Comprehensive or Separate Income Tax? A Sufficient Statistics Approach

Very Preliminary Version

Marie-Noëlle LEFEBVRE* Etienne LEHMANN† Michaël SICSIC‡

CRED(TEPP) Université Paris II Panthéon-Assas§

November 15, 2019

Abstract

In this paper, we investigate how to tax the different sources of income of taxpayers. We consider an optimal nonlinear income tax model with many sources of income. We first exhibit a specification where the optimal tax system consists in a nonlinear schedule that applies to the sum of all income - a *comprehensive* income tax system - and another specification where the optimal tax system consists in a nonlinear schedule specific to each income - a *separate* income tax system. In the more general environment, we specialize the tax schedule to be combination of these two polar systems: the tax system is restricted to be the sum of a comprehensive personal income tax schedule and of income specific tax schedules, we derive an optimal ABC formula for each of these schedules. we also derive a condition expressed in terms of empirically meaningful sufficient statistics under which decreasing the indexation of the personal income tax base on one income and compensating the revenue loss with a lump-sum or a proportional increase in the taxation of that income is socially desirable.

Keywords: Nonlinear Income Taxation. Dual Income Tax, Comprehensive Income Tax

I Introduction

Taxpayers receive different kinds of incomes such as labor income, interest income, dividends, capital gains or losses, business income, rents or imputed rents. There exist two polar systems to tax these different incomes. First, under a *comprehensive* income tax system, tax liability is a function of the sum of all of these incomes, which is close in spirit to what is done in the U.S. or in the U.K. Conversely, under a *separate* income tax system, each income is taxed according to an income-specific schedule. A particular case of separate income tax system is the *dual* income tax where capital income is excluded from the personal income tax base and taxed under a specific proportional schedule. Sweden in 1991, Norway in 1992, Finland in 1993, Spain in 2006 and Germany moved to a dual tax system by exiting a large part of their capital

*email: marie-noelle.lefebvre@u-paris2.fr. Webpage: http://cred.u-paris2.fr/lefebvre.

†email: elehmann@u-paris2.fr. Webpage: http://cred.u-paris2.fr/lehmann. Etienne LEHMANN is also research fellow at CEPR, IZA and CESIfo.

‡email: michael.sicsic@etudiants.u-paris2.fr. Webpage: http://cred.u-paris2.fr/sicsic.

§Université Panthéon-Assas Paris II, 12 Place du Panthéon, 75 234 Paris Cedex 05, France.

income from the personal income tax base. Denmark has now a mixed system (Kleven and Schultz, 2014). In France, dividends were taxed in a dual way from 2007 to 2012, and France has now opted for dual taxation of capital income since 2017. In Netherlands, the 2001 reform moved from the personal income tax from a comprehensive system to a separate system where incomes financial wealth are exited from the personal income tax base and wealth taxed applies to financial wealth (Zoutman, 2018).

There exist informal arguments in favor or against the move to a separate income tax system (e.g. Boadway (2004)).

- On the one hand, a separate income tax system enables the government to shift the burden of taxation to the least responsive tax base. We refer to this argument as the *Ramsey argument* because it focuses on the best way to shift the burden of redistribution across the different tax bases, just as the Ramsey (1927) optimal formula describes the optimal way to shift the burden of taxation across the different commodities. In practice, thanks to dual taxation, Nordic countries succeed in keeping a very progressive personal income tax with high marginal tax rates without harming saving and investment.[1] Intuitively, the Ramsey argument plays in favor of seprating

- On the other hand, a separate income tax system triggers incentives for income shifting, especially for business and self-employed incomes. We thus refer to this argument as the *income-shifting* argument, which intuitively plays in favor of comprehensive tax schedules.

- Moreover, by reducing capital tax rate and keeping high tax rates on labor earnings, moving to a separate income tax benefit high capital earners which is frequently viewed as unfair. We thus refer to this argument as the *equity argument*.

While these reforms generate huge controversies in policy-advising arena and many empirical evaluations, the arguments in favor or against a separate income tax system remain informal. In this paper, we develop a framework to investigate when a (more) separate or a (more) comprehensive tax schedule is desirable.

We start in Sections III and IV by providing two specialization of the model under which we can show the optimal tax schedule is comprehensive (in Section III) or separate (in Section IV).

In Section III, we assume individuals are endowed with weakly separable preferences. This very specific assumption implies that all taxpayers make the same decisions when deciding how to split their efforts across the different tax bases to get a given level of total income. In such a case, we show that a comprehensive tax schedule is sufficient to decentralize the optimal allocation. The argument is similar to the argument in Atkinson and Stiglitz (1976) against

[1] Moreover, a dual income tax system is much simpler to enforce as the tax liability of one income does no longer depend on the other incomes. Enforcement costs is a frequent argument in practice that is not considered in the present paper.

commodity taxation: distorting the choice of efforts across the different tax bases is useless because this does not relax the equity-efficiency trade off under weakly separable preferences.

We then consider in Section IV a specialization where unobserved heterogeneity is one-dimensional and preferences are quasilinear and additively separable. Under these specific assumptions, as high labor income earners are also high capital income earners, whether the burden of taxation should be imposed on labor or on capital is solely an efficient concern, without any equity implication. The one dimensional assumption thus ensures the validity of the above-mentioned Ramsey (1927) argument. In such a situation, the government wants to adapt the marginal tax rate specific to each income to the tax responsiveness specific to this tax base. Moving away from the comprehensive income tax is then necessary to adjust the marginal tax base to its responsiveness. The assumptions of quasilinear and additively separable preferences then guarantees the separate income tax system is sufficient to decentralize the optimal allocation.

These two specializations are much too specific to be empirically plausible. They are useful to formalize under which conditions the pros and cons of each polar system are valid. They are also useful to show that none of these two polar system is generically optimal. Turning to the more general case, to bypass the technical difficulty of multidimensional screening (Mirrlees, 1976, Golosov et al., 2014, Renes and Zoutman, 2017, Spiritus, 2017), we restrict in Section V the tax schedules to be the sum of a comprehensive personal income tax schedule $T_0(\cdot)$ and of n income specific tax schedules (see Equations (8) and (9) below). While theoretically restrictive, this assumption approximate fairly well most of actual tax systems in OECD. Moreover, most of tax reforms on policymakers' agenda can be considered within the class of tax schedules that we consider. We then obtain the following results.

First, we derive optimal marginal tax rate formulas which are specific to each tax base (including the comprehensive taxable income), See Equation (26). These formulas extend the formulas of Diamond (1998) and Saez (2001) for the existence of different tax bases. It clarifies that cross base effects have to be taken into account. This is because a change in the marginal tax rate on one tax base also triggers behavioral responses from the other tax bases, which in turn affect tax revenue. Hence, these cross-base compensated responses have to show up in optimal tax formulas.

Second, we investigate the effect of a reform that changes the indexation of the personal taxable income on a given income. Such a reform is an incremental move towards a more separate or a more comprehensive tax system. On the top of mimicking an uncompensated change in the marginal tax rate, such a reform also affects the personal income tax base, which in turns change the marginal tax rate associated to each tax base, which finally induces compensated responses. We then consider the effects of a incremental change in the indexation parameter of one type of income reform. We derive a formula expressed in terms of empirically meaningful sufficient statistics that compute the amount of lump-sum tax required to compensate the loss in tax revenue. Calibrating this formula on a new exhaustive datasets of French taxpayers, we obtain that the lump-sum compensating tax is lower *i*) the larger the elasticity of capital income with respect to its own marginal net of tax rate, in line with the above-mentioned Ramsey argument and *ii*) the larger the cross-base response of labor income to the capital marginal net of tax rate. If this cross-base response is explained by income-shifting it is negative, so this result is in line with the above-mentioned income-shifting argument.

We also study the case when the marginal exit of one income from the personal income tax base is financed by a proportional tax on this income. This is typically the way policymakers thinks about financing reforms towards a more separate income tax system. If the personal income tax schedule was linear, such effect would be neutral. Conversely, when the personal income schedule is progressive, we derive a formula that states under which conditions in terms of empirically meaningful sufficient statistics such a reform is socially desirable. Our numerical simulations on French data suggest how much the proportional tax should be increased to compensate the reform is relatively insensitive to the behavioral elasticities.

Literature review to be added (very preliminary text)

This paper is organized as follows. The model is presented in Section II. Section III describes a case where the optimal tax schedule is comprehensive. Section IV describes a case where the optimal tax schedule is separate. Finally, Section V consider in the more general case the effects of incremental reforms towards a more separate or a more comprehensive tax system. The numerical application is presented in Section VI.

II The Economy

II.1 Taxpayers

The economy is populated by a unit mass of taxpayers characterized by different types denoted $\mathbf{w}$ belonging to the type space W. Individuals take $n \geq 2$ different actions, which are costly to them. Each action generates a specific income denoted y_i, which is observable by the government. For instance, y_1 can be salary income, y_2 business income, y_3 dividends, etc. Let: $\mathbf{y} = (y_1, ..., y_n)$ denote the vector of incomes or tax bases earned by a taxpayer. The preference of individuals of type $\mathbf{w}$ over after-tax income c (hereafter *consumption*) and tax base $\mathbf{y}$ is described by the utility function $\mathscr{U} : (c, \mathbf{y}; \mathbf{w}) \mapsto \mathscr{U}(c, \mathbf{y}; \mathbf{w})$, which is assumed twice continuously differentiable over $\mathbb{R}_+^{n+1} \times W$. Utility increases in consumption so $\mathscr{U}_c > 0$, decreases in efforts, thereby in income, so $\mathscr{U}_{y_i} < 0$. Let:

$$\mathcal{S}^i(c, \mathbf{y}; \mathbf{w}) \stackrel{\text{def}}{\equiv} -\frac{\mathscr{U}_{y_i}(c, \mathbf{y}; \mathbf{w})}{\mathscr{U}_c(c, \mathbf{y}; \mathbf{w})} \tag{1}$$

denote the marginal rate of substitution between the i^{th} income and consumption. We assume that indifference set are convex. This implies that the matrix[2] $\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j\right]_{i,j}$ is positive definite (see Appendix A).

Types are distributed according to the continuously differentiable density function $f : \mathbf{w} \mapsto f(\mathbf{w})$, which is defined over the convex type space W. Unless otherwise specified, types corresponds to $n \geq 2$ different characteristics denoted $w_1, ..., w_n$, so $\mathbf{w} = (w_1, ..., w_n)$. The type space is denoted W and is a convex.

The government imposes a tax schedule $\mathcal{T} : \mathbf{y} = (y_1, ..., y_n) \mapsto \mathcal{T}(y_1, ..., y_n)$ that depends on each of these incomes. Hence, the after-tax income c of a taxpayer earning tax bases $\mathbf{y}$ is: $c = \sum_{i=1}^n y_i - \mathcal{T}(y_1, ..., y_n)$. Taxpayer of type $\mathbf{w}$ solves:

$$U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,...,y_n)} \quad \mathscr{U}\left(\sum_{k=1}^n y_k - \mathcal{T}(y_1, ..., y_n), \mathbf{y}; \mathbf{w}\right) \tag{2}$$

We assume (see Assumption 1 discussed in II.3) that for each type $\mathbf{w} \in W$, this program admits a single solution denoted $\mathbf{Y}(\mathbf{w}) = (Y_1(\mathbf{w}), ..., Y_n(\mathbf{w}))$. Individuals of type $\mathbf{w}$ consume $C(\mathbf{w}) = \sum_{i=1}^n Y_i(\mathbf{w}) - \mathcal{T}(\mathbf{Y}(\mathbf{w}))$ and enjoy utility level $U(\mathbf{w}) = \mathscr{U}(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w})$. The first order-conditions are:

$$\forall i \in \{1, ..., n\} : \qquad 1 - \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) = \mathcal{S}^i(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w}) \tag{3}$$

Finally, $h(\mathbf{y})$ denotes the joint density of tax bases $\mathbf{y} = (y_1, ..., y_n)$, while for each type of income, $h_i(y_i)$ denotes the unconditional density of the i^{th} income.

II.2 Government

The government faces the following budget constraint:

$$E \leq \mathscr{B} \stackrel{\text{def}}{\equiv} \int_{\mathbf{w}\in W} \mathcal{T}(\mathbf{Y}(\mathbf{w}))\ f(\mathbf{w}) d\mathbf{w} \tag{4}$$

where $\mathscr{B}$ stands for the tax revenue and where $E \geq 0$ is an exogenous amount of public expenditure to finance. The government's objective sums an increasing transformation Φ of taxpayers' individual utility $U(\mathbf{w})$ that may be concave and type-dependent:

$$\mathscr{O} \stackrel{\text{def}}{\equiv} \int_{\mathbf{w}\in W} \Phi(U(\mathbf{w}); \mathbf{w})\ f(\mathbf{w}) d\mathbf{w} \tag{5}$$

When the government is utilitarian, the social transformation is $\Phi(U, \mathbf{w}) = U$ and is linear. When the government has weighted utilitarian preferences, the social transformation takes the form $\Phi(U, \mathbf{w}) = \gamma(\mathbf{w})\ U$. When the government has Bergson-Samuelsonian preferences, the social transformation does not depend on type and is concave in U.

There are different specialization of tax schedules that will be considered in this paper.

[2]We use notation $\left[A_{i,j}\right]_{i,j}$ to denote a square matrix of size n whose term of row i and column j is $A_{i,j}$. Superscript T denotes the transpose operator $\left[A_{i,j}\right]^T_{i,j} = \left[A_{j,i}\right]_{i,j}$. Matrix $\left[A_{i,j}\right]^{-1}_{i,j}$ is the inverse of matrix $\left[A_{i,j}\right]_{i,j}$ and "$\cdot$" denotes the matrix product.

Comprehensive Income Tax system

The tax schedule is said to be *comprehensive* if it takes the form: $\mathcal{T}(\mathbf{y}) = T\left(\sum_{k=1}^{n} y_k\right)$ where $T(\cdot)$ is defined on $\mathbb{R}_+$. The marginal tax rate on each income is then identical, so the first-order conditions (3) simplify to:

$$1 - T'\left(\sum_{k=1}^{n} Y_k(\mathbf{w})\right) = \mathcal{S}^1\left(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w}\right) = ... = \mathcal{S}^n\left(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w}\right) \tag{6}$$

In particular, the marginal rate of substitution $\mathscr{U}_{y_i}/\mathscr{U}_{y_j} = \mathcal{S}^i/\mathcal{S}^j$ between the i^{th} and the j^{th} income is equal to one and is unaffected by taxation under a comprehensive income tax schedule. In other words, the comprehensive tax system does not distort how taxpayers shift their effort among the different tax bases.

Separate Income tax system

The tax schedule is said to be *separate* if it takes the form: $\mathcal{T}(\mathbf{y}) = \sum_{k=1}^{n} T_k(y_k)$ where the $T_k(\cdot)$ schedules are defined on $\mathbb{R}_+$. The marginal tax rate on each income then depends only on this income (i.e. $\mathcal{T}_{y_i y_j} = 0$ if $i \neq j$), so the first-order conditions (3) become:

$$\forall i \in \{1, ..., n\} \qquad 1 - T_i'\left(Y_i(\mathbf{w})\right) = \mathcal{S}^i\left(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w}\right) \tag{7}$$

In other words, with a separate income system, the distortions induced by the tax system on each tax base are independent.

Mixed tax system

We also consider a *mixed* tax system where the tax schedule is assumed to be the sum of a *personal income* tax schedule $T_0(\cdot)$ and of n *income specific* tax schedules $T_i(\cdot)$:

$$\mathcal{T}(\mathbf{y}) = T_0\left(\sum_{k=1}^{n} a_k\, y_k\right) + \sum_{k=1}^{n} T_k(y_k) \tag{8}$$

The personal income tax schedule $T_0(\cdot)$ depends on the *personnal income tax base* or *taxable income* denoted y_0 and we denote $Y_0(\mathbf{w}) = \sum_{k=1}^{n} a_k\, Y_k(\mathbf{w})$. Not all incomes are included in the personal income tax base, and not all income are necessarily fully included in the personal income tax base. For instance, in most OECD countries, it is not primary labor income paid by employers that enters the personal income tax base but labor income after the payment of (employers) social security contributions. Therefore, if y_1 denote primary labor earnings, $a_1 y_1$ denotes taxable labor earnings net of payroll taxes. Similarly, when dividends are included in the personal income tax base, these dividends have previously been taxed though corporate taxation. Hence, if y_2 denotes the primary profits earned by a taxpayer, $a_2 y_2$ denotes taxable dividends, etc. These are the reasons why we consider that the personal income tax base is defined by:

$$y_0 \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k\, y_k \tag{9}$$

where each a_i is a policy instruments that captures how much taxable income y_0 depends on the i^{th} income. Each a_i takes a value between 0 and 1.[3]

Moreover, the tax system is also made of n income-specific tax schedules $T_i(\cdot)$. These schedules add taxes paid by firms and by households on a given tax base.

Under the tax schedule (8), the marginal tax rate on the j^{th} income adds the marginal tax rate $T_j'(y_j)$ of the schedule specific to this income plus a_j times the marginal tax rate $T_0'(y_0)$ of the personal income tax schedule:

$$\mathcal{T}_{y_j}(\mathbf{y}) = T_j'(y_j) + a_j \, T_0' \left(\sum_{k=1}^{n} a_k \, y_k \right) \tag{10}$$

Therefore all incomes affects the j^{th} marginal tax rate through the determination of the taxable income y_0 in (9).

While restrictive, the form of tax schedules in (8) approximate fairly well most of tax systems in OECD economies. It includes the specific case where the tax schedule is purely *comprehensive*, in which case $a_1 = ... = a_n = 1$ and for all i, $y_i \mapsto T_i(y_i) \equiv 0$ and the case where the tax schedule is purely *separate*, in which case $y_0 \mapsto T_0(y_0) \equiv 0$.

II.3 Responses to tax reforms

To analyze the consequences of infinitesimal tax reforms we follow the tax perturbation approach of Golosov et al. (2014).[4] This consists in considering various one-dimensional families of *perturbed* tax schedules called a perturbation.

Definition 1. *A tax perturbation is a twice continuously differentiable mapping $(\mathbf{y}, x) \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x)$ defined over $\mathbb{R}_+^n \times I$, where x denotes the algebraic magnitude of a tax reform and I is an open interval containing* 0 *such that:*

- *For all $\mathbf{y} \in \mathbb{R}_+^n$, one has $\tilde{\mathcal{T}}(\mathbf{y}, 0) = \mathcal{T}(\mathbf{y})$.*
- *After a tax reform of magnitude x, taxpayers face the tax schedule $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x)$*

There are different examples of tax perturbations that are of particular interest. First, the *lump-sum* tax perturbation:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x \tag{11a}$$

consists in a lump-sum transfer x to every taxpayers. We henceforth denote $\frac{\partial Y_i(\mathbf{w})}{\partial \rho}$ the response of the i^{th} income to such lump-sum perturbation.

Second, the *uncompensated* tax perturbation of the j^{th} marginal tax rate:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x \, y_j \tag{11b}$$

[3]There is a normalization issue as for any λ, one can reproduce the same personal income tax with parameter $\hat{a}_i = a_i \lambda$ and personal income tax schedule $\mathbf{y} \mapsto \hat{T}(\sum_{k=1}^n \hat{a}_k y_k)$ defined by $y_0 \mapsto \hat{T}(y_0) \stackrel{\text{def}}{\equiv} T_0(y_0/\lambda)$

[4]See also Hendren (2017). An heuristic version of the tax perturbation approach has been exposed by Piketty (1997) and Saez (2001). An earlier application to linear commodity taxation is exposed in Christiansen (1981).

consists in increasing by x the j^{th} marginal net of tax rate $1 - \mathcal{T}_{y_j}$. We denote $\frac{\partial Y_i{}^u(\mathbf{w})}{\partial \tau_j}$ the response of the i^{th} income to such perturbation to an uncompensated increase in the j^{th} marginal tax rate.

The uncompensated response is due to both income and substitution effects. To isolate the latter, we also consider the *compensated* tax perturbation of the j^{th} marginal tax rate at tax base $\mathbf{Y}(\mathbf{w})$:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x \ \left(y_j - Y_j(\mathbf{w})\right) \tag{11c}$$

which consists in increasing by x the j^{th} marginal net of tax rate $1 - \mathcal{T}_{y_j}$, while leaving unchanged tax liability at $\mathbf{y}$. Such a reform is *compensated* at tax base $\mathbf{Y}(\mathbf{w})$, because tax liability is unchanged at $\mathbf{y} = \mathbf{Y}(\mathbf{w})$, whatever the magnitude x. We denote $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}$ the compensated response of the i^{th} marginal tax rate to the j^{th} marginal tax rate at tax base $\mathbf{Y}(\mathbf{w})$. Appendix B in particular shows that compensated responses are symmetric, i.e. that: $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} = \frac{\partial Y_j(\mathbf{w})}{\partial \tau_i}$.

More generally, a tax perturbation in the *direction* $R : \mathbf{y} \mapsto R(\mathbf{y})$ is defined by:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x \, R(\mathbf{y}) \tag{11d}$$

Finally, a tax perturbation of the i^{th} tax base is defined by:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(y_1, ..., y_{i-1}, (1+x)y_i, y_{i+1}, ..., y_n) \tag{11e}$$

For each tax perturbation, we want to compute the derivative of economic magnitude with respect to x at $x = 0$, i.e. for any economic variable X, computing

$$\left.\frac{\partial X}{\partial x}\right|_{x=0} \stackrel{\text{def}}{\equiv} \lim_{x \mapsto 0} \frac{X|_{\tilde{\mathcal{T}}(\cdot,x)} - X|_{\mathcal{T}(\cdot)}}{x}$$

This notation is obviously meaningful conditional on the tax perturbation $\tilde{\mathcal{T}}(\cdot,\cdot)$ one has in mind. To compute such derivatives, we want to apply the implicit function theorem to the first-order conditions in (3). We thus consider only *unperturbed* tax schedules that verifies the following assumption.[5]

Assumption 1. *The tax schedule* $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ *is such that:*

i) The tax schedule is twice continuously differentiable.

ii) The second-order condition holds strictly, that is matrix $\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j}$ *is positive definite.*

iii) For each type $\mathbf{w} \in W$*, program* (2) *admits a single global maximum.*

Part i) of Assumption 1 ensures that first-order conditions (3) are differentiable in tax base $\mathbf{y}$. It in particular rules out kinks, thereby bunching.[6] Parts i) and ii) of Assumption 1 ensures

[5] Hendren (2017) assumes instead that government's income $\mathscr{B}$ is differentiable in x, which enables some "jumping responses" for zero measure of taxpayers. Golosov et al. (2014) assumes that for each type $\mathbf{w}$, $\mathbf{Y}(\mathbf{w})$ is Lipschitz continuous in x.

[6] In practice, most of real world tax codes are made of different piecewise linear tax schedules with kinks between two consecutive brackets (A noticeable exception being the personal income tax schedule in Germany). In theory, these kinks should induce bunching or gap in the corresponding income distribution. In reality, bunching at convex

that the implicit function theorem can be applied to first-order conditions (3) to ensure that each local maximum of $\mathbf{y} \mapsto \mathscr{U}\left(\sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y}), \mathbf{y}; \mathbf{w}\right)$ is differentiable in type $\mathbf{w}$ and in the magnitude x of a tax perturbation. However, if this mapping admits different global maximum among which a tax payer is indifferent, a small tax reform may trigger a jump of taxpayers with type very close to $\mathbf{w}$ from a bundle close to one of this maximum to a bundle close to another global maximum. Such jumping response prevents $\mathbf{w} \mapsto \mathbf{Y}(\mathbf{w})$ from being differentiable in the magnitude of the tax perturbation and in types. Part *iii*) of Assumption 1 is precisely intended to prevent this kind of "jumping" behavior.

Because the indifference set are convex (See Appendix A), Assumption 1 is automatically satisfied when the tax schedule is linear, or when the tax schedule is weakly convex. It is also satisfied when the tax schedule is not “too” concave, so that $\mathbf{y} \mapsto \sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y})$ is less convex than the indifference set with which it has a tangency point in the $(\mathbf{y}, c)$-space (so that Part *ii)* of Assumption 1 is satisfied) and that this indifference set lies strictly above $\mathbf{y} \mapsto \sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y})$ for all other $\mathbf{y}$ (so that Part *iii)* of Assumption 1 is satisfied). In the same spirit than the first-order mechanism design approach of Mirrlees (1971, 1976), we presume the optimal tax schedule verifies Assumption 1 and derive optimality conditions under this presumption. This presumption has then to be checked ex-post.

We can then define behavioral responses. A tax perturbation affects the first-order conditions (3) through changes in marginal net of tax rates $1 - \mathcal{T}_{y_j}$, which trigger compensated responses, and through the change in tax liability, which triggers income response. Appendix B shows that the behavioral response of the i^{th} income to a tax perturbation is thus given by:

$$\left.\frac{\partial Y_i(\mathbf{w})}{\partial x}\right|_{x=0} = \underbrace{-\frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), 0)}{\partial x}\right|_{x=0}}_{\text{Income responses}} \underbrace{-\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), 0)}{\partial x}\right|_{x=0}}_{\text{Compensated responses}} \tag{12}$$

In particular we get the Slutsky equation relating *uncompensated* to *compensated* response of the i^{th} income to the j^{th} net-of-marginal tax rate:

$$\frac{\partial Y_i{}^u(\mathbf{w})}{\partial \tau_j} = \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} + Y_j(\mathbf{w}) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \tag{13}$$

The tax liability response of a given type of individual can be decomposed into *mechanical* effects, absent any change in tax base, and *behavioral* effects induced by the responses described in (12):

$$\left.\frac{\mathrm{d}\tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\mathrm{d}x}\right|_{x=0} = \underbrace{\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0}}_{\text{Mechanical effects}} + \underbrace{\sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \left.\frac{\partial Y_i(\mathbf{w})}{\partial x}\right|_{x=0}}_{\text{Behavioral effects}} \tag{14}$$

kink points are less frequent than theoretically expected (see however Saez (2010) which can be viewed as a counter-example) and gaps in the income distributions at concave kink points have not been documented empirically, to the best of my knowledge. One interpretation is very low behavioral elasticities. Another explanation that is much more plausible to me is that taxpayers do not optimize with respect to actual tax schedules, but with respect to smooth approximation of actual tax schedules (for instance $\mathbf{y} \mapsto \int \mathcal{T}(\mathbf{y} + \mathbf{u}) \, \mathrm{d}\Psi(\mathbf{u})$ where $\mathbf{u}$ is an n-dimensional random shock on incomes with joint CDF Ψ) which do verify part *i*) of Assumption 1.

Combining the latter Equation with (12), the effect of a tax perturbation on government's revenue (4) is given by:

$$\begin{aligned}\left.\frac{\partial \mathscr{B}}{\partial x}\right|_{x=0} &= \int_{\mathbf{w}\in W}\left\{\left[1-\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\,\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right]\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),0)}{\partial x}\right.\\ &\left.-\sum_{1\le i,j\le n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\,\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\,\left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),0)}{\partial x}\right|_{x=0}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}\end{aligned} \tag{15}$$

Let $\lambda > 0$ denote the shadow cost of public fonds and let:

$$g(\mathbf{w}) \stackrel{\text{def}}{\equiv} \frac{\Phi_U(U(\mathbf{w});\mathbf{w})\,\mathscr{U}_c\left(\sum_{i=1}^{n} Y_i(\mathbf{w})-\mathcal{T}\left(\mathbf{Y}(\mathbf{w})\right),\mathbf{Y}(\mathbf{w});\mathbf{w}\right)}{\lambda} \tag{16}$$

denote the social marginal weight of consumption expressed in monetary term (hereafter the *social weight*) that the government assigns to taxpayer of type $\mathbf{w}$. The government values $g(\mathbf{w})$ Euros the welfare increase of taxpayers of type $\mathbf{w}$ induced by a transfer of one Euro.[7] Applying the envelope theorem to (2), the effect in monetary terms of a tax perturbation on the government's social objective is given by (see Appendix B):

$$\frac{1}{\lambda}\left.\frac{\partial \mathscr{O}}{\partial x}\right|_{x=0} = -\int_{\mathbf{w}\in W} g(\mathbf{w})\,\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),0)}{\partial x}\right|_{x=0} f(\mathbf{w})\mathrm{d}\mathbf{w} \tag{17}$$

The shadow cost of public fonds λ corresponds to the effects on the social objective of a lump-sum transfer to every taxpayers. Combining (15) and (17) for the lump-sum perturbation (11a), the shadow cost of public funds is pined down by:

$$0 = \int_{\mathbf{w}\in W}\left[1-g(\mathbf{w})-\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\,\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w})\mathrm{d}\mathbf{w} \tag{18}$$

A tax perturbation is generically not budget-balanced, unless $\left.\frac{\partial \mathscr{B}}{\partial x}\right|_{x=0} = 0$. Therefore, one wants to evaluate not only the effects of a tax perturbation on welfare, but the effects of a tax perturbation which is combined with the lump-sum rebate of the budget surplus, as the latter perturbation is budget-balanced. The next Lemma, which is proved in Appendix C shows the welfare effect of such combination of tax perturbation is of the same sign as the effect of the initial budget-unbalanced tax perturbation on the government's Lagrangian $\mathscr{L} \stackrel{\text{def}}{\equiv} \mathscr{B} + \mathscr{O}/\lambda$, provided the shadow cost of public funds verifies (18). Combining (15) and (17), this effect is given by:

$$\begin{aligned}\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} &= \int_{\mathbf{w}\in W}\left\{\left[1-g(\mathbf{w})-\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\,\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right]\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),0)}{\partial x}\right|_{x=0}\right.\\ &\left.-\sum_{1\le i,j\le n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\,\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\,\left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),0)}{\partial x}\right|_{x=0}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}\end{aligned} \tag{19}$$

[7] See Saez and Stantcheva (2016) for non-welfarist microfoundations of these welfare weights.

Lemma 1. *If the shadow cost of public funds verifies* (18), *and if* $\left.\frac{\partial \mathcal{L}}{\partial x}\right|_{x=0} >$ *(resp* $<$*)* 0, *then reforming the tax schedule to* $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y}, x)$ *with a small positive* x *(resp. a small negative* x*) and rebating the budget surplus in a lump-sum way is a budget-balanced reform that is socially desirable.*

Lemma 1 provides a condition on behavioral elasticities, type distribution and welfare weights for the a given tax perturbation to be social desirable.

We make the following assumption on preferences:

Assumption 2. *For each bundle* $(c, \mathbf{y})$, *the mapping* $\mathbf{w} \mapsto \left(\mathcal{S}^1(c, \mathbf{y}; \mathbf{w}), ..., \mathcal{S}^n(c, \mathbf{y}; \mathbf{w})\right)$ *is invertible*

This assumption on preferences extends the usual single-crossing condition to the multidimensional context. It is for instance verified when preferences are additively separable of the form:

$$\mathscr{U}(c, \mathbf{y}; \mathbf{w}) = u(c) - \sum_{i=1}^{n} v^i(y_i, w_i) \qquad \text{with :} \qquad u', v^i_{y_i}, v^i_{y_i y_i} > 0 > v^i_{w_i}, v^i_{y_i w_i}$$

Assumption 2 implies that the mapping $\mathbf{y} \mapsto \mathbf{Y}(\mathbf{w})$ is *globally* invertible.[8] We thus get the following relation between the skill density and the tax base density:

$$h(\mathbf{Y}(\mathbf{w})) = \frac{f(\mathbf{w})}{\left|\det \left[\frac{\partial Y_i}{\partial w_j}\right]_{i,j}\right|} \tag{20}$$

III A case where the Optimal Income Tax is Comprehensive

In this section, we exhibit a situation where the optimal allocation can be decentralized by a comprehensive income tax schedule. The Following Proposition is proved in Appendix D.

Proposition 1. *If preferences are weakly separable, i.e. the utility function* $\mathscr{U}$ *takes the form* $\mathscr{U}(c, \mathbf{y}; \mathbf{w}) = \mathcal{U}(c, \mathcal{V}(\mathbf{y}); \mathbf{w})$ *where* $\mathcal{U}_c, \mathcal{U}_{w_i} > 0 > \mathcal{U}_V$, $\mathcal{V}(\cdot)$ *is twice continuously differentiable, increasing in each argument and convex, then the optimal tax is comprehensive.*

The intuition for this result is in the spirit of the theorem of Atkinson and Stiglitz (1976) and of its proof by Laroque (2005) and Gauthier and Laroque (2009). Because of weakly separable preferences, whatever their type, individuals choose how to split their efforts in getting the different tax base to minimize the same aggregation $\mathcal{V}(\cdot)$ of incomes, while the government is only interested in the resources to be shared, i.e. on the sum of all incomes earned by each individual. In particular, the marginal rate of substitution between two different tax bases does not depend on type as it verifies:

$$\frac{\mathscr{U}_{y_i}(c, \mathbf{y}; \mathbf{w})}{\mathscr{U}_{y_j}(c, \mathbf{y}; \mathbf{w})} = \frac{\mathcal{V}_{y_i}(\mathbf{y})}{\mathcal{V}_{y_j}(\mathbf{y})}$$

[8] Assume by contradiction the existence of two types $\mathbf{w}, \mathbf{w}'$ such that that $\mathbf{Y}(\mathbf{w}) = \mathbf{Y}(\mathbf{w}') = \mathbf{y}$. We thus get $C(\mathbf{w}) = C(\mathbf{w}') = \sum_{k=1}^{n} Y_k(\mathbf{w}) - \mathcal{T}(\mathbf{Y}(\mathbf{w})) = c$. According to the first-order conditions (3), we get:

$$(1 - \mathcal{T}_{y_1}(\mathbf{y}), ..., 1 - \mathcal{T}_{y_n}(\mathbf{y})) = \left(\mathcal{S}^1(c, \mathbf{y}; \mathbf{w}), ..., \mathcal{S}^n(c, \mathbf{y}; \mathbf{w})\right) = \left(\mathcal{S}^1(c, \mathbf{y}; \mathbf{w}'), ..., \mathcal{S}^n(c, \mathbf{y}; \mathbf{w}')\right)$$

Assumption 2 therefore implies that $\mathbf{w} = \mathbf{w}'$, which ends the proof that $\mathbf{y} \mapsto \mathbf{Y}(\mathbf{w})$ is *globally* invertible.

Therefore, the government does not need to distort the relative supply of each tax base. A comprehensive tax schedule is therefore optimal.

It is however worth noting that weakly separable preferences does not verify the single crossing assumption 2. When the tax schedule is comprehensive, the program of individuals of type $\mathbf{w}$ can be decomposed into two consecutive stages:

$$\max_{v} \quad \mathcal{U}\left(v - \mathscr{T}(v), \min_{\mathbf{y} \text{ s.t: } \sum_{i=1}^{n} y_i = v} \quad \mathcal{V}(\mathbf{y}); \mathbf{w}\right)$$

Therefore, people earning the same taxable income $v = \sum_{i=1}^{n} y_i$ make the same choice $(y_1, ..., y_n)$. Hence each taxpayers receiving the same amount of the i^{th} tax base also receive the same j^{th} income, a prediction that is clearly counter-factual.

The case of weakly separable preference should thus only be understood as an example illustrating when a comprehensive tax schedule is desirable. Conversely, we guess that when the marginal rate of substitution across different tax base vary with types, as it assumed by Assumption 2, the optimal tax schedule is no longer comprehensive.

IV A case where the Optimal Income tax is Separate

We now consider a different specialization where the optimal tax schedule is separate. The following Proposition is proved in Appendix E.

Proposition 2. *When i) the type space is one-dimensional $W = [\underline{w}, \overline{w}] \subset \mathbb{R}$, ii) along the optimal allocation, each income admits a positive derivative with respect to type and iii) preferences are quasilinear and additively separable of the form:*

$$\mathscr{U}(c, \mathbf{y}; w) = c - \sum_{i=1}^{n} v^i(y_i; w) \qquad \text{with} \qquad v^i_{y_i}(y_i; w), v^i_{y_i, y_i}(y_i; w) > 0 > v^i_{y_i, y_i}(y_i; w)$$

the optimal tax schedule is separate.

Intuitively, when the unobserved heterogeneity is one-dimensional and the different kind of incomes are increasing in type, redistribution is a single dimension problems from high types agents, earning high levels of all incomes, to low types agents earning low levels of all incomes. The government is therefore interested in achieving the same redistributive goal by shifting the burden of redistribution on the least responsive tax base. Under quasilinear preference and additive separable preference, the government can simply achieve this objective of shifting distortions on the least responsive tax base by a separate income tax because with such preference, the choice of each income depends on the tax schedule only through its own marginal tax rate. This because under such preferences, there is neither income effects nor cross base substitution effects.

Again, the assumption of Proposition 2 are very specific. In general, income effects and cross base substitution effects can not be empirically ruled out. Moreover, the one dimensional assumption induces that realized tax bases describes a one-dimensional manifold, which is is clearly counter-factual. So, tis configuration should more be understood as a theoretical curiosity to help understanding when the separate income tax is desirable instead of a relevant policy recommendation.

V Infinitesimal reforms of mixed tax schedules

In this section, we consider general preferences but we restrict the tax schedule to be of the mixed for given in Equation 8. We first characterize the response of taxable income y_0. We show in Appendix F that the responses of the personal income tax base to a lump sum perturbation is given by:

$$\frac{\partial Y_0(\mathbf{w})}{\partial \rho} = \sum_{k=1}^{n} a_k \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \tag{21}$$

while the response of the personal income tax base to a compensated tax change in the j^{th} marginal tax rate is given by:

$$\frac{\partial Y_0(\mathbf{w})}{\partial \tau_j} = \sum_{k=1}^{n} a_k \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \tag{22}$$

and the response of the personal income tax base to an uncompensated tax change in the j^{th} marginal tax rate is given by:

$$\frac{\partial {Y_0}^u(\mathbf{w})}{\partial \tau_j} = \sum_{k=1}^{n} a_k \frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_j} \tag{23}$$

We now consider the effects of a tax perturbation $\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x\ R_i(y_i)$ specific to the i^{th} income, where $R_i(\cdot)$ is the direction of the tax reform. Such a perturbation modifies tax liability by:

$$\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0} = -R_i(Y_i(\mathbf{w}))$$

It modifies the marginal tax rate on the i^{th} marginal tax rate by:

$$\left.\frac{\partial \tilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0} = -R_i'(Y_i(\mathbf{w}))$$

And it does not affect the other marginal tax rate. Using Equation (19), the effect of such perturbation on the Lagrangian is (see Appendix F.1):

$$\begin{aligned} \left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} = & \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_i(Y_i(\mathbf{w})) \right. \\ & + \left. \left[\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\ \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right] R_i'(Y_i(\mathbf{w})) \right\} f(\mathbf{w}) d\mathbf{w} \end{aligned} \tag{24}$$

Equation (24) summarizes the first-order effect of a perturbation of the taxation of the i^{th} income on the government's Lagrangian. For individuals of type $\mathbf{w}$ such reforms induce a change $-R_i(Y_i(\mathbf{w}))$ in tax liability and a change $R_i'(Y_i(\mathbf{w}))$ in the i^{th} net of marginal tax rate. The change in tax liability induces a mechanical effect on tax revenue and on the government's objective, the latter being weighted by the social welfare weight $g(\mathbf{w})$. Hence the mechanical effect is equal to $-(1-g(\mathbf{w}))R_i(Y_i(\mathbf{w}))$ times the density of taxpayers of type $\mathbf{w}$. The change in tax liability also induces income responses $\frac{\partial Y_k}{\partial \rho} R_i(Y_i(\mathbf{w}))$ for all incomes $k \in \{0, ..., n\}$, which trigger a change in tax revenue equal to $\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k}{\partial \rho} R_i(Y_i(\mathbf{w}))$ times the density. Finally, the change $R_i'(Y_i(\mathbf{w}))$ in the i^{th} net of marginal tax rate triggers compensated responses $\frac{\partial Y_k}{\partial \tau_i} R_i'(Y_i(\mathbf{w}))$ for all incomes $k \in \{0, ..., n\}$, which induce a change in tax revenue equal to $\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k}{\partial \tau_i} R_i'(Y_i(\mathbf{w}))$ times the density. Aggregating these effects for all types leads to (24). Importantly not only compensated and income responses of the i^{th} income are taken into account but also "cross base" responses $\frac{\partial Y_k(\mathbf{w})}{\partial \rho}$ and $\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}$ for $k \neq i$, unless the other incomes are not taxed at the margin.

Given the other tax schedules, the tax schedule specific to the i^{th} income is optimal if such income specific tax perturbation triggers no first-order effect on the Lagrangian, whatever the direction $R_i(\cdot)$ of the tax perturbation). Let:

$$\varepsilon_i(y_i) \stackrel{\text{def}}{\equiv} \frac{1 - T'(y_i)}{y_i} \overline{\frac{\partial Y_i}{\partial \tau_i}}\Big|_{Y_i(\mathbf{w})=y_i} \tag{25}$$

denote the average compensated elasticity of the i^{th} income with respect to its own net of marginal tax rate among tax payer earning i^{th} income equal to y_i, where for any variable $X(\mathbf{w})$ and any subset $\Omega \subset W$, $\overline{X(\mathbf{w})}|_{\mathbf{w}\in\Omega}$ stands for the mean of $X(\mathbf{w})$ among skill levels $\mathbf{w}$ for which $\mathbf{w} \in \Omega$. This leads to the following optimal tax formula for the tax schedule specific to the i^{th} income (see Appendix F.1):

$$\begin{aligned} &\frac{T_i'(y_i)}{1 - T_i'(y_i)} \varepsilon_i(y_i)\, y_i h(y_i) + \sum_{0 \le k \le n, k \neq i} \overline{T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}}\Big|_{Y_i(\mathbf{w})=y_i} h_i(y_i) \\ &= \int_{z=y_i}^{\infty} \left\{ 1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=z} - \sum_{k=0}^{n} \overline{T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}}\Big|_{Y_i(\mathbf{w})=z} \right\} h_i(z) \mathrm{d}z \end{aligned} \tag{26}$$

Equation (26) extends to the multidimensional case the optimal ABC tax formula of Diamond (1998) and Saez (2001) for the case with a single income. As Saez (2001), Equation (26) relates optimal marginal tax to empirically estimable sufficient statistics which are behavioral responses, income density and welfare weights. There are however two important differences. First, as the underlying heterogeneity is multidimensional, the sufficient statistics have to be averaged across all types earnings the same level of the i^{th} income.[9] Note that this averaging procedure being along the i^{th} income, it differs from the averaging procedure required for the optimal j^{th} optimal marginal tax rate. Second, and most importantly, not only the compensated elasticity $\varepsilon_i(y_i)$ of the i^{th} income with respect to its own marginal tax rate shows up in the left-hand side of (26). Also the compensated responses of the other bases $\frac{\partial Y_k}{\partial \tau_i}$ to a compensated change in the i^{th} net of marginal tax rate show up.

To understand why, consider, in the spirit of Saez (2001) a reform of the tax schedule specific to the i^{th} income that consists in a small change denoted $\Delta\tau_i$ of the marginal tax rate for

[9]Saez (2001) conjectured his optimal tax formula can be extended to the case with multidimensional unobserved heterogeneity. This have been formally proved only recently (Hendren, 2017, Jacquet and Lehmann, 2017).

taxpayers whose i^{th} income lies in the small interval $[y_i - \delta_{y_i}, y_i]$. This reform triggers a change in tax liability equal to $\Delta\rho = \Delta\tau_i \delta_{y_i}$ for all taxpayers with an i^{th} income above y_i, which induce mechanical and income responses effect equal to the the right-hand side of (26) times $\Delta\rho$. Moreover, for taxpayer earning an i^{th} income between $[y_i - \delta_{y_i}, y_i]$, the tax reform induces compensated response equal to $\frac{\partial Y_k}{\partial \tau_i}$ for all incomes $Y_k(\mathbf{w})$ with $k \in \{0, ..., n\}$, and not only for the i^{th} income. The response of k^{th} income induces a change in the of the k^{th} tax liability equal to $-T_{y_k}(Y_k(\mathbf{w})) \ \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \Delta\tau_i$. As the mass of such taxpayers is $h(y_i)\delta_{y_i}$, summing these effects for all taxpayers with an i^{th} income in $[y_i - \delta_{y_i}, y_i]$ and taking into account the definition of ε_i leads to left-hand side of (26) times $-\Delta\rho = -\Delta\tau_i\delta_{y_i}$. At the optimum, all these first-order effects should compensate each others, which leads to (26). This reasoning in the spirit of Saez (2001) clarifies that not only the compensated elasticity ε_i of the i^{th} income to the change in the i^{th} marginal net tax rate matters, but also the "cross-base" responses to the other tax base $\overline{T'_k(Y_k(\mathbf{w}))\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i})}\Big|_{Y_i(\mathbf{w})=y_i}$ for all $k \neq i$.

We now investigate the effects of reforms of the personal income tax schedule. For this purpose, we consider tax perturbations of the form $\tilde{\mathcal{T}}(\mathbf{y}, x) = \mathcal{T}(\mathbf{y}) - x \ R_0(\sum_{k=1}^{n} y_k)$, where $R_0(\cdot)$ is the direction of the tax reform. Such a perturbation modifies tax liability by:

$$\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0} = -R_0(Y_0(\mathbf{w}))$$

It modifies the marginal tax rate on the j^{th} marginal tax rate by:

$$\left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0} = -a_j R'_j(Y_0(\mathbf{w}))$$

According to (10), the marginal tax rate on the j^{th} income depends not only on the marginal tax rate of its specific tax schedule $T'_j(\cdot)$ but also on the marginal tax rate of the personal income tax schedule discounted by the indexed parameter a_j. Therefore, as shown in Appendix F.2, a compensated reform of the personal income tax schedule generate responses equal to the sum of the j^{th} index parameter a_j times the compensated elasticity of the i^{th} income to a change in the j^{th} net of marginal tax rate.

$$\forall i \in \{0, ..., n\} \qquad \frac{\partial Y_i}{\partial \tau_0} = \sum_{j=1}^{n} a_j \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \tag{27}$$

Combining (9) and (27), the compensated elasticity of taxable income is:[10]

$$\varepsilon_0(y_0) = \frac{1 - T'(y_0)}{y_0} \sum_{1 \le i,j \le n} a_i \ a_j \left.\overline{\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}}\right|_{Y_0(\mathbf{w})=y_0} \tag{28}$$

Given these definitions of the effect of a personal income tax perturbation in the direction $R_0(\cdot)$ are given by Equation (24) with $i = 0$. Consequently, for given income specific tax schedules, the optimal personal income tax schedule verifies (26) with $i = 0$. We thus get:

[10] As the matrix $\left[\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j}$ of compensated responses is positive definite, the compensated elasticity of taxable income is positive unless $a_1 = ... = a_n = 0$.

Proposition 3. *For all $i \in \{0,...,n\}$:*

i) a tax perturbation specific to the i^{th} income affects the government's Lagrangian by (24).

ii) Given the other tax schedules, the optimal tax schedule specific to the i^{th} income is provided by (26).

We can now describe the effects of change in the tax base parameter a_i by considering the effect of the tax perturbation

$$\tilde{\mathcal{T}}(\mathbf{y},x) = T_0\left(\sum_{k=1}^{n} a_k\, y_k - x\, y_i\right) + \sum_{k=1}^{n} T_k(y_k) \tag{29}$$

As shown in Appendix F.3, this perturbation induce three effects for taxpayers of type $\mathbf{w}$.

- First, the tax perturbation decreases tax liability by $Y_i(\mathbf{w})\ T_0'(Y_0(\mathbf{w}))$, which induces a mechanical effect equal to:

$$(g(\mathbf{w})-1)\ Y_i(\mathbf{w})\ T_0'(Y_0(\mathbf{w}))$$

- Second, according to (10) a decrease in a_i reduces the impact of the personal income tax schedule on the marginal tax rate of the i^{th} income. By this effect, the tax perturbation decreases the i^{th} marginal tax by $T_0'(Y_0(\mathbf{w}))$. Combined with the decrease in tax liability by $T_0'(Y_0(\mathbf{w}))\ Y_i(\mathbf{w})$, this mimics an *uncompensated* change in the i^{th} income which triggers a response of the k^{th} income equal to $\frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\ T_0'(Y_0(\mathbf{w}))$ for all $k \in \{0,...,n\}$. These behavioral responses in turn modifies tax revenue by:

$$\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\ T_0'(Y_0(\mathbf{w}))$$

- Finally, according to (9), the perturbation decreases taxable income by $Y_i(\mathbf{w})$. Because of the nonlinearity of the personal income tax schedule, the marginal tax rates on the j^{th} income decreases by $a_j\ Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))$ for all $j \in \{1,...n\}$. These changes in marginal tax rates in turn induce compensated responses $a_j \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\ Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))$ for all incomes $Y_k(\mathbf{w})$ with $k \in \{0,...,n\}$ and for all $j \in \{1,...n\}$, so that the effects on tax revenue is:

$$\left(\sum_{j=1}^{n}\sum_{k=0}^{n} a_j\ T_k'(Y_k(\mathbf{w}))\ \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))$$

Adding these effects, weighting them by the density of taxpayers of type $\mathbf{w}$ and aggregating for all types, the effect on the Lagrangian is:

$$\begin{aligned}\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} = & \int_{\mathbf{w}\in W}\left\{\left[(g(\mathbf{w})-1)\ Y_i(\mathbf{w}) + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\ \frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\right] T_0'(Y_0(\mathbf{w}))\right. \\ & + \left.\left(\sum_{j=1}^{n}\sum_{k=0}^{n} a_j\ T_k'(Y_k(\mathbf{w}))\ \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))\right\} f(\mathbf{w})\,\mathrm{d}\mathbf{w}\end{aligned} \tag{30}$$

According to Lemma 1, Equation (30) evaluates the social desirability of a reform that consists in modifying the tax index parameter and to rebate in a lump sum way the government's net surplus. An alternative way to study this reform is to compute the lump rebate $\ell'(0)$ required to balanced the budget after the tax perturbation defined in (29). Appendix F.3 shows that the compensated lump-sum transfer is given by

$$\ell'(0) = \tag{31}$$
$$\frac{\int\limits_{\mathbf{w}\in W} \left\{ \left[Y_i(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k{}^u(\mathbf{w})}{\partial \tau_i} \right] T_0'(Y_0(\mathbf{w})) - \left(\sum_{j=1}^{n} \sum_{k=0}^{n} a_j T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}}{\int_{\mathbf{w}\in W} \left[1 - \sum_{i=0}^{n} T_i'(Y_i(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] f(\mathbf{w}) \mathrm{d}\mathbf{w}}$$

According to the envelope theorem, taxpayers for which $Y_i(\mathbf{w})T_0'(Y_0(\mathbf{w})) > \ell'(0)$ wins from the reform defined in (29) with a positive x, while the other are the looser of the reform.

Proposition 4. *Reducing a_i and compensating the revenue losses by lump-sum tax $\ell(x)$ requires a lump-sum tax with $\ell'(0)$ given by (31). Taxpayers for which $Y_i(\mathbf{w})T_0'(Y_0(\mathbf{w}))>$ (resp) $\ell'(0)$ wins from this reduction in a_i.*

However, policymakers typically do not consider this way of financing a reform of the personal income tax base (of parameter a_i). The typical way is combining a change in the tax base parameter a_i with a proportional change in tax liability on the i^{th} income, that is with an *uncompensated* change in the in the i^{th} marginal tax rate. If the personal income schedule was linear, i.e. if $T_0(y_0) = t_0 \sum_{k=1}^{n} a_k y_k$, changing the tax parameter a_i would be equivalent to an uncompensated change in the i^{th} marginal tax rate. To see this more clearly, one can compare the effects of perturbing a_i, which is provided by (30) to the effects of a linear tax perturbation specific to the i^{th} income. Plugging $R_i(y_i) = r_i\, y_i$ in (24) and using (13) leads to:

$$\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} = \int_{\mathbf{w}\in W} \left\{ \left[(1 - g(\mathbf{w}))\ Y_i(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\ \frac{\partial Y_k{}^u(\mathbf{w})}{\partial \tau_i} \right] r_i \right\} f(\mathbf{w})\ \mathrm{d}\mathbf{w}$$

The two reforms would therefore be equivalent if $t_0 = r_i$. Otherwise, the change in a_i is not equivalent to an uncompensated change in the tax rate on the i^{th} income for (at least) two reasons. First, the uncompensated change in the tax rate specific to the i^{th} income affects the tax liability of taxpayers proportionally to tax payer's i^{th} income $Y_i(\mathbf{w})$, while the change in a_i affects taxpayers' liability by the personal marginal tax rate $T_0'(Y_0(\mathbf{w}))$ times their i^{th} income $Y_i(\mathbf{w})$. Therefore, whenever the personal income tax schedule is progressive thereby exhibiting increasing marginal tax rates, among taxpayer earning the same i^{th} income $Y_i(\mathbf{w})$, those earnings relatively few other incomes are less affected by the change in the tax base parameter a_i than those earnings relatively more other incomes, because the latter face higher personal income marginal tax rate $T_0'(Y_0(\mathbf{w}))$. For example, exiting capital income from the personal income tax base (reducing a_2) and taxing capital income according to a specific schedule is, among individuals earnings the same capital income, relatively more beneficial to taxpayers whose other incomes also higher, because they face a higher marginal tax rate on their personal income. Second, a reduction in a_i by reducing taxable income decrease the marginal tax rate on taxable income by $Y_i(\mathbf{w})\ T''(Y_0(\mathbf{w}))$. This reduction triggers compensated responses, an effect that does not take place with an uncompensated change in the tax rate specific to the i^{th} income. For example, when exiting capital income from the personal income tax base, marginal tax rate will decrease which typically increases taxable income but also affects capital income in an ambiguous way depending on the cross base effects.

We now evaluate a balanced-budget reform that consists in decreasing the index parameter a_i of the i^{th} income, and in compensating the revenue losses by a proportional increase in the tax rate specific to i^{th} income. This is typically an incremental reform that moves the tax system toward a more separate and a less comprehensive tax system. We thus consider the tax perturbation $\tilde{\mathcal{T}}(\mathbf{y},x) = T_0\left(\sum_{k=1}^n a_k\, y_k - x\, y_i\right) + \sum_{k=1}^n T_k(y_k) + r_i(x)\, y_i$, where $r_i(x)$ is such that $\frac{\partial \mathscr{B}}{\partial x}\Big|_{x=0} = 0$. According to (24) and (30) tax revenues are perturbed by:[11]

$$\begin{aligned}\frac{\partial \mathscr{B}}{\partial x}\Big|_{x=0} &= r_i'(0)\int_{\mathbf{w}\in W}\left\{Y_i(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\,\frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}\\ &- \int_{\mathbf{w}\in W}\left\{\left[Y_i(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w}))\,\frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\right] T_0'(Y_0(\mathbf{w}))\right.\\ &- \left.\left(\sum_{j=1}^{n}\sum_{k=0}^{n} a_j\, T_k'(Y_k(\mathbf{w}))\,\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))\right\} f(\mathbf{w})\ \mathrm{d}\mathbf{w}\end{aligned}$$

If the proportional tax rate on the i^{th} income is on the correct side of the Laffer curve, one has $\int_{\mathbf{w}\in W}\left\{Y_i(\mathbf{w}) - \sum_{k=0}^n T_k'(Y_k(\mathbf{w}))\,\frac{\partial {Y_k}^u(\mathbf{w})}{\partial \tau_i}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w} > 0$ and we must have:

$$\begin{aligned}r_i'(0) &= \frac{\int_{\mathbf{w}\in W}\left\{Y_i(\mathbf{w}) - \sum_{k=0}^n T_k'(Y_k(\mathbf{w}))\,\frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}\right\} T_0'(Y_0(\mathbf{w}))\, f(\mathbf{w})\mathrm{d}\mathbf{w}}{\int_{\mathbf{w}\in W}\left\{Y_i(\mathbf{w}) - \sum_{k=0}^n T_k'(Y_k(\mathbf{w}))\,\frac{\partial Y_k^u \mathbf{w}}{\partial \tau_i}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}}\\ &- \frac{\int_{\mathbf{w}\in W}\left(\sum_{j=1}^{n}\sum_{k=0}^{n} a_j\, T_k'(Y_k(\mathbf{w}))\,\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w})) f(\mathbf{w})\mathrm{d}\mathbf{w}}{\int_{\mathbf{w}\in W}\left\{Y_i(\mathbf{w}) - \sum_{k=0}^n T_k'(Y_k(\mathbf{w}))\,\frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}}\end{aligned} \tag{32}$$

to get a budget-balanced tax perturbation. The effect on the welfare of taxpayers of type $\mathbf{w}$ is then given by:

$$\frac{1}{\lambda}\,\frac{\partial \Phi(U(\mathbf{w});w)}{\partial x}\Big|_{x=0} = \left(T_0'(Y_0(\mathbf{w})) - r_i'(0)\right)\ Y_i(\mathbf{w})\ g(\mathbf{w})$$

and the effect on the social objective is:

$$\frac{1}{\lambda}\,\frac{\partial \mathscr{O}}{\partial x}\Big|_{x=0} = \int_{\mathbf{w}\in W}\left(T_0'(Y_0(\mathbf{w}) - r_i'(0)\right)\ Y_i(\mathbf{w})\ g(\mathbf{w})\ f(\mathbf{w})\mathrm{d}\mathbf{w} \tag{33}$$

[11] The effects of tax revenue can be deducted from the effects on Lagrangian by setting welfare weights $g(\mathbf{w})$ in (24) and in (30).

The sufficient statistics summarizing all the efficiency arguments in favor of exiting i^{th} income from the personal income tax base and taxing in a proportional way is $r_i'(0)$. If $r_i'(0)$ is negative, the personal income tax base is so inefficient that exiting i^{th} income alone increases tax revenue. In such a case, the reform is Pareto improving. Otherwise, a unit decrease in a_i has to be compensated by an increase by $r_i'(0)$ of the proportional tax rate on the i^{th} income to keep the budget balanced, which decrease the welfare of taxpayers. Therefore the lower $r_i'(0)$, the more desirable is this switch to a more separate income tax system. According to Equation (32) this is more likely the case when the effect of a compensated change in the i^{th} marginal tax rate $Y_i(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}$ is lower when the marginal tax rate on personal income is higher, and when the stimulating effect of reducing the taxable income $\sum_{j=1}^{n} \sum_{k=0}^{n} a_j \; T_k'(Y_k(\mathbf{w})) \; \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}$ is stronger.

However, not all taxpayers may benefit from this reform. Among taxpayers earning i^{th} income $Y_i(\mathbf{w})$, those earnings relative more other income typically face a higher marginal tax rate $T_0'(Y_0(\mathbf{w}))$ on their taxable income, thereby benefit relatively more from the exiting of i^{th} income from the personal income tax base than the others. Conversely, the proportional increase in the tax rate specific to the i^{th} income affects identically the welfare of all taxpayers earnings the same i^{th} income. If the all the incomes are perfectly correlated, as it was the case in Section IV, then the effects are the same across all agents earning the same i^{th} income (because they then earn the same taxable income $Y_0(\mathbf{w})$. Conversely, if agents earning the same i^{th} income earn very different taxable income, thereby facing very different marginal tax rate on their personal income, some may win (because $T_0'(Y_0(\mathbf{w})) > r_i'(0)$) from the reform while some other may loose (because $T_0'(Y_0(\mathbf{w})) > r_i'(0)$). Therefore, whether or not the reform is socially desirable depends on the distribution of welfare weights between winners and losers.

Proposition 5. *Reducing a_i and compensating the revenue losses by a proportional tax on the i^{th} income is socially desirable if and only if the expression in (33) is positive.*

VI Numerical Illustration

This section proposes a numerical illustration of the pros and the cons of marginal reforms of capital taxation towards more separate or more comprehensive tax systems. For this purpose, we numerically implement Equations (31) and (32) on POTE, an exhaustive datasets of household tax records in France.

The French tax system is far more is more complex than (8). A first question is how to aggregate the different types of incomes, i.e. the choice of n. Realism calls for n above 100 in the French tax code. However, the larger the n, the more numerous are behavioral responses to calibrate. We hence choose a very parsimonious strategy by aggregating incomes into two categories: labor earnings denoted y_1 and capital (financial) income denoted y_2. The tax records provide taxable labor earnings $a_1 \, y_1$ and taxable capital income $a_2 \, y_2$. As this exercise is a very first illustration, we neglect all other sources of incomes especially rents and capital gains. We also neglect the various tax deductions.

Payroll tax play a very important role in France and generates revenue close to 17% of GDP, while the personal income tax generates only 3% of GDP. Moreover, payroll tax are highly nonlinear, in particular for the lowest part of the income distribution (see e.g. Kramarz and Philippon (2001), Lehmann et al. (2013)). As these non-linearities occur for earnings bellow 1.6 times the minimum wage, a level for which taxpayers declare virtually no capital income, we neglect them in this study. Assuming an employer payroll equal to 40% of posted earnings and an employee payroll tax (including CSG) of 20% and taking into account the fact that salaries benefit from a 10% deduction before being entering the personal income tax base, we take $a_1 = 0.90 \times (1 - (0.2 + 0.4)/(1 + 0.4)) \simeq 51.4\%$ and $T_1(y_1) = (0.2 + 0.4)/(1 + 0.4) \simeq 0.428\, y_1$.

Capital incomes include dividends, interests, and rents from life insurance. These three sort of incomes are taxed differently, but all three are subject to a contribution rate of 15.5% to finance social security before being included in the personal income tax base. On the top of this, dividends benefit from a 40% discount, which is supposed to compensated the 33.3% corporate income tax rate that applies on dividends and financial capital gains, but virtually not on interests. We therefore make the crude approximation of neglecting both the corporate income tax and the 40% discount on dividends. We thus calibrate a_2 and $T_2(\cdot)$ from the social security contribution rate on capital income that was equal to 15.5% in 2017.

Another difficulty is to select behavioral responses. There is huge literature estimating the elasticity of taxable income with respect to marginal net of tax rate (see. Saez et al. (2012)) with no clear consensus on the most plausible value for the long elasticity. There is a smaller literature trying to compare the responsiveness of labor income versus capital income. Extending the approach of Auten and Carroll (1999) and Gruber and Saez (2002) to estimate specifically the elasticity of one income with respect to its net of tax rate, Kleven and Schultz (2014) obtain (See their Table 7) obtain ε_1 close 0.05 and ε_2 close to 0.10. Using reforms of the german income tax system which is comprehensive, Hermle and Peichl (2018) obtain elasticities around 0.13 for wage income, 0.13-0.22 for capital income and 0.3, 0.43 for business income, confirming that capital incomes are more responsive than labor incomes. There has been important tax reforms in France taking place in 2013 which basically induce a sharp and salient increase in marginal tax rates on top earners Boissel and Matray (2019), Guillot (2019), Aghion et al. (2019), Bach et al. (2019), Lefebvre et al. (2019). These reforms leads to a sharp decrease in dividends but not in other types of incomes (Boissel and Matray, 2019, Bach et al., 2019, Lefebvre et al., 2019). Given this, we think it is more realistic to take as a baseline much larger elasticities, with a larger capital income elasticity $_varepsilon_2$ than labor income elasticity ε_1. We take $\varepsilon_1 = 0.15$ and $\varepsilon_2 = 0.6$ for our baseline and conduct sensitivity analysis. Notice that these elasticities are direct, i.e. are those would be valid if the tax schedules were linear (Jacquet and Lehmann, 2017).

Finally, we need to calibrate the cross-base response. We do so by assuming that direct cross base response are given by

$$\frac{\partial y_1^\star}{\partial \tau_2} = \frac{\partial y_2^\star}{\partial \tau_1} = \theta \sqrt{\frac{\partial y_1^\star}{\partial \tau_1} \frac{\partial y_2^\star}{\partial \tau_2}}$$

so, to ensure the matrix of direct responses is positive definite, parameter θ has to lies between -1 and 1. $\theta = 0$ leads to no-cross base response as in Saez and Stantcheva (2017). Income shifting responses are associated to negative value for θ while income responses in a model where taxpayers work in the first period and consume in the first and in the second periods are associated to positive value of θ. In our baseline calibration (see row (3) in Table 1), we assume no cross base response and conduct sensitivity analysis.

	Behavioral parameters			Lump sum rebate		Proportional tax	
	$\varepsilon_1^\star$ (1)	$\varepsilon_2^\star$ (2)	θ (3)	$\ell'(0)$ - (31) (4)	% winners (5)	$r_2'(0)$ - (32) (6)	% winners (7)
(1)	0.15	0.60	-0.50	228.8 €	5.7%	23%	39.5%
(2)	0.15	0.60	-0.25	185.4 €	6.4%	23%	39.5%
(3)	**0.15**	**0.60**	**0.00**	**142.6 €**	**7.3%**	**23%**	**39.5%**
(4)	0.15	0.60	0.25	97.7 €	8.7%	22%	39.5%
(5)	0.15	0.60	0.50	53.7 €	10.9%	19%	39.5%
(6)	0.15	0.80	0.00	106.8 €	8.3%	20%	39.5%
(7)	0.20	0.80	0.00	106.6 €	8.3%	20%	39.5%
(8)	0.20	0.80	0.25	47.6 €	11.4%	10%	39.5%

Table 1: Numerical Illustration

Table 1 present the results of this numerical exercise. Columns (4) and (5) are devoted to the case where a marginal decrease in a_2 is financed by a lump sum tax while Columns (6) and (7) are devoted to the case where the reform is financed by a proportional increase in the capital tax rate. Column (4) implements Equation (31) while Column (6) implements (32). Columns (5) and (7) provide the share of winners under both scenario.

In the case where the reform is financed by a lump sum tax, the lump sum tax amounts to 232.6 €, absent any behavioral responses. In our baseline calibration (row 3) with no cross-base response, the capital income response to the decrease in a_2 trigger *Ramsey effects* on government's revenue so the required lump sum tax is reduced by 39% to 142.6 €. However, the winners are concentrated among high capital earners, so only 7.3% of taxpayers benefit from the reform. Moving from Row (3) to Row (6) illustrate how these figures are sensitive to ε_2. The higher the capital income elasticity, the lower the compensating lump-tax and the slightly larger are the share of winners. Comparing the results in Rows (1) to (5) illustrates the role of cross base responses. The larger the cross-base parameter θ, the more taxpayers respond to the decrease in capital tax rate by a rise in labor earnings. Hence the lump-sum transfer required to balance the budget decreases from 228.8 €when $\theta = -0.5$ in row (1) - a level at which *Ramsey effects* and *income-shifting* effects offset each other so the net effects is close to the one obtained absent any behavioral responses- to 47.6 €when $\theta = 0.5$ a level at which the reform is close to be self-financing. Unsurprisingly the fractions of taxpayers that gain from the reform increases from 5.7% in Row 1 to 10.9% in Row 2. Hence, even in a the most favorable scenario, the reforms seem to benefit to happy fews at the expense of the vast majority. Changing the labor response elasticity ε_1 has virtually no effect when $\theta = 0$, which can be seen by contrasting Rows (6) and (7), while it has in presence of cross base effects, which can be seen by contrasting rows (6) and (8).

In the case where the reform is financed by a proportional tax on capital, the reform essentially consists in shifting the burden of capital taxation from capital earners facing the highest marginal personal income tax rate T_0' to those facing lower marginal personal income tax rate. As most of capital earners lie in the same top tax bracket of the personal income tax schedule, the reform has a tiny impact on tax incentives. The rise in capital tax rate (Column (6)) and the fraction of winners from the reform are quite insensitive to change in behavioral parameters. The compensating tax rate on capital $r_2'(0)$ displayed in Column (6) slightly decreases with the cross base parameter θ (see Rows (1) to (5)) and with the capital income elasticity ε_2 (compare Row (3) to Row (6). Only when θ is different from zero does labor supply elasticity have a huge effect (see Row (6) vs (8)). The fractions of winners (Column (7)) do not change.

VII Conclusion

To be written...

References

Aghion, Philippe, Vlad Ciornohuz, Maxime Gravoueille, and Stefanie Stantcheva, "Reforms and Dynamics of Income: Evidence Using New Panel Data," Working Paper 2019.

Atkinson, A. B. and J. E. Stiglitz, "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 1976, *6* (1-2), 55–75.

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

Bach, Laurent, Antoine Bozio, Brice Fabre, Arthur Guillouzouic, Claire Leroy, and Clément Malgouyres, "Évaluation d'impact de la fiscalité des dividendes," *Rapport IPP*, 2019, *25.*

Boadway, Robin, "The Dual Income Tax System - An Overview," *Ifo DICE Report*, 2004, *2* (3), 03–08.

Boissel, Charles and Adrien Matray, "Higher Dividend Taxes, No problem! Evidence from Taxing Entrepreneurs in France," 2019.

Christiansen, Vidar, "Evaluation of Public Projects under Optimal Taxation," *The Review of Economic Studies*, 1981, *48* (3), 447–457.

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

Gauthier, Stéphane and Guy Laroque, "Separability and public finance," *Journal of Public Economics*, 2009, *93* (11-12), 1168–1174.

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin, "Dynamic Tax Reforms," NBER Working Papers 20 780 2014.

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

Guillot, Malka, "Who Paid the French 75% Tax on Millionaires? Effects on Top Wage Earners and Their Employers," 2019, (3412261).

Hendren, Nathaniel, "Efficient Welfare Weights," *NBER Working Papers*, 2017, (20351). Revise and Resubmit, *Journal of Political Economy*.

Hermle, Johannes and Andreas Peichl, "Jointly Optimal Taxes for Different Types of Income," CESifo Working Paper Series 7248, CESifo Group Munich 2018.

Jacquet, Laurence M. and Etienne Lehmann, "Optimal income taxation with composition effects," CEPR Discussion Papers 12276 2017.

Kleven, Henrik Jacobsen and Esben Anton Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," *American Economic Journal: Economic Policy*, 2014, *6* (4), 271–301.

Kramarz, Francis and Thomas Philippon, "The impact of differential payroll tax subsidies on minimum wage employment," *Journal of Public Economics*, 2001, *82* (1), 115–146.

Laroque, Guy R., "Indirect taxation is superfluous under separability and taste homogeneity: a simple proof," *Economics Letters*, 2005, *87* (1), 141–144.

Lefebvre, Marie-Noëlle, Etienne Lehmann, and Michael Sicsic, "Évaluation de la mise au barème des revenus du capital : Premiers résultats," Technical Report 2019.

Lehmann, Etienne, François Marical, and Laurence Rioux, "Labor income responds differently to income-tax and payroll-tax reforms," *Journal of Public Economics*, 2013, *99* (C), 66–84.

Mirrlees, J. A., "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, 1971, *38* (2), 175–208.

Mirrlees, James, "Optimal tax theory : A synthesis," *Journal of Public Economics*, 1976, *6* (4), 327–358.

Piketty, Thomas, "La Redistribution fiscale contre le chômage," *Revue française d'économie*, 1997, *12* (1), 157–203.

Ramsey, Franck P., "A Contribution to the Theory of Taxation," *The Economic Journal*, 1927, *37* (145), 47–61.

Renes, Sander and Floris Zoutman, "As Easy as ABC? Multidimensional Screening in Public Finance," CESifo Working Paper 6 301 2017.

Saez, Emmanuel, "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 2001, *68* (1), 205–229.

_ , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 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," *Forthcoming in Journal of Public Economics*, 2017.

Saez, Emmanuel Saez, 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.

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

Zoutman, Floris T, "The Elasticity of Taxable Wealth: Evidence from the Netherlands," mimeo 2018.

A Convexity of the Indifference Set

Let $\mathscr{C}(\cdot, \mathbf{y}; \mathbf{w})$ denote the reciprocal of $\mathscr{U}(\cdot, \mathbf{y}; \mathbf{w})$. Tax payers of type $\mathbf{w}$ earning incomes $\mathbf{y}$ should get consumption $c = \mathscr{C}(u, \mathbf{y}; \mathbf{w})$ to enjoy utility $u = \mathscr{U}(c, \mathbf{y}; \mathbf{w})$. We get:

$$\mathscr{C}_u(u, \mathbf{y}; \mathbf{w}) = \frac{1}{\mathscr{U}_c(\mathscr{C}(u, \mathbf{y}; \mathbf{w}), \mathbf{y}; \mathbf{w})} \qquad \mathscr{C}_{y_i}(u, \mathbf{y}; \mathbf{w}) = \mathcal{S}^i(\mathscr{C}(u, \mathbf{y}; \mathbf{w}), \mathbf{y}; \mathbf{w}) \tag{34}$$

For each type $\mathbf{w} \in W$ and each utility level u, we assume the indifference set: $\mathbf{y} \mapsto \mathscr{C}(u, \mathbf{y}; \mathbf{w})$ to be strictly convex. The i^{th} partial derivative of $\mathbf{y} \mapsto \mathscr{C}(u, \mathbf{y}; \mathbf{w})$ is $\mathcal{S}^i(\mathscr{C}(u, \mathbf{y}; \mathbf{w}), \mathbf{y}; \mathbf{w})$, so the Hessian is the matrix $\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j\right]_{i,j} = -\frac{\mathscr{U}_{y_i y_j} + \mathcal{S}^j \mathscr{U}_{c, y_i} + \mathcal{S}^i \mathscr{U}_{c y_j} + \mathcal{S}^i \mathcal{S}^j \mathscr{U}_{cc}}{\mathscr{U}_c}$, which is symmetric. The first-order condition of (2) is given by:

$$0 = (1 - \mathcal{T}_{y_i}(\mathbf{y}))\, \mathscr{U}_c\left(\sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y}), \mathbf{y}; \mathbf{w}\right) + \mathscr{U}_{y_i}\left(\sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y}), \mathbf{y}; \mathbf{w}\right)$$

So, using (3), the matrix of the second-order condition is:

$$\left[\mathscr{U}_{y_i y_j} + \mathcal{S}^j \mathscr{U}_{c y_i} + \mathcal{S}^i \mathscr{U}_{c y_j} + \mathcal{S}^i \mathcal{S}^j \mathscr{U}_{cc} - \mathscr{U}_c \mathcal{T}_{y_i y_j}\right]_{i,j} = -\mathscr{U}_c \left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j}$$

The second-order condition holds strictly for taxpayer of type $\mathbf{w}$ if and only if matrix $\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j}$ is positive definite, that is if and only if, the indifference set $\mathbf{y} \mapsto \mathscr{C}(U(\mathbf{w}), \mathbf{y}; \mathbf{w})$ is more convex than the budget set $\mathbf{y} \mapsto \sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y})$ at $\mathbf{y} = \mathbf{Y}(\mathbf{w})$ and $c = C(\mathbf{w})$.

B Behavioral elasticities

Rewriting Program (2) and the first-order conditions (3) after a general tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x)$ leads to:

$$\max_{\mathbf{y}=(y_1,\dots,y_n)} \quad \mathscr{U}\left(\sum_{k=1}^{n} y_k - \tilde{\mathcal{T}}(\mathbf{y}, x), \mathbf{y}; \mathbf{w}\right) \tag{35}$$

and:

$$\forall i \in \{1, ..., n\}: \qquad \mathcal{S}^i\left(\sum_{k=1}^{n} y_k - \tilde{\mathcal{T}}(\mathbf{y}, x), \mathbf{Y}(\mathbf{w}); \mathbf{w}\right) = 1 - \tilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}), x)$$

Using the implicit function theorem to differentiate these first-order conditions at $y = \mathbf{Y}(\mathbf{w})$ and $x = 0$ leads to:

$$\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j} \cdot \mathrm{d}\mathbf{y}^T = \left\{ -\left(\frac{\partial \tilde{\mathcal{T}}_{y_1}}{\partial x}, ..., \frac{\partial \tilde{\mathcal{T}}_{y_n}}{\partial x}\right)^T + (\mathcal{S}^1_c, ..., \mathcal{S}^n_c)^T \frac{\partial \tilde{\mathcal{T}}}{\partial x} \right\} \mathrm{d}x \tag{36}$$

Under a compensated tax reform of the j^{th} marginal tax rate at income $\mathbf{y} = \mathbf{Y}(\mathbf{w})$ defined in (11c), one gets $\frac{\partial \tilde{\mathcal{T}}}{\partial x} = 0$ and $\frac{\partial \tilde{\mathcal{T}}_{y_k}}{\partial x} = -\mathbb{1}_{j=k}$, so the matrix of compensated responses is given by:

$$\left[\frac{\partial Y_i}{\partial \tau_j}\right]_{i,j} = \left(\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \tag{37}$$

Under the lump-sum tax reform defined in (11a), one has $\frac{\partial \tilde{\mathcal{T}}}{\partial x} = -1$ and $\frac{\partial \tilde{\mathcal{T}}_{y_k}}{\partial x} = 0$, so the vector of income responses is given by:

$$\left(\frac{\partial Y_i}{\partial \rho}\right)^T = -\left(\left[\mathcal{S}^i_{y_j} + \mathcal{S}^i_c \mathcal{S}^j + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \cdot \left(\mathcal{S}^1_c, ..., \mathcal{S}^n_c\right)^T \tag{38}$$

Turnging back to a generic tax perturbation, plugging (37) and (38) into (36) leads to (12). Finally, applying the envelope theorem to (35) leads to:

$$\left.\frac{\partial U(\mathbf{w})}{\partial x}\right|_{x=0} = -\mathscr{U}_c\left(\sum_{k=1}^n Y_k(\mathbf{w}) - \mathcal{T}\left(\mathbf{Y}(\mathbf{w})\right), \mathbf{Y}(\mathbf{w}); \mathbf{w}\right) \frac{\partial \tilde{\mathcal{T}}\left(\mathbf{Y}(\mathbf{w}), 0\right)}{\partial x} \tag{39}$$

which leads to (17). Under an uncompensated tax reform of the j^{th} marginal tax rate defined in (11b), one gets $\frac{\partial \tilde{\mathcal{T}}}{\partial x} = -Y_j(\mathbf{w})$ and $\frac{\partial \tilde{\mathcal{T}}_{y_k}}{\partial x} = -\mathbb{1}_{j=k}$. So, Equation (12) leads to the Slutsky Equation (13).

C Proof of Lemma 1

Let $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x)$ be a tax perturbation and let $\ell(x)$ be the lump-sum rebate such that the tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x) + \ell(x)$ is budget-balanced. We denote $\left.\frac{\partial X}{\partial x}\right|_{x=0}$, the partial derivative of an economic variable X along the tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x)$ while $\left.\frac{\partial X}{\partial x}\right|^\star_{x=0}$ denotes the partial derivative of X along the budget-balanced tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x) + \ell(x)$. We have from the envelope theorem and (17):

$$\begin{aligned} \frac{1}{\lambda}\left.\frac{\partial \mathscr{O}}{\partial x}\right|^\star_{x=0} &= \frac{1}{\lambda}\left.\frac{\partial \mathscr{O}}{\partial x}\right|_{x=0} - \ell'(0)\int_{\mathbf{w}\in W} g(\mathbf{w}) f(\mathbf{w}) d\mathbf{w} \\ &= \frac{1}{\lambda}\left.\frac{\partial \mathscr{O}}{\partial x}\right|_{x=0} - \ell'(0)\int_{\mathbf{w}\in W}\left[1 - \sum_{i=1}^n \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w}) d\mathbf{w} \end{aligned} \tag{40}$$

where the second equality is derived from Equation (18) determining the shadow value of public funds. From (12), we get

$$\left.\frac{\partial Y_i(\mathbf{w})}{\partial x}\right|^\star_{x=0} = \left.\frac{\partial Y_i(\mathbf{w})}{\partial x}\right|_{x=0} - \ell'(0)\frac{\partial Y_i(\mathbf{w})}{\partial \rho}$$

which implies that:

$$0 = \left.\frac{\partial \mathscr{B}}{\partial x}\right|^\star_{x=0} = \left.\frac{\partial \mathscr{B}}{\partial x}\right|_{x=0} + \ell'(0)\int_{\mathbf{w}\in W}\left[1 - \sum_{i=1}^n \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w}) d\mathbf{w} \tag{41}$$

where the first equality is due to $\ell(x)$ being adjusted so that the tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y}, x) + \ell(x)$ is budget-balanced, i.e. $0 = \left.\frac{\partial \mathscr{B}}{\partial x}\right|^\star_{x=0}$. We thus get

$$\ell'(0) = -\frac{1}{\int_{\mathbf{w}\in W}\left[1 - \sum_{i=1}^n \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w}) d\mathbf{w}} \left.\frac{\partial \mathscr{B}}{\partial x}\right|_{x=0} \tag{42}$$

Combining Equations (40) and (41) leads to:

$$\frac{1}{\lambda}\frac{\partial \mathscr{O}}{\partial x}\bigg|_{x=0}^{\star} = \frac{\partial \mathscr{B}}{\partial x}\bigg|_{x=0} + \frac{1}{\lambda}\frac{\partial \mathscr{O}}{\partial x}\bigg|_{x=0} = \frac{\partial \mathscr{L}}{\partial x}\bigg|_{x=0}$$

D Proof of Proposition 1

Following Laroque (2005), the proof consists in stating that for any tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ there exists a mapping $\mathscr{T}(\cdot)$ defined on the positive real line such that each type of individuals gets the same utility under $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ and under $\mathbf{y} \mapsto \mathscr{T}(\sum_{i=1}^{n} y_i)$, but the government's revenues are larger under $\mathbf{y} \mapsto \mathscr{T}(\sum_{i=1}^{n} y_i)$ than under $\mathbf{y} \mapsto \mathcal{T}(\cdot)$.

Let $\mathbf{Y}(\mathbf{w})$ be the solution to:

$$\max_{\mathbf{y}} \quad \mathcal{U}\left(\sum_{i=1}^{n} y_i - \mathcal{T}(\mathbf{y}), \mathcal{V}(\mathbf{y}); \mathbf{w}\right) \tag{43}$$

Let $C(\mathbf{w}) \stackrel{\text{def}}{\equiv} \sum_{i=1}^{n} Y_i(\mathbf{w}) - \mathcal{T}(\mathbf{Y}(\mathbf{w}))$, let $V(\mathbf{w}) \stackrel{\text{def}}{\equiv} \mathcal{V}(Y(\mathbf{w}))$ and let $U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \mathscr{U}(C(\mathbf{w}), \mathbf{Y}(\mathbf{w}); \mathbf{w}) = \mathcal{U}(C(\mathbf{w}), V(\mathbf{w}); \mathbf{w})$.

We first note that if there exist two types $\mathbf{w}^\star \neq \mathbf{w}'$ such that $V(\mathbf{w}^\star) = V(\mathbf{w}')$, then one need to have $C(\mathbf{w}^\star) = C(\mathbf{w}')$. If by contradiction $C(\mathbf{w}^\star) > C(\mathbf{w}')$ (the argument for $C(\mathbf{w}^\star) < C(\mathbf{w}')$ is symmetric), then type $\mathbf{w}'$ would obtain a higher utility by choosing $\mathbf{Y}(\mathbf{w}^\star)$ than $\mathbf{Y}(\mathbf{w}')$ as in such a case: $\mathscr{U}\left(C(\mathbf{w}^\star), \mathbf{Y}(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}^\star), V(\mathbf{w}^\star); \mathbf{w}'\right) > \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}'); \mathbf{w}'\right) = \mathscr{U}\left(C(\mathbf{w}'), \mathbf{Y}(\mathbf{w}'); \mathbf{w}'\right)$ which would contradict that $\mathbf{y} = \mathbf{Y}(\mathbf{w}')$ solves (43) for individuals of type $\mathbf{w}'$.

Next, we define function $\mathcal{R}(\cdot)$ such that, for each real v, either there exists $\mathbf{w}$ such that $v = V(\mathbf{w})$, in which case we define $\mathcal{R}(v) = C(\mathbf{w})$, or $\mathcal{R}(v) = -\infty$. For individuals of type $\mathbf{w}$ solving (43) amounts to solve

$$\max_{v} \quad \mathcal{U}(\mathcal{R}(v), v; \mathbf{w}) \tag{44}$$

As $\mathcal{V}(\cdot)$ is convex, the program

$$V(g) \stackrel{\text{def}}{\equiv} \min_{\mathbf{y}} \quad \mathcal{V}(\mathbf{y}) \quad s.t: \quad \sum_{i=1}^{n} y_i = g \tag{45}$$

is well defined and so is its value $V(\cdot)$. We then define $\mathscr{T}(\cdot)$ by:

$$\mathscr{T} : g \mapsto \mathscr{T}(g) \stackrel{\text{def}}{\equiv} g - \mathcal{R}\left(V(g)\right)$$

Under the tax schedule $\mathbf{y} \mapsto \mathscr{T}\left(\sum_{i=1}^{n} y_i\right)$, one has

$$\sum_{i=1}^{n} y_i - \mathscr{T}\left(\sum_{i=1}^{n} y_i\right) = \mathcal{R}\left(V\left(\sum_{i=1}^{n} y_i\right)\right)$$

Hence, under the tax schedule $\mathbf{y} \mapsto \mathscr{T}\left(\sum_{i=1}^{n} y_i\right)$, taxpayers of type $\mathbf{w}$ solve:

$$\max_{\mathbf{y}} \qquad \mathcal{U}\left(\mathcal{R}\left(V\left(\sum_{i=1}^{n} y_i\right)\right), \mathcal{V}(\mathbf{y}); \mathbf{w}\right)$$

This problem can be solved sequentially. First, solving the dual of (45)

$$\max_{\mathbf{y}} \quad \sum_{i=1}^{n} y_i \quad s.t: \quad \mathcal{V}(\mathbf{y}) = v$$

for given level of subutility v. Second, solving Program (44). The tax schedule $\mathbf{y} \mapsto \mathscr{T}\left(\sum_{i=1}^n y_i\right)$ therefore leads each type of individual to reach the same $V(\mathbf{w})$ and the same utility $U(\mathbf{w})$ than the tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$. However, tax revenues increases if, with the initial tax schedule $\mathcal{T}(\cdot)$, $\mathbf{Y}(\mathbf{w})$ is not solving

$$\max_{\mathbf{y}} \quad \sum_{i=1}^n y_i \qquad s.t.: \quad \mathcal{V}(\mathbf{y}) = V(\mathbf{w}) \qquad \Leftrightarrow \qquad \min_{\mathbf{y}} \quad \mathcal{V}(\mathbf{y}) \qquad s.t.: \quad \sum_{i=1}^n y_i = \sum_{i=1}^n Y_i(\mathbf{w})$$

E Proof of Proposition 2

We need to show that under the assumptions of Proposition 2, the optimal allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$ can be decentralized by a separate income tax. Under the assumptions of Proposition 2, for each $i \in \{1, ..., n\}$, Function $Y_i : w \mapsto Y_i(w)$ is invertible with a reciprocal denoted Y_i^{-1} and defined on $[Y_i(\underline{w}), Y_i(\overline{w})]$.

We first characterize how the separate income tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y}) = \sum_{i=1}^n T_i(y_i)$ should be to decentralize the allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$ and second verify this tax schedule actually decentralize the optimal optimal allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$.

Using the first-order condition (3) on each tax base, we can recover for each type w and each tax base $i \in \{1, ..., n\}$, the i^{th} marginal tax rate from the i^{th} marginal rate of substitution. We thus need to have:

$$T_i'(y_i) = 1 - v_{y_i}^i \left(y_i; Y_i^{-1}(y_i)\right)$$

Let $w^\star$ be a skill level and let $y_i^\star = Y_i(w^\star)$. If the allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$ can be decentralized by a separate income tax, this tax schedule has to verify:

$$\begin{aligned} \mathcal{T}(\mathbf{y}) &= \left(\sum_{i=1}^n Y_i(w^\star)\right) - C(w^\star) + \sum_{i=1}^n T_i(y_i) \\ \text{where :} \qquad T_i(y_i) &= \begin{cases} \int_{y_i^\star}^{y_i} \left[1 - v_{y_i}^i \left(t; Y_i^{-1}(t)\right)\right] \mathrm{d}t & \text{if:} \quad y_i \in [Y_i(\underline{w}), Y_i(\overline{w})] \\ +\infty & \text{if:} \quad y_i \notin [Y_i(\underline{w}), Y_i(\overline{w})] \end{cases} \end{aligned} \tag{46}$$

So up to a constant, each income specific tax schedule is uniquely defined.

We now show that the separate tax schedule (46) induces the allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$. First, as (46) is separate and preferences are additively separable, the n dimensional program (2) of individual of type w can be simplified into n one-dimensional programs:

$$\sum_{i=1}^n \left\{\max_{y_i} \quad y_i - T_i(y_i) - v^i(y_i; w)\right\}$$

From (46), marginal tax rates are given by: $1 - \mathcal{T}_{y_i}(\mathbf{y}) = v_{y_i}^i \left(y_i; Y_i^{-1}(y_i)\right)$. The i^{th} first-order condition is obviously verified when $y_i = Y_i(w)$.

Finally we have to verify that for each skill level w and each tax base, $Y_i(w)$ maximizes $y_i \mapsto y_i - T_i(y_i) - v^i(y_i; w)$. When $y_i \in [Y_i(\underline{w}), Y_i(\overline{w})]$, we get that:

$$y_i - T_i(y_i) = y_i^\star + \int_{y_i^\star}^{y_i} v_{y_i}^i \left(t; Y_i^{-1}(t)\right) \, \mathrm{d}t$$

So, we have:

$$y_i - T_i(y_i) - v^i(y_i; w) = y_i^\star - v^i(y_i^\star; w) + \int_{y_i^\star}^{y_i} \left[v_{y_i}^i \left(t; Y_i^{-1}(t)\right) - v_{y_i}^i (t; w)\right] \mathrm{d}t$$

$$\left[Y_i(w) - T_i(Y_i(w)) - v^i(Y_i(w); w)\right] - \left[y_i - T_i(y_i) - v^i(y_i; w)\right] = \int\limits_{y_i}^{Y_i(w)} \left[v^i_{y_i}\left(t; Y_i^{-1}(t)\right) - v^i_{y_i}(t; w)\right] \mathrm{d}t$$

The latter expression is positive because $v^i_{y_i,w} < 0$ and $Y_i^{-1}(\cdot)$ is increasing.

F Multidimensional case

We first rewrite Equations (14)-(19) when the tax schedule is given by (8). According to (9), we get:

$$\begin{aligned}
\left.\frac{\partial Y_0(\mathbf{w})}{\partial x}\right|_{x=0} &= \sum_{k=1}^{n} a_k \left.\frac{\partial Y_k(\mathbf{w})}{\partial x}\right|_{x=0} \\
&= -\sum_{j=1}^{n}\left(\sum_{k=1}^{n} a_k \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0} - \sum_{k=1}^{n}\left(a_k \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right) \left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}))}{\partial x}\right|_{x=0}
\end{aligned}$$

Equation (12) is thus also verified for taxable income with $i = 0$ as long as the income response and compensated responses of taxable income are respectively defined by (21) and (22).

Given the form of the tax schedule in (8), Equation (14) becomes:

$$\left.\frac{\mathrm{d}\tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\mathrm{d}x}\right|_{x=0} = \left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \left.\frac{\partial Y_k(\mathbf{w})}{\partial x}\right|_{x=0}$$

Combining the latter Equation with (12), we get:

$$\begin{aligned}
\left.\frac{\mathrm{d}\tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\mathrm{d}x}\right|_{x=0} &= \left[1 - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right] \left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} \\
&\quad - \sum_{j=1}^{n}\left[\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right] \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0}
\end{aligned}$$

Using (17), Equation (19), which provides the effect of a tax perturbation on the Lagrangian, becomes:

$$\begin{aligned}
\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} &= \int_{\mathbf{w}\in W} \left\{\left[1 - g(\mathbf{w}) - \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right] \left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), 0)}{\partial x}\right|_{x=0}\right. \\
&\quad \left. - \sum_{j=1}^{n}\left(\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), 0)}{\partial x}\right|_{x=0}\right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}
\end{aligned} \tag{47}$$

F.1 Reforms of the tax schedule specific to the i^{th} income

We consider tax perturbations of the form:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = T_0\left(\sum_{k=1}^{n} a_k\, y_k\right) + \sum_{k=1}^{n} T_k(y_k) - x\, R_i(y_i)$$

which implies:

$$\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} = -R_i(Y_i(\mathbf{w})) \qquad \text{and} \qquad \left.\frac{\partial \tilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} = -R_i'(Y_i(\mathbf{w}))$$

Equation (47) then leads to (24), thereby to part *i*) of Proposition 3. Using the law of iterated expectations to condition type $\mathbf{w}$ on $Y_i(\mathbf{w}) = y_i$ and using (25) leads to:

$$\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} = \int_{y_i \in \mathbb{R}_+} \left\{ \left[\frac{T_i'(y_i)}{1 - T_i'(y_i)}\, \varepsilon_i(y_i)\, y_i + \sum_{0 \le k \le n, k \neq i} \overline{\left. T_k'(Y_k(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right|}_{Y_i(\mathbf{w}) = y_i} \right] R'(y_i) \right.$$
$$\left. - \left[1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=y_i} - \sum_{k=0}^{n} \overline{\left. T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right|}_{Y_i(\mathbf{w})=y_i} \right] R(y_i) \right\} h(y_i)\, \mathrm{d}y_i$$

Integrating the latter equation by parts and using (18) leads to:

$$\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} = \int_{y_i \in \mathbb{R}_+} \left\{ \frac{T_i'(y_i)}{1 - T_i'(y_i)}\, \varepsilon_i(y_i)\, y_i + \sum_{0 \le k \le n, k \neq i} \overline{\left. T_k'(Y_k(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right|}_{Y_i(\mathbf{w}) = y_i} \right.$$
$$\left. - \int_{z=y_i}^{\infty} \left[1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=y_i} - \sum_{k=0}^{n} \overline{\left. T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right|}_{Y_i(\mathbf{w})=y_i} \right] h(z)\mathrm{d}z \right\} R'(y_i)\mathrm{d}y_i$$

If $T_i(\cdot)$ is optimal given the other tax schedules, any perturbation of taxation of the i^{th} income should yield no first-order effect, whatever the direction $R_i(\cdot)$, thereby, whatever $R_i'(\cdot)$. Therefore, the integrand in preceding expression should be zero for all y_i, which leads to (26), thereby to part *ii*) of Proposition 3.

F.2 Reforms of the personal income tax schedule

We consider tax perturbations of the form:

$$\tilde{\mathcal{T}}(\mathbf{y}, x) = T_0 \left(\sum_{k=1}^{n} a_k\, y_k \right) + \sum_{k=1}^{n} T_k(y_k) - x\; R_0 \left(\sum_{k=1}^{n} a_k\, y_k \right)$$

which implies:

$$\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} = -R_0(Y_0(\mathbf{w})) \qquad \text{and} \qquad \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), x)}{\partial x}\right|_{x=0} = -a_j\, R_0'(Y_0(\mathbf{w}))$$

Using (12) leads to:

$$\left.\frac{\partial Y_k(\mathbf{w})}{\partial x}\right|_{x=0} = \sum_{j=1}^{n} a_j \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) \qquad \forall k \in \{1, ..., n\}$$

which implies Equation (27) for $k \in \{1, .., n\}$. Combining the latter equation with (9), (21) and (22) leads to:

$$\begin{aligned} \left.\frac{\partial Y_0(\mathbf{w})}{\partial x}\right|_{x=0} &= \sum_{1 \le k, j \le n} a_k a_j \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \sum_{k=1}^{n} a_k \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) \\ &= \sum_{j=1}^{n} a_j \frac{\partial Y_0(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_0}{\partial \rho}\, R_0(Y_0(\mathbf{w})) \end{aligned}$$

which implies (27) also holds for $k=0$, i.e. with with taxable income. According to Equation (47), one gets:

$$\begin{aligned}
\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} &= \int_{\mathbf{w}\in W} \left\{ \left[\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \left(\sum_{j=1}^{n} a_j \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) \right] R_0'(Y_0(\mathbf{w})) \right. \\
&+ \left. \left[-1 + g(\mathbf{w}) + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_0(Y_0(\mathbf{w})) \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w} \\
&= \int_{\mathbf{w}\in W} \left\{ \left[\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} \right] R_0'(Y_0(\mathbf{w})) \right. \\
&+ \left. \left[-1 + g(\mathbf{w}) + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_0(Y_0(\mathbf{w})) \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w}
\end{aligned}$$

where the second equality uses (27) and corresponds to (24) with $i=0$. Part *i*) of Proposition 3 is therefore also valid for $i=0$, thereby Part *ii*).

F.3 Reforms of the personal income tax base compensated by a lump-sum transfer

The tax perturbation defined in (29) implies:

$$\begin{aligned}
\left.\frac{\partial \tilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),x)}{\partial x}\right|_{x=0} &= -Y_i(\mathbf{w})\; T_0'(Y_0(\mathbf{w})) \\
\left.\frac{\partial \tilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}),x)}{\partial x}\right|_{x=0} &= -T_0'(Y_0(\mathbf{w})) - a_i\; Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w})) \\
\forall j \in \{1,...,n\}, j \neq i \qquad \left.\frac{\partial \tilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),x)}{\partial x}\right|_{x=0} &= -a_j\; Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w}))
\end{aligned}$$

Using (47) leads to:

$$\begin{aligned}
\left.\frac{\partial \mathscr{L}}{\partial x}\right|_{x=0} &= \int_{\mathbf{w}\in W} \left\{ \left[\sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right] T_0'(Y_0(\mathbf{w})) \right. \\
&+ \left(\sum_{j=1}^{n} \sum_{k=0}^{n} a_j\; T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w})) \\
&+ \left. \left[g(\mathbf{w}) - 1 + \sum_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] Y_i(\mathbf{w})\; T_0'(Y_0(\mathbf{w})) \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w}
\end{aligned}$$

Using the Slutsky equation (13), the preceding equation simplifies to (30).

We now combined the tax perturbation defined in (29) with a lump sum perturbation $\ell(x)$ that the tax perturbation $\mathbf{y} \mapsto \tilde{\mathcal{T}}(\mathbf{y},x) + \ell(x)$ is balanced-budget. Taking $g(\mathbf{w}) = 0$ in Equation (30), Equation (42) leads to $\ell'(0) =$

$$\frac{\int\limits_{\mathbf{w}\in W} \left\{ \left[Y_i(\mathbf{w}) - \sum\limits_{k=0}^{n} T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k{}^u(\mathbf{w})}{\partial \tau_i} \right] T_0'(Y_0(\mathbf{w})) - \left(\sum\limits_{j=1}^{n} \sum\limits_{k=0}^{n} a_j T_k'(Y_k(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}}{\int_{\mathbf{w}\in W} \left[1 - \sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] f(\mathbf{w}) \mathrm{d}\mathbf{w}}$$

Plugging (10) and (21) into the last equation leads to: