Comprehensive or Schedular Income Taxation? A General Equilibrium Approach with Nonlinear Taxation*

Preliminary

Laurence JACQUET† Etienne LEHMANN‡

March 22, 2021

Abstract

In a general equilibrium framework, we develop a model of income taxation spanning several types of incomes with multidimensional taxpayer heterogeneity. Starting from any tax schedule, our framework allows one to decide which, of a more comprehensive or a more schedular income tax, is more welfare- *and* efficiency -improving. We express the effects of any tax reform as well as optimal tax formulas in terms of the usual sufficient statistics plus some new ones including mean cross-base responses and general equilibrium effects. These formulas are taken to French data to simulate optimal taxes with two main sources of income, labor and capital. The optimal tax system consists in U-shape marginal tax rates on the sum of both sources of income (without deduction) and a negative tax rate on capital income.

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

I Introduction

In the beginning of the twentieth century, a hot political debate shook the French Third Republic, regarding the relevance of designing a unique, global, progressive income tax - in the vein of the Prussian *Einkommensteuer*.[1] First and foremost among the proponents of this reform was French statesman Joseph Caillaux, whose first project was to replace the tax on financial assets and the property tax (e.g., the window tax) by a single tax of the aforementioned type. In 1907, in an effort to rally more supporters to his initial project, Caillaux proposed to add proportional schedular taxes for a set of specific incomes. The project was adopted in 1909 by the French Parliament, but due to lengthy debates in the French Senate, the final Law about

*We thank Marcus Berliant, Kate Cuff, Kory Kroft, Adam Lavecchia, Johannes Schmieder, Joel Slemrod, Michael Smart, Elina Tuominen and participants at various conferences and seminars for helpful comments and suggestions. The usual disclaimers apply.

†CY Cergy Paris Université, CNRS, THEMA, F-95000, Cergy, France email: laurence.jacquet@cyu.fr. Webpage https://sites.google.com/site/econlaurencejacquet/home

‡CRED (TEPP) Université Paris II Panthéon-Assas, 12 Place du Panthéon, 75 234 Paris Cedex 05, France. email: elehmann@u-paris2.fr. Webpage: http://cred.u-paris2.fr/lehmann. Etienne LEHMANN is also research fellow at CEPR, IZA and CESIfo.

[1] In the 19th century, the Prussian income tax system had already served as a model for the introduction of similar systems in other German States and, later, in 1920, for the first federal German income tax, see Ketterle (1994) and Bartels (2019).

the global tax was not adopted before July, 15th, 1914. Further delays induced mainly by the start of the First World War led to the Law about complementary schedular taxes being voted on July, 31st, 1917 only.

This hybrid of a tax system that applies the same tax schedule to the sum of all incomes– hereafter a comprehensive tax system– and of a tax system where each type of income is subject to a specific tax schedule– hereafter a schedular tax system– still characterizes most current tax systems in the world today. For instance, the United States have a mixed system with an important comprehensive part where labor and capital incomes are taxed together according to a nonlinear tax schedule, while schedular taxes are applied to e.g. dividends and capital gains and losses.[2] Another major example of a mixed tax system akin to a schedular one is the ("Nordic") dual income tax system that prevails in many European countries, with its proportional tax on capital income and its progressive tax schedule on other incomes together, capital income being totally excluded from the the latter base (Boadway, 2004).[3] As for the controversy regarding the merits and flaws of comprehensive vs schedular systems, it is still ongoing, among taxation scholars and practitioners alike (e.g., Burns and Krever (1998), Benoteau and Meslin (2017), Bastani and Waldenström (2020)). Various arguments have been put forward in favor of a more schedular or of a more comprehensive tax system, but so far they have not been systematically studied within an integrated framework. In this paper, we provide such a framework with multiple tax bases, nonlinear income tax schedules, multidimensional heterogeneity in taxpayers and general equilibrium effects.

This framework allows us to define sufficient conditions for the optimal tax system to be schedular or comprehensive. These conditions rely on strong hypotheses which point out to a mixed tax system as being the most likely candidate for an optimal tax system in real-world economies. We then determine the optimal mixed tax system – defined as a nonlinear tax on the sum of all personal incomes, with income-specific discount rates, combined with (nonlinear) income-specific taxes – in terms of sufficient statistics. We also derive closed-form formulas for the first-order effects of arbitrary local perturbations of the mixed tax system on government revenue and welfare. We obtain a sufficient condition, written in terms of sufficient statistics, to determine whether any local budget-balanced tax reform is socially desirable or not.

Despite the mathematical difficulties that arise in multidimensional screening models, we are able to solve the tax problem with a tax perturbation approach using the fact that the set of available tax instruments is a finite sum of one-dimensional tax schedules. Methodologically,

[2]See Ordower (2014) for an exposition of the embedded schedularity in the U.S. federal income tax.

[3]Denmark was the first to introduce the dual tax system in 1987. Sweden did it in 1991, Norway in 1992, Finland in 1993, Spain in 2006, Germany in 2008 and finally France, first from 2008 to 2012, and again from 2018 onwards.

on top of this first contribution, we also build up a general equilibrium model and we define new sufficient statistics.

We obtain a non-linear income tax formula that is valid for each income tax base and that extends the Diamond (1998) and (Saez, 2001) ABC formula by showing that some of the usual sufficient statistics need to be slightly modified and that new ones matter. First, due to multidimensional heterogeneity, each relevant sufficient statistic is obtained by an average across taxpayers with the same amount of a given specific income. Second, we define sufficient statistics for cross-base responses and for the impact of general equilibrium effects.

Our sufficient statistics for cross-base responses incorporate all cross-base responses whatever their micro-foundations. Our *macroeconomic price spillover statistics* capture the magnitude of general equilibrium effects on government tax revenue and social welfare. General equilibrium effects take place because any tax reform modifies the supply of production factors, which impacts their marginal productivities, hence their prices according to the aggregate demands of production factors. This in turn implies new changes in the supply of production factors, and so on. In the tax formulas, the macro price spillover statistics are simply added to the relevant marginal tax rates. The formula points out that a positive (negative) macro price spillover statistic on a certain income tends to reduce (increase) the optimal marginal tax on this income.[4]

Moreover, we characterize how a reduction of the level of a specific income in the personal income tax base (because its deduction rate increases) impacts tax revenue and welfare. From this, we can then provide a sufficient condition describing when one should move towards a slightly more (or less) comprehensive budget-balanced income tax system. We show two new channels that matter when one shifts towards a more or less comprehensive tax system. A reduction of the personal tax base (i) automatically reduces the level of tax on personal income, hence individual tax liability, and (ii) modifies the marginal tax rate on personal income, since the personal income tax schedule is nonlinear. First, the reduction in tax liability generates mechanical loss in tax revenue, mechanical gains in welfare and positive wealth responses from every income source.[5] Second, with a U-shape marginal personal income tax profile, taxpayers with relatively low levels of personal income face an increase of their marginal personal income tax rate while taxpayers with relatively high personal income face a reduction. Therefore, there is a transfer of deadweight losses from high to low personal income earners. On top of this, for

[4]These macro price spillover statistics therefore include the redistributive "trickle-down" effects (Rothschild and Scheuer, 2013).

[5]Wealth response is the usual income effect here with several sources of income. To avoid any confusion, we choose the terminology "wealth response from a source of income" rather than "income effect/response from a source of income".

each source of income taxed both via the personal income tax and via a nonlinear schedular tax, the marginal tax rate is also impacted by a change in the personal tax base. Extra behavioral responses then occur. The sum of all these effects and responses on tax revenue an welfare is ambiguous. Therefore, one needs to implement the condition and the optimal formulas on real data to conclude.

We calibrate our model on French *Enquête Revenus Fiscaux Sociaux* (ERFS) data. We consider two production factors, labor and capital and we group all source of income in two categories, labor and capital income. We use estimates of Lefebvre et al. (2019) to calibrate the elasticity of labor income to 0.1 and the elasticity of capital income to 0.65. We consider the mixed tax system where the personal income, which is nonlinearly taxed, contains labor income plus capital income, the latter being submitted to a linear deduction rate. Beside this, capital income is also linearly taxed. One of the sub-cases of this mixed tax system is a schedular tax which, more precisely in this case, is a dual tax. Another sub-case of this mixed tax schedule is a comprehensive tax, when capital and labor income are entirely and exclusively taxed according to the nonlinear personal income tax.

We find that the optimal tax system consists in entirely including capital and labor income in the personal income tax base with U-shape marginal tax rates and to provide a subsidy to capital income earners (i.e. a negative tax rate on capital income). In all our simulations, it is always socially improving to include more and more capital income in the personal income tax base. More precisely, all our simulations emphasize (i) a positive impact on tax revenue and welfare implied by mechanical effects and wealth responses due to the increase in the level of personal income tax and (ii) a negative impact from behavioral responses induced by the change in the marginal personal income tax rates. All income sources are therefore taxed together according to the usual U-shaped marginal tax rates and to limit the deadweight loss from capital income implied by the high marginal personal income tax rates, a negative tax rate on capital income is optimal. This result is robust whatever the value of the capital-labor elasticity of substitution. Including general equilibrium effects reduces the capital income subsidy. Our results remain also valid with income-shifting (which increases the subsidy towards capital earners).

We also compare the optimal tax system with the optimal dual tax (i.e. when one constraints the personal income tax base to contain only labor income) and with the optimal comprehensive tax system (i.e. when one constraints capital and labor income to be taxed exclusively through the nonlinear personal income tax schedule).

Under a dual tax, the tax rate on capital income is positive and the optimal marginal tax rates on labor income are U-shaped and as high as the marginal tax rates that prevail on the sum of labor and capital income in the optimal tax system. Intuitively, the optimal positive tax rate on capital income prevails with a dual tax, contrary to the optimal tax system, because capital income is taxed only once. With the optimal tax system, capital income is taxed twice: once via the nonlinear personal income tax and a second time through the separate capital tax. However, the latter is negative to compensate for too high marginal tax rates on personal incomes with a large proportion of capital. With the dual tax, the U-shape profile of optimal marginal tax rates (which are inversely related to the mean elasticities of income) can be as high as in the optimal tax system because this profile does not apply to more elastic capital income. This also explains why the dual tax does not need to provide a subsidy to capital income earners.

When the tax system is comprehensive, the optimal marginal tax rates on high personal income levels are up to 10 percentage points lower than in the optimal tax system and in the optimal dual tax one. Intuitively, in the comprehensive system, including all incomes in the personal income tax base increases the mean elasticities where personal income contains a large share of capital.

Our article contributes to the literature on optimal taxation, where agents differ along multiple characteristics on which taxes cannot be conditioned (as in Jacquet and Lehmann (2021)) and where these agents earn several sources of income (as in Mirrlees (1976, 1986), Kleven et al. (2007, 2009), Golosov et al. (2014), Hermle and Peichl (2018), Lehmann et al. (2020)).[6] It also builds up on a large literature on labor and capital taxation that sheds light on the fact that multidimensional heterogeneity –i.e. heterogeneity in earnings abilities and in an additional attribute– makes capital taxation useful. Among the attributes that make capital taxation optimal, the literature has emphasized returns on investments (Gahvari and Michelleto, 2016, Kristjánsson, 2016, Gerritsen et al., 2020), time preferences for consumption (Saez, 2002, Diamond and Spinnewijn, 2011, Golosov et al., 2013) and endowments (Cremer et al., 2003).[7] Intuitively, if these additional attributes correlate with individuals' earnings abilities, taxes on capital become useful as indirect means to tax people with high ability. We extend this analysis by determining the optimal tax bases as well as the optimal nonlinear income tax schedule for every tax base. The capital taxation literature has begun to connect theories of optimal capital taxation and sufficient statistics, see Saez and Stantcheva (2018). We also follow an approach in

[6] Rothschild and Scheuer (2013) and Gomes et al. (2017) study the optimal sector-specific tax schedules where agents also differ along several dimensions but they earn income from a single sector.

[7] Another argument for taxing capital is the uncertainty agents face about their labor productivity profiles over time, e.g. Albanesi and Sleet (2006), Golosov et al. (2016).

terms of sufficient statistics and extend this analysis by determining the optimal tax bases and structure of a tax system, the optimal nonlinear tax schedule for each tax base and by introducing general equilibrium effects.[8] We also contribute to the tax incidence analysis (Sachs et al. (2020) as well as Fullerton and Metcalf (2002), Kotlikoff and Summers (1987) for surveys) since we also characterize the first-order effects on government revenue and social welfare of locally reforming a given, potentially suboptimal, tax system.

II The Economy

II.1 Firms

We consider an economy with a unit-mass of taxpayers and a representative firm that produces a numeraire good using n inputs denoted $(\mathcal{X}_1, ..., \mathcal{X}_n)$. The production function is denoted by $\mathcal{F} : (\mathcal{X}_1, ..., \mathcal{X}_n) \mapsto \mathcal{F}(\mathcal{X}_1, ..., \mathcal{X}_n)$. The function $\mathcal{F}$ is increasing in its arguments, with partial derivatives denoted by $\mathcal{F}_{\mathcal{X}_i}$. Its second partial derivatives are negative, i.e. $\mathcal{F}_{\mathcal{X}_i\mathcal{X}_i} < 0$. Assuming perfect competition, the firm chooses its inputs to maximize its profit:

$$\max_{\mathcal{X}_1,...,\mathcal{X}_n} \qquad \mathcal{F}\left(\mathcal{X}_1, ..., \mathcal{X}_n\right) - \sum_{i=1}^{n} p_i \, \mathcal{X}_i$$

where $p_i \in \mathbb{R}_+$ stands for the price of the i^{th} input. From the first-order condition of this maximization, the price i is equal to the marginal productivity of the i^{th} input, that is:[9]

$$p_i = \mathcal{F}_{\mathcal{X}_i}\left(\mathcal{X}_1, ..., \mathcal{X}_n\right). \tag{1}$$

The inverse (aggregate) demand function for input i is defined thanks to Equation (1). This equation summarizes the (inputs) demand side of the economy. Prices are endogenous whenever production factors are imperfect substitutes. For instance, suppose a production function with two inputs, capital and labor which are imperfectly substitute. A tax cut in capital income will encourage effort to generate capital. Capital becomes more abundant, which reduces its marginal productivity (due to the diminishing marginal productivities of input factors) and its price (because of (1)). Whenever the second-order cross-derivative $\mathcal{F}_{\mathcal{X}_i\mathcal{X}_j}$ is positive, this will also raise the marginal productivity of labor, hence raise its price (because of (1)), .

An interesting limiting case is the one where the production function is linear, i.e.

$$\mathcal{F}\left(\mathcal{X}_1, ..., \mathcal{X}_n\right) = \sum_{i=1}^{n} \gamma_i \, \mathcal{X}_i$$

[8]General equilibrium effects have been studied in optimal nonlinear tax models in e.g. Rothschild and Scheuer (2013), Scheuer (2014), Ales and Sleet (2016), Ales et al. (2017), Scheuer and Werning (2017), Sachs et al. (2020).

[9]We do not consider the taxation of intermediate inputs, which we think would be irrelevant given our assumption of a representative firm.

(with $\gamma_i > 0$), all inputs are perfect substitutes and prices become exogenous with $p_i = \gamma_i$. In this case, without loss of generality, we can normalize actions with $\gamma_i = 1$ (hence prices are also normalized to one) so that:

$$\mathcal{F}\left(\mathcal{X}_1, ..., \mathcal{X}_n\right) = \sum_{i=1}^{n} \mathcal{X}_i. \tag{2}$$

In appendix K, we show that incorporating the taxation of production factors does not modify our results and it simply requires an adequate re-scaling of the income tax function $\mathcal{T}(\cdot)$. We therefore assume zero taxation of production factors without loss of generality.

II.2 Taxpayers

Each taxpayer is characterized by different individual characteristics summarized in their vector of type $\mathbf{w} = (w_1, ..., w_n)$. Unless otherwise specified, $n \geq 2$. 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.

Each taxpayer takes $n \geq 2$ different actions denoted by $\mathbf{x} = (x_1, ..., x_n)$. These actions are, for instance, the amount of effective units of labor, the amount of investment units in capital and the amount of business income. The generation of each action x_i, that are the n inputs of the representative firm, comes with effort costs that depend on the vector of type $\mathbf{w}$ according to the utility function $(c, \mathbf{x}; \mathbf{w}) \mapsto \mathscr{U}(c, \mathbf{x}; \mathbf{w})$, where c denotes after-tax income. The utility function is assumed twice continuously differentiable over $\mathbb{R}_+^{n+1} \times W$, in the first argument, with partial derivative denoted $\mathscr{U}_c > 0$ and decreasing in each action, $\mathscr{U}_{x_i} < 0$.

The i^{th} action x_i generates income y_i according to $y_i = p_i\, x_i$ where endogenous input prices p_i are taken as given by the taxpayers. The prices p_i are the macroeconomic return of taxpayers' i^{th} actions and the prices the firm faces for its i^{th} inputs. These prices are summarized by the vector $\mathbf{p} = (p_1, ..., p_n)$. For instance, if x_1 denotes effective labor, the price p_1 denotes the wage per unit of effective labor, then y_1 denotes labor income. If x_2 denotes savings and p_2 denotes the gross return of saving, then y_2 denotes capital income, etc.

The government taxes incomes according to the nonlinear tax schedule that (possibly) depends on the kind of income:

$$\mathcal{T} : \mathbf{y} = (y_1, ..., y_n) \mapsto \mathcal{T}(\mathbf{y}) = \mathcal{T}\left(y_1, ..., y_n\right).$$

Consumption is $c = \sum_{i=1}^{n} y_i - \mathcal{T}\left(y_1, ..., y_n\right)$. For a $\mathbf{w}$-taxpayer, we denote her marginal rate of substitution between the i^{th} action and consumption by:

$$\mathcal{S}^i(c, \mathbf{x}; \mathbf{w}) \stackrel{\text{def}}{\equiv} -\frac{\mathscr{U}_{x_i}(c, \mathbf{x}; \mathbf{w})}{\mathscr{U}_c(c, \mathbf{x}; \mathbf{w})}. \tag{3}$$

We assume that the indifference sets are convex. This implies that the matrix $\left[\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j\right]_{i,j}$ is positive definite, as shown in Appendix A.[10] A $\mathbf{w}$-taxpayer chooses her actions $\mathbf{x}$ to solve:

$$U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \max_{\mathbf{x}=(x_1,\ldots,x_n)} \qquad \mathscr{U}\left(\sum_{k=1}^{n} p_k\, x_k - \mathcal{T}\left(p_1\, x_1, \ldots, p_n\, x_n\right), \mathbf{x}; \mathbf{w}\right) \tag{4}$$

For a $\mathbf{w}$-taxpayer, this is equivalent to choosing incomes $\mathbf{y}$ to solve:

$$U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,\ldots,y_n)} \qquad \mathscr{U}\left(\sum_{k=1}^{n} y_k - \mathcal{T}\left(y_1, \ldots, y_n\right), \frac{y_1}{p_1}, \ldots, \frac{y_n}{p_n}; \mathbf{w}\right) \tag{5}$$

We assume (see Assumption 3 discussed in Section IV) that for each taxpayer of type $\mathbf{w} \in W$, this program admits a single solution with actions denoted by $\mathbf{X}(\mathbf{w}) = (X_1(\mathbf{w}), \ldots, X_n(\mathbf{w}))$ and incomes denoted by $\mathbf{Y}(\mathbf{w}) = (Y_1(\mathbf{w}), \ldots, Y_n(\mathbf{w}))$ where $Y_i(\mathbf{w}) = p_i\, X_i(\mathbf{w})$. The utility achieved by these taxpayers is $U(\mathbf{w}) = \mathscr{U}\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)$ and the first order-conditions are:

$$\forall i \in \{1, \ldots, n\}: \qquad 1 - \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) = \frac{1}{p_i}\mathcal{S}^i\left(C(\mathbf{w}), \frac{Y_1(\mathbf{w})}{p_1}, \ldots, \frac{Y_n(\mathbf{w})}{p_n}; \mathbf{w}\right) \tag{6}$$

where $C(\mathbf{w}) = \sum_{i=1}^{n} Y_i(\mathbf{w}) - \mathcal{T}\left(\mathbf{Y}(\mathbf{w})\right)$. For each type $i = 1, \ldots, n$ of income, the left-hand side represents the marginal retention rate of the i^{th} income that gives the gain in terms of after tax income when the i^{th} before tax income y_i increases by one euro. The right-hand side corresponds to the marginal rate of substitution between the i^{th} action and after tax income, i.e. the marginal cost in monetary terms of increasing by one euro the i^{th} before tax income y_i.

II.3 Equilibrium

Our equilibrium concept is defined as follows:

Definition 1 (Equilibrium). *Given a tax schedule $\mathbf{y} \mapsto T(\mathbf{y})$, an equilibrium is a set of price $\mathbf{p} = (p_1, \ldots, p_n)$, of incomes $\mathbf{Y}(\mathbf{w})$ for each type $\mathbf{w}$ of taxpayers and of aggregate incomes $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ such that:*

i) Given price $\mathbf{p}$, incomes $\mathbf{Y}(\mathbf{w})$ maximize $\mathbf{w}$-taxpayers utility according to (5).

ii) Aggregate incomes $(\mathcal{Y}_1, \ldots, \mathcal{Y}_n)$ sum individual incomes according to:

$$\mathcal{X}_i \stackrel{def}{\equiv} \int_{\mathbf{w}\in W} X_i(\mathbf{w})\, f(\mathbf{w})\mathrm{d}\mathbf{w} \qquad \text{and} \qquad \mathcal{Y}_i \stackrel{def}{\equiv} \int_{\mathbf{w}\in W} Y_i(\mathbf{w})\, f(\mathbf{w})\mathrm{d}\mathbf{w} = p_i\, \mathcal{X}_i, \tag{7}$$

that is the input markets clear.

iii) Prices are given by inverse demand functions (1) with $\mathcal{X}_i = \mathcal{Y}_i / p_i$.

We denote the joint income density of tax bases $\mathbf{y} = (y_1, \ldots, y_n)$ by $h(\mathbf{y})$ and the unconditional density of the i^{th} income by $h_i(y_i)$.

[10] $A_{i,j}$ is a term of matrix A for which the row is i and the column is j.

II.4 Two policy-relevant examples

The economy we have described is very general. It allows one to study any taxation problem where taxpayers can earn different kinds of income. To illustrate the generality of our framework, we now provide two examples of tax problems that one can easily solve in our framework: the two-period model with labor supply and savings and a model of income-shifting between distinct tax bases. For each of these models, we explain what x_i, y_i and w_i ($\forall i$) represent and how to reinterpret utility function $\mathscr{U}(c, \mathbf{x}; \mathbf{w})$ so that the interpretation of the results will be straightforward.

Example 1: The two-period model

It is useful to begin with an intertemporal setting in order to focus on capital taxation for which the literature has largely emphasized the relevance of the Atkinson-Stiglitz theorem (Atkinson and Stiglitz (1976) and Boadway (2012, Chapter 3), for a nice survey (see also Farhi and Werning (2010) for the reinterpretation of the two period model to estate taxation). Adopting a two-period setting suffices to make the point. We denote the first period or stage by s and the second period or stage by $s+1$. Taxpayers are characterized by $\mathbf{w} = (w_1, w_2)$ where w_1 is their individual labor productivity (skill) and w_2 is their individual inherited wealth.

In the first period, $\mathbf{w}$-taxpayers inherit w_2, save x_2 and consume:

$$c_s = w_2 - x_2$$

In the second period, taxpayers have capital income that is denoted by y_2 with $y_2 = p_2\, x_2$ where p_2 is the (endogenous) return of saving. In the second period, taxpayers also work. They supply x_1 efficient units of labor that depend on their productivity w_1 with market wage rate p_1 so that labor income is $y_1 = p_1\, x_1$. The consumption in second period is the sum of both capital and labor incomes minus taxes $T(y_1, y_2)$, i.e.

$$c_{s+1} = y_1 + y_2 - T(y_1, y_2)$$

which corresponds to our definition of after-tax income c in the general framework. We denote $\mathbf{w}$-agents preferences over first period consumption c_a, second period consumption c_b and efficient units of labor x_1 by $(c_a, c_b, x_1) \mapsto \mathcal{U}(c_a, c_b, x_1; w_1)$. From this lifetime utility, we can retrieve the utility function of the general framework, by a change of variables, as follows:

$$\mathscr{U}(c, x_1, x_2; \mathbf{w}) \stackrel{\text{def}}{\equiv} \mathcal{U}\left(\underbrace{w_2 - x_2}_{=c_a}, \underbrace{c}_{=c_b}, x_1; w_1\right). \tag{8}$$

Example 2: The income-shifting model

Our framework can be consistent with many form of income-shifting and cross-base responses. The example below exhibits one of them.

Consider linear production function (2) with two inputs so that $\gamma_1 = \gamma_2 = p_1 = p_2 = 1$ which implies that $x_1 = y_1$ and $x_2 = y_2$. Assume $\mathbf{w}$-taxpayers have preferences $(d, z_1, z_2) \mapsto \mathcal{U}(d, z_1, z_2; \mathbf{w})$ over consumption d, a first kind of income z_1 and a second kind of income z_2 with $\mathcal{U}_d > 0 > \mathcal{U}_{z_1}, \mathcal{U}_{z_2}$. As an illustration, we may think of self-employed business-owners, where z_1 stands for their effective labor income and z_2 stands for the return on their business. In this context $\mathbf{w} = (w_1, w_2)$ where w_1 is their labor productivity and w_2 is their ability in generating return on their business.

Assume that with some monetary cost $S(\sigma; \mathbf{w})$, taxpayers can shift an amount of income $\sigma \gtreqless 0$ from their first kind of income z_1 to their second kind of income z_2. Reported incomes are then $y_1 = x_1 = z_1 - \sigma$ and $y_2 = x_2 = z_2 + \sigma$. One subtracts the monetary cost $S(\sigma; \mathbf{w})$ from after-tax incomes $c = y_1 + y_2 - T(y_1, y_2)$ to obtain consumption d, i.e. $d = c - S(\sigma; \mathbf{w})$. Assume the cost function S is convex in σ for all $\mathbf{w}$-taxpayers. The determination of how much income to shift is a subprogram for which the value function enables us to retrieve the utility function of the general framework as follows:

$$\mathscr{U}(c, x_1, x_2; \mathbf{w}) \overset{\text{def}}{\equiv} \max_{\sigma} \qquad \mathcal{U}(c - S(\sigma; \mathbf{w}), x_1 + \sigma, x_2 - \sigma; \mathbf{w}) \tag{9}$$

where $\mathcal{U}(c - S(\sigma; \mathbf{w}), x_1 + \sigma, x_2 - \sigma; \mathbf{w}) = \mathcal{U}(d, z_1, z_2; \mathbf{w})$. The indirect utility function associated to this program allows one to be back in our general framework.

Note that one can also interpret y_2 as income invested in tax heavens, in which case this income is constrained to induce no revenue for the domestic government.

II.5 Government

The government acts as a Stackelberg leader in choosing the tax policy, taking into account how its choice is affecting the above-defined equilibrium. In doing so, the government faces the following budget constraint:

$$E \leq \mathscr{B} \overset{\text{def}}{\equiv} \int_{\mathbf{w} \in W} \mathcal{T}(\mathbf{Y}(\mathbf{w})) \; f(\mathbf{w}) \mathrm{d}\mathbf{w} \tag{10}$$

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

$$\mathscr{W} \overset{\text{def}}{\equiv} \int_{\mathbf{w} \in W} \Phi(U(\mathbf{w}); \mathbf{w}) \; f(\mathbf{w}) \mathrm{d}\mathbf{w}. \tag{11}$$

This social objective includes many different specific social objectives. When the government is utilitarian, the social transformation is linear with $\Phi(U, \mathbf{w}) = U$. When the government has weighted utilitarian preferences, the social transformation takes the form $\Phi(U, \mathbf{w}) = \gamma(\mathbf{w})\, U$. A particular case is the maximin (Rawlsian) social welfare where the weights are nil, except for taxpayers with the lowest utility level. Finally, when the government has Bergson-Samuelson preferences, the social transformation does not depend on type and is concave in U.

We assume the government maximizes a linear combination of tax revenue $\mathscr{B}$ and of social welfare $\mathscr{W}$ that we call the government's Lagrangian:

$$\mathscr{L} \stackrel{\text{def}}{\equiv} \mathscr{B} + \frac{1}{\lambda}\mathscr{W} \tag{12}$$

where the Lagrange multiplier $\lambda > 0$ represents the marginal cost of public funds. It is worth noting that we choose to express the Lagrangian in monetary units instead of utility units.

II.6 Taxation regimes

Given the focus of our paper, we consider three main tax regimes: the comprehensive income tax, the separate income tax and a tax regime which is in between those two and that we call the mixed income tax system.[11]

Comprehensive Income Tax system

The tax schedule $\mathcal{T}(\mathbf{y})$ is said to be *comprehensive* if it bears on the sum of all incomes, i.e.:

$$\mathcal{T}(\mathbf{y}) = T\left(\sum_{k=1}^{n} y_k\right)$$

where $T(\cdot)$ is defined on $\mathbb{R}_+$. The marginal tax rate on each income is then identical, so the first-order conditions (6) simplify to:

$$1 - T'\left(\sum_{k=1}^{n} Y_k(\mathbf{w})\right) = \frac{\mathcal{S}^1\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)}{p_1} = \dots = \frac{\mathcal{S}^n\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)}{p_n} \tag{13}$$

Intuitively, since all incomes are put together, the comprehensive tax system does not distort how taxpayers shift their effort among the different incomes. Indeed, 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 the relative price p_i/p_j and it does not depend on taxation.

[11] In a different framework, "mixed taxation" is used to define commodity taxes in the presence of labor income tax, as in Mirrlees (1976).

Schedular Income tax system

The tax schedule $\mathcal{T}(\mathbf{y})$ is said to be *schedular* if different tax schedules bear on distinct kinds of income, i.e.:

$$\mathcal{T}(\mathbf{y}) = \sum_{k=1}^{n} T_k(y_k)$$

where the $T_k(\cdot)$ schedules are defined on $\mathbb{R}_+$. The tax $T_k(\cdot)$ being specific to the kind of income y_k, the marginal tax rate on income y_k depends only on this income (i.e. $\mathcal{T}_{y_i y_j} = 0$ if $i \neq j$), so the first-order conditions (6) become:

$$\forall i \in \{1, ..., n\} \qquad 1 - T_i'(Y_i(\mathbf{w})) = \frac{\mathcal{S}^i(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w})}{p_i} \tag{14}$$

With a schedular income tax system, the marginal tax rate on one kind of income does not depend on the tax on other incomes. With a schedular tax system, taxpayers face incentives to shift some kind of income towards another kind of income that is less taxed.[12]

Mixed tax system

The *mixed* tax system incorporates both the comprehensive tax and the schedular tax system. It consists in adding n *income-specific* tax schedules denoted $T_i(\cdot)$, specific to each income, to a *personal income* tax schedule denoted $T_0(\cdot)$. The personal income tax schedule applies to the sum of all incomes with possible deductions. More specifically, we denote $a_i(y_i)$ the i^{th} taxable income, i.e. the i^{th} income after deductions, with $0 \leq a_i(y_i) \leq y_i$.[13] The net-of deduction functions $a_i(.)$ are assumed increasing and differentiable. Hence, the personal income tax base is equal to $\sum_{k=1}^{n} a_k(y_k)$. The mixed tax schedule is:

$$\mathcal{T}(\mathbf{y}) = T_0\left(\sum_{k=1}^{n} a_k(y_k)\right) + \sum_{k=1}^{n} T_k(y_k) \tag{15}$$

where:

$$y_0 \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k(y_k) \tag{16}$$

[12] An important tax issue around the world is how one should tax incomes from distinct members of the same household. In this context, a comprehensive income tax corresponds to the regime of *joint* or *family* taxation systems where the combined income of married couples and in some cases whole families is taxed as one single unit as in e.g., France, Liechtensein, Luxembourg and Portugal. One can wonder whether the schedular tax system we study in this paper corresponds to the *individual* taxation system which prevails in e.g., Belgium, the Netherlands, Sweden and the UK and under which, the incomes of individuals are taxed separately regardless of marital status or family circumstance. The separate tax system is distinct from the individual taxation system. Indeed the individual taxation system has a unique tax schedule while the separate tax system allows one to apply a distinct income tax schedule to the income of each member of the household (as advocated by (Alesina et al., 2011)). As far as we know, no real tax system has distinct tax schedules for husbands and wives.

[13] There is a normalization issue here. For any $\lambda > 0$, one can reproduce the same personal income tax with deduction functions $\hat{a}_i(y) = \lambda\, a_i(y)$ and personal income tax schedule $\mathbf{y} \mapsto \hat{T}_0(\sum_{k=1}^{n} \hat{a}_k(y_k))$ defined by $y_0 \mapsto \hat{T}_0(y_0) \stackrel{\text{def}}{\equiv} T_0(y_0/\lambda)$. Note that (17) would be unaffected by such a re-normalization.

The personal income tax base $y_0 \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k(y_k)$, personal income for short hereafter, can (partially) exclude some incomes. Incomes that are totally excluded fall under $\sum_{k=1}^{n} T_k(y_k)$. "Partial exclusion" means that part of the income is included in y_0 thanks to the deduction functions $a_k(\cdot)$, whereas the part that is excluded may or may not fall under $\sum_{k=1}^{n} T_k(y_k)$. For instance, in most OECD countries, it is not the cost of labor for the employers that enters the personal income tax base but labor income after payment of social security contributions. Therefore, if y_1 denotes labor cost, $a_1(y_1)$ denotes taxable labor income net of social security contributions. Similarly, when dividends are included in the personal income tax base, these dividends are net of corporate taxation. Denoting y_2 before tax profits accruing to a shareholder, $a_2(y_2)$ denotes taxable dividends net of corporate tax. Virtually, all sources of income can be subject to this kind of deduction.

The mixed tax system encapsulates both the comprehensive and schedular tax systems as specific cases. Indeed assuming $a_1(y) \equiv ... \equiv a_n(y) \equiv y$ and for all i, $y_i \mapsto T_i(y_i) \equiv 0$ and substituting them in (15) yields the comprehensive tax system while $y_0 \mapsto T_0(y_0) \equiv 0$ yields the schedular one.

When one derives both sides of (15) with respect to income y_j, we can see that 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 the marginal deduction rate that applies to this income $a_j'(y_j)$ times the marginal tax rate of the personal income tax schedule $T_0'(y_0)$ that applies to the total personal income y_0:

$$\mathcal{T}_{y_j}(\mathbf{y}) = T_j'(y_j) + a_j'(y_j)\, T_0'\left(\sum_{k=1}^{n} a_k(y_k)\right). \tag{17}$$

The j^{th} marginal tax rate obtained from the schedule $\mathcal{T}(\mathbf{y})$ is then impacted by all incomes through the determination of the taxable personal income y_0 in (16).

III Self-clearing cases

In this section, we present specifications that directly lead to recommend either a separate or a comprehensive tax schedule. These cases are the simplest frameworks that, for purposes of illustration, lead to unambiguous policy recommendations.

III.1 Cases where the Optimal Income Tax is Schedular

In this subsection, we present two economic environments where the optimal tax is schedular.

Proposition 1. *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:*

$$\mathscr{U}(c, \mathbf{x}; w) = c - \sum_{i=1}^{n} v^i(x_i; w) \qquad with \qquad v^i_{x_i}, v^i_{x_i,x_i} > 0 > v^i_w, v^i_{x_i,w} \tag{18}$$

then, the optimal tax is schedular.

The proof can be found in Appendix B. Intuitively, when the unobserved heterogeneity is one-dimensional and the different kinds of income are increasing in type w, redistribution is a single dimension problem from high-types taxpayers, i.e. earning high amounts of each type of income, to low-types taxpayers, who earn low amounts of each type of income. To say it differently, a high income of any kind signals a high type since incomes are perfectly correlated. Due to the separability in actions x_i in the utility function (18), the tax rate on a specific income y_i impacts only the effort to generate this income. There is no cross-base substitution effects. Moreover, due to the quasilinearity in consumption, there is no income effect. The government can therefore simply shift distortions on the least responsive tax bases in the vein of an inverse elasticity rule see e.g., Ramsey (1927) and Baumol and Bradford (1970). This is made possible with a schedular income tax system.

Proposition (Atkinson-Stiglitz). *When, in two-periods model with endogenous labor supply and savings, the i) preferences (8) are weakly separable between efficient labor, x_1, and consumption bundles (c_s and c_{s+1}), i.e.:*

$$\mathcal{U}(c_s, c_{s+1}, x_1; w_1) = \mathsf{U}\left(V(c_s, c_{s+1}), x_1; w_1\right) \qquad with \qquad \mathsf{U}_V, V_{c_s}, V_{c_{s+1}} > 0$$

with $V(\cdot)$ twice continuously differentiable and increasing in each argument and ii) individuals have the same initial wealth w_2 and heterogeneous productivity w_1, then the optimal tax is schedular.

In the above proposition, taxpayers are heterogeneous along a single dimension, their labor productivity, w_1. Combined with the weak separability of the utility function, all assumptions of the Atkinson and Stiglitz (1976) theorem are satisfied. We know from the latter theorem that capital should therefore not be taxed at the optimum. Indeed taxing capital will not improve equity in comparison to the non-linear tax on labor earnings, while additionally distorting savings. In our framework, this requires to exclude capital from the personal income tax base so that a schedular tax system is optimal.

III.2 A case where the Optimal Income Tax is Comprehensive

In this subsection, we describe a situation where the optimal tax system is comprehensive. The following Proposition is proved in Appendix C. Our proof is constructed on a similar reasoning than in one found in Konishi (1995), Laroque (2005) and Kaplow (2008) but is valid with general tax instruments and multidimensional incomes.[14]

Proposition 2. *If preferences are weakly separable, i.e. the utility function $\mathscr{U}$ takes the form $\mathscr{U}(c, \mathbf{x}; \mathbf{w}) = \mathcal{U}(c, \mathcal{V}(\mathbf{x}); \mathbf{w})$ 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 and if the production function exhibits perfect substitution as in (2) then, the optimal tax is comprehensive.*

Since preferences are weakly separable, whatever their type, individuals choose how to split their actions in getting the different kinds of income to minimize the same aggregation $\mathcal{V}(\cdot)$ of actions. Moreover, the government is only interested in the sum of all incomes earned by each individual. Indeed actions being weakly separable from after-tax income in the utility function, two taxpayers who have the same aggregate effort $\mathcal{V}(\mathbf{x})$ but differ in their type $\mathbf{w}$ or in their consumption c will choose the same bundle of actions $\mathbf{x}$. This will be the case for a person of a given type mimicking the income vector $\mathbf{y}$ of a person with another type. This incentive constraint cannot be weakened by imposing schedular taxation. It can only make all taxpayers worse off. Indeed, the marginal rate of substitution between two different actions does not depend on type as it verifies:

$$\frac{\mathscr{U}_{x_i}(c, \mathbf{x}; \mathbf{w})}{\mathscr{U}_{x_j}(c, \mathbf{x}; \mathbf{w})} = \frac{\mathcal{V}_{x_i}(\mathbf{x})}{\mathcal{V}_{x_j}(\mathbf{x})}$$

Therefore, a modification of the action vector $\mathbf{X}(\mathbf{w})$ assigned to $\mathbf{w}'$-taxpayers affects their utility in the same way as the utility of $\mathbf{w}$-taxpayers mimicking $\mathbf{w}'$-taxpayers. The government does not need to distort the relative supply of each action. A comprehensive tax schedule is therefore optimal.

Even if the realism of the cases presented in 1-2 is questionable,[15] they are helpful to em-

[14] These authors show that a linear indirect tax is useless when a nonlinear labor income tax prevails. Indeed, despite the fact that the agents choose the same allocation under both tax systems, the government's revenue is proven to be larger with a zero indirect tax rate than with a positive one.

[15] Proposition 1 builds upon taxpayers who differ along a single dimension as standard in the Mirrlees (1971) literature. This is not very convincing empirically, in particular with different kinds of income. Under the weakly separable preferences used in Proposition 2, people who earn the same taxable income $v = \sum_{i=1}^{n} x_i$ choose the same actions $(x_1, ..., x_n)$. Taxpayers who earn the same level of one kind of income must earn the same levels of income for all other sources of income, which is also not very convincing empirically. Indeed the program of individuals of type $\mathbf{w}$ can be decomposed into two consecutive stages. In the first stage, taxpayers choose how to split their actions $\mathbf{x}$ to earn a given taxable income $v = \sum_{i=1}^{n} x_i$:

$$\min_{\mathbf{x} \text{ s.t: } \sum\limits_{i=1}^{n} x_i = v} \quad \mathcal{V}(\mathbf{x}).$$

In the second stage, taxpayers choose what taxable income to choose:

$$\max_{v} \quad \mathcal{U}\left(v - \mathcal{T}(v), \min_{\mathbf{x} \text{ s.t: } \sum\limits_{i=1}^{n} x_i = v} \quad \mathcal{V}(\mathbf{x}); \mathbf{w}\right).$$

phasize the mechanisms that lead to recommend either a schedular or a comprehensive tax system.

IV Computing the effects of tax reforms

In this section, we characterize how the equilibrium (See Definition 1) is impacted by tax reforms. For this purpose, we first study in subsection IV.1 taxpayers' responses to a set of tax reforms and to price changes. These responses – that we decompose into wealth responses, compensated responses and price responses – take into account that a tax reform or a price change can simultaneously impact several income bases. We compute taxpayers' responses to any possible tax reform by differentiating taxpayers' first-order conditions associated to program (5). We hence obtain *micro* responses that occur when prices are taken as given, as in the usual framework.[16] We also obtain responses to price changes.

Second, in subsection IV.2, we characterize, using the firms' demand equations (1), how any micro response to tax reforms impulses general-equilibrium effects through changes in the input prices. Micro responses change aggregate actions, that change input prices through (1), which in turn induce taxpayers responses to price changes, and so on. We then define sufficient statistics, which we call *macro spillover* statistics, that summarize this process. Starting from a given initial, potentially suboptimal, tax schedule, we then give a general formula describing the impact of tax reforms on welfare taking into account general equilibrium effects. This formula is expressed in terms of behavioral responses and sufficient statistics which are empirically meaningful.

IV.1 Taxpayers' responses to tax reforms and price changes

We begin by defining a tax reform.

Definition 2. *A tax reform replaces the tax schedule* $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ *by a new twice continuously differentiable tax function* $(\mathbf{y}, t) \mapsto \widetilde{\mathcal{T}}(\mathbf{y}, t)$ *defined over* $\mathbb{R}_+^n \times I$, *where the scalar* $t \gtrless 0$ *is a measure of the magnitude of the tax reform and* I *is an open interval containing* 0 *such that for all* $\mathbf{y} \in \mathbb{R}_+^n$, *one has* $\widetilde{\mathcal{T}}(\mathbf{y}, 0) = \mathcal{T}(\mathbf{y})$ *so that* $\widetilde{\mathcal{T}}(\mathbf{y}, 0)$ *is the initial tax schedule.*

Consider an arbitrary reform that replaces the initial tax schedule by $\mathbf{y} \mapsto \widetilde{\mathcal{T}}(\mathbf{y}, t)$. We denote the utility level of $\mathbf{w}$-taxpayers by $\widetilde{U}(\mathbf{w}, t)$, their i^{th} income by $\widetilde{Y}_i(\mathbf{w}, t)$ and the i^{th} price

The first stage's problem is type-independent so that taxpayers who earn the same i^{th} income also receive the same j^{th} income.

[16] We call them *micro* responses (see also Kroft et al. (2020)) because in *micro*econometrics, if a tax reform affects only a treatment group and not a control group and both groups face the same prices, the usual empirical strategies, such as difference-in-differences, would only identify micro responses ignoring the effects of changes in prices.

by $\widetilde{p}_i(t)$. Incomes generated by a $\mathbf{w}$-taxpayer $\widetilde{\mathbf{Y}}(\mathbf{w},t) = (\widetilde{Y}_1(\mathbf{w},t),,...,\widetilde{Y}_n(\mathbf{w},t))$ solve:

$$\widetilde{U}(\mathbf{w},t) \stackrel{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,...,y_n)} \quad \mathscr{U}\left(\sum_{i=1}^{n} y_i - \widetilde{\mathcal{T}}(\mathbf{y},t), \frac{y_1}{\widetilde{p}_1(t)},..., \frac{y_n}{\widetilde{p}_n(t)}; \mathbf{w}\right) \tag{19}$$

In a similar way, we denote $\widetilde{\mathscr{B}}(t)$, $\widetilde{\mathscr{W}}(t)$ and $\widetilde{\mathscr{L}}(t) \stackrel{\text{def}}{\equiv} \widetilde{\mathscr{B}}(t) + \frac{1}{\lambda}\widetilde{\mathscr{W}}(t)$, the government's tax revenue (defined in (10)), the social objective (defined in (11)) and the government's Lagrangian (defined in (12)) when the tax schedule is perturbed according to $(\mathbf{y},t) \mapsto \widetilde{\mathcal{T}}(\mathbf{y},t)$.[17]

We now explain how the economy adjusts to tax reforms by presenting, for each type of taxpayers $\mathbf{w}$, the responses of each kind of income that are due to (i) behavioral responses and (ii) endogenous prices. As in the case of exogenous prices and a single kind of income (Saez, 2001), any tax reform can imply wealth responses, compensated responses and uncompensated responses. It is worth stressing that all these responses are *total* responses of incomes, as in Jacquet et al. (2013), Scheuer and Werning (2017) and Sachs et al. (2020). They take into account the nonlinearity of the tax schedules hence the circular process that occurs with nonlinear tax schedules: A change in income by $\mathbf{w}$-taxpayers creates endogenously a change in the marginal tax rate on their income so that they further adjust their income.

IV.1.a Behavioral responses

Wealth responses

We define the wealth responses as the behavioral responses to a small change ρ in the tax liability of $\mathbf{w}$-taxpayers so that the tax schedule becomes:

$$\widetilde{\mathcal{T}}(\mathbf{y},\rho) = \mathcal{T}(\mathbf{y}) - \rho. \tag{20a}$$

We denote $\frac{\partial Y_i(\mathbf{w})}{\partial \rho}$ how $\mathbf{w}$-taxpayers modify their i^{th} income after this *lump-sum* tax perturbation and call it wealth responses.

Compensated responses

We now study a tax reform that impacts the individual first-order conditions only through substitution effects, shutting down wealth responses. We denote $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}$ the *compensated response* of a $\mathbf{w}$-taxpayer in terms of her i^{th} income $Y_i(\mathbf{w})$ to a change in the j^{th} marginal net-of-tax rate by a constant amount τ_j around income $Y_j(\mathbf{w})$, while leaving unchanged the level of tax at initial incomes $\mathbf{Y}(\mathbf{w})$. That is, after a compensated tax reform, the tax schedule becomes:

$$\widetilde{\mathcal{T}}(\mathbf{y},\tau_j) = \mathcal{T}(\mathbf{y}) - \tau_j\left(y_j - Y_j(\mathbf{w})\right). \tag{20b}$$

[17]Note that we define the perturbed Lagrangian $\widetilde{\mathscr{L}}(t)$, keeping unchanged the weight $1/\lambda$ put on the social objective $\widetilde{\mathscr{W}}(t)$. This will appear convenient in Proposition **??** below.

The response and reform are said *compensated* in the sense that the tax level is unchanged at $y = Y(\mathbf{w})$, whatever the magnitude τ_j.

Due to substitution effects, this change in the j^{th} marginal tax rate can modify every kind of income $Y_i(\mathbf{w})$ $(i = 1, ..., n)$. Indeed, as in the Mirrleesian model (with a single kind of income), due to substitution effects, when one modifies the marginal tax rate on income j, the taxpayer modifies her effort to earn the j^{th}-income, hence the level of $Y_j(\mathbf{w})$. Differing from the single kind of income model, substitution effects can also take place between the distinct kinds of incomes, e.g. because of income shifting. Due to these cross-income responses, a reform of the j^{th} marginal tax rate can possibly impact every other income $i = 1, ..., n$.

Uncompensated responses

We denote $\frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_j}$ the *uncompensated* response of the i^{th} income to a change in the j^{th} marginal net-of-tax rate by a constant amount τ_j, when one relaxes the assumption of constant tax liability. After an uncompensated tax reform, the tax schedule becomes:

$$\widetilde{\mathcal{T}}(\mathbf{y}, \tau_j) = \mathcal{T}(\mathbf{y}) - \tau_j\, y_j, \tag{20c}$$

As one can expect, if prices are held constant, the *compensated* and *uncompensated* responses of the i^{th} income to the j^{th} marginal tax rate are related by the Slutsky equation according to:

$$\frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_j} = \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} + Y_j(\mathbf{w}) \frac{\partial Y_i(\mathbf{w})}{\partial \rho}, \tag{20d}$$

where, in the right-hand side, the compensated response of a $\mathbf{w}$-taxpayer in terms of income i is added to her wealth response times her chosen quantity of income j.

Price responses

Finally, let $\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}$ denote the taxpayer's behavioral response in terms of her specific income Y_i caused by a 1% increase in the j^{th} price. Any change in the price of a given input can impact individual effort to generate this input hence the level of aggregate income associated to this input. The change in the price of this input can also impact the effort to generate another input (hence the level of associated aggregate income) whenever inputs are not perfect substitutes (since, when they are, the marginal productivities and input prices are fixed).

IV.1.b Effects of tax reforms

Effects on incomes

We now detail the behavioral adjustments of each kind of income to a tax reform of magnitude t. We denote $\left.\frac{\partial A}{\partial t}\right|_{t=0}$, the partial derivative of an economic variable A along the tax perturbation $\mathbf{y} \mapsto \widetilde{\mathcal{T}}(\mathbf{y}, t)$ at $t = 0$. Using the behavioral responses defined above, we can explicit $\frac{\partial \widetilde{Y}_i(\mathbf{w},t)}{\partial t}\Big|_{t=0}$. As shown in Appendix D, this leads to the following expression:[18]

$$\begin{aligned}
\left.\frac{\partial \widetilde{Y}_i(\mathbf{w}, t)}{\partial t}\right|_{t=0} = & \underbrace{-\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0}}_{\text{Compensated responses}} - \underbrace{\frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0}}_{\text{Wealth responses}} \\
& + \underbrace{\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} \left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0}}_{\text{Prices responses}}
\end{aligned} \tag{21}$$

From the individual first-order conditions (6), a tax perturbation affects these conditions through three channels. First, changes in the marginal tax rates $\mathcal{T}_{y_j}$ in the left-hand side of (6) create *compensated* responses from all income sources. Second, the change in the tax liability induces *wealth* responses. Third, prices responses $\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}$ occur as soon as the price of an input changes due to behavioral responses. In our framework, prices changes result from general equilibrium effects which will be detailed in Section IV.2. As long as one focuses on micro responses only, $\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\Big|_{t=0} = 0$. Equation (21) highlights the extent to which our model encompasses all possible behavioral responses.

Effects on tax liability

The impact of a tax reform on the tax liability of $\mathbf{w}$-taxpayers $\widetilde{\mathcal{T}}\left(\widetilde{\mathbf{Y}}(\mathbf{w}, t), t\right)$ can be decomposed into *mechanical* and *behavioral* effects, as follows:

$$\left.\frac{\mathrm{d}\widetilde{\mathcal{T}}\left(\widetilde{\mathbf{Y}}(\mathbf{w}, t), t\right)}{\mathrm{d}t}\right|_{t=0} = \underbrace{\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0}}_{\text{Mechanical effects}} + \underbrace{\sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \left.\frac{\partial \widetilde{Y}_i(\mathbf{w}, t)}{\partial t}\right|_{t=0}}_{\text{Behavioral effects}}. \tag{22}$$

The first term on the right hand side of (22) is the mechanical effect of the tax reform, i.e., the mechanical change in individual tax liability assuming that the individual decisions in the levels of the different kinds of income as well as the different input prices remain constant (in other words, assuming that individual total income $Y(\mathbf{w})$ remains constant). The second term is the behavioral effects of the reform. Each behavioral response that modifies a kind of income induces a change in tax liability proportional to the associated marginal tax rate $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))$.

Plugging Equation (21) into (22) allows one to decompose the impact of a tax perturbation in terms of the effects induced by the changes in tax liabilities (i.e. mechanical effects and wealth effects), by the changes in marginal tax rates (compensated effects) and by the (log of)

[18] We derive (21) using the implicit function theorem thanks to Assumption 3 in Appendix D.

price changes.

$$\left.\frac{\mathrm{d}\widetilde{\mathcal{T}}\left(\widetilde{\mathbf{Y}}(\mathbf{w},t),t\right)}{\mathrm{d}t}\right|_{t=0} = \left[1-\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial\rho}\right]\left.\frac{\partial\widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \tag{23}$$

$$-\sum_{1\leq i,j\leq n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial\tau_j}\left.\frac{\partial\widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}+\sum_{1\leq i,j\leq n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial\log p_j}\left.\frac{\partial\log\widetilde{p}_j(t)}{\partial t}\right|_{t=0}.$$

Effects on welfare

Since $\lambda>0$ denotes the shadow cost of public funds, the marginal social welfare weight in monetary units associated with taxpayers $\mathbf{w}$, is defined as:

$$g(\mathbf{w})\stackrel{\text{def}}{\equiv}\frac{\Phi_U(U(\mathbf{w});\mathbf{w})\,\mathscr{U}_c\,(C(\mathbf{w}),\mathbf{X}(\mathbf{w});\mathbf{w})}{\lambda} \tag{24}$$

The marginal social welfare weight is the social value of giving one extra unit of consumption to taxpayers with type $\mathbf{w}$, assuming prices are held constant.

The effects, in monetary terms, of a tax reform on the social welfare of a $\mathbf{w}$-taxpayer can be obtained by adding the mechanical effects on her tax liability to the effects of the reform on individual utilities, weighting the sum by the marginal social welfare weights $g(\mathbf{w})$, as follows:

$$\frac{1}{\lambda}\left.\frac{\partial\Phi\left(\widetilde{U}(\mathbf{w},t);\mathbf{w}\right)}{\partial t}\right|_{t=0}=\left(-\left.\frac{\partial\widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}+\sum_{j=1}^{n}\left(1-\mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right)Y_j(\mathbf{w})\left.\frac{\partial\log\widetilde{p}_j(t)}{\partial t}\right|_{t=0}\right)g(\mathbf{w}) \tag{25}$$

The proof, where the envelope theorem is applied to the individual maximization program (19), is relegated to Appendix D. For each taxpayer, the tax reform has a direct effect on welfare through the change in tax liability as the term $-\left.\frac{\partial\widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}$ points out. On top of this *mechanical effect*, the tax reform can also modify behaviors and prices. Changes in behaviors only induce second-order effects on welfare. Indeed taxpayers' decisions are perturbed from their optimum and since they choose their incomes $\mathbf{y}$ to maximize their utility, they are indifferent to small changes in their incomes $\mathbf{y}$ to a first-order approximation. This "envelope" argument is well understood since Saez (2001). However, it does not apply to prices changes because taxpayers take prices as given. Applying the envelope theorem to (4), a one-percent increase in the j^{th} price, $\left.\frac{\partial\log\widetilde{p}_j(t)}{\partial t}\right|_{t=0}$, has an impact on the taxpayer's utility that is identical to a mechanical increase of consumption by the amount $(1-\mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w})))Y_j(\mathbf{w})$ (see Appendix D for details).

Effects on government's Lagrangian

We now obtain the impact of a tax reform on the government's Lagrangian (12). We sum, across all types $\mathbf{w}$, the impact on their tax liability (23) and on their welfare (25). This yields:

$$\begin{aligned}
\frac{\partial \widetilde{\mathscr{L}}}{\partial t} = & \int_{\mathbf{w}\in W} \left\{ \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \right. \\
& - \sum_{1\le i,j\le n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left. \frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \\
& \left. + \sum_{j=1}^{n} \left[\left(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w})\, g(\mathbf{w}) + \sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} \right] \left. \frac{\partial \log \widetilde{p}_j(t)}{\partial t} \right|_{t=0} \right\} f(\mathbf{w}) d\mathbf{w}.
\end{aligned} \tag{26}$$

This formula expresses in terms of empirically estimable sufficient statistics and social welfare weights whether a given tax reform is socially desirable. We now study how this formula has to be modified to take into account the endogenous changes in prices.

IV.2 General equilibrium

We now focus on the impact of a tax reform on the equilibrium (see Definition 1)In general equilibrium, taxpayers' decisions depend on prices, and prices are determined by firms' inverse demand equations. A tax reform impacts this general equilibrium because it impacts taxpayers decisions, through what we call *micro responses*. Combining Equations (7) and (21) where one has put to zero the term that contains the prices responses $\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} \left. \frac{\partial \log \widetilde{p}_j(t)}{\partial t} \right|_{t=0}$, the micro responses of the i^{th} aggregate income to tax reforms are defined by:

$$\left. \frac{\partial \widetilde{\mathcal{Y}}_i(t)}{\partial t} \right|_{t=0}^{Micro} = - \int_{\mathbf{w}\in W} \left\{ \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} + \sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left. \frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \right\} f(\mathbf{w}) d\mathbf{w}. \tag{27}$$

With endogenous prices, these micro responses of the aggregate incomes $\mathcal{Y}_i(t)$ modify all input levels, thereby the marginal products of each input, and eventually the factors' prices, according to aggregate input demand equations (1). In turn, each taxpayer responds to these price changes according to (21). Therefore, all aggregate incomes $(\widetilde{\mathcal{Y}}_1(t), ..., \widetilde{\mathcal{Y}}_n(t))$ are impacted, which in turn feeds back into the prices, and so on. At equilibrium, for each reform's magnitude t, prices $(\widetilde{p}_1(t), ..., \widetilde{p}_n(t))$ have to verify the following fixed-point conditions in the prices' adjustment:

$$\forall t, \forall j \in \{1, ..., n\} \qquad \widetilde{p}_i(t) = \mathcal{F}_{\mathcal{X}_i} \left(\frac{\widetilde{\mathcal{Y}}_1(t)}{\widetilde{p}_1(t)}, ..., \frac{\widetilde{\mathcal{Y}}_n(t)}{\widetilde{p}_n(t)} \right). \tag{28}$$

Let Ξ denote the matrix of inverse demand elasticities in which the term in the i^{th} line and j^{th} column is the inverse input's demand elasticity of the i^{th} price p_i with respect to the j^{th} input factor $\mathcal{X}_j$:

$$\Xi_{i,j} \stackrel{\text{def}}{\equiv} \frac{\mathcal{X}_j\, \mathcal{F}_{\mathcal{X}_i \mathcal{X}_j}}{\mathcal{F}_{\mathcal{X}_i}}. \tag{29a}$$

Let Σ denote the matrix of i^{th} aggregate income elasticity with respect to price p_j, i.e. the matrix in which the term in the i^{th} line and the j^{th} column is given by:

$$\Sigma_{i,j} \stackrel{\text{def}}{\equiv} \left.\frac{\partial \log \mathcal{Y}_i}{\partial \log p_j}\right|_{t=0} = \frac{1}{\mathcal{Y}_i} \int_{\mathbf{w}\in W} \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \tag{29b}$$

The percentage change in aggregate income i when the price p_j changes is made of the sum, across taxpayers, of the percentage changes of the individual incomes i generated by all taxpayers when price j is modified. We denote I_n the n-identity matrix and we make the following assumption:

Assumption 1. *The matrix $I_n + \Xi - \Xi \cdot \Sigma$ is invertible.*

The matrix $I_n + \Xi - \Xi \cdot \Sigma$ shows up when one log-differentiate (28). Thanks to Assumption 1, Equation (28) is invertible and one can apply the implicit function theorem to ensure that equilibrium prices are differentiable with respect to the magnitude t of the tax perturbation. When the production function is linear (as in (2)), matrix Ξ is nil hence Assumption 1 is automatically verified. Therefore, by continuity, Assumption 1 remains satisfied as long as the elasticities of substitution between input factors are sufficiently high.

IV.2.a Macroeconomic price spillovers

In Appendix D, we describe how a tax reform impacts prices. A tax reform first implies micro responses at given prices. These responses induce prices' changes through the demand side of the economy that depend on matrix Ξ. These prices' changes in turn induce prices' responses through the supply side of the economy, according to matrix Σ, which in turn imply prices' responses from the demand side, and so on. We thus have:

$$\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0} = \sum_{i=1}^{n} \Pi_{j,i} \left.\frac{\partial \widetilde{\mathcal{Y}}_i(t)}{\partial t}\right|_{t=0}^{Micro} \quad where: \; \Pi = (I_n + \Xi - \Xi \cdot \Sigma)^{-1} \cdot \Xi \cdot \begin{pmatrix} \frac{1}{\mathcal{Y}_1} & 0 \\ 0 & \frac{1}{\mathcal{Y}_2} \end{pmatrix} \tag{30}$$

with Matrix $[A]^{-1}$ the inverse of matrix A. Matrix $\Pi_{j,i}$ describes how micro responses to tax reforms translate into a log-change of prices through this process (see Appendix D). We refer to Π as the matrix of price multipliers. The term $\Pi_{j,i}$ provides the relative change in the j^{th} price to an aggregate micro response of the i^{th} income to any tax reform. As a limit case, under the linear production function (2), the matrix of inverse demand elasticities Ξ simplifies to the nil matrix according to (29a), in which case the price multiplier $\Pi_{j,i}$ are also nil and the process of prices' adjustments vanishes.

For each type $i \in \{1, ...n\}$ of income, we now define a sufficient statistic, μ_i, which indicates how any micro response of the i^{th} income impacts the Lagrangian (12) through changes in prices, whatever the tax reform that triggers this micro response, and whatever the types of workers who respond. We call it *macroeconomic price spillover* statistic, hereafter *macro price spillover* statistic. On top of the general equilibrium aspect, the term *macro* emphasizes that the impact of this change is not specific to a particular tax reform, nor to a type of taxpayers $\mathbf{w}$. The terms *price spillover* stresses that the tax reform impacts prices via firms' decisions and taxpayers' responses to prices' changes. We get: $\forall i \in \{1, ..., n\}$:

$$\mu_i \stackrel{\text{def}}{\equiv} \sum_{j=1}^{n} \Pi_{j,i} \int_{\mathbf{w}\in W} \left[\left(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w})\, g(\mathbf{w}) + \sum_{k=1}^{n} \mathcal{T}_{y_k}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j} \right] f(\mathbf{w}) \mathrm{d}\mathbf{w} \tag{31}$$

From (30), the price multipliers $\Pi_{j,i}$ capture how micro responses in incomes result in changes in prices in general equilibrium. According to (26), the integral in (31) corresponds to the impact on the Lagrangian of a one-percent increase in the j^{th} price. Plugging Equations (27), (30) and (31) into (26) leads to the impact of a tax reform on the government Lagrangian formulated as:

$$\begin{aligned} \left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{\mathbf{w}\in W} \left\{ \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} \left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) + \mu_i\right) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \right. \\ &- \left. \sum_{1\leq i,j\leq n} \left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) + \mu_i\right) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \end{aligned} \tag{32}$$

This tax formula is expressed as a function of behavioral responses, the macro price spillover statistic μ_i and other sufficient statistics that can be estimated empirically, as we show with French data in Section VI. Being easily implementable empirically, this formula can be used to evaluate the impact of any tax reform in terms of tax revenue and welfare. We now provide economic intuitions for each of its terms.

Absent any behavioral response, the tax reform mechanically impacts the government tax receipts and the social welfare weight as reflected by $1 - g(\mathbf{w})$ in the first line of Equation (32). Then, for each $\mathbf{w}$-taxpayer and each type i of income, behavioral responses and price responses have to be taken into account. First behavioral (wealth and compensated) responses modify $Y_i(\mathbf{w})$ by $\Delta Y_i(\mathbf{w})$ so that tax liability is affected by $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\Delta Y_i(\mathbf{w})$.

In addition, the presence of endogenous prices and the implied general equilibrium effects encompassed in μ_i modify prices along the process described by Equation (30). Incorporating these price spillover effects amounts to correcting (i.e. increasing or decreasing) the marginal tax rates $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))d\mu_i$. When $\mu_i > 0$ ($\mu_i < 0$), the change in aggregate incomes $\mathcal{Y}_i$, implied by the tax reform and all individual behavioral responses, increases (decreases) the government Lagrangian, via the above-mentioned process.

More intuition on μ_i can be obtained in the classical general equilibrium context, with capital indexed by 1 and labor indexed by 2, and with tax revenue and social welfare mainly fed by labor income y_1. Lowering marginal tax rates on capital income will encourage to supply more capital, which will reduce the marginal productivity of capital (assuming diminishing marginal returns) and its price p_2. Assuming capital and labor are imperfect substitutes, it will also raise, the marginal productivity of labor and its price p_1. Therefore, to guarantee this positive spillover effect, one expects a positive macro spillover statistic for capital $\mu_2 > 0$ to correct downwards the marginal tax rate on capital income $\mathcal{T}_{y_2}(\mathbf{Y}(\mathbf{w}))$. Symmetrically, one expects a negative macro spillover statistic for labor $\mu_1 < 0$ to correct upwards the marginal tax rate on labor income $\mathcal{T}_{y_1}(\mathbf{Y}(\mathbf{w}))$. Thanks to this correction, the detrimental reduction of the labor price p_1 induced by spillover effects (from micro responses which increase $\mathcal{Y}_1$) will be limited.

In this context, in general equilibrium, lowering marginal tax rates on capital income (with a positive macro price spillover statistic) will encourage to supply more capital, which will reduce the marginal productivity of capital (assuming diminishing marginal ~~products~~ returns) and its price. Assuming capital and labor are imperfect substitutes, it will raise, the marginal productivity of labor and its price.

In an economy where labor is the main source of income, as in France for instance, the aggregate amount of labor income being larger than the aggregate amount of capital, a change in capital has a stronger impact on input prices than a change in labor. Hence, we can expect, in absolute terms, μ_1 to be lower than μ_2. Although these results are obtained under very simplifying assumptions, the mechanisms we highlight are more general and will help us to understand the values obtained in the numerical simulations.

IV.2.b Effects of balanced tax reforms

Equation (32) allows policy advisers to determine the effects of a tax reform. However, let us stress that these tax reforms are not budget-balanced, unless $\left.\frac{\partial \widetilde{\mathscr{B}}}{\partial t}\right|_{t=0} = 0$. It is very important for policy makers to be able to choose the best tax reform among those that are self-financed. It is well known that one way to easily balance any tax reform is to use a (positive or negative) lump-sum transfer, see Sandmo (1998) and Jacobs (2018). In the next proposition, we characterize the impact in terms of welfare of any tax reform balanced in a lump-sum way.

The effects on social welfare of a lump-sum transfer to every taxpayer corresponds to the shadow cost of public funds λ. Applying Equation (32) to the lump-sum reform (20a), the shadow cost of public funds is pined down by:[19]

$$0=\int_{\mathbf{w}\in W}\left[1-g(\mathbf{w})-\sum_{i=1}^{n}\left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))+\mu_i\right)\frac{\partial Y_i(\mathbf{w})}{\partial\rho}\right]f(\mathbf{w})d\mathbf{w}, \tag{33}$$

with λ included into $g(\mathbf{w})$, see Equation (24).

Proposition 3. *If the shadow cost of public funds verifies* (33) *and if* $\frac{\partial\widetilde{\mathscr{L}}}{\partial t}\Big|_{t=0}$ *defined in* (32) *is positive (negative), then reforming the tax schedule to* $\mathbf{y}\mapsto\widetilde{\mathcal{T}}(\mathbf{y},t)$ *with a small positive* t *(a small negative* t*) and rebating the net budget surplus in a lump-sum way is a budget-balanced reform that is socially desirable.*

According to Proposition 3, which is proved in Appendix E, the welfare impact of a tax reform balanced thanks to a lump-sum transfer has the same sign as the effect of the initial tax reform on the government's Lagrangian. In light of the tax formula (32), one can describe how to self-finance any tax reform (in a lump-sum way) and conclude whether this reform is socially desirable or not.

In the rest of the paper, in order to define income densities, we make the following assumption on preferences:

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

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

Assumption 2 implies that the mapping $\mathbf{y}\mapsto\mathbf{Y}(\mathbf{w})$ is *globally* invertible.[20] Let $\left[\frac{\partial Y_i(\mathbf{w})}{\partial w_j}\right]_{i,j}$ denote the Jacobian matrix of this mapping for $\mathbf{w}$-taxpayers. We thus get the following relationship between the joint income density and the type density:

$$h(\mathbf{Y}(\mathbf{w}))=\frac{f(\mathbf{w})}{\left|\det\left[\frac{\partial Y_i(\mathbf{w})}{\partial w_j}\right]_{i,j}\right|}. \tag{34}$$

[19]We need to assume that, taking into account all behavioral responses, one has:

$$1-\sum_{k=1}^{n}\int_{\mathbf{w}\in W}\left(\mathcal{T}_{y_k}(\mathbf{Y}(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial\rho}f(\mathbf{w})d\mathbf{w}>0$$

i.e. that a lump-sum transfer to taxpayers reduces government's tax revenue.

[20]Assume by contradiction the existence of two types $\mathbf{w},\mathbf{w}'$ such that that $\mathbf{Y}(\mathbf{w})=\mathbf{Y}(\mathbf{w}')=\mathbf{y}$ and therefore $\mathbf{X}(\mathbf{w})=\mathbf{X}(\mathbf{w}')=\mathbf{x}$. 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 (6), we get:

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

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

V Optimal Taxation under mixed tax schedules

Having illustrated in Section III that a comprehensive or a schedular tax system is hardly optimal in reality, we can now study the effects of tax reforms within the family of mixed tax functions described in Equations (15) and (16). We first study the effects of reforming the income-specific tax schedules $T_i(\cdot)$ and personal income tax schedules $T_0(\cdot)$ to derive optimal tax schedules' formulas. Second, we consider reforms of the personal income tax base to discuss whether or not it is socially desirable to reform the system towards a slightly more or a slightly less schedular tax system.

V.1 Optimal mixed Tax schedules

Effect of tax reforms on personal income

We first describe wealth responses, compensated responses, uncompensated responses and price responses (described in Subsection IV.1.a) of the personal income y_0 defined in (16). Each of these responses is the weighted sum of the responses of each specific income y_k, for $k \in \{1, ..., n\}$, the weights being the marginal deduction rates $a'_k(y_k)$. We first characterize how behavioral responses to a tax reform modify the personal income tax base y_0 defined in (16). We combine the taxpayers' responses (20a), (20b), (20c) and the changes in prices induced by general equilibrium described in (30) with the definition of personal income in (16). For any reform, the impact on personal income y_0 consists in the sum of the induced changes in each specific income k, each of these income changes being multiplied by its marginal deduction rate $a'_k(y_k)$. Formally, the wealth response is:

$$\frac{\partial Y_0(\mathbf{w})}{\partial \rho} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a'_k(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}. \tag{35a}$$

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

$$\frac{\partial Y_0(\mathbf{w})}{\partial \tau_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a'_k(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}. \tag{35b}$$

The uncompensated response of personal income tax base to an (uncompensated) tax change in the j^{th} marginal tax rate is:

$$\frac{\partial Y_0^u(\mathbf{w})}{\partial \tau_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a'_k(y_k) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_j}, \tag{35c}$$

and the price response of personal income tax base to a relative change in the j^{th} price is:

$$\frac{\partial Y_0}{\partial \log p_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a'_k(y_k) \frac{\partial Y_k}{\partial \log p_j}. \tag{35d}$$

Note that, since personal income tax y_0 does not correspond to an input factor, we normalize $p_0 = 1$ and $\mu_0 = 0$.

Effect of reforming the tax that applies to a specific income

Consider a reform of the tax schedule on a specific income i, for any $i \in \{1, ..., n\}$. This reform replaces the initial tax schedule by the new tax function $\widetilde{\mathcal{T}}(\mathbf{y}, t) = \mathcal{T}(\mathbf{y}) - t\ R_i(y_i)$ where $R_i(\cdot)$ is the direction of the reform. This reform modifies the individual tax liability by:

$$\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_i(Y_i(\mathbf{w})). \tag{36}$$

It does modify the i^{th} marginal tax rate by:

$$\left.\frac{\partial \widetilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_i'(Y_i(\mathbf{w})). \tag{37}$$

Note that it does not modify the other marginal tax rates. Substituting (36) and (37) in (32), we obtain that the effect, on the Lagrangian, of reforming the tax schedule that prevails on the i^{th} income, is:

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = & \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_i(Y_i(\mathbf{w})) \right. \\ & + \left. \left[\sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right] R_i'(Y_i(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}.\end{aligned} \tag{38}$$

The economic intuition behind Equation (38) is similar to the one we gave for Equation (32). However, Equation (38) is expressed in terms of the $n+1$ marginal tax rates $T_k'(Y_k(\mathbf{w}))$ associated to the one-dimensional schedules $y_k \mapsto T_k(Y_k(\mathbf{w}))$ and not in terms of the partial derivatives $\mathcal{T}_{y_k}(Y_1(\mathbf{w}), ..., Y_n(\mathbf{w}))$ of the overall n-dimensional tax schedule $(y_1, ..., y_n) \mapsto \mathcal{T}(y_1, ..., y_n)$. Thus, for individuals of type $\mathbf{w}$, a reform of the taxation of the i^{th} income induces a change $-R_i(Y_i(\mathbf{w}))$ in tax liability and a change $-R_i'(Y_i(\mathbf{w}))$ in the i^{th} 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 wealth responses $\frac{\partial Y_k}{\partial \rho} R_i(Y_i(\mathbf{w}))$ for all incomes $k \in \{0, ..., n\}$.

Behavorial responses then come into play: the change $R_i'(Y_i(\mathbf{w}))$ in the i^{th} marginal net-of-tax rate creates compensated responses $\frac{\partial Y_k}{\partial \tau_i} R_i'(Y_i(\mathbf{w}))$ for all incomes $k \in \{0, ..., n\}$. All these responses modify tax liability by a factor equal to the marginal tax rate $T_k'(Y_k(\mathbf{w}))$ and modify also prices through changes in the k^{th} input factor in the production process.

The latter channel is taken into account by the macro spillover sufficient statistic μ_k. Aggregating these effects for all types leads to Equation (38). Importantly, not only does Equation (38) take into account compensated and wealth responses of the i^{th} income, it also encompasses cross-base responses that are denoted by $\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}$ for $k \neq i$.

Effect of reforming the personal income tax schedule

We now investigate the effects of any reform of the personal income tax schedule $T_0(\cdot)$. This reform replaces the initial tax schedule by $\widetilde{\mathcal{T}}(\mathbf{y}, t) = \mathcal{T}(\mathbf{y}) - t\ R_0(\sum_{k=1}^{n} a_k(y_k))$, where $R_0(\cdot)$ is the direction of the tax reform. This reform modifies individual tax liability by:

$$\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_0(Y_0(\mathbf{w})). \tag{39}$$

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

$$\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -a_j'(Y_j(\mathbf{w}))\ R_j'(Y_0(\mathbf{w})) \tag{40}$$

In Equation (40), one can see that the marginal deduction rate that applies to the j^{th} income shows up when one reforms the personal income tax. This differs from (37), which was obtained from reforming a specific income tax schedule.

Now, according to (17), 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 marginal discount factor a_j'. Therefore, as shown in Appendix H, a compensated personal income tax reform generates responses equal to the weighted sum of the compensated responses of the i^{th} income to a change in the j^{th} marginal net-of-tax rate, the weights being the j^{th} marginal discount rates a_j':

$$\forall i \in \{0, ..., n\} \qquad \frac{\partial Y_i}{\partial \tau_0} = \sum_{j=1}^{n} a_j'(Y_j(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}. \tag{41}$$

Given these definitions, the effect of a personal income tax reform in the direction $R_0(\cdot)$ on the government's Lagrangian is also given by Equation (38) with $i = 0$, as shown in Appendix H. Equation (38) therefore summarizes the first-order effects, on the government's Lagrangian, of a reform of both the personal income tax and a specific income tax.

Optimal specific or personal income tax schedule

The tax schedule specific to the i^{th} income is optimal if its reform does not imply first-order effects on the Government's Lagrangian, whatever the direction $R_i(\cdot)$ of the tax perturbation and whatever the other tax schedules. This reasoning also applies to the optimal personal income ($i = 0$) tax profile. To obtain the optimal tax formulas either for the personal or any specific income, we then equalize (38) to zero. In preamble, to make this tax formula easy to implement on data, we define a set of sufficient statistics that one can substitute in it.

For any variable $Z(\mathbf{w})$ and for any $i = 0, ..., n$, we denote $\overline{Z(\mathbf{w})}|_{Y_i(\mathbf{w})=y_i}$ the mean of $Z(\mathbf{w})$ among types $\mathbf{w}$ for which $Y_i(\mathbf{w}) = y_i$. We denote $\varepsilon_i(y_i)$ the mean compensated elasticity of the i^{th} income with respect to its own marginal net-of-tax rate. This mean is calculated among $\mathbf{w}$-taxpayers who earn their i^{th} income equal to y_i. We formally define this elasticity as:

$$\varepsilon_i(y_i) \stackrel{\text{def}}{\equiv} \frac{1-T_i'(y_i)}{y_i} \overline{\frac{\partial Y_i}{\partial \tau_i}\Big|}_{Y_i(\mathbf{w})=y_i}. \tag{42}$$

We denote $\varepsilon_0(y_0)$ the mean compensated elasticity of personal income with respect to the personal marginal net-of-tax rate τ_0. This mean is calculated among $\mathbf{w}$-taxpayers for which $Y_0(\mathbf{w}) = y_0$. Mathematically, combining (35b) and (41) allows us to define this elasticity as:[21]

$$\varepsilon_0(y_0) = \frac{1-T_0'(y_0)}{y_0} \overline{\frac{\partial Y_0}{\partial \tau_0}\Big|}_{Y_0(\mathbf{w})=y_0} = \frac{1-T_0'(y_0)}{y_0} \sum_{1\le i,j\le n} \overline{a_i'(Y_i(\mathbf{w}))\, a_j'(Y_j(\mathbf{w}))\, \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\Big|}_{Y_0(\mathbf{w})=y_0}. \tag{43}$$

The compensated elasticity of the personal income tax with respect to its own net-of-marginal tax rate depends on all incomes compensated responses $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}$ to changes in all net-of-marginal tax rates τ_j for $i, j \in \{1, ..., n\}$, weighted by the net-of-marginal discount rates $a_i'(Y_i(\mathbf{w}))$ and $a_j'(Y_j(\mathbf{w}))$.

Proposition 4. *Under a mixed tax schedule, and for all $i \in \{0, ..., n\}$:*

i) A tax perturbation specific to the i^{th} income in the direction $R_i(\cdot)$ with a positive (negative) t combined with a lump-sum rebate is socially desirable if (38) is positive (negative).

ii) Given the other (arbitrary or optimal) tax schedules and deduction functions $a_k(\cdot)$, the optimal nonlinear tax schedule specific to the i^{th} income is provided by:

$$\frac{T_i'(y_i)+\mu_i}{1-T_i'(y_i)}\, \varepsilon_i(y_i)\, y_i\, h_i(y_i) + \sum_{0\le k\le n, k\ne i} \overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\Big|}_{Y_i(\mathbf{w})=y_i} h_i(y_i) \tag{44}$$
$$= \int_{z=y_i}^{\infty} \left\{ 1 - \overline{g(\mathbf{w})}\big|_{Y_i(\mathbf{w})=z} - \sum_{k=0}^{n} \overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\Big|}_{Y_i(\mathbf{w})=z} \right\} h_i(z) \mathrm{d}z.$$

iii) Given the other (arbitrary or optimal) tax schedules and deduction functions $a_k(\cdot)$, the optimal linear tax rate denoted t_i specific to the i^{th} income is provided by:

$$\frac{t_i+\mu_i}{1-t_i} \int_{\mathbf{w}\in W} \varepsilon_i^u(\mathbf{w})\, Y_i(\mathbf{w})\, f(\mathbf{w}) \mathrm{d}\mathbf{w} + \int_{\mathbf{w}\in W} \sum_{k=0, k\ne i}^{n} \left(T_k'(Y_k(\mathbf{w}))+\mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} f(\mathbf{w}) \mathrm{d}\mathbf{w}$$
$$= \int_{\mathbf{w}\in W} [1-g(\mathbf{w})]\, Y_i(\mathbf{w})\, f(\mathbf{w}) \mathrm{d}\mathbf{w}. \tag{45}$$

where ε_i^u denotes the uncompensated elasticity of the i^{th} income with respect to $1-t_i$, i.e.:

$$\varepsilon_i^u(\mathbf{w}) \stackrel{def}{\equiv} \frac{1-t_i}{Y_i(\mathbf{w})} \frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_i}.$$

[21] 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 $a_1 = ... = a_n = 0$.

The proof of (i) and (ii) can be found in Appendix G for $i = 1, ..., n$ and in Appendix H for $i = 0$. The proof of (iii) is in Appendix I. Equation (44) generalizes to an economy with many incomes, multidimensional types and general equilibrium effects, the optimal ABC tax formula derived by Diamond (1998) and Saez (2001) with a single income. Equation (44) relates optimal marginal tax rates to empirically estimable sufficient statistics which are behavioral responses, income density, macro spillover statistics and welfare weights.

To grasp the intuition behind each term of the above tax formula, one can heuristically derive it as in Saez (2001). To do so, consider the effects of a small increase in the marginal tax rate on the i^{th} income around income y_i and a uniform increase in tax liability for all i^{th} income above y_i.[22] Given the other tax schedules, the tax schedule specific to the i^{th} income is optimal if these reforms do not imply first-order effects on the Lagrangian. The left-hand side of Equation (44) describes the impact of the change in the marginal tax rate and its right-hand side details the effects due to the change in tax liability.

A rise in the i^{th} marginal tax rate around income y_i implies compensated responses $\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}$. First, there is a direct response of the i^{th} income, $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_i}$, which is proportional to the mean compensated elasticity ε_i of the i^{th} income with respect to its own marginal net-of-tax rate (as emphasized in Equation (42)). This response is encapsulated into the first term of Equation (44) left-hand side. On top of this response, which is already present in Saez (2001), prevail the (compensated) cross-base tax responses of all other tax bases $\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}$ for $k \in \{0, ..., n\} \setminus \{i\}$. These responses are in the second term of Equation (44) left-hand side. Another difference with the standard one-dimensional formula is that all these compensated responses have to be averaged across all taxpayers with the same i^{th} income y_i so that composition effects take place (Jacquet and Lehmann, 2021).[23] A third difference is that these compensated responses not only have a direct impact on the Lagrangian by modifying the tax revenue proportionally to marginal tax rates $T'_k(y_k)$, but also induce prices' changes in general equilibrium. All *micro* compensated responses produce prices' changes (see (1)) which imply responses of taxpayers to these price changes, and so on. The sufficient statistics that summarize these general equilibrium effects through prices' changes are given by μ_k which are equal to zero in Saez (2001).

A rise in the tax liability above income y_i implies mechanical gains in terms of tax revenue and mechanical welfare losses that are emphasized by the aggregation of $1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=z}$ for

[22] The effects are obviously symmetric when the tax marginal rate and tax liability are reduced.

[23] Saez (2001) conjectures his optimal tax formula can be extended to the case with multidimensional unobserved heterogeneity. This has been formally proved only recently (Hendren, 2020, Jacquet and Lehmann, 2021). In our framework with heterogeneous types of income, when one needs to take the mean of a variable, the latter is averaged across taxpayers who earn the i^{th} income at level y_i. This differs from the model with a single income and multidimensional types where the means are taking across sufficient statistics of agents who earn the same level of the unique income y.

all $z \geq y_i$ in the right-hand side of (44). It also creates wealth responses $\frac{\partial y_i(\mathbf{w})}{\partial \rho}$. Another key difference compared to the Mirrleesian framework is that these wealth responses not only have a direct impact on the Lagrangian by modifying the tax revenue proportionally to marginal tax rates $T'_k(y_k)$, they also create macro spillover effects which show up in the formula thanks to the sufficient statistics μ_k. An additional difference with the single-income, one-dimensional framework lies in in the averaging of the mechanical losses and responses to wealth change across all taxpayers with the same i^{th} income.

We now heuristically discuss the determinants of optimal marginal tax rates. First, the optimal marginal tax rate on the i^{th} income at income y_i is decreasing in the average of the welfare weights assigned to taxpayers earning an i^{th} income above y_i, since all these incomes are mechanically impacted by any change in the optimal marginal tax rate on the i^{th} income. It also depends on the i^{th} income distribution. Ceteris paribus, it decreases with the product of income and income density, since the larger $y_i h_i(y)$, the larger the impact of compensated responses. It also, ceteris paribus, increases with the number of taxpayers $1 - H_i(y_i))$ with i^{th} incomes larger than y_i since the larger this number, the larger the mechanical and income effects.

Second, the optimal marginal tax rate on the i^{th} income at income y_i is, ceteris paribus, increasing when the mean compensated elasticity ε_i decreases. The inverse elasticity rule remains valid. From Equations (43) and (44), this implies that the optimal marginal tax rate on *personal* income $T'_0(y_0)$ decreases when incomes which are the most responsive to tax reforms are withdrawn from the definition of the personal income tax base. For instance, if the most responsive tax base is capital income, then, the mean compensated elasticity of personal income ε_0 is lower with a separate tax on capital income than with a more comprehensive tax system. This leads to a more progressive personal income tax schedule with a more schedular tax system. This might explain why Scandinavian countries have implemented the dual tax system (Boadway, 2004, Sørensen, 2009) in the early nineties.

Third, our formula also highlights the role played by cross-base responses $\frac{\partial Y_k}{\partial \tau_i}$ for $k \neq i$. Consider a rise $\Delta T'_{y_i}$ in the i^{th} marginal tax rate around income level y_i. This induces compensated responses of each k^{th} income that is given by $\Delta Y_k = -\Delta T'_{y_i} \frac{\partial Y_k}{\partial \tau_i}$ (where the increase $\Delta T'_{y_i}$ corresponds to a reduction $\Delta \tau_i$ of the i^{th} marginal net-of-tax rate which explains the minus sign). Each compensated cross-base response, impacts the government's Lagrangian by $-(T'(Y_k) + \mu_k) \frac{\partial Y_k}{\partial \tau_i} \Delta T'_{y_i}$. Hence, whenever $T'(Y_k) + \mu_k > 0$, the less positive or the more negative is the cross-base response $\frac{\partial Y_k}{\partial \tau_i}$, i.e. the lower the reduction of the personal income tax basis due to ΔY_k, the less costly or the more beneficial is the response of the k^{th} income for the government. In this context, one can then recommend a higher i^{th} optimal marginal tax rate. In particular, lower income-shifting leads to more negative cross-base responses $\frac{\partial Y_k}{\partial \tau_i}$ and so to higher optimal marginal tax rate on the i^{th} income, provided that $T'(Y_k) + \mu_k > 0$. This leads Saez and Zucman (2019) to argue in favor of a comprehensive tax system.

Fourth, the macro spillover statistics μ_k magnify the compensated responses and the wealth responses. In particular, a larger macro spillover statistic on the i^{th} income μ_i tends to, ceteris paribus, reduce the i^{th} optimal marginal tax rate. To understand why, consider a rise in the i^{th} marginal tax rate around income y_i. This induces compensated responses that reduce the i^{th} income of taxpayers concerned by this tax reform. These responses imply a detrimental reduction in tax liability (whenever $T'(y_i) > 0$) in terms of tax revenue. Moreover, these compensated responses, by decreasing the i^{th} aggregate income $\mathcal{Y}_i$ in turn induce change in price that affects the government's Lagrangian. The larger the macro spillover statistics μ_i on the i^{th} income, the more detrimental are the consequences of these compensated responses through changes in prices, so the lower the i^{th} optimal marginal income tax rate. In particular, in an economy with capital income and labor income, the more positive is the macro spillover statistic on capital, the lower are the optimal marginal tax rates on capital income. Intuitively, the micro responses that increase aggregate capital income increase, in general equilibrium, the marginal productivity of labor hence its price. These "trickle down" effects reduce optimal capital tax rates.

When the tax schedule on the i^{th} income is restricted to be linear, with no restriction on the other tax schedules, similar intuitions apply, with the following particularities. First, under a linear tax schedule, wealth effects and compensated effects can be combined and substituted with the uncompensated responses, as can been verified using the Slutsky Equations (20d). Second, in the optimal linear tax formula (45), the means of sufficient statistics over the whole population appear instead of the means of sufficient statistics at a given income level. Last, as expected from the optimal linear tax formula (see e.g. Piketty and Saez (2013)), the means of welfare weights and uncompensated elasticities are income-weighted. Conversely, the mean of uncompensated cross-base responses $\frac{\partial Y_k^u}{\partial \tau_i}$ for $k \neq i$ are not income-weighted because these responses are expressed in terms of derivatives and not in terms of elasticities.

V.2 Toward a more or less schedular tax system

V.2.a How much of each type of income in the personal tax base?

Moving toward a more schedular (a more comprehensive) tax system with taxpayers having less income y_i which is part of their personal income is equivalent to increasing (decreasing) the discounting of the i^{th} income $a_i(y_i)$, as follows:

$$\widetilde{\mathcal{T}}(\mathbf{y},t) = T_0\left(\sum_{k=1}^{n} a_k(y_k) - t\, y_i\right) + \sum_{k=1}^{n} T_k(y_k) \tag{46}$$

with $t > 0$ ($t < 0$). The j^{th} marginal tax rate, for $j \neq i$, is now equal to:

$$\widetilde{\mathcal{T}}_{y_j}(\mathbf{y},t) = a_j'(y_j)\; T_0'\left(\sum_{k=1}^{n} a_k(y_k) - t\, y_i\right) + T_j'(y_j). \tag{47}$$

The i^{th} marginal tax rate is equal to:

$$\widetilde{\mathcal{T}}_{y_i}(\mathbf{y},t) = \left(a_i'(y_i) - t\right)\, T_0'\left(\sum_{k=1}^{n} a_k(y_k) - t\, y_i\right) + T_i'(y_i). \tag{48}$$

When the tax system becomes more schedular, thanks to an increase of the discounting of the i^{th} income $a_i(y_i)$, the personal tax base $y_0(\mathbf{w})$ is reduced by $Y_i \Delta t$. Proposition 5 describes the impact on the Lagrangian and states when this budget-balanced reform is socially desirable.

Proposition 5. *(i) Under a mixed tax schedule, a small reduction of the personal income tax base, described by* (48)*, modifies the government's Lagrangian as follows:*

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = & \int_{\mathbf{w}\in W}\Bigg\{\left[\left(g(\mathbf{w})-1\right)\, Y_i(\mathbf{w}) + \sum_{k=0}^{n}\left(T_k'(Y_k(\mathbf{w})) + \mu_k\right)\, \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}\right] T_0'(Y_0(\mathbf{w})) \\ & + \sum_{k=0}^{n}\left(T_k'(Y_k(\mathbf{w})) + \mu_k\right)\, \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w}))\Bigg\} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \end{aligned} \tag{49}$$

ii) A reform that consists in combining a discount of the i^{th} income from the taxable income according to (46) *(with $t > 0$) with a lump-sum transfer that makes the overall perturbation budget-balanced is socially desirable if Equation* (49) *is positive.*

The proof is in Appendix J. This result allows us to highlight arguments that have been ignored until now but need to be taken into account in the debate on schedular versus comprehensive tax systems. A reduction of the personal tax base automatically reduces the level of tax on personal income $T_0(\cdot)$ hence individual tax liability and modifies the marginal tax rate on personal income, since $T_0(\cdot)$ is nonlinear. The impact of a reduction of $y_0(\cdot)$ is twofold: there are effects conveyed by $T_0'(\cdot)$ in the first line of Equation (49) and other effects are propagated with $T_0''(\cdot)$ in the second line of Equation (49). These two channels had hitherto not been studied in the literature.

First, the amount of income y_i which is withdrawn from the personal tax base is not taxed anymore through $T_0(\cdot)$. The reduction in the amount of tax paid is equal to $T_0'(Y_0(\mathbf{w}))Y_i(\mathbf{w})\Delta t$. This reduction of tax liability generates a mechanical loss in tax revenue and a mechanical welfare gain, $\int_{\mathbf{w}\in W}[g(\mathbf{w})-1]T_0'(Y_0(\mathbf{w}))Y_i(\mathbf{w})\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}$ that are in the first line of Equation (49).

This reduction in tax liability also creates wealth responses from all income sources. Indeed this reduction in tax liability is equivalent to a lump-sum transfer to every worker who earns the source of income y_i. These wealth responses, that occur for each source of income, modify tax revenue and welfare (due to general equilibrium effects) as follows:

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} Y_i(\mathbf{w})\ T_0'(Y_0(\mathbf{w}))\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}, \tag{50}$$

where price general equilibrium effects are encapsulated into the sufficient statistics μ_k. Moreover, the withdrawal of some income y_i from the personal tax base modifies the marginal tax rate of the i^{th} income $\mathcal{T}_{y_i}(\cdot)$ since the latter not only depends on $T_i'(\cdot)$ (which is not modified) but also on $T_0'(\cdot)$, as emphasized in Equation (48). The i^{th} marginal tax rate is reduced by $T_0'(Y_0(\mathbf{w}))\ \Delta t$. This reduction in the i^{th} marginal tax rate creates (cross-base and within-base) compensated responses from all sources of income. These responses increase tax revenue and also welfare (due to general equilibrium effects in μ_k) by

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\ T_0'(Y_0(\mathbf{w}))\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}. \tag{51}$$

Using the Slutsky equation (20d), the impact, on welfare and government tax revenue, of these wealth and compensated responses is equivalent to the impact of uncompensated responses on welfare and tax revenue, i.e.

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}\ T_0'(Y_0(\mathbf{w}))\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}, \tag{52}$$

that one can see in the first line of Equation (49).

Second, because of the nonlinearity of the personal income tax schedule, the j^{th} marginal tax rate $\mathcal{T}_{y_j}(\cdot)$ is also modified by $a_j'(Y_i(\mathbf{w}))Y_i(\mathbf{w})T_0''(Y_0(\mathbf{w}))\Delta t$ (from (47)) where the curvature of the personal income tax matters as emphasized by $T_0''(Y_0(\mathbf{w}))$. This change in the marginal personal income tax rate $T_0'(Y_0(\mathbf{w}))$ creates compensated responses from the i^{th} income that modify tax revenue and welfare by

$$\int_{\mathbf{w}\in W} \left(T_i'(Y_i(\mathbf{w})) + \mu_i\right) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}.$$

where (41) has been used. On top of this, the modification of $T_0'(Y_0(\mathbf{w}))$ also modify all other marginal tax rates, $\forall j \in \{1, ..., n\}$ $(j \neq i)$, by

$$a_j'(Y_k(\mathbf{w}))\ Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))\Delta t.$$

Therefore, all compensated responses from other sources of income modify tax revenue and welfare by

$$\int_{\mathbf{w}\in W} \sum_{k\neq i, k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w}))\Delta t f(\mathbf{w})\mathrm{d}\mathbf{w}.$$

In summary, the compensated responses from the i^{th} income and the compensated responses from other incomes are equal to:

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w})\, T_0''(Y_0(\mathbf{w}))\Delta t f(\mathbf{w}) \mathrm{d}\mathbf{w}, \tag{53}$$

an expression that one can find in the last line of (49).

With a U-shape personal income marginal tax schedule, $T_0''(Y_0(\mathbf{w}))$ is negative for relatively low personal incomes $Y_0(\mathbf{w})$ and positive for relatively high $Y_0(\mathbf{w})$. When one withdraws some income y_i from the personal income tax base, it therefore increases the marginal personal income tax rates of low earners of income y_0 (whose $T_0''(Y_0(\mathbf{w}))<0$) and decreases the ones of richer earners (whose $T_0''(Y_0(\mathbf{w}))>0$). The deadweight losses associated to compensated responses due to the change of $T_0'(Y_0(\mathbf{w}))$ are therefore transferred from high to low personal income earners.

Adding Equations (50), (52) and (53), we obtain the effect on the Lagrangian described in (49).

VI Numerical Illustrations

In this Section, we numerically implement the optimal mixed tax formula, so as to verify, on real data, whether it is desirable to tax distinct sources of income comprehensively or schedularly or whether a compromise between both tax systems is preferable as one may expect from our theory. We emphasize the relative magnitude of the mechanical and behavioral effects, highlighted in Equation (49), that matter when determining the optimal size of the personal income tax base. We illustrate that it is the impact in the level of the marginal tax rates $T_0'(\cdot)$ (second line of (49)) that offsets the impact of a change in the levels of the personal income tax $T_0(\cdot)$ (first line of (49)).

For our numerical exercises, we assume two production factors and group income from different sources we observe, in the administrative tax return data for France, into two categories: (1) labor income, y_1 and capital income, y_2. For simplicity, labor income includes all incomes we have in our data set except incomes from capital. In labor income, there are all wages earned in France (and abroad, whenever applicable) by French taxpayers, two-third of each income self-employed individuals report from their business activity[24] as well as unemployment benefits and copyrights. Capital income comprises interests (from bonds and other sources), dividends, financial gains and one-third of self-employed reported income.

[24] In the French Tax records, income of self-employed are declared either as *Benefices Industriels et Commerciaux* (BIC), *Benefices Non Commerciaux* (BNC) or Benefices Agricoles (BA). We choose 2/3 to approximate the share of these different sources of income that falls under the labor income tax since it is the usual and relatively stable national income share for labor income vs capital income.

Since in France, labor income is included, after possible deductions, into the personal income tax base and since there is no other specific tax schedule on labor income, y_1 is taxed according to the nonlinear personal income tax schedule $T_0(\cdot)$ and $T_1(y_1) \equiv 0$. Since a dual tax prevails in France, we also consider a flat tax rate on capital income, $T_2(y_2) = t_2\, y_2$. The mixed tax schedule (15) becomes:

$$\mathcal{T}(y_1, y_2) = T_0(y_1 + a_2\, y_2) + t_2\, y_2 \qquad a_2 \in [0, 1]. \tag{54}$$

On top of the nonlinear personal income tax schedule $T_0(\cdot)$, the government has two instruments a_2 and t_2. This simple framework is sufficiently rich to include, as specific cases, the comprehensive tax system when $a_2 = 1$ and $t_2 = 0$ and the schedular tax system when $a_2 = 0$. The schedular tax system is the dual one since $T_2(y_2) = t_2\, y_2$.

VI.1 Specification and calibration

We calibrate our model on French *Enquête Revenus Fiscaux Sociaux* (ERFS)) data available through the Quetelet network.[25] This dataset merges a part of the French Labor Force survey with a set of variables extracted from the respondents' tax records. We build our capital income variable by summing the different sources of financial income included in the personal income tax base, financial incomes taxed at a flat rate ("Prélèvement Forfaitaire Libératoire", which applies essentially to life insurance and specific savings contracts known as PEA) and 1/3 of the distinct sources of income from the self-employed. Importantly, we do not observe capital gains and losses in ERFS. Moreover, we choose to exclude pensions and social transfers since they are exogenous income levels. Since we focus on a two factor production function, we also exclude rents and real estate income.

The sample consists of 27,804 tax units with positive labor and capital income. Average labor income equals € 36,578 (with the median at € 30,398) and average capital income equals € 3,017 (with the median at € 310).

Utility is assumed quasilinear, which is standard since Diamond (1998), and with a constant direct elasticity of (gross) income y_i with respect to its net-of-marginal tax rate e_i, as in Diamond (1998) and (2001):

$$\mathscr{U}(c, x_1, x_2; w_1, w_2) = c - \frac{\varepsilon_1}{1+e_1} x_1^{\frac{1+e_1}{e_1}}\, w_1^{-\frac{1}{e_1}} - \frac{e_2}{1+e_2} x_2^{\frac{1+e_2}{e_2}}\, w_2^{-\frac{1}{e_2}} \tag{55}$$

We take the estimates of Lefebvre et al. (2019), who rely on an extended version of our dataset (POTE), to calibrate the direct elasticity of labor income $e_1 = 0.10$ and the direct elasticity of

[25] We have also calibrated our model with U.S. Current Population Survey (CPS) data and we have found similar results to the ones we describe here.

capital income $e_2 = 0.65$. With this specification, w_1 and w_2 stand respectively for labor and capital income in the no-tax economy.

Lefebvre et al. (2019) do not find any income-shifting[26] so that we assume away cross-base responses in our baseline scenario.

From the ERFS data and an approximation of the actual tax schedule, we recover for each of observation, its labor w_1 and capital w_2 abilities. The income density is estimated using a biweight kernel with a bandwidth of € $89,028$.[27]

The production function is a CES:

$$\mathcal{F}(\mathcal{X}_1, \mathcal{X}_2) = \left[A_1\, \mathcal{X}_1^{1-\gamma} + A_2 \mathcal{X}_2^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \tag{56}$$

where A_1 and A_2 are the scale parameters of inputs and $1/\gamma$ is the elasticity of substitution. We take $\gamma = 0$ as the baseline value and conduct sensitivity analysis with respect to γ. Without loss of generality, we normalize the price p_1 and p_2 to be equal to 1 in the no-tax economy in order to pin down A_1 and A_2.

We consider a maximin social objective which is equivalent to maximizing tax revenue.

VI.2 The optimal system

In our numerical simulations, we determine the optima under three distinct tax regimes: the comprehensive tax system by setting $a_2 = 1$ and $t_2 = 0$ in (54), the (schedular) dual tax system by setting $a_2 = 0$ and the mixed tax system by letting a_2 and t_2 being optimized.

We first present the results in the benchmark case, i.e. with $\varepsilon_1 = 0.10$, $\varepsilon_2 = 0.65$, $\gamma = 0$ and no income-shifting. In Figure 1, we compare the optimal marginal tax rates on personal income under the comprehensive tax system (dashed red line), the dual tax system (alternate blue lines) and the mixed tax system (black lines). Whatever the tax regime, the optimal marginal tax profile is U-shaped. The optimal marginal tax rate decreases with income from zero to about € $100,000$ and then increases with income.

We directly see that the optimal mixed tax system is neither comprehensive, nor dual since the optimal mixed tax does coincide neither with the dual tax schedule nor with the mixed tax schedule. In our simulations, the mean elasticity ε_0 of personal income y_0 is smaller under the dual tax regime (See (43)) than under the comprehensive tax regime, so the optimal marginal tax rates are higher (See (44)) under the dual tax. The difference is especially important for

[26] Lefebvre et al. (2019) find this result using the suppression, in 2013, by the newly elected Hollande government, of the optional flat tax for dividends that was available, forcing all dividends to be taxed under the progressive tax schedule.

[27] The biweight kernel $K(x) = (15/16)(1 - x^2)^2$ eases the computation (since it is a 4th degree polynomial) and provides differentiable estimated densities, since $K(\cdot)$ is differentiable with zero derivatives at $x = -1, 0, 1$).

income levels above € 100, 000 since the share of capital income in total income is larger for these levels of income.

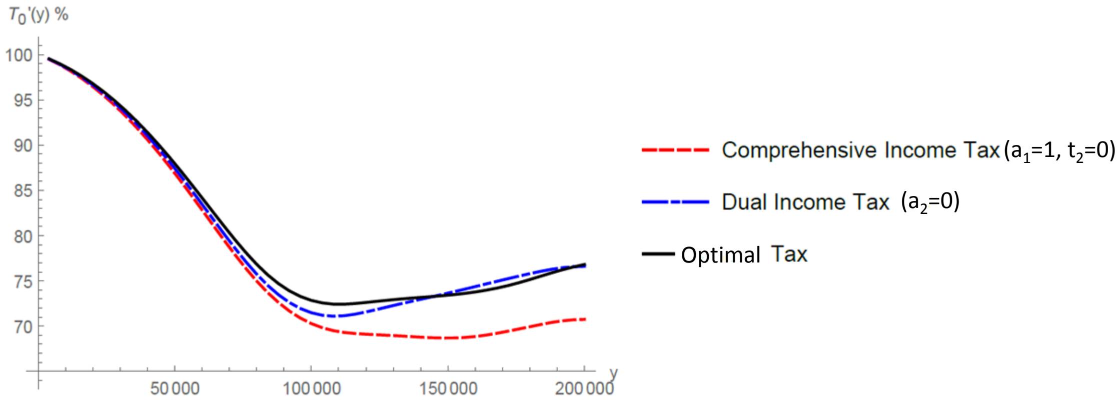

Figure 1: Optimal marginal tax rates $T_0'(y_0)$ under the comprehensive income tax regime (where $a = 1$, $t_2 = 0$), under the dual tax regime (where $a = 0$, optimal $t_2 = 60.6\%$), and under the mixed tax regime (optimal $a = 1$, optimal $t_2 = -22.8\%$)

The capital income net-of-discount rate, a_2, that gives the proportion of capital income included into the personal income tax base, is constrained under both the comprehensive and the dual tax regimes. When a_2 increases, the cross-base responses imply a reduction in capital income. To countervail this reduction, the optimal tax rate on capital income decreases with a_2. In our simulations, when a_2 increases from 0, which is required under the dual tax system, to its optimal value $a_2 = 1$ in the mixed tax system, the linear tax rate t_2 decreases from $t_2 = 60.6\%$[28] to a negative value of -22.8%. This result is rather unexpected yet intuitive.

With a comprehensive tax, the tax base consists in a very unelastic tax base (labor income y_1) and a very elastic tax base (net-of-discount rate capital income a_2y_2). In order to keep a large enough personal income tax base y_0, the government can only impose relatively low marginal tax rates T_0' at the top of the distribution of y_0 (i.e., above € 100, 000, a point from which the share of capital income in total income increases with total income),[29] otherwise the loss in tax revenue due to the high elasticity of y_2 may overcome the gain obtained by increasing the marginal tax on the inelastic income y_1. In contrast, when all instruments are unconstrained, the marginal tax rates T_0' bear on the full capital income y_2 as the optimal value of a_2 is 1. In

[28]This optimal Laffer rate with a dual tax is obtained from (45) which, in the baseline scenario, is $t_2 = 1/(1 - \varepsilon_2)$.

[29]When the total income is lower than € 100, 000, the average share of capital income in total income is around 7.4% whereas it is around 14.4% when the total income is higher than € 100, 000.

this situation, the government raises the marginal tax rates T_0' to a high level at the top of the distribution, which extracts a high tax revenue from labor income and would also extract a high tax revenue from capital income if it were less elastic. To counterbalance the behavioral responses that reduce y_2, the government offers a transfer toward capital earners $t_2 y_2 < 0$ with $t_2 = -22.6\%$. This transfer is high, but does not offset the gains derived from the raise in T_0' when moving from the comprehensive tax system to the mixed tax system.

Let us first describe why all capital income enters the personal income tax base, i.e. $a_2 = 1$, at the optimum. We know from (49) that increasing a_2 implies two main and ambiguous effects on the Lagrangian.

First, increasing a_2, i.e. increasing personal income $y_0 = y_1 + a_2 y_2$, automatically increases the level of tax on personal income $T_0(\cdot)$. This generates a mechanical gain in tax revenue. This increase in tax liability also creates wealth responses from all sources of income which reduce tax revenue. Moreover, the inclusion of some extra capital income in the personal income tax base modifies the marginal tax rate of capital income, the latter depending on $T_0'(\cdot)$, as emphasized in $\mathcal{T}_{y_2} = a_2 \; T_0'(y_0) + t_2$. The increase in a_2 creates (within-base) compensated responses from capital incomes that reduces tax revenue. In our simulations, the impact of all these effects on the Lagrangian, that are identified by the first line of (49), are positive.

Second, because of the nonlinearity of $T_0(\cdot)$, increasing a_2 also modifies the marginal tax schedule on capital income $\mathcal{T}_{y_2}(y_1, y_2) = a_2 \; T_0'(y_1 + a_2 \; y_2) + t_2$ because $\mathcal{T}_{y_2 y_2}(\cdot) = a_2^2 \; T_0''(.)$. Due to the U-shape of the tax function (see Figure 1), increasing a_2 increases the marginal tax rates $T_0'(y_0)$ for taxpayers with personal income above € $100,000$. This rise in their marginal personal income tax rates increases their marginal tax rate on capital income $\mathcal{T}_{y_2}(\cdot)$. This induces compensated responses from rich earners who reduce their capital income hence, their personal tax base. Moreover, due to the U-shaped $T_0(\cdot)$, the same increase of a_2 reduces the marginal tax rate on capital income $\mathcal{T}_{y_2}(\cdot)$ of taxpayers with personal income below € $100,000$ which, in turn, increases their capital income hence, their personal tax base. The second line of Equation (49) add up these effects. In our simulations, This second line is negative and the sum of both lines of (49) is also negative. Therefore the effect on the Lagrangian of increasing a_2 is negative in all our simulations.

At the optium, this negative impact of $a_2 = 1$ is counterbalanced by a negative tax on capital income $t_2 < 0$, in all our simulations. The tax authority taxes heavily capital income with $a_2 = 1$ and relatively large marginal tax rates on high personal income levels where the proportion of capital income is larger and, at the same time, prevents an erosion of capital income with subsidies $t_2 y_2$ that target capital earners. This optimal negative t_2 allows to more than offset the loss in welfare and government tax revenue implied by having increased a_2 until $a_2 = 1$.

	Calibration	Dual System	Optimal system	
		t_2	a_2	t_2
(1)	Baseline scenario	60.6%	100%	– 22.8%
(2)	$\gamma = 0.75$	58.5%	100%	– 25.2%
(3)	With income shifting	62.3%	100%	– 21.4%

Table 1: Optimal t_2 under the optimal dual tax and optimal a_2 and t_2 in the optimal tax system, w/o and w/ general equilibrium effects ($\gamma = 0.75$) and w/o and w/ income shifting.

VI.3 Sensitivity analyses

General equilibrium effects

For our sensitivity analyses, we first depart from the baseline calibration by assuming that labor and capital are imperfect substitutes so that general equilibrium effects can occur. Instead of $\gamma = 0$ in (56), we experiment the case where $\gamma = 0.75$, i.e. a capital/labor elasticity of 1.33. The personal income tax schedules obtained under our three tax regimes are very close to those displayed in Figure 1 for the baseline case. In contrast, t_2, with the dual tax, is reduced by 2.1 percentage points (see Row (2) of Table 1). In the optimal scenario, the transfer t_2 (taken in absolute value) increases by 2.3 percentage points (see Row (2) of Table 1). This is due to the influence of macro price spillover statistics on capital income which do not exist in the baseline scenario and are now equal to 5.4% under the Dual Tax system and to 6.5% under the optimal tax system.[30] Compared to the baseline scenario, the introduction of general equilibrium effects makes socially desirable to decrease the tax on capital income in order to boost capital income, thereby the marginal product of labor, hence, labor supply and eventually labor income.

Cross-base responses

Second, we emphasize the impact of cross-base responses in making income-shifting possible. We use the linear technology (in (2)) of our benchmark and add a quadratic cost of income-shifting to the utility function (55). Using (9), the utility function is:

$$\mathscr{U}(c, y_1, y_2; w_1, w_2) \stackrel{\text{def}}{\equiv} \max_{x_1, x_2, \sigma} \quad c - \frac{e_1}{1+e_1} x_1^{\frac{1+e_1}{e_1}} w_1^{-\frac{1}{e_1}} - \frac{e_2}{1+\varepsilon_2} x_2^{\frac{1+e_2}{e_2}} w_2^{-\frac{1}{e_2}} - \frac{\sigma^2}{2\,\Gamma(w_1, w_2)}$$
$$s.t \quad : \quad y_1 = x_1 + \sigma \qquad and \qquad y_2 = x_2 - \sigma,$$

which implies that the amount of shifted income verifies $\sigma = \Gamma(w_1, w_2)(\mathcal{T}_{y_2} - \mathcal{T}_{y_1})$. Row (3) of Table 1 corresponds to an economy where we calibrate the scale parameter Γ to 10% of

[30]Macro price spillover statistics on labor are equal to -0.4% under the dual tax system and -0.5% under the optimal tax system.

the minimum between labor income and capital income, in the no-tax scenario. When ones allows for income-shifting, there is little effect on marginal tax rates that apply on personal income, in the optimal case. In contrast, under the dual tax (where only labor income enters the personal income tax base), the capital income tax rate is relatively lower than the marginal tax rates on personal income –i.e. labor income–. Due to income-shifting, the Laffer tax rate on capital income (from Equation (45)) increases by 1.7 percentage points to give more incentives to supply labor. In the fully optimal system, for the same reason, the subsidy towards capital is reduced by 1.4 percentage points (Row 3 of Table 1).

VII Conclusion

[to be completed]

References

Albanesi, Stefania and Christopher Sleet, "Dynamic Optimal Taxation with Private Information," *The Review of Economic Studies*, 2006, *73* (1), 1–30.

Ales, Laurence and Christopher Sleet, "Taxing Top CEO Incomes," *American Economic Review*, November 2016, *106* (11), 3331–66.

_ , Antonio Andrés Bellofatto, and Jessie Jiaxu Wang, "Taxing Atlas: Executive compensation, firm size, and their impact on optimal top income tax rates," *Review of Economic Dynamics*, 2017, *26*, 62–90.

Alesina, Alberto, Andrea Ichino, and Loukas Karabarbounis, "Gender-Based Taxation and the Division of Family Chores," *American Economic Journal: Economic Policy*, 2011, *3* (2), 1–40.

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.

Bartels, Charlotte, "Top Incomes in Germany, 1871?2014," *The Journal of Economic History*, 2019, *79* (3), 669–707.

Bastani, S. and Daniel Waldenström, "How should capital be taxed," *Journal of Economic Survey*, 2020.

Baumol, William J. and David F. Bradford, "Optimal Departures From Marginal Cost Pricing," *The American Economic Review*, 1970, *60* (3), 265–283.

Benoteau, Isabell and Olivier Meslin, "Les prélèvements obligatoires sur le capital des ménages : comparaisons internationales," 2017.

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

_ , *From Optimal Tax Theory to Tax Policy, Reprospective and Prospective Views*, CES, The MIT Press Cambridgen Massachusetts, London, England, 2012.

Burns, Lee and Rick Krever, "Individual income tax," in Victor Thuronyi, ed., *Tax Law Design and Drafting, Volume 2*, International Monetary Fund, 1998, chapter 14, pp. 495 – 563.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet, "Capital income taxation when inherited wealth is not observable," *Journal of Public Economics*, 2003, *87* (11), 2475–2490.

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

_ **and Johannes Spinnewijn**, "Capital Income Taxes with Heterogeneous Discount Rates," *American Economic Journal: Economic Policy*, November 2011, *3* (4), 52–76.

Farhi, Emmanuel and Iván Werning, "Progressive Estate Taxation*," *The Quarterly Journal of Economics*, 2010, *125* (2), 635–673.

Fullerton, Don and Gilbert Metcalf, "Tax incidence," in A. J. Auerbach and M. Feldstein, eds., *Handbook of Public Economics*, 1 ed., Vol. 4, Elsevier, 2002, chapter 26, pp. 1787–1872.

Gahvari, F. and L.M. Michelleto, "Capital income taxation and the Atkinson-Stiglitz theorem," *Economics Letters*, 2016, *147*, 86–89.

Gerritsen, Aart, Bas Jacobs, Alexandra V. Rusu, and Kevein Spiritus, "Optimal Taxation of Capital Income with Heterogeneous Rates of Return," CESifo Working Paper, Munich: CESifo 6599 2020.

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

_ **, Maxim Troshkin, Aleh Tsyvinski, and Matthew Weinzierl**, "Preference heterogeneity and optimal capital income taxation," *Journal of Public Economics*, 2013, *97*, 160–175.

_ **,** _ **, and** _ , "Redistribution and Social Insurance," *American Economic Review*, February 2016, *106* (2), 359–86.

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan, "Differential Taxation and Occupational Choice," *The Review of Economic Studies*, 2017, *85* (1), 511–557.

Hendren, Nathaniel, "Measuring Economic Efficiency Using Inverse-Optimum Weights," *Journal of Public Economics*, 2020, (187).

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

Jacobs, Bas, "The marginal cost of public funds is one at the optimal tax system," *International Tax and Public Finance*, 2018, *25* (4), 883–912.

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

Jacquet, Laurence M. and Etienne Lehmann, "Optimal income taxation with composition effects," *Journal of the European Economic Association*, 2021, *19* (2).

Kaplow, Louis, *The Theory of Taxation and Public Economics*, Princeton University Press, 2008.

Ketterle, J., *Die Einkommensteuer in Deutschland. Modernisierung und Anpassung einer direkten Steuer von 1890.91 bis 1920*, Botermann und Botermann Verlag Koln, 1994.

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez, "The Optimal Income Taxation of Couples as a Multi-Dimensional Screening Problem," CESifo Working Paper 2092, CESifo 2007.

— , — , **and** — , "The Optimal Income Taxation of Couples," *Econometrica*, 2009, *77* (2), 537–560.

Konishi, Hideo, "A Pareto-improving commodity tax reform under a smooth non-linear income tax," *Journal of Public Economics*, 1995, *56* (1), 413–446.

Kotlikoff, Laurence J. and Lawrence H. Summers, "Chapter 16 Tax incidence," in "Handbook of Public Economics," Vol. 2 of Handbook of Public Economics, Elsevier, 1987, pp. 1043–1092.

Kristjánsson, A.S., "Optimal taxation with endogenous return to capital," Working paper Oslo University 06 2016.

Kroft, Kory, Kavan Kucko, Etienne Lehmann, and Johannes Schmieder, "Optimal Income Taxation with Unemployment and Wage Responses: A Sufficient Statistics Approach," *American Economic Journal: Economic Policy*, 2020, *12* (1), 254–292.

Laroque, Guy, "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," CRED Working paper 2020-09 2019.

Lehmann, Etienne, Renes Sander, Kevin Spiritus, and Floris T. Zoutman, "Optimal Taxation with Multiple Incomes and Types," Mimeo, Department of Business and Management Science, Norwegian School of Economics 2020.

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

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

— , "The theory of optimal taxation," in K. J. Arrow and M.D. Intriligator, eds., *Handbook of Mathematical Economics*, Vol. 3, Elsevier, 1986, chapter 24, pp. 1197–1249.

Ordower, Henry, "Schedularity in U.S. Income Taxation and its Effect on Tax Distribution," *Northwestern University Law Review*, 2014, *108* (3), 905–24.

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

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

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

Sachs, Dominik, Aleh Tsyvinski, and Nicolas Werquin, "Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium," *Econometrica*, 2020, *88* (2), 469–493.

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

— , "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 2002, *83* (2), 217–230.

— **and Gabriel Zucman**, *The Triumph of Injustice: How the Rich Dodge Taxes and How to Make Them Pay*, WW Norton, 2019.

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

Sandmo, Agnar, "Redistribution and the marginal cost of public funds," *Journal of Public Economics*, 1998, *70* (3), 365–382.

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

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

Sørensen, Peter Birch, "Dual Income Taxes: A Nordic Tax System," EPRU Working Paper Series 2009-10, Economic Policy Research Unit (EPRU), University of Copenhagen. Department of Economics 2009.

A Convexity of the Indifference Set

Let $\mathscr{C}(\cdot, \mathbf{x}; \mathbf{w})$ denote the reciprocal of $\mathscr{U}(\cdot, \mathbf{x}; \mathbf{w})$. Tax payers of type $\mathbf{w}$ making actions $\mathbf{x}$ should get consumption $c = \mathscr{C}(u, \mathbf{x}; \mathbf{w})$ to enjoy utility $u = \mathscr{U}(c, \mathbf{x}; \mathbf{w})$. Using (3), we obtain:

$$\mathscr{C}_u(u, \mathbf{x}; \mathbf{w}) = \frac{1}{\mathscr{U}_c(\mathscr{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w})} \qquad \mathscr{C}_{x_i}(u, \mathbf{x}; \mathbf{w}) = \mathcal{S}^i(\mathscr{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}) \tag{57}$$

For each type $\mathbf{w} \in W$ and each utility level u, we assume the indifference sets: $\mathbf{y} \mapsto \mathscr{C}(u, \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w})$ to be strictly convex. The i^{th} partial derivative of $\mathbf{y} \mapsto \mathscr{C}(u, \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w})$ being $\dfrac{\mathcal{S}^i(\mathscr{C}(u, \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w})}{p_i}$, the Hessian is matrix

$$\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j}\right]_{i,j} = \left[-\frac{\mathscr{U}_{x_i x_j} + \mathcal{S}^j \mathscr{U}_{c,x_i} + \mathcal{S}^i \mathscr{U}_{cx_j} + \mathcal{S}^i \mathcal{S}^j \mathscr{U}_{cc}}{p_i\, p_j\, \mathscr{U}_c}\right]_{i,j}$$

which is symmetric. Finally, the latter matrix is obviously positive definite if and only if matrix $\left[\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j\right]_{i,j}$ is positive definite as well.

The first-order condition of (5) is given by:

$$0 = (1 - \mathcal{T}_{y_i}(\mathbf{y}))\, \mathscr{U}_c\left(\sum_{k=1}^n y_k - \mathcal{T}(\mathbf{y}), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right) + \frac{1}{p_i} \mathscr{U}_{x_i}\left(\sum_{k=1}^n y_k - \mathcal{T}(\mathbf{y}), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right)$$

Therefore, using (6), the matrix of the second-order condition is:

$$\left[\frac{\mathscr{U}_{x_i x_j} + \mathcal{S}^j \mathscr{U}_{cx_i} + \mathcal{S}^i \mathscr{U}_{cx_j} + \mathcal{S}^i \mathcal{S}^j \mathscr{U}_{cc}}{p_i\, p_j} - \mathscr{U}_c \mathcal{T}_{y_i y_j}\right]_{i,j} = -\mathscr{U}_c \left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}$$

Hence, for taxpayers of type $\mathbf{w}$, the second-order condition holds strictly if and only if the matrix $\left[\dfrac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}$ is positive definite, i.e. if and only if the indifference set $\mathbf{y} \mapsto \mathscr{C}(U(\mathbf{w}), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w})$ is strictly more convex than the budget set $\mathbf{y} \mapsto \sum_{k=1}^n y_k - \mathcal{T}(\mathbf{y})$ at $\mathbf{y} = \mathbf{Y}(\mathbf{w})$.

B Proof of Proposition 1

The proof contains two steps. Under the assumptions of Proposition 1, we first characterize the separate income tax system $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y}) = \sum_{i=1}^{n} T_i(y_i)$ that is *necessary* to decentralize the allocation $w \mapsto (C(w), Y_1(w), ..., Y_n(w))$. Second, we proof that this tax schedule is *sufficient* to decentralize the optimal allocation $w \mapsto (C(w), Y_1(w), ..., Y_n(w))$.

Under the assumptions of Proposition 1, for each $i \in \{1, ..., n\}$, the function $Y_i : w \mapsto Y_i(w)$ is invertible with a reciprocal denoted Y_i^{-1} and defined on $[Y_i(\underline{w}), Y_i(\overline{w})]$. Under quasilinear and additively separable utility function (18), the i^{th} marginal rate of substitution defined in (3) simplifies to $\mathcal{S}^i(c, \mathbf{x}; w) = v^i_{x_i}(x_i, w)$. Using the first-order condition (6) on each income, we can recover for each type w and each $i \in \{1, ..., n\}$, the i^{th} marginal tax rate from the i^{th} marginal rate of substitution. We have:

$$T_i'(y_i) = 1 - \frac{1}{p_i} v^i_{x_i}\left(\frac{y_i}{p_i}; Y_i^{-1}(y_i)\right) \tag{58}$$

To determine the separate tax schedule that decentralizes the optimal allocation, one simply needs to integrate (58). Let $w^\star$ be a given skill level. 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:

$$\mathcal{T}(\mathbf{y}) = \left(\sum_{i=1}^{n} Y_i(w^\star)\right) - C(w^\star) + \sum_{i=1}^{n} T_i(y_i) \tag{59}$$

$$\text{where :} \quad T_i(y_i) = \begin{cases} \displaystyle\int_{Y_i(w^\star)}^{y_i} \left[1 - \frac{1}{p_i} v^i_{x_i}\left(\frac{z}{p_i}; Y_i^{-1}(z)\right)\right] \mathrm{d}z & \text{if} \quad y_i \in [Y_i(\underline{w}), Y_i(\overline{w})] \\ +\infty & \text{if} \quad y_i \notin [Y_i(\underline{w}), Y_i(\overline{w})] \end{cases}$$

This tax schedule assigns to taxpayers earning $(y_1, ..., y_n) = (Y_1(w^\star), ..., Y_n(\mathbf{w}^\star))$ a level of tax liability equal to $\sum_{i=1}^{n} Y_i(w^\star) - C(w^\star)$, which corresponds to the tax intended for $w^\star$-taxpayers. For all other income levels $\mathbf{y} = (y_1, ..., y_n)$ that are reached by the optimal allocation to decentralize, (i.e. for which $y_i \in [Y_i(\underline{w}), Y_i(\overline{w})]$) , the tax liability is computed by integrating for each type i of income the marginal tax rate in (58) between $Y_i(w^\star)$ and y_i. Otherwise, the tax liability is infinite.

We now show that the separate tax schedule (59) is sufficient to decentralize the allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$. As (59) is separate and preferences are additively separable, the $n-$dimensional program (5) of w-individuals can be simplified into the following n one-dimensional programs:

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

Whenever $y_i \in [Y_i(\underline{w}), Y_i(\overline{w})]$, we get from (59) that:

$$y_i - T_i(y_i) = Y_i(w^\star) + \frac{1}{p_i} \int_{Y_i(w^\star)}^{y_i} v^i_{x_i}\left(\frac{z}{p_i}; Y_i^{-1}(z)\right) \mathrm{d}z.$$

So, we have:

$$y_i - T_i(y_i) - v^i\left(\frac{y_i}{p_i}; w\right) = Y_i(w^\star) - v^i\left(\frac{Y_i(w^\star)}{p_i}; w\right) + \frac{1}{p_i} \int_{Y_i(w^\star)}^{y_i} \left[v^i_{x_i}\left(\frac{z}{p_i}; Y_i^{-1}(z)\right) - v^i_{x_i}\left(\frac{z}{p_i}; w\right)\right] \mathrm{d}z.$$

The derivative of the latter expression with respect to y_i is:

$$\frac{\partial\left(y_i - T_i(y_i) - v^i\left(\frac{y_i}{p_i}; w\right)\right)}{\partial y_i} = \frac{1}{p_i}\left[v^i_{x_i}\left(\frac{y_i}{p_i}; Y_i^{-1}(y_i)\right) - v^i_{x_i}\left(\frac{y_i}{p_i}; w\right)\right].$$

Since $w \mapsto Y_i(w)$ is strictly increasing and $v^i_{x_i,w} < 0$, this derivative is nil for $y_i = Y^i(w)$, positive for $y_i < Y^i(w)$ and negative for $y_i > Y^i(w)$. Hence, under the tax schedule defined in (59), a $\mathbf{w}$-taxpayer chooses $y_i = Y_i(\mathbf{w})$ which ends the proof of Proposition (1).

C Proof of Proposition 2

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 taxpayer makes the same decision and gets the same utility under the initial tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ and under the comprehensive tax schedule $\mathbf{y} \mapsto \mathscr{T}(\sum_{i=1}^n y_i)$, but the government's revenue is larger under the comprehensive tax system $\mathbf{y} \mapsto \mathscr{T}(\sum_{i=1}^n y_i)$ than under $\mathbf{y} \mapsto \mathcal{T}(\cdot)$. The reasoning is similar to the one found in Konishi (1995), Laroque (2005) and Kaplow (2008)[31] Our proof is constructed on a similar reasoning but is valid with general tax instruments and multidimensional incomes.

Under the linear production function (2), the inverse demand equations (1) simplify to $p_i = 1$. Let $\mathbf{X}(\mathbf{w}) = \mathbf{Y}(\mathbf{w})$ be the solution to:

$$\max_{\mathbf{y}} \quad \mathcal{U}\left(\sum_{i=1}^n y_i - \mathcal{T}(\mathbf{y}), \mathcal{V}\left(y_1, ..., y_n\right); \mathbf{w}\right). \tag{60}$$

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

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: $\mathscr{U}\left(C(\mathbf{w}^\star), \mathbf{X}(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}^\star), V(\mathbf{w}^\star); \mathbf{w}'\right) > \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}'); \mathbf{w}'\right) = \mathscr{U}\left(C(\mathbf{w}'), \mathbf{X}(\mathbf{w}'); \mathbf{w}'\right)$ which would contradict that $\mathbf{y} = \mathbf{Y}(\mathbf{w}')$ solves (60) for individuals of type $\mathbf{w}'$.

Next, we define the expenditure function $\mathcal{R}(\cdot)$ such that, for each subdisutility level v, either there exists $\mathbf{w}$ such that $v = V(\mathbf{w})$, in which case we define $\mathcal{R}(v) = C(\mathbf{w})$, or $\mathcal{R}(v) = -\infty$. Note also that $\mathcal{R}$ is increasing over the set of attained subdisutility. Otherwise, there would exist w and w' such that $v = V(\mathbf{w}) < v' = V(\mathbf{w}')$ and $\mathcal{R}(v) = C(\mathbf{w}) \geq \mathcal{R}(v) = C(\mathbf{w}')$. This would lead to $\mathcal{U}(C(\mathbf{w}), V(\mathbf{w}); \mathbf{w}') > \mathcal{U}(C(\mathbf{w}'), V(\mathbf{w}'); w')$, which would contradict that $\mathbf{y} = \mathbf{Y}(\mathbf{w}')$ solves (60) for individuals of type $\mathbf{w}'$.

For individuals of type $\mathbf{w}$ solving (60) amounts to solve

$$\max_{v} \quad \mathcal{U}(\mathcal{R}(v), v; \mathbf{w}). \tag{61}$$

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

$$V(g) \stackrel{\text{def}}{\equiv} \min_{\mathbf{y}} \quad \mathcal{V}\left(y_1, ..., y_n\right) \quad s.t: \quad \sum_{i=1}^n y_i = g \tag{62}$$

is well defined and so is its value $V(\cdot)$. In particular, $V(\cdot)$ is increasing since $\mathcal{V}$ is increasing in each argument. In (62), g is the sum of the different kinds of income y_i $(i = 1, ..., n)$ when these income levels are chosen to minimize the subdisutility $\mathcal{V}$ of all actions together. We then define $\mathscr{T}(\cdot)$ by:

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

[31] These authors show that a linear indirect tax is useless when a nonlinear labor income tax prevails. Indeed, despite the fact that the agents choose the same allocation under both tax systems, the government's revenue is proven to be larger with a zero indirect tax rate than with a positive one.

which is g minus the value of consumption reached when the subdisutility of all actions is minimized. Under the comprehensive 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:

$$\max_{\mathbf{y}} \qquad \mathcal{U}\left(\mathcal{R}\left(V\left(\sum_{i=1}^{n} y_i\right)\right), \mathcal{V}\left(y_1, ..., y_n\right); \mathbf{w}\right).$$

This problem can be solved sequentially. First, one solves the dual program of (62)

$$\max_{\mathbf{y}} \quad \sum_{i=1}^{n} y_i \quad s.t: \quad \mathcal{V}\left(y_1, ..., y_n\right) = v$$

for a given level of subdisutility v since $\mathcal{R}$ and V are increasing mappings. Second, one solves Program (61). The tax schedule $\mathbf{y} \mapsto \mathscr{T}\left(\sum_{i=1}^{n} y_i\right)$ therefore leads each type of taxpayer to make the same decisions and to reach the same $V(\mathbf{w})$ as well as the same utility $U(\mathbf{w})$ than under the tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$.

However, the tax revenue is under under the initial tax schedule, since $\mathbf{Y}(\mathbf{w})$ is solving

$$\max_{\mathbf{y}} \quad \sum_{i=1}^{n} y_i - \mathcal{T}(\mathbf{y}) \qquad s.t.: \quad \mathcal{V}\left(y_1, ..., y_n\right) = V(\mathbf{w})$$

instead of solving:

$$\max_{\mathbf{y}} \quad \sum_{i=1}^{n} y_i \qquad s.t.: \quad \mathcal{V}\left(y_1, ..., y_n\right) = V(\mathbf{w})$$

the latter program having the same solution as:

$$\min_{\mathbf{y}} \quad \mathcal{V}\left(y_1, ..., y_n\right) \qquad s.t.: \quad \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} Y_i(\mathbf{w}).$$

D Responses to tax reforms

To be able to apply the implicit function theorem to the first-order condition associated to the individual maximization program, we make the following assumption.

Assumption 3. *The initial tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ is such that:*

i) The initial tax schedule is twice continuously differentiable.

ii) The second-order condition associated to the individual maximization program (5) *holds strictly, i.e. the matrix* $\left[\dfrac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i \, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}$ *is positive definite.*

iii) For each type $\mathbf{w} \in W$, program (5) *admits a unique global maximum.*

Part (i) of Assumption 3 ensures that first-order conditions (6) are differentiable in incomes $\mathbf{y}$. It rules out kinks in the tax function, thereby bunching.[32] Parts (i) and (ii) of Assumption 3

[32] In practice, most of real world tax schedules are piecewise linear. In theory, bunching should occur at convex kink points and gaps in the income distribution should occur at concave kink points. In practice, bunching is very rare (with the noticeable exception of Saez (2010)) and gaps as well. This discrepancy between theory and reality is plausibly due to the fact that taxpayers do not optimize with respect to the exact tax schedule but with respect to some smooth approximation of it, e.g. $\mathbf{y} \mapsto \int \mathcal{T}(\mathbf{y} + \mathbf{u}) \, d\Psi(\mathbf{u})$ where $\mathbf{u}$ is an n-dimensional random shock on incomes with joint CDF Ψ, which does verify part i) of Assumption 3.

together ensure that the implicit function theorem can be applied to first-order conditions (6) to ensure that each local maximum of $\mathbf{y} \mapsto \mathscr{U}\left(\sum_{k=1}^{n} y_k - \widetilde{\mathcal{T}}(\mathbf{y}, t), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right)$ is differentiable in type $\mathbf{w}$, in price $\mathbf{p}$ and in the tax perturbation's magnitude t. If this mapping admits several global maxima among which taxpayers are indifferent, any small tax reform may then lead to a jump in taxpayer's choice from one maximum to another one. Part (*iii*) prevents this situation and ensures the allocation changes in a differentiable way with the magnitude of the tax reform and with types.

Because the indifference set is convex (See Appendix A), Assumption 3 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 function $\mathbf{y} \mapsto \sum_{k=1}^{n} y_k - \mathcal{T}(\mathbf{y})$ 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 3 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 3 is satisfied). In the same spirit as the first-order mechanism design approach of Mirrlees (1971, 1976), we presume the optimal tax schedule verifies Assumption 3.[33] We derive optimality conditions and verify ex-post whether this presumption is validated by the obtained solution.

Derivation of Equation (21) with exogenous and endogenous prices and of Equations (20d) and (30)

Since taxpayers take the price $\mathbf{p} = (\widetilde{p}_1(t), ..., \widetilde{p}_n(t))$ as given, they solve, under the tax schedule $\mathbf{y} \mapsto \widetilde{\mathcal{T}}(\mathbf{y}, t)$, the following program which depends on the magnitude t of the tax perturbation and on the price vector $\mathbf{p}$:

$$\widehat{U}(\mathbf{w}; t, \mathbf{p}) \stackrel{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,...,y_n)} \quad \mathscr{U}\left(\sum_{i=1}^{n} y_i - \widetilde{\mathcal{T}}(\mathbf{y}, t), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right). \tag{63}$$

This individual maximization programs summarizes the supply side of our model since it gives all individual supplies of income. The first-order conditions are:

$$\forall i \in \{1, ..., n\}: \qquad \frac{1}{p_i}\mathcal{S}^i\left(\sum_{k=1}^{n} y_k - \widetilde{\mathcal{T}}(\mathbf{y}, t), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right) = 1 - \widetilde{\mathcal{T}}_{y_i}(\mathbf{y}, t). \tag{64}$$

Let $\widehat{\mathbf{Y}}(\mathbf{w}, t, \mathbf{p}) = (\widehat{Y}_1(\mathbf{w}, t, \mathbf{p}), ..., \widehat{Y}_n(\mathbf{w}, t, \mathbf{p}))$ denotes the solution, i.e. the individual supply of income. At equilibrium where $p_j = \widetilde{p}_j(t)$, one obviously has $\widetilde{Y}_i(\mathbf{w}, t) \equiv \widehat{Y}_i(\mathbf{w}, \widetilde{\mathbf{p}}(t))$ for all $i \in \{1, ..., n\}$ and $\widetilde{U}(\mathbf{w}, t) \equiv \widehat{U}(\mathbf{w}, \widetilde{\mathbf{p}}(t))$.

Under Assumption 3, the implicit function theorem ensures that the solution $\widehat{Y}(\mathbf{w}, t, \mathbf{p})$ to program (63) is differentiable with respect to t and to $\mathbf{p}$ and its partial derivatives at $\mathbf{p} = (\widetilde{p}_1(0), ..., \widetilde{p}_n(0))$ and $t = 0$ can be obtained by differentiating Equations (64) at $y = \mathbf{Y}(\mathbf{w})$. This leads to, $\forall i \in \{1, ..., n\}$:

$$\sum_{j=1}^{n}\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right] \mathrm{d}y_j = \left[-\frac{\partial \widetilde{\mathcal{T}}_{y_i}}{\partial t} + \frac{\mathcal{S}^i_c}{p_i}\frac{\partial \widetilde{\mathcal{T}}}{\partial t}\right] \mathrm{d}t + \sum_{j=1}^{n}\left(\mathbb{1}_{i=j}(1 - \mathcal{T}_{y_i}) + \frac{\mathcal{S}^i_{x_j}\, y_j}{p_i\, p_j}\right)\frac{\mathrm{d}p_j}{p_j}.$$

This differentiation can be rewritten in matrix form as:

$$\begin{aligned}\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j} \cdot \mathbf{dy}^T \quad &= \quad \left\{-\left[\frac{\partial \widetilde{\mathcal{T}}_{y_1}}{\partial t}, ..., \frac{\partial \widetilde{\mathcal{T}}_{y_n}}{\partial t}\right]^T_i + \left[\frac{\mathcal{S}^1_c}{p_1}, ..., \frac{\mathcal{S}^n_c}{p_n}\right]^T_i \frac{\partial \widetilde{\mathcal{T}}}{\partial t}\right\} \mathrm{d}t \\ &\quad + \left[\mathbb{1}_{i=j}(1 - \mathcal{T}_{y_i}) + \frac{\mathcal{S}^i_{x_j}\, y_j}{p_i\, p_j}\right]_{i,j} \cdot \left(\frac{\mathrm{d}p_1}{p_1}, ..., \frac{\mathrm{d}p_n}{p_n}\right)^T.\end{aligned} \tag{65}$$

[33]Conversely, Golosov et al. (2014) do assume that the income function is locally Lipschitz continuous in tax reforms, while Hendren (2020) does assume that aggregate tax revenue varies smoothly in response to changes in the tax schedule.

where superscript T denotes the transpose operator $\left[A_{i,j}\right]_{i,j}^{T} = \left[A_{j,i}\right]_{i,j}$ and "$\cdot$" denotes the matrix product. Under a compensated tax reform of the j^{th} marginal tax rate at income $\mathbf{y} = \mathbf{Y}(\mathbf{w})$, as defined in (20b), one gets $\frac{\partial \widetilde{\mathcal{T}}}{\partial t} = 0$ and $\frac{\partial \widetilde{\mathcal{T}}_{y_k}}{\partial t} = -\mathbb{1}_{j=k}$. Hence, according to (65), the matrix of compensated responses is given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} = \left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1}. \tag{66a}$$

Under the lump-sum tax reform defined in (20a), one has $\frac{\partial \widetilde{\mathcal{T}}}{\partial t} = -1$ and $\frac{\partial \widetilde{\mathcal{T}}_{y_k}}{\partial t} = 0$. Hence, according to (65), the vector of wealth responses is given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right]_i^T = -\left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \cdot \left(\mathcal{S}^1_c, ..., \mathcal{S}^n_c\right)^T. \tag{66b}$$

Finally, according to (65), the responses to changes in log prices are given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}\right]_{i,j} = \left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c \mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \cdot \left[\mathbb{1}_{i=j}(1 - \mathcal{T}_{y_i}) + \frac{\mathcal{S}^i_{x_j}\, y_j}{p_i\, p_j}\right]_{i,j}. \tag{66c}$$

Consider a tax perturbation as defined in Definition 2, plugging (66a) and (66b) into (65) yields:

$$\frac{\partial \widehat{Y}_i(\mathbf{w}, t = 0, \mathbf{p})}{\partial t} = -\underbrace{\frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), 0)}{\partial t}\right|_{t=0}}_{\text{Wealth responses}} - \underbrace{\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), 0)}{\partial t}\right|_{t=0}}_{\text{Compensated responses}} \tag{67}$$

which, with exogenous prices, leads to (21).

Under an uncompensated tax reform of the j^{th} marginal tax rate as defined in (20c), one gets $\frac{\partial \widetilde{\mathcal{T}}}{\partial x} = -Y_j(\mathbf{w})$ and $\frac{\partial \widetilde{\mathcal{T}}_{y_k}}{\partial x} = -\mathbb{1}_{j=k}$. So, Equation (67) leads to the Slutsky Equation (20d).

Finally, applying the envelope theorem to (63) leads to:

$$\frac{\partial \widehat{U}(\mathbf{w}, t = 0, \mathbf{p})}{\partial t} = -\mathscr{U}_c\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \left.\frac{\partial \widetilde{\mathcal{T}}\left(\mathbf{Y}(\mathbf{w}), t\right)}{\partial t}\right|_{t=0} \tag{68a}$$

$$\begin{aligned}\frac{\partial \widehat{U}(\mathbf{w}, t = 0, \mathbf{p})}{\partial \log p_j} &= -\mathscr{U}_{x_j}\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \frac{Y_j(\mathbf{w})}{p_j} \\ &= \mathscr{U}_c\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \left(1 - T_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w})\end{aligned} \tag{68b}$$

where the last equality follows from (3) and (6).

To compute the responses of prices to a tax reform, define the aggregate i^{th} income as function of the price $\mathbf{p}$ and of the magnitude t of the tax perturbation $\mathbf{y} \mapsto \widetilde{\mathcal{T}}(\mathbf{y}, t)$ as follows:

$$\mathcal{Y}_i(t, \mathbf{p}) \stackrel{\text{def}}{\equiv} \int_{\mathbf{w} \in W} \widehat{Y}_i(\mathbf{w}, t, \mathbf{p})\, f(\mathbf{w}) d\mathbf{w}$$

From the inverse demand equations (1), prices $\widetilde{\mathbf{p}}(t) = (\widetilde{p}_1(t), ..., \widetilde{p}_n(t))$ have to solve:

$$\forall t, \forall j \in \{1, ..., n\} \qquad p_i = \mathcal{F}_{\mathcal{X}_i}\left(\frac{\mathcal{Y}_1(t, \mathbf{p})}{p_1}, ..., \frac{\mathcal{Y}_n(t, \mathbf{p})}{p_n}\right).$$

Log-differentiating the latter equation leads to:

$$\begin{aligned}
\left[\frac{\mathrm{d}p_i}{p_i}\right]_i &= \Xi\cdot\left[\frac{\mathrm{d}\mathcal{X}_i}{\mathcal{X}_i}\right] = \Xi\left(\left[\frac{\mathrm{d}\mathcal{Y}_i}{\mathcal{Y}_i}\right]_i - \left[\frac{\mathrm{d}p_i}{p_i}\right]_i\right)\\
(I_n+\Xi)\cdot\left[\frac{\mathrm{d}p_i}{p_i}\right]_i &= \Xi\cdot\left[\frac{\mathrm{d}\mathcal{Y}_i}{\mathcal{Y}_i}\right]_i = \Xi\cdot\left(\left[\frac{1}{\mathcal{Y}_i}\frac{\partial\widetilde{\mathcal{Y}}_i(t)}{\partial t}\bigg|_{t=0}^{Micro}\right]_i + \Sigma\cdot\left[\frac{\mathrm{d}p_i}{p_i}\right]_i\right)\\
(I_n+\Xi-\Xi\cdot\Sigma)\cdot\left[\frac{\mathrm{d}p_i}{p_i}\right]_i &= \Xi\cdot\begin{pmatrix}\frac{1}{\mathcal{Y}_1} & 0\\ 0 & \frac{1}{\mathcal{Y}_2}\end{pmatrix}\cdot\left[\frac{\partial\widetilde{\mathcal{Y}}_i(t)}{\partial t}\bigg|_{t=0}^{Micro}\right]_i
\end{aligned}$$

Hence, under Assumption 1, one can apply the implicit function theorem to ensure that the vector of prices is differentiable with respect to t and that Equation (30) holds. Adding these price responses to Equation (67) and using (66c) leads to Equation (21). Combining Equations (24), (68a) and (68b) lead to (25).

E Proof of Proposition 3

Let $\mathbf{y}\mapsto\widetilde{\mathcal{T}}(\mathbf{y},t)$ be a tax perturbation and let $\ell(t)$ be the lump-sum rebate such that the tax perturbation $\mathbf{y}\mapsto\widetilde{\mathcal{T}}(\mathbf{y},t)+\ell(t)$ guarantees a balanced budget. Denote $\frac{\partial A}{\partial t}\Big|_{t=0}^{\star}$ the partial derivative of A along the budget-balanced tax perturbation $\mathbf{y}\mapsto\widetilde{\mathcal{T}}(\mathbf{y},t)+\ell(t)$. We thus get $\frac{\partial\mathscr{B}}{\partial t}\Big|_{t=0}^{\star}=0$ and so:

$$\frac{1}{\lambda}\frac{\partial\widetilde{\mathscr{W}}}{\partial t}\bigg|_{t=0}^{\star} = \frac{\partial\widetilde{\mathscr{L}}}{\partial t}\bigg|_{t=0}^{\star}$$

Let $\frac{\partial A^{\rho}}{\partial t}\Big|_{t=0}$ be the partial derivative of A along the lump-sum perturbation (20a). According to (32), we get:

$$\frac{\partial\widetilde{\mathscr{L}}}{\partial t}\bigg|_{t=0}^{\star} = \frac{\partial\widetilde{\mathscr{L}}}{\partial t}\bigg|_{t=0} + \ell'(0)\,\frac{\partial\widetilde{\mathscr{L}}^{\rho}}{\partial t}\bigg|_{t=0}.$$

Equation (33) implies:

$$\frac{\partial\widetilde{\mathscr{L}}^{\rho}}{\partial t}\bigg|_{t=0} = 0.$$

Combing these three equations leads to:

$$\frac{1}{\lambda}\frac{\partial\widetilde{\mathscr{W}}}{\partial t}\bigg|_{t=0}^{\star} = \frac{\partial\widetilde{\mathscr{L}}}{\partial t}\bigg|_{t=0}.$$

Since $\lambda>0$, $\frac{\partial\widetilde{\mathscr{W}}(t)}{\partial t}\Big|_{t=0}^{\star}$ is positive, i.e. the budget-balanced reform improves welfare if and only if $\frac{\partial\widetilde{\mathscr{L}}}{\partial t}\Big|_{t=0}>0$.

F Responses of taxable income under a mixed tax schedule

According to (16) and (21), we get:

$$\left.\frac{\partial \widetilde{Y}_0(\mathbf{w},t)}{\partial t}\right|_{t=0} = \sum_{k=1}^{n} a_k'(y_k)\left.\frac{\partial \widetilde{Y}_k(\mathbf{w},t)}{\partial t}\right|_{t=0}$$

$$= -\sum_{1\le j,k\le n} a_k'(y_k)\,\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} - \sum_{k=1}^{n} a_k'(y_k)\,\frac{\partial Y_k(\mathbf{w})}{\partial \rho}\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}$$
$$- \sum_{1\le j,k\le n} a_k'(y_k)\,\frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0}$$

Equation (21) is thus also verified for taxable income with $i=0$ as long as the income response, the compensated responses and the response to relative price changes are respectively defined by (35a), (35b) and (35d). Moreover, for $z=\rho,\tau_j,\log p_j$, we obtain:

$$\begin{aligned}
\sum_{i=1}^{n}\left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))+\mu_i\right)\frac{\partial Y_i(\mathbf{w})}{\partial z} &= \sum_{k=1}^{n}\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial z} + T_0'(Y_0(\mathbf{w}))\sum_{k=1}^{n} a_k'(Y_k(\mathbf{w}))\frac{\partial Y_k(\mathbf{w})}{\partial z}\\
&= \sum_{k=1}^{n}\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial z} + T_0'(Y_0(\mathbf{w}))\frac{\partial Y_0(\mathbf{w})}{\partial z}\\
&= \sum_{k=0}^{n}\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial z}
\end{aligned}$$

where the first equality is obtained by using Equations (16) and (17) and by inverting subscripts i and k. The second equality is obtained using Equations (35a), (35b) and (35d). The last equality holds because we have normalized $\mu_0=0$. Equation (32) then becomes:

$$\begin{aligned}
\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{\mathbf{w}\in W}\left\{\left[1-g(\mathbf{w})-\sum_{k=0}^{n}\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right]\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}\right.\\
&\quad\left.-\sum_{j=1}^{n}\left(\sum_{k=0}^{n}\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right)\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}\right\} f(\mathbf{w})\mathrm{d}\mathbf{w}.
\end{aligned} \tag{69}$$

G Reforms of the tax schedule specific to the i^{th} income and its optimal shape (with arbitrary or optimal other taxes)

We consider tax perturbations of the form:

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

which implies (36) and (37). Plugging these equations into (69) leads to Equation (38). Applying our proof of Proposition 3 (Appendix E), it is therefore straightforward to proof part i) of Proposition 4.

Using the law of iterated expectations to condition types $\mathbf{w}$ on $Y_i(\mathbf{w})=y_i$ and using (42), we obtain:

$$\begin{aligned}
\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{y_i\in\mathbb{R}_+}\left\{\left[\frac{T_i'(y_i)+\mu_i}{1-T_i'(y_i)}\,\varepsilon_i(y_i)\,y_i + \sum_{0\le k\le n,k\neq i}\overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\left.\frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\right|_{Y_i(\mathbf{w})=y_i}}\right] R'(y_i)\right.\\
&\quad\left.-\left[1-\overline{g(\mathbf{w})}\big|_{Y_i(\mathbf{w})=y_i} - \sum_{k=0}^{n}\overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right)\left.\frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right|_{Y_i(\mathbf{w})=y_i}}\right] R(y_i)\right\} h_i(y_i)\,\mathrm{d}y_i
\end{aligned}$$

Integrating the latter equation by parts and using (33) leads to:

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{y_i \in \mathbb{R}_+} \left\{ \frac{T_i'(y_i)+\mu_i}{1-T_i'(y_i)}\, \varepsilon_i(y_i)\, y_i\, h_i(y_i) + \sum_{0\le k\le n, k\neq i} \overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\Big|_{Y_i(\mathbf{w})=y_i}}\, h_i(y_i) \right. \\ &\quad \left. - \int_{z=y_i}^{\infty} \left[1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=z} - \sum_{k=0}^{n} \overline{\left(T_k'(Y_k(\mathbf{w}))+\mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\Big|_{Y_i(\mathbf{w})=z}} \right] h_i(z)\mathrm{d}z \right\} R'(y_i)\mathrm{d}y_i\end{aligned}$$

If $T_i(\cdot)$ is optimal whatever the other tax schedules, any reform 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 the preceding expression should be zero for all y_i, which leads to (44) and thereby, to part *ii)* of Proposition 4.

H Reforms of the personal income tax schedule

We consider tax perturbations of the following form:

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

which implies (39) and (40). Using (21), one obtains the impact of this type of reform of the personal income tax on all types of income, $\forall k \in \{1,...,n\}$:

$$\left.\frac{\partial \widetilde{Y}_k(\mathbf{w},t)}{\partial t}\right|_{t=0} = \sum_{j=1}^{n} a_j'(Y_j(\mathbf{w}))\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) + \sum_{j=1}^{n} \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\, \left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0}. \tag{70}$$

Combining $\widetilde{\mathcal{T}}(\mathbf{y},t) = \mathcal{T}(\mathbf{y}) - t\, R_i(y_i)$ with a compensated tax reform of the personal income described in Equation (20b), so that, in (70) one has $R_0(\cdot) = 0$, $R_0'(\cdot) = -1$ and $\left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0} = 0$, we obtain (41). Given (41), for $k \in \{1,..,n\}$, Equation (70) simplifies to:

$$\left.\frac{\partial \widetilde{Y}_k(\mathbf{w},t)}{\partial t}\right|_{t=0} = \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0}\, R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) + \sum_{j=1}^{n} \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\, \left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0}.$$

Combining the latter equation with (16), (35a), (35b) and (41) for $i = k = 0$ leads to:

$$\begin{aligned}\left.\frac{\partial \widetilde{Y}_0(\mathbf{w})}{\partial t}\right|_{t=0} &= \sum_{1\le k,j\le n} a_k'(Y_k(\mathbf{w}))\, a_j'(Y_j(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\, R_0'(Y_0(\mathbf{w})) \\ &\quad + \sum_{k=1}^{n} a_k'(Y_k(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) + \sum_{1\le k,j\le n} a_k'(Y_k(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\, \left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0} \\ &= \sum_{j=1}^{n} a_j'(Y_j(\mathbf{w}))\, \frac{\partial Y_0(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_0(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) \\ &\quad + \sum_{1\le k,j\le n} a_k'(Y_k(\mathbf{w}))\, \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\, \left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0}.\end{aligned}$$

We can conclude that (21) also holds for $j = 0$, i.e. with taxable personal income tax reforms. According to Equation (69), one gets:

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{\mathbf{w}\in W} \left\{ \left[\sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \left(\sum_{j=1}^{n} a_j'(Y_j(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) \right] R_0'(Y_0(\mathbf{w})) \right.\\ &+ \left. \left[-1 + g(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_0(Y_0(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w} \\ &= \int_{\mathbf{w}\in W} \left\{ \left[\sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} \right] R_0'(Y_0(\mathbf{w})) \right.\\ &+ \left. \left[-1 + g(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_0(Y_0(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}\end{aligned}$$

where the second equality uses (41). We thus get (38) with $i = 0$. Part (*i*) of Proposition 4 is therefore also valid for $i = 0$, thereby also its Part (*ii*).

I Optimal linear tax schedule

Rewriting Equation (38) with the uncompensated tax perturbation of the i^{th} income defined in (20c) (i.e. taking $R_i(Y_i(\mathbf{w})) = Y_i(\mathbf{w})$ and $R'(Y_i(\mathbf{w})) = 1$) and using the Slutsky equations (20d) leads to:

$$\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1\right] Y_i(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}.$$

Assuming that the i^{th} income is taxed at the linear rate t_i, so that $T_i(y_i) = t_i\, y_i$ leads to:

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1\right] Y_i(\mathbf{w}) + (t_i + \mu_i) \frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_i} \right.\\ &+ \left. \sum_{k=0, k\neq i}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}.\end{aligned}$$

Equating to zero the latter expression, where one substitutes the uncompensated elasticity $\varepsilon_i^u(\mathbf{w}) \stackrel{\text{def}}{\equiv} \frac{1-t_i}{Y_i(\mathbf{w})} \frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_i}$, and rearranging terms, leads to (45).

J Proof of Proposition 5

The reform of the i^{th} deduction rate defined in (46) implies:

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

where (48) and (47) have been used for the second and third equation, respectively. Combining these expressions with (69) leads to:

$$\begin{aligned}
\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = & \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k \right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] Y_i(\mathbf{w})\ T_0'(Y_0(\mathbf{w})) \right. \\
& + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k \right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\ T_0'(Y_0(\mathbf{w})) \\
& + \left. \left(\sum_{j=1}^{n}\sum_{k=0}^{n} a_j'(Y_j(\mathbf{w})) \left(T_k'(Y_k(\mathbf{w})) + \mu_k \right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \right) Y_i(\mathbf{w})\ T_0''(Y_0(\mathbf{w})) \right\} f(\mathbf{w})d\mathbf{w}.
\end{aligned}$$

Using the Slutsky equation (20d) and Equation (41), the preceding equation simplifies to (49), which, combined with Proposition 3, ends the proof of Proposition 5.

K Input Taxation is superfluous

In this appendix, we show that the taxation of production inputs can be replicated by an adequate re-scaling of the income tax function $\mathcal{T}(\cdot)$. Hence our assumption that there is no tax taxation of inputs is without loss of generality. Assume input i is taxed at rate $\alpha_i < 1$.

For each $i \in \{1, ..., n\}$, let p_i denote producers' input prices and let $q_i = p_i(1-\alpha_i)$ denote suppliers' prices. The i^{th} market income is $y_i = p_i\ x_i$ while the i^{th} *taxable* income is equal to $q_i\ x_i = (1-\alpha_i)x_i$. The tax schedule is a function of the vector of taxable income $(q_1\ x_1, ..., q_n\ x_n) = ((1-\alpha_1)y_1, ..., (1-\alpha_n)y_n)$. Hence, after-tax income c verifies:

$$c = \sum_{i=1}^{n} q_i\ x_i - \mathcal{T}(q_1\ x_1, ..., q_n\ x_n) = \sum_{i=1}^{n} y_i - \alpha_i y_i - \mathcal{T}\left((1-\alpha_1)y_1, ..., (1-\alpha_n)y_n\right)$$

In the presence of input taxation, instead of (5), a $\mathbf{w}$-taxpayer solves:

$$U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \max_{y_1,...,y_n} \quad \mathscr{U}\left(\sum_{k=1}^{n} y_k - \alpha_i\ y_k - \mathcal{T}\left((1-\alpha_1)y_1, ..., (1-\alpha_n)y_n\right), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w} \right)$$

Definition 1 of the equilibrium is otherwise unchanged. Since the inverse demand equations (1) and the market clearing conditions (7) are unchanged, the same equilibrium $\mathbf{p} = (p_1, ..., p_n)$, $\mathbf{w} \mapsto \mathbf{Y}(\mathbf{w})$ and $(\mathcal{Y}_1, ..., \mathcal{Y}_n)$ is obtained with input price vectors $(\alpha_1, ..., \alpha_n)$ and income tax $\mathcal{T}(\cdot)$ or without any input taxation and the renormalized income tax schedule $(y_1, ..., y_n) \mapsto \mathcal{T}\left((1-\alpha_1)y_1, ..., (1-\alpha_n)y_n\right) + \sum_{i=1}^{n} \alpha_i\ y_i$.

L Numerical simulations

The algorithm iterates different operations on the "real" dataset made of 27,804 observations from ERFS and a grid of 50 taxable income y_0 levels between € 4,000 and € 200,000 that we refer to the "virtual dataset".

The real dataset initially contains for each observation a labor and a capital income level from ERFS, the ERFS weight of the observation as well as an approximation of the marginal tax rate on labor income and of capital income. We use taxpayers' first order conditions (6) to assign a type (w_1, w_2) to each observation of the real dataset.

Each node of the virtual dataset is made by a personal income tax level y_0 and a marginal tax rate of the personal income tax schedule $T_0'(y_0)$. The personal marginal tax schedule $y_0 \mapsto T_0'(\cdot)$ is approximated by a linear interpolation. To extrapolate marginal tax rate above € 200,000, we suppose marginal tax rate are constant above € 200,000.

a_2 is between a minimum of 0 the (dual case) and a maximum value of 1. The different steps in the numerical process for a given a are the following.

1. Given the linear interpolation of personal marginal tax rate obtained from the personal income tax base and an intercept for the personal income tax schedule, compute for each observation of the real dataset the solution of taxpayer's program (5), i.e. their labor income and capital income. Deduce from this solution compensated elasticities and utility levels. Compute the macro spillover terms using (31). Compute the statistics that show up in (44) (for $i = 0$) and in (45) (for $i = 2$).[34]

2. For each observation of the virtual dataset, we compute personal income density by a kernel density estimation, and we compute the mean of the sufficient statistics that show up in (44) by kernel regression on the real dataset.

3. Unlike for the comprehensive tax system, Evaluate (45) by summing the statistics that show up in (45).

4. Go back to step 1

Once marginal personal income tax rates and linear tax rate across two successive iteration differ by less than 0.1 percentage point, the algorithm Evaluate (49) on the real dataset. If (49) is positive (negative), update a_2 to be between its minimum (maximum) value and its preceding one and update the bounds of a_2. The program stops when the difference between the minimum and maximum value of a_2 is sufficiently low.

[34]Under Maximin one has $g(\mathbf{w}) = 0$. Moreover, with quasilinear preferences, wealth effects are nil $\frac{\partial Y_k}{\partial \rho} = 0$. Finally, we take $a_1(y_1) = y_1$, so $a_1'(y_1) = 1$, $T_1(y_1) = T_1'(y_1) = 0$.