Y}(\tau_0)$, for groups with relatively large θ. Conversely, we have $\overline{y}(\theta,\tau_L)/\overline{y}(\theta,\tau_0) > \overline{Y}(\tau_L)/\overline{Y}(\tau_0)$, i.e. $\overline{y}(\theta,\tau_L)/\overline{Y}(\tau_L) > \overline{y}(\theta,\tau_0)/\overline{Y}(\tau_0)$, for groups with relatively low θ. Since these ratios intervene as weights in the calculation of the aggregate earnings elasticity e, the θ-groups with relatively large elasticities θ matter less in this calculation than those with relatively low elasticities θ. This change in the composition of the population when moving from τ_0 to τ_L indubitably reduces the elasticity e, hence pushing further up the optimal linear tax rate τ_L. In other words, ignoring the endogeneity of the elasticity of aggregate earnings due to composition effects leads one to understimate the revenue maximizing linear tax rate.[10]

As a back-of-the-envelope numerical illustration, consider the case where the economy is made of two groups, a high elasticity one with $\theta_H = 0.8$ and a low elasticity one with $\theta_L = 0.1$. Assume both groups are of equal size $\mu(\theta_L) = \mu(\theta_H) = 0.5$ and are characterized by the same average income in the actual economy where the tax rate is $\tau_0 = 0.3$. Then, ignoring the heterogeneity in the elasticity of labor supply, one obtains a revenue maximizing linear tax rate equal to $1/(1 + 0.45) \simeq 69.0\%$. By contrast, taking into account composition effects (using (14)) leads to a revenue maximizing linear tax rate which rockets to 75.3%.

V Tax perturbation and sufficient statistics with multidimensional heterogeneity

In this section, we provide all the required elements to use a tax perturbation approach when individual heterogeneity is multidimensional and we derive the sufficient statistics that will show up in the optimal tax formula. The tax perturbation approach is valid only under some circumstances that Saez (2001) did not clarify. This is the reason why Saez (2001) needed to show that his tax formula was consistent with the one of Mirrlees (1971). However, he did so only when individual heterogeneity is one-dimensional. In this section, we first state sufficient conditions for using a tax perturbation method when individual characteristics are multidimensional. We then give the relevant, empirically measurable sufficient statistics in terms of which we will express the optimal tax formulas. We will derive the optimal nonlinear tax schedule under multidimensional heterogeneity in the next section. In that section, we will also enlighten the reader about composition effects. The mechanism design approach, which

---

[10]A symetric argument prevails when assuming $\tau_L < \tau_0$. In this case $(1-\tau_L)/(1-\tau_0) > 1$ so that $\left(\frac{1-\tau_L}{1-\tau_0}\right)^\theta$ increases with θ. Following the same reasoning as in the case of $\tau_L > \tau_0$, composition effects increase the endogenous elasticity of aggregate earnings hence further reduce τ_L. In other words, ignoring composition effects leads to overestimating the optimal linear tax rate.

is the alternative method to obtain the optimal tax system, will be presented in Section VIII. In the latter section, we will bridge the gap between both approaches.

V.1 Sufficient conditions for a tax perturbation

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

$$U^R(m; w, \theta) \overset{\text{def}}{\equiv} \max_y \ \mathscr{U}\left(y - T(y) + m\,R(y), y; w, \theta\right) \tag{15}$$

We denote $Y^R(m; w, \theta)$ the income of workers of types (w, θ) after the reform and her consumption becomes $C^R(m; w, \theta) = Y^R(m; w, \theta) - T(Y^R(m; w, \theta)) + m\,R(Y^R(m; w, \theta))$. 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 (15), we get:

$$\frac{\partial U^R}{\partial m}(m; w, \theta) = \mathscr{U}_c\left(C^R(m; w, \theta), Y^R(m; w, \theta); w, \theta\right)\,R(y) \tag{16}$$

Using (3), the first-order condition associated to (15) equalizes to zero the following expression:

$$\mathscr{Y}^R(y, m; w, \theta) \overset{\text{def}}{\equiv} 1 - T'(y) + m\,R'(y) - \mathscr{M}\left(y - T(y) + m\,R(y), y; w, \theta\right). \tag{17}$$

For simplicity, we drop the superscript R and write $\mathscr{Y}_y(Y(w, \theta); w, \theta)$ for $\mathscr{Y}_y^R(Y(w, \theta), 0; w, \theta)$ since at $m = 0$, $\mathscr{Y}_y^R$ does no longer depend on the direction R of the tax reform. To use the tax perturbation method, one needs the following assumptions:

Assumption 2. *Sufficient conditions for a tax perturbation.*

  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 (17) is differentiable. Part *ii)* guarantees it is invertible in income y. Under *i)* and *ii)*, one can apply the implicit function theorem to (17) to describe how a local maximum of the individual maximization program (15) changes after a tax reform. Part *iii)* ensures that after an incremental tax reform or change in skill, the maximum remains global. Indeed 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 jump in individual's choice from one maximum to another one (which is associated to a jump in the supply of effort). Part *iii)* prevents this situation and ensures the allocation changes in a differentiable way with the magnitude m of a tax reform.

Let us emphasize 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 assumed 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). 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 when this indifference curve is strictly above $y \mapsto y - T(y)$ for all other y (so that Part *iii)* of Assumption 2 is satisfied). In a nutshell, Assumption 2 is satisfied whenever the marginal tax rate does not decrease too rapidly with income.

Assumptions 2 implies the property that income is differentiable with respect to m after a tax reform in the direction $R(\cdot)$. This property plays a role similar to the assumption of (local) Lipschitz continuity of the income function in Golosov et al. (2014) or the assumption of Hendren (2017) that aggregate tax revenue, $\iint_{\theta \in \Theta, w \in \mathbb{R}_+} T(Y(w, \theta)) f(w|\theta) dw \, d\mu(\theta)$ varies smoothly in response to changes in the tax schedule. Hendren (2017)'s assumption allows for discrete changes in individual behavior in response to small tax changes, which is more general than our property of differentiable income or than the Lipschitz continuity assumption of Golosov et al. (2014). However, the assumptions of Golosov et al. (2014) and of Hendren (2017) are about the way economic outcomes respond to tax reforms, which is rather ad-hoc since these responses are endogenous. Conversely, we obtain the property that each income decision varies smoothly after a tax reform by considering only tax functions that verify Assumption 2. A strength of our approach is then to give micro-foundations to the property of smooth responses to tax reforms. We can also note that Assumption 2 bears on tax functions that are endogenous objects. Considering only tax functions that verify this assumption is a restriction similar to that which consists in considering only smooth allocation with no bunching, as done in the first-order mechanism design approach. We will demonstrate this in Proposition 2.

V.2 Perturbations and elasticities

In this subsection, we derive the behavioral responses to small perturbations of the tax function. We use them to formulate the way income reacts to any tax reform. The elasticities we obtain are also helpful to define the sufficient statistics relevant to the optimal tax formula with multidimensional heterogeneity (see Subsection V.4 and onwards).

Applying the implicit function theorem to $\mathscr{Y}^R(y, m; w, \theta) = 0$ at $(y = Y^R(m; w, \theta), m = 0; w, \theta)$, we obtain:

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

with:

$$\mathcal{Y}_y^R(y,m;w,\theta) = -T''(y) - \mathcal{M}_y(y - T(y) + m\,R(y), y; w, \theta) \tag{18a}$$
$$- \mathcal{M}(y - T(y) + m\,R(y), y; w, \theta)\,\mathcal{M}_c(y - T(y) + m\,R(y), y; w, \theta)$$
$$\mathcal{Y}_m^R(y,m;w,\theta) = R'(y) - R(y)\,\mathcal{M}_c(y - T(y) + m\,R(y), y; w, \theta). \tag{18b}$$

We consider two types of tax perturbations to derive the behavioral elasticities. First, we analyze the effects of small changes in the marginal tax rates to capture substitution effects. Second, the income effects are isolated thanks to a uniform transfer to all workers.

Consider first a change in 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. This reform is therefore given the label *compensated*. Formally, the direction of this compensated reform is $R(y) = y - Y(w,\theta)$; 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 $R'(Y(w,\theta)) = 1$. Using (2) and (3), we define the *compensated elasticity* of earnings with respect to the marginal retention rate $1 - T'(.)$ as:

$$\varepsilon(w;\theta) \overset{\text{def}}{\equiv} \frac{1 - T'(Y(w,\theta))}{Y(w,\theta)} \frac{\partial Y^c}{\partial m} = \frac{\mathcal{M}(C(w,\theta), Y(w,\theta); w, \theta)}{-Y(w,\theta)\,\mathcal{Y}_y(Y(w,\theta); w, \theta)} > 0 \tag{19a}$$

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 $R(y) \equiv 1$. Define the *income effect* as:

$$\eta(w;\theta) \overset{\text{def}}{\equiv} \frac{\partial Y^I}{\partial m} = \frac{\mathcal{M}_c(C(w,\theta), Y(w,\theta); w, \theta)}{\mathcal{Y}_y(Y(w,\theta); w, \theta)} \tag{19b}$$

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

Combining (19a) and (19b) with (18b), the way income of individuals (w,θ) reacts to any tax reform $R(\cdot)$ is given by:

$$\frac{\partial Y^R}{\partial m}(0; w, \theta) = \varepsilon(w,\theta)\frac{Y(w,\theta)}{1 - T'(Y(w,\theta))}R'(Y(w,\theta)) + \eta(w,\theta)\,R(Y(w,\theta)) \tag{19c}$$

where the compensated elasticity and income effect show up.

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 (17) with respect to skill w. Note that this ensures that income $Y(\cdot,\theta)$ is a continuously differentiable function in skill. Using $\mathcal{Y}_w^R = -\mathcal{M}_w$, the elasticity of earnings with respect to skill w is defined as:

$$\alpha(w;\theta) \overset{\text{def}}{\equiv} \frac{w}{Y(w,\theta)}\dot{Y}(w,\theta) = \frac{w\,\mathcal{M}_w(C(w,\theta), Y(w,\theta); w, \theta)}{Y(w,\theta)\,\mathcal{Y}_y(Y(w,\theta); w, \theta)} > 0 \tag{19d}$$

which is positive from Assumption 1. Note that Assumptions 1 and 2 rule out bunching.[11]

---

[11]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.

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.

V.3 Total vs direct elasticities and income responses

Jacquet et al. (2013) propose to use *total* elasticities and income responses in order to include the circular process induced by the endogeneity of marginal tax rates. This helps streamline tax formulas. As Scheuer and Werning (2017), we follow this strategy here. Our definitions of elasticities and income response (19a)-(19d) then 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 $\mathcal{Y}_Y$ (see Equation (18a)), emphasizes the role played by the local curvature of the tax schedule. By contrast, the empirical literature estimates responses that do not take into account the local curvature of the tax function. We refer to these responses as *direct* responses.[12] 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 (18a) and thereby in (19a)-(19d). These would be the relevant concepts if the tax function were linear.

To better grasp the distinction between direct and total responses, consider an exogenous change in w, or a tax reform inducing a *direct* change in earnings $\Delta_1 y$ proportional to the direct response $\varepsilon^\star(w;\theta)$, $\eta^\star(w;\theta)$ and $\alpha^\star(w;\theta)$. When the tax schedule is nonlinear, this 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 (Saez, 2001). The income level 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:

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

$$= \Delta_1 y \frac{1-T'(Y(w,\theta))}{1-T'(Y(w,\theta))+Y(w,\theta)\,T''(Y(w,\theta))\,\varepsilon^\star(w,\theta)}$$

---

[12]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 allows empirical researchers to identify the causal effect of tax on individual behavior is the exogenous shock induced by a tax reform. This shock makes some taxpayers face a change in their marginal tax rate while others do not (Feldstein, 1995, Auten and Carroll, 1999, Gruber and Saez, 2002), see Saez et al. (2012). By using two-stages least squares estimators, this approach implicitly assumes that marginal tax rates do locally not depend on taxable income. It thus identifies *direct* behavioral responses. The second method 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 (20c) 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.

where the ratio in the latter 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:

$$\varepsilon(w,\theta) = \frac{1-T'(Y(w,\theta))}{1-T'(Y(w,\theta))+Y(w,\theta)\,T''(Y(w,\theta))\,\varepsilon^\star(w,\theta)}\,\varepsilon^\star(w,\theta) \tag{20a}$$

$$\eta(w,\theta) = \frac{1-T'(Y(w,\theta))}{1-T'(Y(w,\theta))+Y(w,\theta)\,T''(Y(w,\theta))\,\varepsilon^\star(w,\theta)}\,\eta^\star(w,\theta) \tag{20b}$$

$$\alpha(w,\theta) = \frac{1-T'(Y(w,\theta))}{1-T'(Y(w,\theta))+Y(w,\theta)\,T''(Y(w,\theta))\,\varepsilon^\star(w,\theta)}\,\alpha^\star(w,\theta) \tag{20c}$$

Equalities (20a)-(20c) are obtained from the definitions of elasticities, income responses and from (2).[13] In the real world, most of income tax schedules are piecewise linear. In this case, direct and total responses differ at the kinks of tax schedules which makes the distinction between both particularly relevant.[14]

V.4 Sufficient statistics

To derive the relevant sufficient statistics, we need to define the conditional income density. Denote $h(y|\theta)$ the conditional income density within group θ at income y and $H(y|\theta) \overset{\text{def}}{\equiv} \int_{z=0}^{y} h(z|\theta)dz$ the corresponding conditional income CDF. To obtain the tax formula, we need the link between the conditional income density and the conditional skill density. According to (19d) 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 (19d) leads to:

$$h(Y(w,\theta)|\theta) = \frac{f(w|\theta)}{\dot{Y}(w,\theta)} \qquad \Leftrightarrow \qquad Y(w,\theta)\,h(Y(w,\theta)|\theta) = \frac{w\,f(w|\theta)}{\alpha(w,\theta)} \tag{21}$$

Let $W(\cdot,\theta)$ now denote the reciprocal of $Y(\cdot,\theta)$ so that, within each group θ, individuals of type $(w = W(y,\theta),\theta)$ earn income 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:

$$\hat{h}(y) \overset{\text{def}}{\equiv} \int_{\theta\in\Theta} h(y|\theta)\,d\mu(\theta) \tag{22a}$$

---

[13]From (18a) and (19a) 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 (20a). The same goes for Equations (20b) and (20c).

[14]Within tax brackets, the linearity of the tax implies $T''(y) = 0$ so that total and direct responses are identical. At kinks, two possibilities can occur. First, the marginal tax rate can increase, i.e. $T''(y) = +\infty$, which implies that total responses are nil and bunching prevails (since $\alpha(w,\theta) = 0$ from (20c)). Second, the marginal tax rate can decrease at kinks, i.e. $T''(y) = -\infty$, so that the second-order condition which is necessary for using tax perturbations (see Assumption 2) is no longer satisfied. Intuitively, one has $\alpha(w,\theta) = \infty$ and income jumps.

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

$$\hat{\varepsilon}(y) = \int_{\theta \in \Theta} \varepsilon \left(W(y, \theta), \theta \right) \frac{h(y|\theta)}{\hat{h}(y)} d\mu(\theta) \tag{22b}$$

where each within-group total elasticity is timed by the relative proportion $h(y|\theta)/\hat{h}(y)$ of individuals in the corresponding group among individuals who earn y. The mean *total* income effect at income level y is:

$$\hat{\eta}(y) = \int_{\theta \in \Theta} \eta \left(W(y, \theta), \theta \right) \frac{h(y|\theta)}{\hat{h}(y)} d\mu(\theta) \tag{22c}$$

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

$$\hat{g}(y) = \int_{\theta \in \Theta} g \left(W(y, \theta), \theta \right) \frac{h(y|\theta)}{\hat{h}(y)} d\mu(\theta) \tag{22d}$$

VI Optimal tax schedule and composition effects

In this section, after stating the desirable tax reforms, we derive the optimal nonlinear tax schedule under multidimensional heterogeneity using the tax perturbation approach detailed in Section V. We express the optimal tax formulas in terms of the empirically measurable sufficient statistics that we also defined in Section V. The formulas we obtain let us shed the light on the crucial role played by *composition effects*, the importance of which cannot be overstated. Indeed, we intuitively explain why neglecting composition effects may lead to severe biases in the derivation of the optimal tax schedule. We give numerical illustrations of this bias in Section VII.

VI.1 Desirable tax reforms

Having defined the general tax reforms and described their impact on individual income in Section V, we now study when a given tax reform is desirable. To do so, we locally perturb the tax system in a direction $R(y)$ with magnitude m. The initial tax system can be optimal or suboptimal. 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. *Under Assumptions 1 and 2, reforming the tax schedule in the direction $R(\cdot)$ triggers first-order effects on the Lagrangian* (7) *equal to:*

$$\begin{aligned}
\frac{\partial \mathscr{L}^R}{\partial m} &= \int_{y=0}^{\infty} \left\{ \left[\hat{g}(y) - 1 + T'(y) \, \hat{\eta}(y) \right] \hat{h}(y) - \frac{d}{dy} \left[\frac{T'(y)}{1 - T'(y)} \hat{\varepsilon}(y) \, y \, \hat{h}(y) \right] \right\} R(y) \, dy \quad (23) \\
&+ \lim_{y \to \infty} \frac{T'(y)}{1 - T'(y)} \hat{\varepsilon}(y) \, y \, \hat{h}(y) \, R(y) - \lim_{y \to 0} \frac{T'(y)}{1 - T'(y)} \hat{\varepsilon}(y) \, y \, \hat{h}(y) \, R(y)
\end{aligned}$$

The proof is relegated to Appendix A.1. 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 $R(\cdot)$ which is close to

what is proposed in Golosov et al. (2014). However, what they propose is not applicable in our framework: with their method, individuals must take at least as many actions as they have characteristics. With our method, one can solve models with one action and many types.[15]

An important point to notice is that, in general, implementing a reform with direction $R(\cdot)$ 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 (23). One can then define a balanced-budget tax reform with magnitude m and direction $R(\cdot)$ by combining it 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 with magnitude m and direction $R(\cdot)$ on the Lagrangian. Expression (23) is therefore useful to determine which tax reforms are desirable. If (23) is positive, it is socially desirable to implement a tax reform with direction $R(\cdot)$ and a positive magnitude m and to combine this reform with a lump-sum rebate to keep the government's budget balanced. Symmetrically, if Expression (23) is negative, it is socially desirable to implement a tax reform with direction $-R(\cdot)$ and positive magnitude m combined with a lump-sum transfer.

VI.2 Optimal tax schedule and composition effects

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:*

$$\frac{T'(y)}{1-T'(y)} = \frac{1}{\hat{\varepsilon}(y)} \frac{1-\hat{H}(y)}{y\,\hat{h}(y)} \left(1 - \frac{\int_y^\infty [\hat{g}(z) + \hat{\eta}(z)\,T'(z)]\,\hat{h}(z)dz}{1-\hat{H}(y)}\right) \tag{24a}$$

$$1 = \int_0^\infty [\hat{g}(z) + \hat{\eta}(z)\,T'(z)]\,\hat{h}(z)dz. \tag{24b}$$

If income effects were assumed away, Equation (24b) 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 $\hat{\eta}(z) \cdot T'(z)$.

The optimal tax rate given in Equation (24a) 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 inverse local Pareto parameter $\frac{1-\hat{H}(y)}{y\,\hat{h}(y)}$ of the income distribution and *iii)* the social preferences and income effects $1 - \frac{\int_y^\infty [\hat{g}(z)+\hat{\eta}(z)\cdot T'(z)]\hat{h}(z)dz}{1-\hat{H}(y)}$. This term indicates the distributional benefits of increasing the tax

---

[15]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 (21) and (22a).

16

liability by one unit for all workers with incomes above y. Diamond (1998) and Saez (2001) discuss 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.[16] It must be stressed here that the averaging procedure is a far cry from using the same corrective term in the averaging process. 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 (20a)-(20b) and (22b)-(22d)).

The optimal tax formulas in Equations (24a) and (24b) are functions of sufficient statistics $\hat{h}(y)$, $\hat{\varepsilon}(y)$, $\hat{\eta}(y)$ and $\hat{g}(y)$ which are endogenous to the tax schedule. Therefore, estimating these sufficient statistics in the actual economy and simply plugging the obtained estimates into the optimal tax formulas (24a) and (24b) may lead to biased results. This is because these sufficient statistics typically take different values in the actual economy and in the optimal one. This well-known limit of sufficient statistics formulas already prevails when unobserved heterogeneity is one-dimensional. However, when heterogeneity is multidimensional, additional mechanisms appear which exacerbate the discrepancies between sufficient statistics in both economies. We now describe these additional mechanisms that we call *composition effects*. In a nutshell, composition effects stem, at every income level, from the prevalence of a distinct composition of population in the actual and optimal economies.

According to Equation (22a)-(22d), sufficient statistics $\hat{\varepsilon}(y)$, $\hat{\eta}(y)$ and $\hat{g}(y)$ at a given income level y are weighted averages of group-specific sufficient statistics $\varepsilon(y|\theta)$, $\eta(y|\theta)$ and $g(y|\theta)$, where the weights are given by the relative proportion $h(y|\theta)/\hat{h}(y)$ of individuals of group θ among individuals who earn income y. Composition effects take place since these relative proportions are also endogenous to the tax schedule. Therefore, among individuals earning a given income y, there may be relatively fewer (more) individuals with lower compensated elasticity $\varepsilon(y|\theta)$, income response $\eta(y|\theta)$ or welfare weight $g(y|\theta)$ in the optimal economy than in the actual one. Therefore, $\hat{\varepsilon}(y)$, $\hat{\eta}(y)$ or $\hat{g}(y)$ rise (shrink) when moving from the actual to the optimal economy. Consider for instance that, at each income level, individuals differ neither in income response (e.g. because individual preferences are quasilinear in consumption so that $\eta(y|\theta) = 0$) nor in welfare weight (e.g. because the government's objective is maximin, i.e. $g(y|\theta) = 0$) but in compensated elasticity $\varepsilon(y|\theta)$. Assume also that marginal tax rates are larger at the optimum than in the actual economy (which is very likely, under maximin, in the U.S.). In this case, taxpayers respond to the rise of marginal tax rates (from their actual to their optimal levels) by reducing their incomes. These responses are larger in groups where the

---

[16]This is more intuitive than using the direct elasticity and income effect, which implies to encapsulate the circularity (described by (20a)-(20c)) in a so-called "virtual density" as in Saez (2001), Equation (13).

compensated elasticity $\varepsilon(y|\theta)$ is larger. Consequently, the income distribution $h(\cdot|\theta)$ of high-elasticity groups is much more shifted to the left than the income distributions of low-elasticity groups. At low income levels, the relative proportion $h(y|\theta)/\hat{h}(y)$ of low- (high-) elasticity groups tends to decrease (increase). This tends to increase the mean compensated elasticity $\hat{\varepsilon}(y)$ of low income earners. It then reduces the optimal marginal tax rate $T'(y)$ they face. For high income earners, composition effects play in the opposite direction. The relative proportion $h(y|\theta)/\hat{h}(y)$ of low (high) elasticity groups increases (decreases). This pushes down the mean compensated elasticity $\hat{\varepsilon}(y)$ which, in turn, increases the optimal marginal tax rate $T'(y)$. In a nutshell, when optimal marginal tax rates are larger than the current ones, composition effects tend to decrease optimal marginal tax rates for low income levels and to increase marginal tax rates for high income levels. In the next section, we numerically illustrate that the optimal marginal tax rates may be highly sensitive to composition effects.

VII   Numerical illustrations of composition effects

Numerical simulations are key to assess the importance of composition effects. To highlight their role on compensated elasticities, we need to downplay the impact the composition of population may have on social welfare weights. To do so, we assume maximin social preferences. In addition, we rule out income effects by assuming quasilinear preferences as given in Equation (9). This specification entails a constant direct compensated elasticity of earnings with respect to the marginal retention rate, $\varepsilon^\star(w,\theta) = \theta$, and a direct elasticity of earnings with respect to skill normalized to 1, $\alpha^\star(w,\theta) = 1$. The first-order condition (10) still applies with $\tau = T'(y)$. At the optimum, a taxpayer of group θ earning income y is endowed with skill:

$$W(y,\theta) = (1 - T'(y))^{-\theta}\, y \tag{25}$$

Under these specifications, Appendix A.3 shows that the optimal nonlinear tax formula (24a) simplifies to:

$$\frac{T'(y)}{1 - T'(y)} = \frac{1}{\int\limits_{\theta\in\Theta} \theta\, p\left(W(y,\theta)|\theta\right)\, d\hat{\mu}(y,\theta)} \tag{26}$$

where $p(w|\theta)$ is the local Pareto parameter of the conditional skill distribution within group θ:

$$p(w|\theta) = \frac{w\, f(w|\theta)}{1 - F(w|\theta)} \tag{27}$$

and where $\hat{\mu}(y,\theta)$ is the distribution of groups θ among individuals earning an income larger than y. According to (25) the mass of individuals who earn an income larger than y in group θ is $1 - H(y|\theta) = 1 - F\left(W(y,\theta)|\theta\right)$. Using (25), the distribution of groups θ among individuals who earn an income larger than y at the optimum is thus described by the cumulative distribution function:

$$\hat{\mu}(y,\theta) \overset{\text{def}}{\equiv} \frac{\int\limits_{\theta'\in\Theta, \theta'\leq\theta} \left(1 - F\left((1 - T'(y))^{-\theta'}\, y \,\middle|\, \theta'\right)\right)\, d\mu(\theta')}{\int\limits_{\theta'\in\Theta} \left(1 - F\left((1 - T'(y))^{-\theta'}\, y \,\middle|\, \theta'\right)\right)\, d\mu(\theta')} \tag{28}$$

In Equation (26), the optimal marginal tax rate at any income level depends negatively on a weighted average, across groups, of the products of the direct compensated elasticity θ and the local Pareto parameter $p\left(W(y,\theta)|\theta\right)$ within each group. It is clear that the marginal tax rate $T'(y)$ affects these sufficient statistics through the skill level $W(y,\theta)$ at which they need to be evaluated. This makes sufficient statistics distinct in the actual and optimal situations. As already mentioned, this first mechanism of endogeneity is already well established in the standard model where individuals differ along a single dimension (their skills), i.e. where there is a single θ-group. In this model, the optimal marginal tax rate is given by:

$$\frac{T'(y)}{1 - T'(y)} = \frac{1}{\theta \, p\left(W(y)\right)} \tag{29}$$

where the sufficient statistic $p\left(W(y)\right)$ is endogenous to $T'(y)$ (Chetty, 2009).

In tax formula (26), a second mechanism of endogeneity appears, because the distribution of groups among individuals who earn an income larger than y, $\hat{\mu}(y,\theta)$, is endogenous to $T'(y)$ (see Equation (28)). As we move from the actual to the optimal economy, the composition of population behind the same income level is different. Indeed $T'(y)$ is distinct in both economies and individuals who present the same income in the actual economy while coming from different groups have distinct behavioral responses to the change in $T'(y)$.

To know whether one can reasonably neglect composition effects, it is crucial to look at their importance numerically. To illustrate the quantitative role composition effects have on optimal marginal tax rates, we compare the relevant sufficient statistics and optimal marginal tax rates in two economies. In the *single-group* economy, workers differ only along their skills so that composition effects are excluded. This corresponds to the standard unidimensional case with a single source of endogeneity through $p(W(y))$, as described in Equation (29). In the *multidimensional* economy, workers differ in terms of skills w and elasticities θ. Comparing these two economies allows us to conclude about the relevance of composition effects. We perform this comparison according to four distinct scenarios.

VII.1 Calibration

It is well established that male and female workers respond differently to changes in the tax rate (see e.g. Bargain and Peichl (2017)). We therefore perform our calibrations with gender-specific elasticities. This characterizes our multidimensional economy, in which agents differ both in skills and in elasticities. Based on the estimates of male and female labor supply elasticities in Blau and Kahn (2007), we take $\theta_L = 0.1$ for men (with the subscript "L" standing for low-elasticity) and $\theta_H = 0.8$ (with "H" standing for high-elasticity) for women. Gender is an easily observable characteristic for the tax authority but we assume that taxation cannot be conditioned on it for legal reasons.[17] In the single-group economy, we take a mean elasticity

---

[17]Avoiding the use of tagging based on gender seems to us a reasonable assumption even though a few countries had (or still have) tax systems that treat men and women differently (Stotsky, 1997). For instance, the Netherlands moved from granting a higher tax-free allowance to a married man than to a married woman, to an equal basic tax allowance in 1984.

	single-group economy (one group, $\theta = 0.45$)	multidimensional economy (two groups, $\theta_L = 0.1$, $\theta_H = 0.8$)
Scenario 1	Pareto with $p = 1.5$	Pareto distributions with $p_L = p_H = 1.5$
Scenario 2	Pareto with $p = 1.5$	Pareto distributions with $p_L = 1.3$, $p_H = 2$
Scenario 3	kernel + Pareto with $p = 1.5$ if $w \geq 152,314$	kernel on pooled income data + Pareto distribution with $p_L = p_H = 1.5$ if $w \geq 133,401$ for men and if $w \geq 173,963$ for women
Scenario 4	kernel + Pareto with $p = 1.5$ if $w \geq 260,597$	kernel on men income data + Pareto with $p_L = 1.3$ if $w \geq 133,212$ kernel on women income data + Pareto with $p_H = 2$ if $w \geq 180,092$

Table 1: Summary of numerical exercises' calibrations

$\theta = 0.45$ for the single group (pooling both genders).

We now present the four numerical scenarios according to which we contrast the single-group and the multidimensional economies. These scenarios depend on how the remaining parameters of the model (i.e. the local Pareto parameter $p(\cdot|\theta)$ and the conditional skill distributions) vary. In the first two scenarios, we calibrate skill distributions from income assuming the latter is Pareto distributed, whereas in the other two scenarios, we infer skill distributions non-parametrically from income data. The scenarios are summed up in Table 1.

In our first two numerical scenarios, we focus on the top of the income distribution. Our assumption that income is Pareto distributed is in line with empirical evidence at the top. The actual tax schedule is a linear one with a constant marginal tax rate of 40% as in Saez (2001, 2002). This entails that the conditional skill distributions for men and women are also Pareto so that $p(\cdot|\theta)$, in Equation (26), is constant. Therefore, composition effects are the only source of endogeneity. In the first scenario, the single-group economy features a single income distribution, which is specified as Pareto with parameter $p = 1.5$.[18] In the multidimensional economy men and women have their own income distributions, which in this scenario are assumed to be identical (both are specified as Pareto with $p = 1.5$). One then uses the first-order condition of the workers' program (Equation (25)) and these identical income distributions to infer the skill distributions. Because men and women have distinct elasticities, the skill distributions will be different.

In the second scenario, the single-group economy is identical to the one in the first scenario. In the multidimensional economy, men and women's income distributions have distinct Pareto coefficients, with $p_L = 1.3$ for men and $p_H = 2$ for women. The group of male workers then has a fatter upper tail than the group of female workers, which is consistent with the estimates obtained by Atkinson et al. (2016). Figure 1 shows this second specification is consistent, above

---

[18]From estimates in Diamond and Saez (2011, p.170), and Piketty and Saez (2013, p. 424), we know that the top of the income distribution in the U.S. is extremely well approximated by a Pareto distribution and that the implicit Pareto parameter is 1.5.

the \$200,000 threshold and even more so above the \$ 400,000 threshold, with the observed pattern of local Pareto parameters computed by Piketty and Saez (2013).
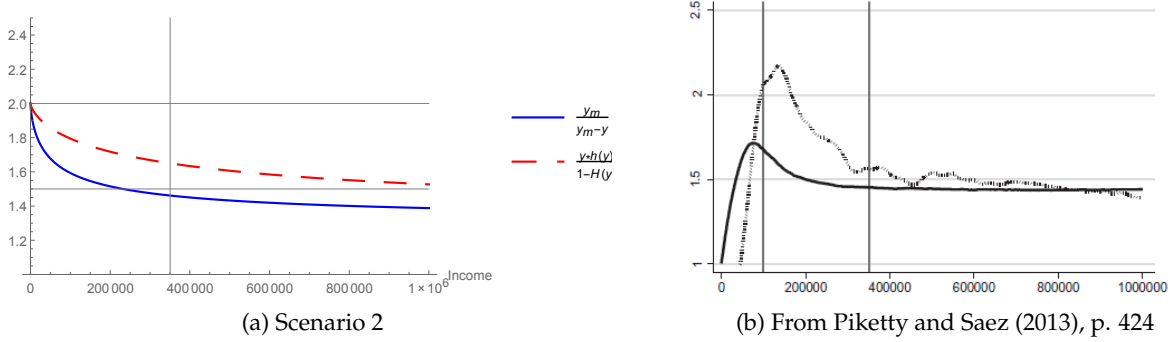

(a) Scenario 2   (b) From Piketty and Saez (2013), p. 424

Figure 1: $y \mapsto \frac{y_m(y)}{y_m(y)-y}$ (solid lines) where $y_m(y)$ is the average income above y and $y \mapsto \frac{y\,\hat{h}_0(y)}{1-\hat{H}_0(y)}$ (dotted lines) where the subscript zero stands for actual economy. The gray vertical lines depict the 99th percentile threshold.

In our third and fourth scenarios, we rely on CPS income data (2016), the observed five different tax brackets[19] and the first-order condition of the workers' program to infer the skill level from each observation of income, using a Gaussian kernel. We consider only individuals who are single, so as to highlight composition effects in the simplest setting without the complexity of interrelated labor supply decisions within families.

In the third scenario, in the single-group economy, we pool together CPS income data for men and women. Since high-income earners are top-coded in the CPS, we expand (in a continuously differentiable way) our kernel estimation of the pooled skill distribution using a Pareto with parameter $p = 1.5$. The extension of the distribution takes place at skill $w = \$152,314$ which correspond to \$133,819 of annual gross income. In the multidimensional economy, we infer men and women skill distributions from the same income distribution. We expand the estimated skill distributions assuming Pareto distributions with the same parameter value $p_L = p_H = 1.5$ for men and women. As in the first scenario, the skill distributions of men and women are distinct because of the distinct elasticities θ. The extensions take place at skill $w = \$133,401$ for men and skill $w = \$173,963$ for women. These skill levels correspond respectively to \$129,618 and \$138,199 of annual gross earnings.

In the fourth scenario, the single-group economy relies on the same estimated and expanded skill distribution as in the third scenario. In the multidimensional economy, we estimate group-specific skill distributions from CPS income data for men and women separately. We expand the estimated skill distributions using group-specific Pareto distributions with parameters $p_L = 1.3$ for men and $p_H = 2$ for women (these values are based on the above-mentioned empirical evidence). These extensions take place at skill $w = \$133,212$ for men

---

[19]Taken from the OECD Taxdatabase.

and skill $w = \$180,092$ for women. These skill levels correspond respectively to $\$129,435$ and $\$143,068$ of annual gross earnings.

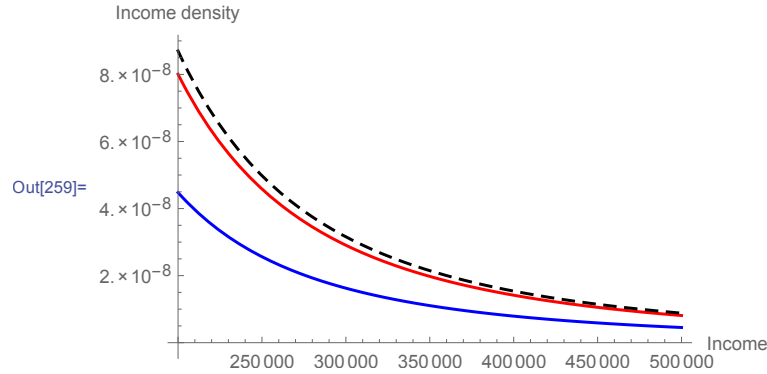## VII.2 Scenarios 1 and 2: Top tax rates and composition effects

Figure 2: Income densities in the actual economy (dotted black curve) and in the (optimal) multidimensional economy for men (red curve) and for women (blue curve) in Scenario 1.

In the first scenario, men and women have identical income distributions but distinct elasticities (lower for men, higher for women) in the multidimensional economy. Whether in the single-group economy or in the multidimensional one, we observe that the top tax rate increases when going from the actual economy to the optimum. Behavioral responses therefore shift income densities to the left. In the single-group economy, going from the actual economy to the optimum makes the optimal tax rate on high incomes rise from 40% to $1/(1+1.5 \times 0.45) \simeq$ 59.7% (from (29))[20]. In the multidimensional economy, since the elasticity of women is higher than the one of men, the income distribution of women shifts more to the left than the one of men, as can be seen on Figure 2. This reduces the mean compensated elasticity from 0.45 to 0.39. In the multidimensional economy, the optimal tax rate[21] then reaches $1/(1+1.5 \times 0.39) \simeq 63\%$. In this scenario, the difference (more than 3 p.p.) between the single-group and the multidimensional economies can be entirely attributed to composition effects. In other words, taking into account the composition effects that arise in the multidimensional case leads to a higher tax rate not only than in the actual economy but also than at the usual unidimensional optimum. Even with hypothetical identical income distributions for men and women and a simplistic linear tax scheme in the actual economy, composition effects modify the optimal top tax rate. One can expect this discrepancy to increase with more complex actual tax schemes and real income data. This will be confirmed with Scenarios 3 and 4.

With the second scenario, the Pareto parameters of the density of male and female workers' income distributions are distinct in the multidimensional economy as empirically observed. When income goes to infinity, the group of male (i.e. low-elasticity) workers whose density has

---

[20]Note that the behavioral responses modify neither the elasticity θ nor the local Pareto parameter in (29).

[21]Since the Pareto coefficient is identical for men and women, from (26), one knows that the optimal nonlinear tax rate does not depend on income.

the lowest Pareto coefficient (i.e. whose density has the fattest upper tail) is the only one left. Therefore, the elasticity and Pareto coefficient of these workers are the only ones that matter to compute the asymptotic optimal marginal tax rate. At the very top, the tax rate is equal to $1/(1+1.3 \times 0.1) \simeq 88.5\%$. For the sake of comparison, the very top tax rate when heterogeneity is one-dimensional (dashed red curve in Figure 3) is constant at $1/(1+1.5 \times 0.45) \simeq 59.7\%$ (i.e., 29 p.p. lower). Then taking or not taking into account composition effects is a key issue in the design of the marginal tax rates that should be faced by the richest workers.

Interestingly, the impact of composition effects does not only concern the very top. For large income levels, the share of men (i.e. workers with a low elasticity and a low Pareto parameter) increases with income. More precisely, in the actual economy, the share of men is calibrated to be equal to 0.5 among individuals who earn an income higher than the top 1% threshold of $350,500$ and it increases to 1 when the threshold goes to infinity. Differing from Scenario 1, p_L is distinct from p_H and $\hat{\mu}(y, \theta)$ (in (28)) does vary with income in the multidimensional economy. The optimal nonlinear tax rate does therefore depend on income (see Equation (26)). Since the share of men (i.e. low-elasticity workers) increases with income, we expect optimal marginal tax rates to increase with income, in the multidimensional economy. Top optimal marginal tax rates are displayed, as a function of income levels, on the left panel of Figure 3 and, in terms of percentiles, on the right panel of the same figure. Note that the top 1, 0.5 and 0.1 percentiles correspond to substantially distinct income levels: $350,050$, $537,100$ and $1,528,500$, respectively. The solid (blue) curves highlight that indeed top optimal marginal tax rate undergo a drastic change, up to 21.3 percentage points, at the top of the income distribution. It rises from 61.0% for the top 1% of the population to 68.7% for the top 0.5% and it reaches 82.3% for the top 0.1%. The differences are also outstanding when one compares these tax rates on the very rich to the 60% obtained in the single-group economy where one neglects composition effects. Indeed, in the multidimensional economy, the income distribution of women (in blue) shifts substantially to the left while the one of men is barely affected, as can be seen on Figure 4. These composition effects drive that, going from the actual to the optimal economy, marginal tax rates decrease below the 99th percentile (around $350,050$) while they substantially increase above this threshold. These results again emphasize the crucial role played by composition effects on rich workers.

To have a clearer picture of the magnitude of the composition effects at play here, we compute the mean (direct) elasticities both in the single-group economy (removing, in doing so, the endogeneity due to composition effects) and in the multidimensional economy. We provide these elasticities for top incomes on Figure 5. Comparing both elasticities allows one to isolate pure composition effects. In the multidimensional economy, the (optimal) mean compensated elasticity with composition effects (solid blue curve), written as:

$$[\theta_L h(y \mid \theta_L) \mu(\theta_L) + \theta_H h(y \mid \theta_H) \mu(\theta_H)] / [h(y \mid \theta_L) \mu(\theta_L) + h(y \mid \theta_H) \mu(\theta_H)],$$

decreases with income due to the left-shift of women density when marginal tax rates increase,

Figure 3: Optimal marginal tax rates with (solid blue curve) composition effects, i.e. Scenario 2 for the multidimensional economy (Pareto income distributions when $p_L = 1.3$ and $p_H = 2$ and $\theta_L = 0.1$ and $\theta_H = 0.8$).The dashed red line is the optimal tax rate without composition effect, i.e. the single-group economy in Scenarios 1 and 2 ($\theta = 0.45$ and $p = 1.5$). The gray vertical line correspond to the 99th percentile.
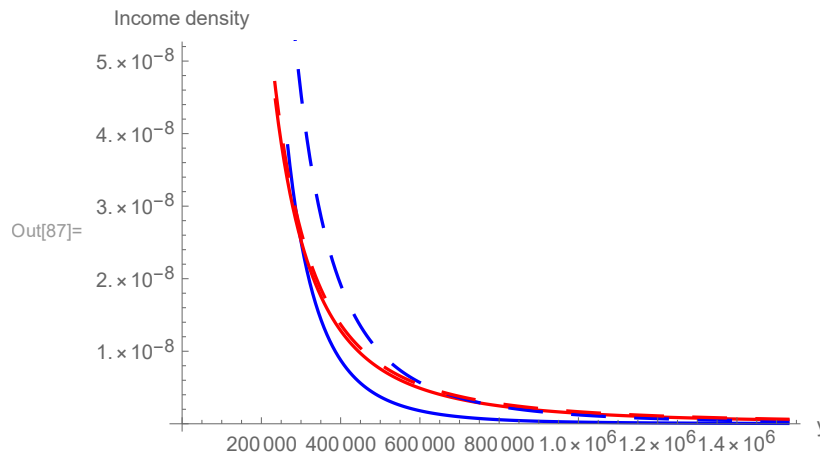

Figure 4: Income densities in the actual economy (dotted curves) and in the (optimal) multidimensional economy (plain curves) for men (in red) and for women (in blue) in Scenario 2.

as explained in Section VI.2. The mean elasticity with composition effects is larger (lower) than the elasticity found in the single-group economy (which is equal to 0.45) for income levels below (above) about \$350.000. More precisely, at \$150.000, the mean elasticity with composition effect is about 30% larger than the elasticity without composition effect. At the 99.5th percentile, the mean elasticity with composition effect is 43% lower than the elasticity without composition effect. Even worse, the mean elasticity with composition effects that takes place at the 99.9 percentile is 65% times lower than the elasticity which excludes composition effects. There is no doubt that incorporating composition effects drastically modifies this crucial sufficient statistic.

Composition effects further affect the study of top tax rates. In the literature, it is generally thought that, for this purpose, a linear tax above a threshold and a nonlinear tax will give equivalent results. Piketty and Saez (2013) (Figure 3, p. 91, reported in Panel (b) of Figure 1 in
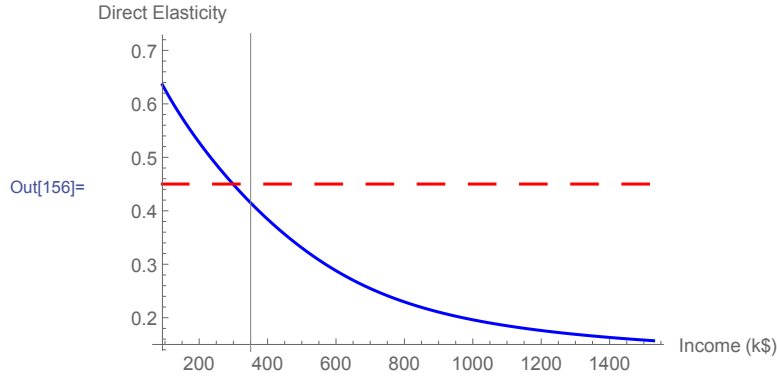
Figure 5: Mean (direct) elasticities with and without composition effects, Scenario 2. The solid blue curve corresponds to the mean elasticities, at the optimum, with composition effects ($\theta_L = 0.1$, $\theta_H = 0.8$, $p_L = 1.3$ and $p_H = 2$). The dashed red line is the elasticity at the optimum with a single group, i.e. without composition effect ($\theta = 0.45$ and $p = 1.5$). The gray vertical line correspond to the 99th percentile.

the present paper) show, under unidimensional individual heterogeneity, that both approaches converge for the highest percentiles of the income distribution. In Appendix A.4, we show that it is not always the case when individual heterogeneity is multidimensional. We highlight that, in this case, major differences may prevail when one calculates the top tax rates with a linear tax above a threshold (as done in e.g. Saez (2001)) instead of a nonlinear tax as done here (see Equation (26)). In both cases, the optimal tax rates are weighted averages of the group-specific products $\theta_i \ p_i$. However, the weights differ. With a linear tax rate above a threshold y, the weights are given by the mass of *excess* incomes above the threshold. By contrast, with a nonlinear income tax, the weights are given by the mass of incomes above y. Indeed the tax perturbation one can use to derive the tax formula with a linear tax above a threshold relies on a constant increase in the tax rate above this threshold. In each group, this change in tax rate applies to each individual income in excess of the threshold and not to the entire individual income as it is the case with a nonlinear tax.

VII.3 Scenarios 3 and 4: Composition effects with CPS data

To check the robustness of our numerical results on the crucial role played by composition effects, we now rely on CPS data and on the actual nonlinear tax schedule. We proceed as in Scenarios 1 and 2. We first assume hypothetical identical income densities in the actual economy for male and female workers in Scenario 3 and we relax this assumption in Scenario 4. Assuming identical income densities for men and women is pedagogically helpful to see how the changes in conditional income densities modify the optimal tax schedule through composition effects. Directly using the observed distinct densities will not allow us to clearly emphasize this. Real data also allows us to enlarge the scope of our results to the entire income distribution. Figure 6 displays the optimal marginal tax rates obtained in the third scenario. Marginal tax rates with composition effects (in solid, blue line) are below the ones obtained without composition effects (in red dotted lines) for incomes below about $50,000$. Conversely,

marginal tax rates with composition effects are above the ones without composition effects, for incomes above $50,000. This confirms the results already obtained with Scenario 2. Moreover, the very top marginal tax rate increases from $1/(1+1.5 \times 0.45) \simeq 59.7\%$ (from (29)) in the single-group economy without composition effects to 67.9% (from (26)) in the economy with composition effects, a substantial difference of 8.2 percentage points.
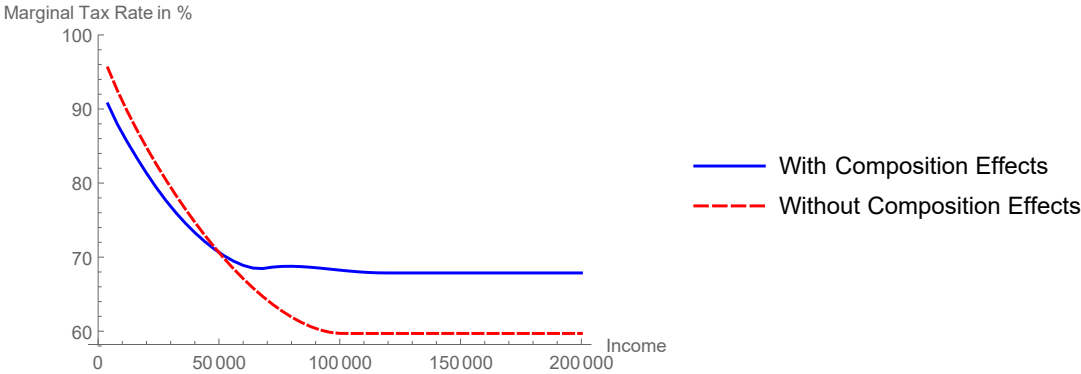

Figure 6: Optimal marginal tax rates with (solid blue line) and without composition effects (dashed red lines) in Scenario 3 with CPS (2016) income distribution extended by a Pareto distribution.
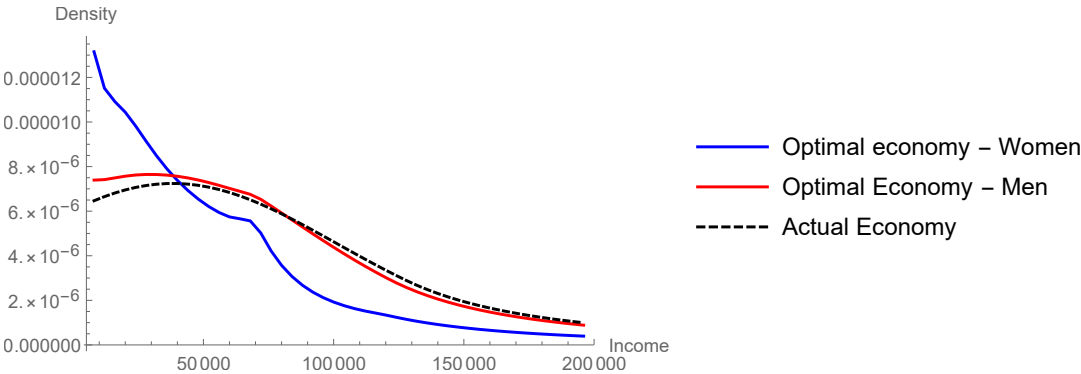

Figure 7: Densities in the actual economy in Scenario 3 with CPS (2016) income distribution extended by a Pareto distribution and optimal income densities in the multidimensional economy

The dotted black curve in Figure 7 gives the conditional income densities which are assumed identical for men and women in Scenario 3. Figure 7 also displays how the conditional income densities of men (red curve) and women (solid blue curve) diverge in the multidimensional economy where the tax policy has switched from the actual tax schedule to the optimal one. The conditional density of men is almost not modified compared to the actual economy which is not surprising since their elasticity $\theta_L = 0.1$ is quite low. In contrast, we see on Figure 7 that more women end up at the bottom of the income distribution in the optimal economy.[22]

---

[22]The conditional income density of women (see Figure 7) and the profile of direct elasticity of women (see Figure 8) exhibit a sort of kink around $70,000. This income level corresponds to the skill level at which the Pareto expansion takes place for women. Although this expansion is realized in a continuously differentiable way, the first derivative of the density varies very rapidly around the expansion. This in turns implies locally a rapid change in
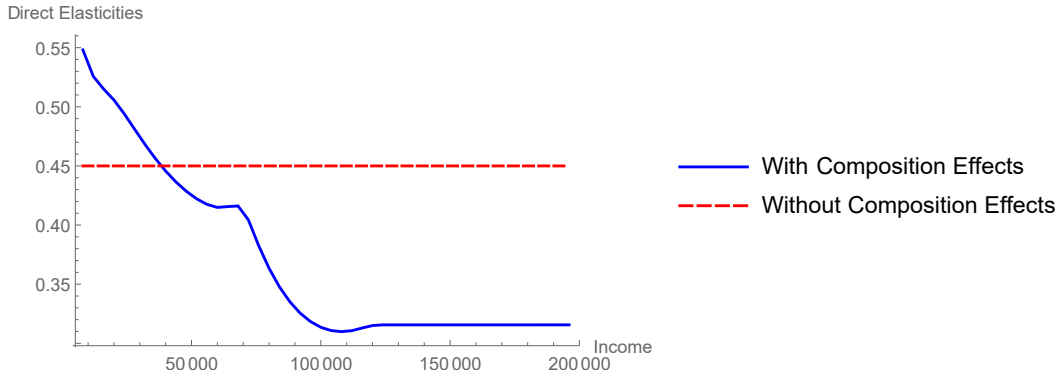
Figure 8: Mean (direct) elasticities with and without composition effects, Scenario 3

Because of this important change in the income density of women, when incorporating composition effects, the mean (direct) elasticity increases for incomes below $40,000$ while it decreases beyond this level, as can be seen in Figure 8. A similar impact was observed in Scenario 2 (see Figure 5). In Figure 8, we isolate pure composition effects by comparing the mean (direct) compensated elasticities with and without composition effects. The economy with one group has a direct elasticity of 0.45 as discussed earlier. In the optimal economy with two groups, the mean compensated elasticity

$$[\theta_L h\left(y\,|\theta_L\right)\mu(\theta_L) + \theta_H h\left(y\,|\theta_H\right)\mu(\theta_H)] \,/\, [h\left(y\,|\theta_L\right)\mu(\theta_L) + h\left(y\,|\theta_H\right)\mu(\theta_H)]$$

drastically decreases with income. It starts at 0.66 for very low incomes and reaches 0.31 for high income levels. The difference between mean direct elasticities with and without composition effects is definitely not negligible. As expected from Equation (26), a reduction (rise) of the optimal mean elasticity goes hand in hand with a rise (decrease) of the marginal tax rate. Indeed, when going from the single-group to the multidimensional economy, the marginal tax rates increase above the $40,000$ threshold, while they substantially decrease below this income level, as can be seen in Figure 6. This scenario confirms the decisive role played by composition effects.

In Scenario 4, we relax the assumption of Scenario 3 according to which income densities for men and women are identical under the actual tax schedule. Instead both conditional income densities are calibrated using the gender specific income distribution of men and women from CPS data and the gender specific calibration of Pareto parameters (see Table 1). As can be seen on Figure 9, we again find that composition effects reduce marginal tax rates (by as much as 6 p.p.) below an income threshold (around $60,000$) and increases them above this threshold. As in Scenario 3, this goes hand in hand with the fact that the income density of women (with a relatively high elasticity) is much more concentrated to the left than that of men (whose

---

the local Pareto parameter for women and in the marginal tax rate. It induces a sudden surge in the corrective term (see Equation (20c)). Note that this change in the corrective term increases with the difference between men and women elasticities. From Equation (21), the rapid rise in the corrective term induces a substantial increase of the women income density around $70,000$.
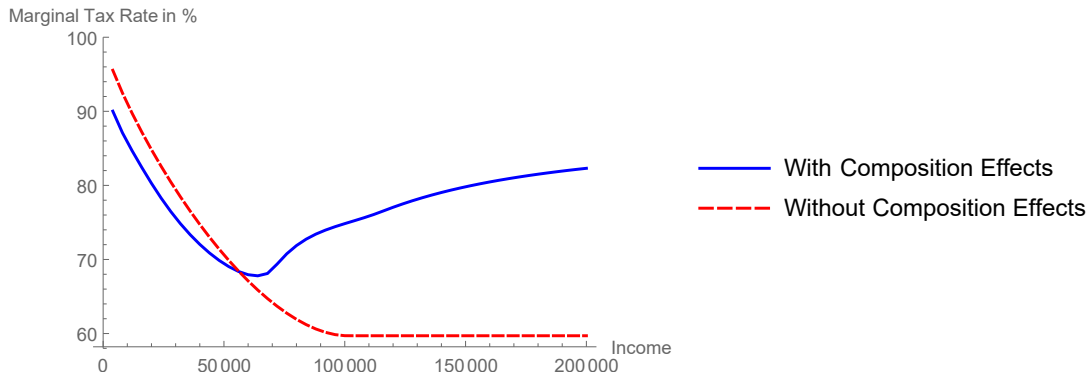
Figure 9: Optimal marginal tax rates with (solid blue line) and without composition effects (dashed red lines) in Scenario 4 with CPS (2016) income distribution extended by a Pareto distribution
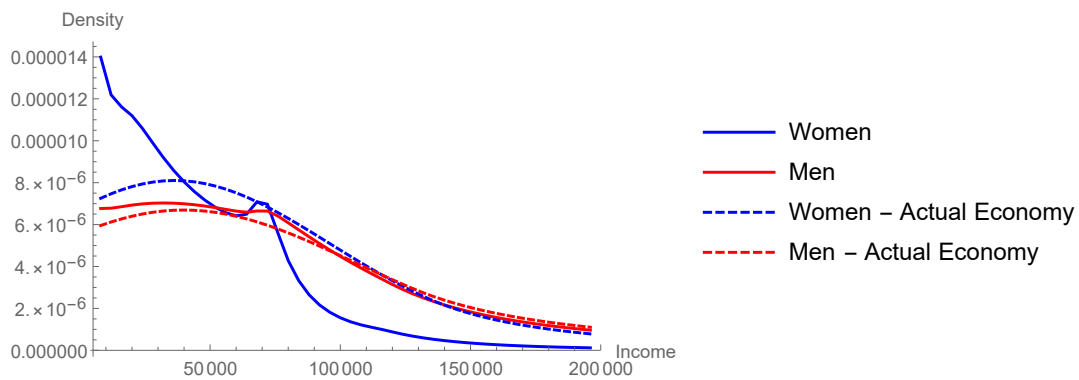

Figure 10: Densities in the actual economy in Scenario 4 with CPS (2016) income distribution extended by a Pareto distribution (dashed red and blue lines) and optimal income densities in the multidimensional economy (red and blue lines)

elasticity is relatively low), as can be seen on Figure 10.[23] Compared to Scenario 3, the fatter upper tail for men than for women increases the discrepancy between optimal marginal tax rates at the top with and without composition effects. Marginal tax rates on high incomes are 10 to 20% larger when one takes composition effects into account. On Figure 11, we compare the mean (direct) compensated elasticities with and without composition effects. The difference between mean elasticities with and without composition effects are again non-negligible.

These numerical results put the stress on the need for (more) empirical studies on sufficient statistics conditional on income to derive the shape of optimal marginal income tax.

VIII   Tax perturbation method versus mechanism design approach

In this section, we show the equivalence between the tax perturbation approach (which relies on the sufficient conditions in Assumption 2) and the mechanism design approach, assuming individual characteristics are multidimensional. This equivalence is established under

---

[23]Note that the sort of kinks that occur around $70,000 for the conditional income density of women (see Figure 10) and the profile of direct elasticity of women (see Figure 11) prevail for the same reason as in Scenario 3. This is detailed in Footnote 22.
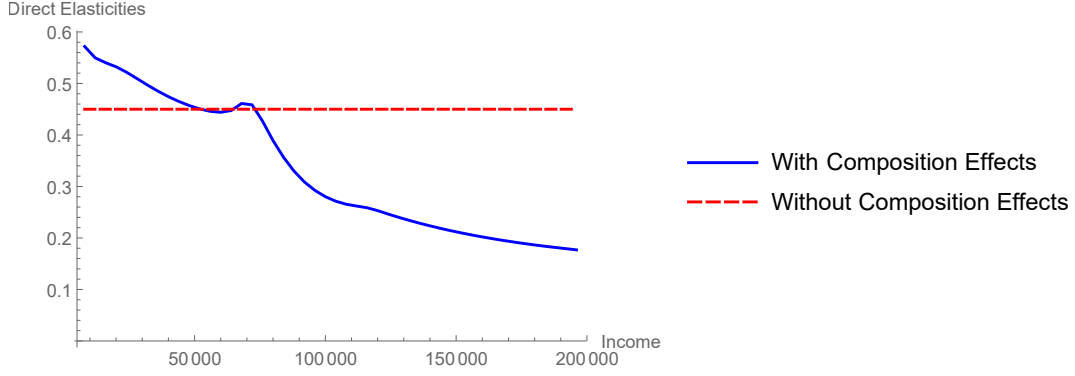
Figure 11: Mean (direct) elasticities with and without composition effects, Scenario 4

the within-group single-crossing condition (Assumption 1).

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 allocation $(w, \theta) \mapsto (C(w, \theta), Y(w, \theta))$ that verifies the incentive constraints,

$$\forall w, \theta, w', \theta' \in (\mathbb{R}_+^* \times \Theta)^2 \qquad \mathscr{U}\left(C(w, \theta), Y(w, \theta); w, \theta\right) \geq \mathscr{U}\left(C(w', \theta'), Y(w', \theta'); w, \theta\right). \quad (30)$$

According to (30), individuals of type (w, θ) are better off 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 (Salanié, 2005). Then, incentive constraints (30) imply the first-order incentive constraints, i.e.

$$\forall (w, \theta) \in \mathbb{R}_+^* \times \Theta \qquad \dot{U}(w, \theta) = \mathscr{U}_w\left(C(w, \theta), Y(w, \theta); w, \theta\right) \qquad (31)$$

These first-order incentive constraints are necessary but not sufficient to verify the incentive constraints (30). A sufficient condition is that the allocation also verifies a monotonicity constraint according to which in each group, $Y(\cdot, \theta)$ is nondecreasing in skill. We adopt a slightly more restrictive assumption.

Assumption 3. *The allocation $(w, \theta) \mapsto (C(w, \theta), Y(w, \theta))$ is smooth if and only if it is continuously differentiable, it verifies (30) 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 3 in the first-order mechanism design approach. The proof is in Appendix A.5.

Proposition 2. *Under Assumption 1,*

i) *Any tax schedule $y \mapsto T(\cdot)$ verifying Assumption 2 (i.e. the conditions for the tax perturbation) induces a smooth allocation that verifies Assumption 3.*

29

ii) Any smooth allocation verifying Assumption 3 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 3 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 $Y(\cdot, \theta)$ admitting a strictly positive derivative in skill as required in Assumption 3. In the mechanism design approach, the latter condition is related to the second order incentive constraints. 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 3.

Thanks to Proposition 2, first-order mechanism design and tax perturbation approaches are analog. The (first-order) mechanism design approach consists in choosing, among the allocations that verify Assumption 3, the one that maximizes the social objective (6) subject to the budget constraint (5). It involves computing the first-order effect, on the Lagrangian (7), of a small perturbation of the optimal allocation within the set of smooth and incentive compatible allocations. Since the allocation after perturbation has to verify Assumption 3, it is decentralized by a tax schedule that has to verify Assumption 2. Therefore, as stated in Proposition 2, the effects of a perturbation of the allocation that preserves Assumption 3 are equivalent to the responses of the allocation to a perturbation of the tax function that preserves Assumption 2. In other words, the mechanism design approach focuses on the effects of an allocation perturbation whereas the tax perturbation approach focuses on the effects of the tax reform that decentralizes this perturbation of the allocation. For this reason, the mechanism design approach and the tax perturbation approach are the two faces of the same coin.

In the literature where the unobserved heterogeneity is unidimensional, the mechanism design approach can be developed under less restrictive assumptions than Assumption 3. 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. This is excluded with the tax perturbation because of Assumption 2 but has been largely studied with the mechanism design approach. Note that the alternative "pathology" where individuals may be indifferent between two levels of income 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).

IX Concluding Comments

In this paper, we provide formulas to calculate sufficient statistics in the presence of multidimensional individual heterogeneity. Multidimensional heterogeneity generates a new channel

through which sufficient statistics differ in the optimal and actual economies. We call this additional channel "composition effects". These effects are due to the modification of the average behavioral response at each income level. Using optimal tax policy as the field of choice to illustrate our point, we emphasize the key role of composition effects in the calculation of sufficient statistics. We show that neglecting composition effects entails a potentially large bias on sufficient statistics hence on optimal tax schedules. We quantify this bias through a few numerical examples. Our results stress the need for empirical studies on labor supply elasticities and distribution parameters for different demographic groups e.g., according to age, ethnicity and gender. This paper is clearly a call for more empirical evidence within sub-groups of population at distinct levels of income to clarify the importance of composition effects in the design of optimal tax schedules. This is part of our research agenda.

As a second main contribution, we prove the equivalence, when unobserved heterogeneity is multidimensional, between the tax perturbation method and (first-order) mechanism design. Both methods have been used separately to solve optimal income tax problems. While the latter method is widely used in various fields in economics, the former is more specific to the optimal taxation literature (although it may find applications in industrial organization as well). Having ascertained their equivalence (and sufficient conditions of this equivalence) is therefore an important result.

To illustrate the generality of our results in this concluding section, we now provide examples of tax problems that one can easily solve in our framework. For each of them, we explain what y, w, θ represent so that the interpretation of the results is straightforward. In the vein of this paper, we choose optimal tax problems but our framework even extends beyond optimal taxation, e.g., to nonlinear pricing problems where consumers differ along several unobserved dimensions.

Optimal joint taxation of labor and non-labor income

Consider individuals have two sources of taxable income: a non-labor income z and a labor income $y - z$. Those incomes are jointly taxed and the tax function does not distinguish between both incomes. This applies, for instance, in countries like France where incomes received from renting property are jointly taxed with labor income. As explained in Scheuer (2014), a single nonlinear tax schedule is the system that is in place for employed workers and self-employed small business owners in many countries, including the U.S.. In this case, y is the total taxable income and we interpret θ as the ability to earn non-labor income z and w as the skill. Individuals of type (w, θ) solve:

$$\max_{y,z} \quad \mathcal{U}\left(y - T(y), y - z, z; w, \theta\right)$$

where two decision variables appear instead of one variable in the core of our paper. This program can be solved sequentially, the first step being the choice of non-labor income z for a given taxable income y which leads to $\mathscr{U}(c, y; w, \theta) \overset{\text{def}}{=} \max_{z} \mathcal{U}(c, y - z, z; w, \theta)$. The second step is the choice of y as in Equation (1). In the process, one simply needs to ensure the semi-

indirect utility function $\mathscr{U}(\cdot, \cdot; w, \theta)$ verifies Assumption 1 to apply our tax formulas.

Optimal joint income taxation of couples

The joint income taxation of couples is a variant of the previous application, in which $y - z$ is the labor income of one individual and z is the one of his/her partner. The tax does not distinguish between $y - z$ and z and only depends on the sum of both incomes, y (as in France, Germany and the US). We redefine w and θ as the respective skill level of each member of the couple. The optimal tax schedules derived in this paper are then interpreted as the optimal tax schedules when the couple is the tax unit and each partner decides along the intensive margin. Our model opens a new avenue for research on this topic (e.g., Brett (2007), Kleven et al. (2009) and Cremer et al. (2012)).

Optimal income taxation with tax avoidance

In this application, w is the skill and θ is the ability to avoid taxation. We assume that tax enforcement (penalty, monitoring, etc.) is given. We denote z the sheltered labor income (i.e. income that is not taxed at all) and $y + z$ the (total) labor income. The tax only depends on the taxable income y. Consumption becomes $c + z$, with $c = y - T(y)$ being the after-tax income. All results obtained in this paper are valid in this context when one simply makes sure that Assumption 1 holds.

A   Appendix

A.1   Proof of Lemma 1

Let $\mathscr{L}^R$ be the Lagrangian that results from applying a reform with a direction R and magnitude m on the Lagrangian (7):

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

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

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

We use (8), (16) and (19c) to obtain the first equality. We use (21) for the change of variable from skill w to income y in the second equality. Note the role of the within-group single-crossing

condition (Assumption 1) behind this change of variable. It implies that in each group, income is an increasing function of skill with a strictly positive derivative. Therefore, in each group, the income density is continuous without any mass point nor hole. We use (22a)-(22d) for the third equality. Integrating by parts the integral of $\frac{T'(y)}{1-T'(y)} y \, \hat{\varepsilon}(y) \, \hat{h}(y) \, R'(y)$ leads to (23).

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

$$U^{\tilde{R}}(m; w, \theta) \overset{\text{def}}{=} \max_{y} \quad \mathscr{U}\left(y - T(y) + m\,R(y) + \ell(m), y; w, \theta\right) \tag{32}$$

Applying the envelope theorem to (32) at $m = 0$ yields:

$$\frac{\partial U^{\tilde{R}}}{\partial m}(0; w, \theta) = \left(R(y) + \ell'(0)\right) \, \mathscr{U}_c\left(C(w, \theta), Y(w, \theta); w, \theta\right) \tag{33}$$

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 (18b), (19b) and (19c) leads to:

$$\frac{\partial Y^{\tilde{R}}}{\partial m}(0; w, \theta) = \frac{\partial Y^{R}}{\partial m}(0; w, \theta) + \eta(w, \theta)\,\ell'(m) \tag{34}$$

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 with magnitude m. We get

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

We then obtain:

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

Using (34), we can then write:

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

so that:

$$\ell'(0) = \frac{1}{1 - \iint\limits_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} T'(Y(w, \theta))\,\eta(w, \theta)\,f(w|\theta)dw\,d\mu(\theta)} \frac{\partial \mathscr{B}^{R}}{\partial m}(0) \tag{35}$$

Finally, using (33), we get:

$$
\begin{aligned}
\frac{\partial \mathscr{S}^{\tilde{R}}}{\partial m}(0) &= \frac{\partial \mathscr{S}^R}{\partial m}(0) + \ell'(0) \iint_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} \Phi'_u(U(w,\theta);w,\theta)\, \mathscr{U}_c(C(w,\theta),Y(w,\theta);w,\theta)\, f(w|\theta)dw\, d\mu(\theta) \\
&= \frac{\partial \mathscr{S}^R}{\partial m}(0) + \frac{\displaystyle\iint_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} \Phi'_u(U(w,\theta);w,\theta)\, \mathscr{U}_c(C(w,\theta),Y(w,\theta);w,\theta)\, f(w|\theta)dw\, d\mu(\theta)}{1 - \displaystyle\iint_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta)dw\, d\mu(\theta)}\, \frac{\partial \mathscr{B}^R}{\partial m}(0) \\
&= \lambda\, \frac{\partial \mathscr{L}^R}{\partial m}(0)
\end{aligned}
\tag{36}
$$

where the latter equality holds if and only if

$$
\lambda = \frac{\displaystyle\int_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} \Phi'_u(U(w,\theta);w,\theta)\, \mathscr{U}_c(C(w,\theta),Y(w,\theta);w,\theta)\, f(w|\theta)dw\, d\mu(\theta)}{1 - \displaystyle\int_{(w,\theta)\in\mathbb{R}^*_+\times\Theta} T'(Y(w,\theta))\, \eta(w,\theta)\, f(w|\theta)dw\, d\mu(\theta)}
\tag{37}
$$

A.2 Proof of Proposition 1

An optimal tax system implies that any tax reform $R(.)$ does not yield any first-order effect on the Lagrangian (7). That is (23) is nil at $m = 0$ for any direction $R(\cdot)$. 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 any income y, we have:

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

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 (24a). Making y tends to 0 in (24a) and using $\lim_{y\mapsto 0}\frac{T'(y)}{1-T'(y)}\hat\varepsilon(y)\, y\, \hat h(y) = 0$ leads to (24b).

A.3 Derivation of Equation (26)

Under maximin ($\hat g(y) = g(w,\theta) = 0$), without income effects ($\hat\eta(y) = \eta(w,\theta) = 0$) and using Equations (22a)-(22d), Equation (24a) can be rewritten as:

$$
\frac{T'(y)}{1-T'(y)}\int_{\theta\in\Theta} \varepsilon(W(y,\theta),\theta)\frac{y\, h(y|\theta)}{1-H(y|\theta)}d\hat\mu(y,\theta) = 1
\tag{38}
$$

where $d\hat\mu(y,\theta) = \frac{1-H(y|\theta)}{1-\hat H(y)}d\mu(\theta)$, the latter being derived from $\hat\mu(y,\theta)$ in Equation (28).

In the optimal economy, from (25), we have $H(y|\theta) \equiv F\left((1-T'(y))^{-\theta}y\,\middle|\,\theta\right)$. Differentiating both sides of this equality with respect to income y leads to:

$$
\begin{aligned}
h(y|\theta) &= \left(1 + \frac{y\, T''(y)\,\theta}{1-T'(y)}\right)(1-T'(y))^{-\theta}\, f\left((1-T'(y))^{-\theta}y\,\middle|\,\theta\right) \\
y\, h(y|\theta) &= \left(\frac{1-T'(y)+y\, T''(y)\,\theta}{1-T'(y)}\right)W(y,\theta)\, f\left(W(y,\theta)|\theta\right) \\
\varepsilon(y,\theta)\, y\, h(y|\theta) &= \theta\, W(y,\theta)\, f\left(W(y,\theta)|\theta\right) \\
\varepsilon(y,\theta)\frac{y\, h(y|\theta)}{1-H(y|\theta)} &= \theta\frac{W(y,\theta)\, f\left(W(y,\theta)|\theta\right)}{1-F(W(y,\theta)|\theta)} = \theta\, p\left(W(y,\theta)|\theta\right)
\end{aligned}
\tag{39}
$$

where the third equality uses (20a) and $\varepsilon^\star(y,\theta) = \theta$ (under quasilinear and isoelastic individual preferences) and the latter equality uses $H(y|\theta) = F(W(y,\theta)|\theta)$. Plugging (39) into (38) and using the definition of te Pareto parameter $p(w|\theta)$ (Equation (41)) leads to (26).

34

A.4 Top tax rate calculated with a linear versus a nonlinear tax schedule

In this appendix, we emphasize that taking composition effects into account leads to major differences between top tax rates calculated with a linear tax function for all incomes above a given threshold and top tax rates calculated with a nonlinear tax schedule as done in e.g. Section 4 of Saez (2001). This exercise finds its motivation in the fact that the literature highlights that a linear tax above a threshold and a nonlinear tax will generally give equivalent results (Piketty and Saez, 2013, Figure 3, p. 91, reported in Panel (b) of Figure 1) which we show is not the case with multidimensional heterogeneity (with different Pareto parameters and elasticities across groups).

To make the comparison, we first study the optimal linear tax rate τ for workers with income above threshold y. As in the core of the paper, a variable with the subscript zero is in the actual economy and a variable with an asterisk is at the optimum. In the actual economy, the upper part of the income density within each group with elasticity θ_i is Pareto:

$$h_0(y|\theta_i) = k_i p_i \, y^{-(1+p_i)} \qquad \Leftrightarrow \qquad 1 - H_0(y|\theta_i) = k_i \, y^{-p_i} \text{ with } i := L, H \tag{40}$$

where k_i is the scale parameter and p_i is here the local Pareto parameter of the conditional *income* distribution within group θ:

$$p_i = p(y|\theta_i) = \frac{y \, h_0(y|\theta_i)}{1 - H_0(y|\theta_i)} \tag{41}$$

Also note that the mean of incomes above income y within the group with elasticity θ_i is labeled $y_m(y|\theta_i)$ and it simplifies to:[24]

$$y_m(y|\theta_i) = \frac{p_i}{p_i - 1} y. \tag{42}$$

The optimal top (linear) tax rate for income levels above the threshold y solves:

$$\tau_* = \frac{1}{1 + \widetilde{\mu}_L(y, \tau_0, \tau_*) \, \theta_L \, p_L + \widetilde{\mu}_H(y, \tau_0, \tau_*) \, \theta_H \, p_H} \qquad \text{where} \tag{43}$$

$$\widetilde{\mu}_i(y, \tau_0, \tau_*) \stackrel{\text{def}}{\equiv} \frac{\dfrac{\mu(\theta_i) \, k_i}{p_i - 1} \left(\dfrac{1 - \tau_*}{1 - \tau_0}\right)^{\theta_i p_i} y^{-p_i}}{\dfrac{\mu(\theta_L) \, k_L}{p_L - 1} \left(\dfrac{1 - \tau_*}{1 - \tau_0}\right)^{\theta_L p_L} y^{-p_L} + \dfrac{\mu(\theta_H) \, k_H}{p_H - 1} \left(\dfrac{1 - \tau_*}{1 - \tau_0}\right)^{\theta_H p_H} y^{-p_H}} \quad \text{with } i := L, H.$$

Proof of Equation (43)

Using Equation (10) that we have obtained with the optimal linear tax rate in Section IV, we can write the income earned in the actual economy by an individual who earns y in the optimal economy as:

$$\tilde{Y}_0(y, \theta_i) = \left(\frac{1 - \tau_0}{1 - \tau_*}\right)^\theta y. \tag{44}$$

---

[24]Using (40), the mean of incomes above income y within group with elasticity θ_i is:

$$y_m(y|\theta_i) = \frac{\int_y^{+\infty} z p_i z^{-1-p_i} dz}{\int_y^{+\infty} p_i z^{-1-p_i} dz} = \frac{\left[\frac{z^{1-p_i}}{1-p_i}\right]_z^{+\infty}}{\left[\frac{z^{-p_i}}{-p_i}\right]_y^{+\infty}} = \frac{p_i}{p_i - 1} y.$$

which is the right-hand side of (42).

From this equation, we can write

$$H_*(y|\theta_i) = H_0(\tilde{Y}_0(y,\theta_i)|\theta_i). \tag{45}$$

Differentiating both sides of (45) in y and using (40) and (44), we obtain:

$$h_*(y|\theta_i) = k_i\, p_i \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_i\, p_i} y^{-(1+p_i)} = \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_i\, p_i} h_0(y|\theta_i) \text{ with } i := L, H.$$

and

$$1 - H_*(y|\theta_i) = k_i \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta\, p_i} y^{-p_i} = \left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta\, p_i} (1 - H_0(y|i)). \tag{46}$$

To obtain the optimal τ, we can determine the usual mechanical and behavioral effects following a small increase in the tax rate. The total mechanical effect consists in summing, across groups, the extra taxes that individuals who earn incomes larger than the threshold y pay when one incrementally modifies the linear tax rate that prevails on all incomes above the threshold:

$$M = \sum_{i:=L,H} (1 - H(z|\theta i))\, \mu(\theta_i)[y_m(y|\theta_i) - y] \text{ with } i := L, H.$$

All individuals above y mechanically pay extra taxes. For each worker, the change in the linear tax rate applies to the amount of her income which is above the threshold. The mechanical effect is therefore proportional to the mass $1 - H(y|\theta i)$ of workers who face the increase in tax rate (i.e. all workers whose incomes are above the threshold) times the difference between the mean of incomes above threshold y and this threshold, i.e. $y - y_m(y|\theta_i) = \frac{y}{p_i-1}$, using Equation (42).

This extra tax payment also creates behavioral responses from everyone above the threshold. This leads to an additional change in collected taxes. In the group of elasticity θ_i, the mass of workers above the threshold, $1 - H(y|\theta_i)$, has a behavioral response which is proportional to $-\frac{\tau}{1-\tau}y_m(y|\theta)\theta_i(1 - H(y|\theta_i))$. One can note that this response is proportional to the mean income above the threshold. Aggregating behavioral responses across all top bracket taxpayers yields:

$$B = \frac{\tau}{1-\tau}\sum_{i:=L,H} \theta_i y_m(y|\theta_i)(1 - H(z|\theta i))\,\mu(\theta_i) \text{ with } i := L, H.$$

At the optimum, $M + B = 0$. Using (40), (42) and the fact that the mass of workers above the threshold can be rewritten (thanks to (46)) as $1 - H(y|\theta i)) = k_i\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta\, p_i} y^{-p_i}$ gives:

$$\frac{\tau_*}{1-\tau_*} = \frac{\frac{1}{p_L-1}\mu(\theta_1)\,k_L\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_L p_L} y^{-p_L} + \frac{1}{p_H-1}\mu(\theta_H)\,k_H\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_H p_H} y^{-p_H}}{\mu(\theta_L)\,k_L\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_L p_L} y^{-p_L}\theta_L\frac{p_L}{1-p_L} + \mu(\theta_H)\,k_H\left(\frac{1-\tau_*}{1-\tau_0}\right)^{\theta_H p_H} y^{-p_H}\theta_H\frac{p_H}{1-p_H}}$$

which can be rewritten as (43). $\square$

We can now compare the top tax rate formula when a linear tax prevails above a threshold, Equation (43), to the tax formula obtained when a nonlinear tax prevails, Equation (26). In both cases, the optimal tax rates are weighted averages of the group-specific products $\theta_i\, p_i$. However, the weights associated with an optimal nonlinear income tax rate are different from those associated with a linear income tax rate above a threshold. When computing optimal nonlinear income tax at income y, one can see from Equations (26) and (28) that the weights are given by the share of individuals in group θ_i above income y, i.e. the mass of incomes above y. This stands in stark contrast with the optimal tax rate that is linear above a threshold, for

which the weights are given, as we have just stated, by the mass of *excess* incomes above the threshold. Intuitively, this is because the tax perturbation one can use to derive this tax formula relies on a constant increase in the tax rate above threshold y. In each group θ_i, this change in tax rate applies to each individual income in excess of the threshold, $y_m(y|\theta_i) - y$.
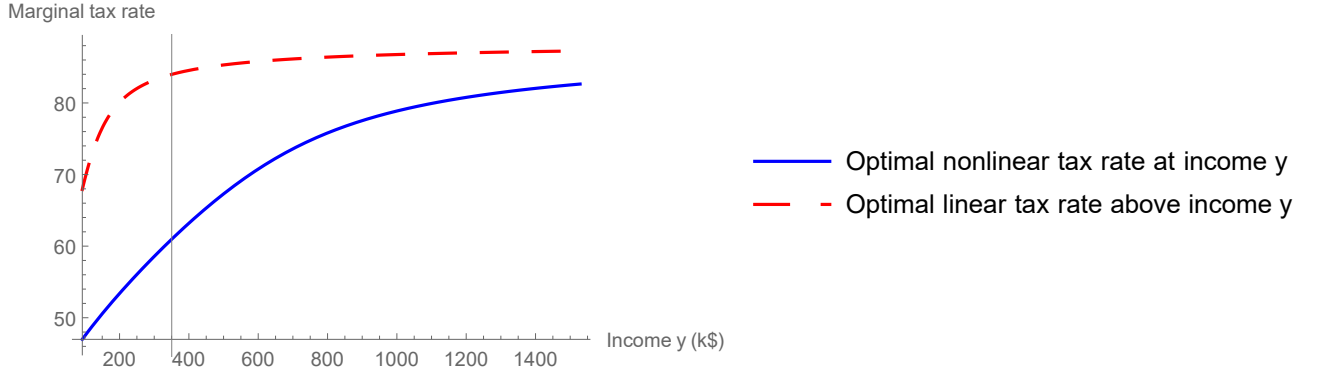

Figure 12: Nonlinear vs linear above a threshold tax schedules.

We now numerically illustrate the important difference in optimal tax rates when one uses the linear-above-a-threshold rather than the nonlinear tax schedule to calculate top tax rates. Figure 12 depicts the optimal nonlinear tax schedule (blue curve) and the optimal linear-above-a-threshold tax schedule (hyphenated red curve). To obtain the latter, we use the income level at which the tax rate is calculated as the threshold beyond which a linear tax rate prevails. The linear tax rate increases with the threshold. More precisely, it rises from 84% with a threshold set at $375,0000$ (which corresponds to the 99th percentile) to 87% with a threshold set at $1,000,0000$. The optimal linear-above-a-threshold tax rate is then quite sensitive to the choice of threshold beyond which one calculates the linear tax rate. Moreover, when one compares the optimal nonlinear and linear-above-a-threshold tax rates at the top of the income distribution, one sees drastic differences. At the 99th percentile, the linear-above-a-threshold tax rate is 25 percentage points higher than the nonlinear tax rate. One needs to calculate tax rates beyond $1,400,000$ to see this difference eventually reduced to 6 percentage points, which is still a significant difference. When one determines the optimal top tax rates, assuming a linear tax at the top (i.e., above a threshold) yields a substantially different tax rate than the one obtained with a nonlinear tax schedule.

A.5 Proof of Proposition 2

Part *i*) of Proposition 2.

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 (19d). 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 latter inequality leads to (30). Therefore the allocation $w \mapsto (C(\cdot, \theta), Y(\cdot, \theta))$ induced by $T(\cdot)$ verifies (30), thereby Assumption 3.

Part *ii*) of Proposition 2

Let $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$ be a mapping defined over $\mathbb{R}_+^* \times \Theta$ which verifies Assumption 3. 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 off with the bundle $(C(w'),Y(w'))$ designed for individuals of type (w',θ'), which would be in contradiction with (30). 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:

$$\forall y \in \mathbb{Y} \qquad T(y) \stackrel{\text{def}}{\equiv} Y(w,\theta) - C(w,\theta) \qquad \text{where } (w,\theta) \text{ are such that: } y = Y(w,\theta) \qquad (47)$$

Given this tax schedule, Program (1) of individuals of type (w,θ) is equivalent to:

$$\max_{(w',\theta')\in\mathbb{R}_+^*\times\Theta} \mathscr{U}\left(C(w',\theta'), Y(w',\theta'); w,\theta\right),$$

the solution of the latter is (w,θ) since $(w,\theta) \mapsto (C(w,\theta), Y(w,\theta))$ verifies the incentive constraints (30). Therefore, the tax schedule $T(\cdot)$ defined by (47) decentralizes the given allocation.[25]

We now 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 with a strictly positive derivative. 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 (47) can be rewritten on $\mathbb{Y}(\theta)$ by:

$$T(y) = y - C\left(Y^{-1}(y,\theta),\theta\right) \qquad (48)$$

Moreover, we get that $\mathbb{Y} = \cup_{\theta\in\Theta}\mathbb{Y}(\theta)$ and is therefore an open set. Hence, for each income $y \in \mathbb{Y}$, there exists a group θ such that $T(\cdot)$ verifies (48) in the neighborhood of y.

To show that $T(\cdot)$ verifies Part *i*) of Assumption 2, note that from (48), $T(\cdot)$ is continuously differentiable as $Y^{-1}(\cdot,\theta)$ and $C(\cdot,\theta)$ are continuously differentiable. Moreover, from (2), we have:

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

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. Therefore, $T'(\cdot)$ is continuously differentiable and $T(\cdot)$ verifies Part *i*) of Assumption 2.

To show that $T(\cdot)$ verifies Part *ii*) of Assumption 2,[26] note that the first-order condition (17) can be rewritten as $\mathscr{Y}(Y(w,\theta);w,\theta) \equiv 0$ for all skill levels. Differentiating this equality with respect to skill leads to: $\mathscr{Y}_y(Y(w,\theta);w,\theta)\dot{Y}(w,\theta) + \mathscr{Y}_w(Y(w,\theta);w,\theta) = 0$. As $\mathscr{Y}_w(Y(w,\theta);w,\theta) = -\mathscr{M}_w(C(w,\theta), Y(w,\theta);w,\theta)$ which is positive from Assumption 1 and $\dot{Y}(w,\theta) > 0$ from Assumption 3, then one must have $\mathscr{Y}_y(Y(w,\theta);w,\theta) < 0$, 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 consumption

---

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

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

an individual of type (w, θ) should get to enjoy utility u while earning 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 $\mathscr{C}_w = -\mathscr{U}_w/\mathscr{U}_c$. Let us 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 they earn income y'. Therefore, $\mathscr{Q}(w)$ is a 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 (from which they obtain utility $U(w, \theta)$) and the utility they would get by earning income y' and consuming $y' - T(y')$. We have by assumption $\mathscr{Q}(w^*) = 0$. We obtain:

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

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) with 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

Arkolakis, Arnaud Costinot, and Andres Rodriguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Atkinson, A. B., A. Casarico, and S. Voitchovsky, "Top Incomes and the Gender Divide," Melbourne Institute Working Paper Series Working Paper 27/16 2016.

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

Bargain, Olivier and Andreas Peichl, "Steady-State Labor Supply Elasticities: An International Comparison," *IZA Journal of Labor Economics*, 2017, *Forthcoming.*

Bierbrauer, Felix J. and Pierre C. Boyer, "Politically feasible reforms of non-linear tax systems," CEPR Discussion Paper 13059 2018.

Blau, F. and L. Kahn, "Changes in the labor supply behavior of married women: 1980-2000," *IZA Journal of Labor Economics*, 2007, *25*, 393–438.

Blumkin, Tomer, Efraim Sadka, and Yotam Shem-Tov, "International tax competition: zero tax rate at the top re-established," *International Tax and Public Finance*, 2015, *22* (5), 760–776.

Boadway, R. and L. Jacquet, "Optimal Marginal and Average Income Taxation under Maximin," *Journal of Economic Theory*, 2008, *143*, 425–441.

Boadway, Robin, Katherine Cuff, and Maurice Marchand, "Optimal Income Taxation with Quasi-linear Preferences Revisited," *Journal of Public Economic Theory*, 2000, *2* (4), 435–60.

__, **Maurice Marchand, Pierre Pestieau, and María del Mar Racionero**, "Optimal Redistribution with Heterogeneous Preferences for Leisure," *Journal of Public Economic Theory*, 2002, *4* (4), 475–498.

Brett, Craig, "Optimal Nonlinear Taxes for Families," *International Tax and Public Finance*, 2007, *14*, 225–261.

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

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

Choné, Philippe and Guy Laroque, "Negative Marginal Tax Rates and Heterogeneity," *American Economic Review*, 2010, *100* (5), 2532–47.

Christiansen, V., "Which commodity taxes should supplement the income tax?," *Journal of Public Economics*, 1984, *24(2)*, 195–220.

Cremer, Helmuth, Jean-Marie Lozachmeur, and Pierre Pestieau, "Income taxation of couples and the tax unit choice," *Journal of Population Economics*, 2012, *25* (2), 763–778.

Cuff, Katherine, "Optimality of Workfare with Heterogeneous Preferences," *The Canadian Journal of Economics*, 2000, *33(1)*, 149–174.

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

__ **and Emmanuel Saez**, "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 2011, *25* (4), 165–90.

__ **and James Mirrlees**, "Optimal Taxation and Public Production II: Tax Rules," *American Economic Review*, 1971, *61* (3), 261–278.

Ebert, Udo, "A Reexamination of the Optimal Nonlinear Income Tax," *Journal of Public Economics*, 1992, *49* (1), 47–73.

Feldstein, Martin, "Behavioral Responses to Tax Rates: Evidence from TRA86," *American Economic Review*, 1995, *85*, 170–174.

Fleurbaey, Marc and François Maniquet, *Theory of Fairness and Social Welfare*, Cambridge: Cambride University Press, 2011.

__ **and __**, "Optimal Taxation Theory and Principles of Fairness," *Journal of Economic Literature*, 2017, *Forthcoming*.

Gerritsen, Aart, "Optimal nonlinear taxation: the dual approach," *mimeo*, 2016.

Golosov, Mikhail, Tsyvinski, and Nicolas Werquin, "A Variational Approach to the Analysis of Tax Systems," Mimeo (older version available as NBER working paper 20780) 2014.

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan, "Differential Taxation and Occupational Choice," *The Review of Economic Studies*, 2017, *Forthcoming.*

Gruber, Jon and Emmanuel Saez, "The Elasticity of Taxable Income: Evidence and Implications," *Journal of Public Economics*, 2002, *84*, 1–32.

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

__ **and Jean-Jacques Laffont**, "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*, 1984, *25*, 329–369.

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

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

Hendren, Nathaniel, "Efficient Welfare Weights," NBER Working Papers issued in 2014, re-

vised in October 2017 20351, 2017.

Hornstein, Andreas, Per Krusell, and Giovanni L. Violante, "Frictional Wage Dispersion in Search Models: A Quantitative Assessment," *American Economic Review*, 2011, *101* (7), 2873–2898.

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden, "Optimal redistributive taxation with both extensive and intensive responses," *Journal of Economic Theory*, 2013, *148* (5), 1770–1805.

Kleven, H. and C. Kreiner, "The marginal cost of public funds: Hours of work versus labor force participation," *Journal of Public Economics*, 2006, *90*, 1955 – 1973.

Kleven, Henrik J, "Sufficient Statistics Revisited," mimeo 2018.

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

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez, "The Optimal Income Taxation of Couples," *Econometrica*, 2009, *77* (2), 537–560.

Lehmann, Etienne, Laurent Simula, and Alain Trannoy, "Tax Me If You Can! Optimal Nonlinear Income Tax between Competing Governments," *Quarterly Journal of Economics*, 2014, *129* (4), 1995–2030.

Lockwood, Benjamin B. and Matthew Weinzierl, "De Gustibus non est Taxandum: Heterogeneity in Preferences and Optimal Redistribution," *Journal of Public Economics*, 2015, *124*, 74–80.

Lollivier, S. and J.-C. Rochet, "Bunching and Second-Order Conditions: A Note on Optimal Tax Theory," *Journal of Economic Theory*, 1983, *31*, 392–400.

Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," *The Review of Economic Studies*, 1971, *38*, 175–208.

Piketty, Thomas, "La redistribution fiscale face au chômage," *Revue Française d'Economie*, 1997, *12*, 157–201.

_ **and Emmanuel Saez**, "Optimal Labor Income Taxation," in Martin Feldstein Alan J. Auerbach, Raj Chetty and Emmanuel Saez, eds., *Handbook of public economics*, Vol. 5, Elsevier, 2013, chapter 7, pp. 391 – 474.

Ramsey, Frank, "A Contribution to the Theory of Taxation," *Economic Journal*, 1927, *37*, 47–61.

Roberts, Kevin, "A Reconsideration of the Optimal Income Tax," in P. Hammond and G. Myles, eds., *Incentives and Organization: Essays in Honour of Sir James Mirrlees*, 2000.

Rochet, Jean-Charles and A. Stole Lars, "Nonlinear Pricing with Random Participation," *The Review of Economic Studies*, 2002, *69* (1), 277–311.

Rothschild, Casey and Florian Scheuer, "Redistributive Taxation in the Roy Model," *The Quarterly Journal of Economics*, 2013, *128* (2), 623–668.

_ **and** _ , "A Theory of Income Taxation under Multidimensional Skill Heterogeneity," Mimeo (older version available as NBER working paper 19822 2014.

_ **and** _ , "Optimal Taxation with Rent-Seeking," *The Review of Economic Studies*, 2016, *83* (3), 1225–1262.

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin, "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," NBER Working Papers 22646, 2016.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *The Review of Economic Studies*, 2001, *68*, 205–229.

_ , "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses," *Quarterly Journal of Economics*, 2002, *117*, 1039–1073.

_ , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, 2

(1), 180–212.

_ **and Stefanie Stantcheva**, "Generalized Social Marginal Welfare Weights for Optimal Tax Theory," *American Economic Review*, 2016, *106* (1), 24–45.

_ **and** _ , "A Simpler Theory of Optimal Capital Taxation," *Journal of Public Economics*, 2018, *162*, 120–142.

_ **, Joel Slemrod, and Seth H. Giertz**, "The Elasticity of Taxable Income with Respect to Marginal Tax Rates: A Critical Review," *Journal of Economic Literature*, 2012, *50* (1), 3–50.

Salanié, Bernard, *The Economics of Contracts*, Cambridge, MA: MIT Press, 2005.

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

_ , "Entrepreneurial Taxation with Endogenous Entry," *American Economic Journal: Economic Policy*, 2014, *6* (2), 126–63.

_ **and Iván Werning**, "Mirrlees meets Diamond-Mirrlees: Simplifying Nonlinear Income Taxation," Mimeo (older version available as NBER working paper 22076) 2016.

_ **and Ivan Werning**, "The Taxation of Superstars," *Quarterly Journal of Economics*, 2017, *132*, 211–270.

Spiritus, Kevin, "Characterizations for Optimal Multidimensional Taxation with Multidimensional Heterogeneity of Agents," mimeo 2017.

Stotsky, Janet G., "How Tax Systems Treat men and Women Differently," *Finance and Development*, 1997, *34(1)*, 30–33.

Werning, Iván, "Pareto Efficient Income Taxation," MIT Working Paper 2007.