THEMA

théorie économique, modélisation et applications

THEMA Working Paper n°2021-19
CY Cergy Paris Université, France

How to Tax Different Incomes?

Laurence JACQUET, Etienne LEHMANN

October, 2021

CERGY PARIS UNIVERSITÉ

cnrs

How to Tax Different Incomes?*

Laurence JACQUET† Etienne LEHMANN‡

October 4, 2021

Abstract

We study the optimal tax system when taxpayers earn different kinds of income by supplying different inputs. Imperfect substitution between inputs allows for general equilibrium effects. We consider any type of cross-base responses to tax changes such as income-shifting. Formalizing the tax schedule as the sum of many one-dimensional schedules, we express optimal marginal tax rate on any kind of income in terms of sufficient statistics, including new ones for cross-base responses and general equilibrium effects. We also identify the conditions under which making the personal income tax marginally more schedular is socially desirable. The comprehensive and schedular (dual, in particular) income taxes being recurring proposals in the public debate, we derive sufficient conditions under which each form of tax is optimal. We stress how empirically restrictive these conditions are. Using a new algorithm on French tax return data, we characterize the optimal combination of a nonlinear tax schedule on personal income and a linear tax rate on capital income. We find that one should include, without any deduction, all income sources in the personal income base and subsidize the source of income which is more elastic. We find that cross-base responses have little effects on the personal nonlinear income tax schedule but increases by 5.9 to 6.9 percentage points the capital tax rate. General equilibrium effects also increases this tax rate by around 4.5 percentage points.

Keywords: nonlinear income taxation, several income sources, cross-base responses, endogenous prices, dual income tax, comprehensive income tax

JEL codes: H21, H22, H24

*We thank Marcus Berliant, Pierre Boyer, Kate Cuff, Antoine Ferey, Aart Gerritsen, Kory Kroft, Adam Lavecchia, Stéphane Robin, Johannes Schmieder, Joel Slemrod, Michael Smart, Elina Tuominen and participants at various conferences and seminars for helpful comments and suggestions. A previous version of this paper circulated under the title "Comprehensive or Schedular Income Tax". 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: https://sites.google.com/view/etiennelehmann/. Etienne LEHMANN is also research fellow at CEPR, IZA and CESIfo.

I Introduction

Taxpayers earn income from different sources such as salaries, dividends, business income, etc. How these distinct sources should be taxed is the object of sustained interest and controversy in the public debate. Views typically revolve around an equity-efficiency trade-off that considers who earns each source of income, how strongly would each source react to a tax change and how a tax change on one source of income spills over to other sources of income. Income-shifting and other cross-base responses are typically at the core of the debate as well as general equilibrium effects due to imperfect substitution between inputs.

We investigate how all these interactions impact the optimal income tax system. We assume that taxpayers are endowed with $n > 1$ distinct unobserved characteristics. Each taxpayer supplies n different inputs with endogenous prices that give her n distinct sources of income. We conduct a tax incidence analysis where we embrace cross-bases responses and general equilibrium effects thanks to new sufficient statistics. To derive optimal tax formulas in this general multidimensional environment, we consider that taxation consists of one schedule for each source of income and one schedule for personal income. The latter is the sum of all sources of income from which any income source can be partly or totally deducted. This system, besides being general, approximates very well most real-world tax schedules. Using tax perturbations, we characterize each optimal tax schedule.

We design and implement a new algorithm to take our optimal tax formulas to French tax return data. We split all reported incomes so that they fall either into the "labor income" or "capital income" category. The supply elasticity of capital is set at a higher value than the one of labor (Lefebvre et al., 2021). Our algorithm succeeds in computing the optimal combination of a nonlinear personal income tax schedule and a linear tax rate on capital income. The assumption of a linear tax rate on capital income is primarily made for computational tractability, but is also connected to proposals of a flat tax on capital income in the policy debate. We obtain three sets of analytical and numerical findings.

First, our optimal income tax formulas depend on estimates of cross-base responses to tax changes. Departing from the previous literature, our approach is totally amenable to any micro-foundation that would lie behind cross-base responses. In the paper, we present two micro-founded environments as illustrations. First, cross-base responses may stem from entrepreneurs who have the opportunity to shift income between distinct bases. Second, our model is consistent to capture, in a reduced-form way, the mechanisms that arise in macro dynamic models where saving generates capital and decreases consumers utility by reducing prior-consumption. The behavioral responses can take any value so that an infinite elasticity of capital to taxes, as one has in infinite horizon neoclassical models (Ramsey, 1928, Judd, 1985, Chamley, 1986, Straub and Werning, 2020), is possible. We show that optimal tax rates on one source of income are higher when other income sources respond less to an increase of these tax rates, e.g. because of lower income-shifting. Numerically, we find that income-shifting has very little effects on the personal nonlinear income tax schedule but it increases by 5.9 to 6.9 percentage points the linear capital tax rate.

Second, our tax formulas depend on the new *macroeconomic spillover* statistics that quantify general equilibrium adjustments. Let a tax reform induce a rise in one source of income. The subjacent input increases. This initial response impacts the marginal products of inputs and, thereby, prices. Taxpayers in turn modify their input supplies, which further affect prices, and so on. The spillover statistic associated to an input measures the impact of increasing this input through these general equilibrium adjustments. A positive (negative) spillover statistic means that an increase in the corresponding specific input creates prices changes that are socially beneficial (detrimental). It is then optimal to improve (reduce) incentives to earn the corresponding income through lower (higher) marginal tax rates. Departing from the previous literature, our formulas are extremely flexible regarding micro-foundations. They remain valid with perfect competition, monopoly pricing, rent-seeking behaviors, etc. One only needs the matrix of inverse demand elasticities and the matrix of taxpayers' responses to prices variations, whatever the micro-foundations behind these matrices. Numerically, we find that moving from an infinite to 0.67 elasticity of substitution between capital and labor increases the tax rate on capital by around 4.5 percentage points.

Third, there is a long-standing and controversial debate between defenders of a *comprehensive* tax system (e.g. Saez and Zucman (2019)) and proponents of *schedular* tax systems. Several contributions in the literature have discussed the pros and cons of each tax system, e.g., Burns and Krever (1998), Boadway (2004), Bastani and Waldenström (2020). A comprehensive tax system applies the same tax schedule to the sum of all sources of income. By contrast, under schedular taxation, each source of income is subject to a specific schedule. Among schedular tax systems, the *dual* income tax, with a linear tax on capital income and a nonlinear and progressive tax schedule on labor income, is popular and prevails in many European countries (Benoteau and Meslin, 2017). We draw policy implications for this debate. We derive a formula which quantifies whether a marginal reform towards a more comprehensive or more schedular system is desirable. Equating this formula to zero also gives the optimal rate by which an income should be deducted from the personal income tax base. We also provide sufficient conditions under which the optimal tax is schedular, and an other under which it is comprehensive. Importantly, we obtain these sufficient conditions without assuming that the tax system is the sum of income-specific schedules and a personal income tax schedule. These conditions emphasize the fact that neither tax schedule is systematically optimal.

Numerically, the optimal tax system is neither comprehensive nor dual. It instead consists in including all capital income in the personal income tax base and applying a negative rate on capital income. The high marginal tax rates on personal incomes create disincentives to earn capital income that the negative tax on capital income mitigates. Furthermore, all our simulations show that, under dual and comprehensive taxation, the optimal marginal tax rates on personal income are U-shaped. However, optimal marginal tax rates diverge beyond 100,000€. For instance, at 200,000€, the optimal marginal tax rate is 5.9 points larger under dual taxation. Intuitively, personal income contains only labor income under dual taxation, which implies a lower income elasticity than under comprehensive taxation. This elasticity difference is especially important for top income earners who concentrate most of capital incomes. This result might validate the strategy of Nordic countries that adopt a dual tax with a highly progressive income tax system (Boadway, 2004, Sørensen, 2009).

Our paper contributes to several strands of the literature. First, many papers investigate whether capital income should be taxed in two-period models. While the Atkinson and Stiglitz (1976) theorem leads to zero tax on capital (Mankiw et al., 2009), this recommendation relies on very restrictive assumptions, especially a single dimension of unobserved heterogeneity. It thus does not seem to be very policy relevant (Diamond and Saez, 2011). Among the attributes that make capital taxation desirable, the literature has emphasized heterogeneity in returns on investments (Gahvari and Micheletto, 2016, Kristjánsson, 2016, Saez and Stantcheva, 2018, Gerritsen et al., 2020), in time preferences (Saez, 2002, Diamond and Spinnewijn, 2011, Golosov et al., 2013), in parental altruism (Farhi and Werning, 2010, Piketty and Saez, 2013b) or in wealth (Cremer et al., 2003). Intuitively, if these additional attributes correlate with individuals' earnings abilities, taxes on capital become useful as indirect means to tax people with high ability.[1] We take for granted the desirability of capital taxation. As in Saez and Stantcheva (2018), we express optimal tax rates as a function of empirically meaningful sufficient statistics, independent of specific micro-foundations. We depart from their paper in three important directions. First, we allow for cross-base responses in a very flexible way.[2] In particular, we allow for cases where a larger tax on capital increases labor income, for instance because of income-shifting. But we also allow for situations where a larger tax on capital income decreases labor income. This happens, for instance, when individuals work to consume later on. In this case, a larger tax on savings reduces the reward from working to save and hence reduces the labor supply. Second, we allow for imperfect substitution between labor and capital. Hence, our formula takes into account that capital taxation may affect investment, therefore labor demand and, eventually, labor income. Last, we not only derive optimal tax rates for the different income tax schedules but we also derive a formula for the optimal deduction rate of capital income from the personal income tax base.[3]

Proponents of dual taxation argue that it has the advantage of keeping the progressivity of the personal income tax schedule while reducing the tax rate on the most responsive tax basis (Boadway,

[1]Another argument for taxing capital is the uncertainty agents face about their labor productivity profiles over time, e.g. Golosov et al. (2003), Albanesi and Sleet (2006)

[2]Saez and Stantcheva (2018, Proposition 4) consider the possibility of income-shifting, but they restrict their analysis to linear tax schedules and utility linear in consumption.

[3]Ferey et al. (2021) also express optimal tax rates on labor and capital incomes in terms of sufficient statistics. Unlike Saez and Stantcheva (2018), they allow for rich cross-base responses. Contrary to the present paper, their main results are obtained under one-dimensional unobserved heterogeneity and perfect substitution between labor and capital.

2004, Sørensen, 2009). Critics however stress that it mostly advantages high income earners (Sandmo, 2005, Hermle and Peichl, 2018) and encourages income-shifting between tax bases (Saez and Zucman, 2019). Compliance and administration costs also differ between both systems (Slemrod and Gillitzer, 2014) and as far as we know, there is no consensus regarding the cheapest one. Our contribution to the literature on the merits and flaws of comprehensive versus dual taxation (e.g., Burns and Krever (1998), Benoteau and Meslin (2017) or Bastani and Waldenström (2020)) goes beyond showing that neither tax schedule is systematically optimal. Our optimal deduction formula quantifies when it is socially desirable to move towards a more schedular or a more comprehensive tax system. In addition, we propose a new algorithm that allows one to obtain the optimal nonlinear tax system with several sources of income. This allows us to draw the optimal tax system on French data: the personal income comprehends all capital income and the linear tax on capital is negative.

Our paper finally contributes to the multidimensional optimal tax literature. First we provide the optimal tax formulas with n sources of income when the tax schedule is the sum of many one-dimensional schedules. This schedule is very flexible, relevant empirically and we are confident that if one were able to obtain the optimal general n-dimension tax schedule of Mirrlees (1976), it would be very close to our optimum.[4] Second, our environment includes (endogenous) tax incidence. Therefore, our formulas take into account general equilibrium effects. These effects have been studied in environments with one source of income in Rothschild and Scheuer (2013) and Sachs et al. (2020). Differing from the previous literature, we propose macro spillover statistics than offer the advantage of being directly computed with empirical variables.[5] Our approach has also the advantage of being agnostic regarding the micro-foundations of the general equilibrium effects. The price of each input does not need to be equal to its marginal productivity, it can be any function of the aggregate supply of inputs. Third, our optimal tax formulas capture cross-base responses, whatever their micro-foundations, again differing from the previous literature (e.g. Saez and Stantcheva (2018)).

The paper is organized as follows. In the next section we present the model. In Section III, we exhibit specific restrictions on the primitives of the model to ensure the overally optimal $n-$dimensional tax schedule coincides either with a schedular tax system or with a comprehensive one. In Section IV, we compute the effects of tax reforms, taking into account general equilibrium effects. In Section V, we derive the optimal tax formula and the optimal deduction formula under the restriction that the tax schedule is the sum of a personal income tax schedule and n income-specific tax schedules. We implement these formulas numerically in Section VI and conclude in the last section.

[4] To find the optimal n-dimension tax schedule, one needs to solve a nonlinear partial differential equation (Mirrlees, 1976). Despite several attempts (e.g., Kleven et al. (2007)), this equation has not been theoretically solved yet. Its numerical solution has recently been proposed in Spiritus et al. (2021).

[5] In Sachs et al. (2020), the additional terms that show up in their tax formulas are the complex solution of a their integral Equation (9).

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 $(\mathcal{X}_1, ..., \mathcal{X}_n)$. Unless otherwise specified, $n \geq 2$. 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 non-positive, i.e. $\mathcal{F}_{\mathcal{X}_i \mathcal{X}_i} \leq 0$. Assuming perfect competition, the firm, which is price-taker, chooses its inputs to maximize its profit:

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

where $p_i \in \mathbb{R}_+$ stands for the price of the i^{th} input. From the first-order conditions, we obtain the inverse demand equations:

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

The input price p_i is equal to the marginal productivity of the i^{th} input, $\mathcal{F}_{\mathcal{X}_i}$. Prices are then endogenous whenever inputs are imperfect substitutes. For instance, suppose a production function with two inputs, capital and labor, which are imperfect substitutes. 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 inputs) and therefore 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)).

In contrast, when the production function is linear, i.e.

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

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

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

Importantly, the equality of the price of the i^{th} input to its marginal product, as given in Equation(1), is not a requirement. Our framework remains valid as long as p_i is determined by any general function of the aggregate supply of inputs $(\mathcal{X}_1, ..., \mathcal{X}_n)$. Departing from perfect competition, we can therefore assume monopolistic competition as done in the large literature that builds on Dixit and Stiglitz (1977). We can alternatively consider that compensation is decided by on-the-job bargaining between an employer and an employee so that the price of labor can be distinct from its marginal product, see e.g. Hungerbühler et al. (2006) and Piketty et al. (2014) where compensation-bargaining responses prevail. Our framework is valid insofar as each input price depends on the aggregate supply of inputs, whatever the micro rationale for the inverse demand equations (1).

In appendix K, we show that incorporating the taxation of inputs does not modify our results.[6] We can therefore assume zero taxation of inputs without loss of generality.

II.2 Taxpayers

Each taxpayer is characterized by different individual characteristics summarized in their vector of type $\mathbf{w} = (w_1, ..., w_n)$. 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 denoted W.

Each taxpayer supplies $x_i \geq 0$ units of input $\mathcal{X}_i$ and her supply is denoted by $\mathbf{x} = (x_1, ..., x_n)$. For instance, a taxpayer can supply an amount of effective units of labor x_1, an amount of investment units in capital x_2, etc. Each 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. To taxpayers, price p_i is the macroeconomic return of the i^{th} input they supply. To the firm, it is the price of this input. Input prices are summarized by the vector $\mathbf{p} = (p_1, ..., p_n)$. For instance, if x_1 denotes effective labor, price p_1 is the wage per unit of effective labor and y_1 is labor income. If x_2 denotes savings, p_2 is the gross return on savings and y_2 is capital income, and so on.

Each supply of input comes with effort or a utility cost that depends on the vector of type $\mathbf{w}$ according to 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 $\mathbb{R}_+^{n+1} \times W$, increasing in the after-tax income (with partial derivative denoted $\mathscr{U}_c > 0$) and decreasing in the supply of each input (with partial derivative denoted $\mathscr{U}_{x_i} < 0$).

The government taxes the n incomes according to the nonlinear tax schedule:

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

Consumption is $c = \sum_{i=1}^{n} y_i - \mathcal{T}\left(y_1, ..., y_n\right)$. We denote the marginal rate of substitution between a $\mathbf{w}$-taxpayer's supply of input x_i and her 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 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.[7] A $\mathbf{w}$-taxpayer chooses her supply of inputs $\mathbf{x}$ to solve:

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

This is equivalent to choosing incomes $\mathbf{y}$ to solve:

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

[6] Taxing inputs is equivalent to re-scaling the income tax function $\mathcal{T}(\cdot)$.
[7] $A_{i,j}$ is a term of matrix A for which the row is i and the column is j.

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 supplies of inputs denoted by $\mathbf{X}(\mathbf{w}) = (X_1(\mathbf{w}), ..., X_n(\mathbf{w}))$ and incomes denoted by $\mathbf{Y}(\mathbf{w}) = (Y_1(\mathbf{w}), ..., Y_n(\mathbf{w}))$ where $Y_i(\mathbf{w}) = p_i\, X_i(\mathbf{w})$. Aggregating the individual supplies of input $X_i(\mathbf{w})$, we obtain its total amount $\mathcal{X}_i$ used in the production process, i.e. $\mathcal{X}_i \stackrel{\text{def}}{\equiv} \int_{\mathbf{w}\in W} X_i(\mathbf{w})\, f(\mathbf{w}) d\mathbf{w}$. The utility achieved by the taxpayers is $U(\mathbf{w}) = \mathscr{U}\,(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w})$ and the first order-conditions are:

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

where $C(\mathbf{w}) = \sum_{i=1}^{n} Y_i(\mathbf{w}) - \mathcal{T}\,(\mathbf{Y}(\mathbf{w}))$. For each kind $i = 1, ..., n$ of income, the left-hand side is the marginal net-of-tax rate of the i^{th} income. It corresponds to the marginal gain, in terms of after-tax income, of the i^{th} pretax income y_i. The right-hand side is the marginal rate of substitution between supply of input x_i and after-tax income. It corresponds to the marginal cost of supplying the i^{th} pretax income, in monetary terms.

II.3 Equilibrium

Our concept of equilibrium is the following:

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

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

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

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

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

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 represent and how to reinterpret utility function $\mathscr{U}\,(c, \mathbf{x}; \mathbf{w})$.

Example 1: The two-period model

It is useful to begin with an intertemporal setting in order to focus on capital taxation. The literature has largely emphasized the relevance of the two-period model (Atkinson and Stiglitz (1976), see also Boadway (2012, Chapter 3) for a nice survey and Farhi and Werning (2010) for the reinterpretation of this model to estate taxation). The two-period model captures the essential mechanisms by which taxation affects individual behaviors in most macroeconomic models. Capital is accumulated thanks to savings, at the cost of foregone consumption. Savings and foregone consumption are crucial for the characterization of the steady state in the neo-classical growth model (Ramsey, 1928) and of the steady-state(s) in the overlapping generation model (Diamond, 1965). The two-period model encapsulates these two determinants. Furthermore, models in the vein of in Chamley (1986) and Judd (1985) rely on an infinite elasticity of supply of capital. In our framework, the elasticity of supply of any input with respect to its tax rate can take any value, including infinity.

We denote the first period by a and the second period by b. Taxpayers are characterized by $\mathbf{w} = (w_1, w_2)$ where w_1 is their individual labor productivity and w_2 is their initial wealth, e.g. their inherited wealth. In the first period, $\mathbf{w}$-taxpayers inherit w_2, save x_2 and consume $c_a = w_2 - x_2$ In the second period, taxpayers work and earn capital income $y_2 = p_2\, x_2$ where p_2 is the (endogenous) return on savings. Depending on their productivity w_1, they supply x_1 efficient units of labor so that their labor income is $y_1 = p_1\, x_1$. Their consumption in the second period is the sum of both their capital and labor incomes minus taxes $T(y_1, y_2)$, i.e. $c_b = y_1 + y_2 - T(y_1, y_2)$. This corresponds to our definition of after-tax income c in the general framework. Let $(c_a, c_b, x_1) \mapsto \mathcal{U}(c_a, c_b, x_1; w_1)$ be the preferences of $\mathbf{w}$-agents over first period consumption c_a, second period consumption c_b and efficient units of labor x_1. From this lifetime utility, we retrieve the utility function of the general framework, using the following change of variables:

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

Example 2: The income-shifting model

Our framework can be consistent with many forms of income-shifting. 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 $\mathcal{U}_d > 0 > \mathcal{U}_{z_1}, \mathcal{U}_{z_2}$. For instance, taxpayers can be self-employed and business-owners with z_1 for their effective labor income and z_2 for the return on their business. Their labor productivity is w_1 and their ability to generate return on their business is w_2.

With some monetary cost $S(\sigma; \mathbf{w})$, taxpayers can shift an amount of income $\sigma \gtrless 0$ 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 income $c = y_1 + y_2 - T(y_1, y_2)$ 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:

$$\mathscr{U}(c, x_1, x_2; \mathbf{w}) \overset{\text{def}}{\equiv} \max_{\sigma} \quad \mathcal{U}\left(\underbrace{c - S(\sigma;\mathbf{w})}_{=d}, \underbrace{x_1 + \sigma}_{=z_1}, \underbrace{x_2 - \sigma}_{=z_2}; \mathbf{w}\right) \tag{9}$$

The indirect utility function associated to this program lets us return to our general framework.

The second kind of income y_2 can also be invested in tax heavens which implies no tax revenue for the domestic government. This can easily be taken into account in the government's budget constraint, without modifying our general framework.

II.5 Government

The government is a Stackelberg leader. When it chooses the tax policy, it knows how its choice impacts the above-defined equilibrium. It 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. The government maximizes 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 general specification includes many different social objectives. The objective is utilitarian when $\Phi(U, \mathbf{w}) = U$ and weighted utilitarian when $\Phi(U, \mathbf{w}) = \gamma(\mathbf{w})\ U$. One obtains maximin or the maximization of tax revenue when $\gamma(\mathbf{w})$ equal zero for every taxpayer except those with the lowest utility level. When $\Phi(U, \mathbf{w})$ does not depend on type and is concave in U, one has Bergson-Samuelson preferences.

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

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

where Lagrange multiplier $\lambda > 0$ represents the social value of public funds. We choose to express the Lagrangian in monetary units instead of utility units.

II.6 Taxation regimes

We now define the comprehensive income tax, the schedular income tax and a mix of those two which we call the mixed income tax system.[8]

Comprehensive Income Tax system

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

$$\mathcal{T}(\mathbf{y}) = T\left(\sum_{k=1}^{n} y_k\right) \tag{13}$$

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

$$1 - T'\left(\sum_{k=1}^{n} Y_k(\mathbf{w})\right) = \frac{\mathcal{S}^1\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)}{p_1} = \ldots = \frac{\mathcal{S}^n\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)}{p_n} \tag{14}$$

Intuitively, the comprehensive tax system does not differentiate between distinct sources of income hence, it does not distort how taxpayers allocate their effort across 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. It does not depend on taxation.

Schedular Income tax system

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

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

where the $T_k(\cdot)$ schedules are defined on $\mathbb{R}_+$. Since tax $T_k(\cdot)$ is specific to 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 $i \neq j$). Therefore, the first-order conditions (6) become:

$$\forall i \in \{1, \ldots, n\} \qquad 1 - T_i'\left(Y_i(\mathbf{w})\right) = \frac{\mathcal{S}^i\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right)}{p_i} \tag{16}$$

From (16), one sees that the marginal tax rate on one kind of income does not depend on tax rates on other incomes. Moreover, taxpayers face incentives to shift income towards the source of income less taxed.[9]

Among the class of schedular tax systems, those with linear $T_i(\cdot)$ are defined as *dual* tax systems.

[8] In a different framework, "mixed taxation" is used to define commodity taxes in the presence of labor income tax, as in Mirrlees (1976). This is distinct from our definition of mixed tax schedule.

[9] A related issue is how distinct members of the same household should be taxed. In France, Luxembourg or Portugal, one has *joint* or *family* taxation. The combined income of the whole family is taxed as one single unit. This corresponds to comprehensive taxation. In Belgium, the Netherlands, Sweden or the UK, each household member is taxed separately with the same schedule. This *individual* taxation is a restrictive form of schedular taxation, e.g. gender-based taxation (Alesina et al., 2011).

Mixed tax system

The *mixed* tax system consists in adding n income-specific tax schedules $T_i(\cdot)$ to a *personal income* tax schedule $T_0(\cdot)$. Personal income is the sum of all incomes with possible deductions. Let $a_i(y_i)$ denote the i^{th} income after deductions, with $0 \leq a_i(y_i) \leq y_i$ and $0 \leq a_i'(y_i) \leq 1$.[10] The net-of-deduction functions $a_i(.)$ are assumed increasing and differentiable. The personal income tax base, personal income hereafter, is equal to $\sum_{k=1}^{n} a_k(y_k)$ and the mixed tax schedule is:

$$\mathcal{T}(\mathbf{y}) = T_0\left(\sum_{k=1}^{n} a_k(y_k)\right) + \sum_{k=1}^{n} T_k(y_k) \tag{17}$$

where:

$$y_0 \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k(y_k) \tag{18}$$

Depending on the value taken by $a_k(y_k)$, the income y_k can be partially or fully deducted from personal income. The specific tax schedule $T_k(\cdot)$ can apply to the amount of income subtracted from personal income. An income that is totally deducted from personal income is taxed via its specific schedule. For instance, in most OECD countries, labor costs borne by employers enter personal income only after payment of social security contributions. In this example, y_1 is labor cost and $a_1(y_1)$ is taxable labor income net of social security contributions. Similarly, when dividends are included in personal income, corporate taxes have been deducted. Pretax profit to the shareholder is y_2 and $a_2(y_2)$ is her (taxable) dividends net of corporate tax. Moreover, there are examples of tax exempt income such as imputed rents or capital gains on main residence. Denote y_3 these examples, one has $a_3(y_3) = 0$.

The mixed tax system encapsulates both the comprehensive and the schedular tax systems as specific cases. When one substitutes $a_1(y) \equiv ... \equiv a_n(y) \equiv y$ and $y_i \mapsto T_i(y_i) \equiv 0$, for all i, into the mixed tax schedule (17), one obtains the comprehensive tax system (13). When one substitutes $y_0 \mapsto T_0(y_0) \equiv 0$ into (17), one obtains the schedular tax system (15).

The j^{th} marginal tax rate is obtained by deriving both sides of (17) with respect to y_j:

$$\mathcal{T}_{y_j}(\mathbf{y}) = T_j'(y_j) + a_j'(y_j)\, T_0'\left(\sum_{k=1}^{n} a_k(y_k)\right). \tag{19}$$

The (effective) marginal tax rate on the j^{th} income is the sum of the marginal income-specific tax rate $T_j'(y_j)$ and of the marginal income-specific net-of-deduction rate $a_j'(y_j)$ times the marginal personal income tax rate $T_0'(y_0)$. The j^{th} marginal tax rate then depends on all incomes through the determination of personal income y_0 in (18).

[10] 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 (19) would be unaffected by such a re-normalization.

III Self-clearing cases

In this section, we present specifications that directly lead to recommend either a schedular or a comprehensive income tax schedule. These specifications are summarized in Propositions 1-3 below. The realism of the cases presented in these propositions is questionable, but they are helpful to emphasize the mechanisms that lead to recommend either a schedular or a comprehensive tax system. Moreover, they allow us to emphasize that recommending one system or the other is far from straightforward, as such a recommendation is possible only in very specific economic conditions and under restrictive assumptions.

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:*

$$\mathscr{U}(c, \mathbf{x}; w) = c - \sum_{i=1}^{n} v^i(x_i; w) \qquad \text{with} \qquad v^i_{x_i}, v^i_{x_i,x_i} > 0 > v^i_w, v^i_{x_i,w} \tag{20}$$

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. Due to the separability in the supply of x_i in the utility function (20), 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). This is made possible with a schedular income tax system.

Note, however, that Proposition 1 relies on the assumption that taxpayers differ along a single dimension, as is standard in the Mirrlees (1971) literature. This is not very convincing empirically, in particular with different kinds of income.

Proposition 2. *When, in two-periods model with endogenous labor supply and savings, the preferences (8) are weakly separable between efficient labor, x_1, and consumption bundles (c_a and c_b), i.e.:*

$$\mathcal{U}(c_a, c_b, x_1; w_1) = \mathsf{U}\left(V(c_a, c_b), x_1; w_1\right) \qquad \text{with} \qquad \mathsf{U}_V, V_{c_a}, V_{c_b} > 0$$

with $V(\cdot)$ twice continuously differentiable and increasing in each argument and when 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, again, 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, zero capital taxation 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.[11] Our proof is similar to the proof Atkinson and Stiglitz (1976)'s theorem by Konishi (1995), Laroque (2005) and Kaplow (2008) but is valid with general tax instruments and multidimensional incomes.[12] The following Proposition is proved in Appendix C.

Proposition 3. *If preferences are weakly separable in x_i, 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 minimize the same aggregation $\mathcal{V}(\cdot)$ of inputs supplies when they choose their inputs supplies to obtain incomes of different kinds. Moreover, the government is only interested in the sum of all incomes earned by each individual. Indeed inputs supplies 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 inputs supplies $\mathbf{x}$. 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 the supplies of two different inputs 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 inputs supplies 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 input. A comprehensive tax schedule is therefore optimal.

[11]This result is in the vein of Armstrong (1996) where the optimal nonlinear multi-product tariff is cost-based.

[12]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.

Under the weakly separable preferences assumed in Proposition 3, people who earn the same taxable income $v = \sum_{i=1}^{n} x_i$ choose the same inputs supplies $(x_1, ..., x_n)$. If different taxpayers earn the same level of one kind of income, they must earn the same levels of every other income, which is not very convincing empirically.[13]

IV Tax reforms

In this section, we characterize how tax reforms impact the equilibrium (see Definition 1). For this purpose, we first study, in Subsection IV.1, the taxpayers' responses to a set of tax reforms and to prices changes. In these responses – which we decompose into income responses, compensated responses and price responses – we account for the simultaneous impacts that a tax reform or a price change may have on multiple incomes. We compute the taxpayers' responses to any possible tax reform by differentiating the 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.[14] We also obtain responses to prices changes.

Second, in subsection IV.2, we characterize, using the firm's demand equations (1), how any micro response to tax reforms has general-equilibrium effects through changes in the prices of inputs. Micro responses modify aggregate inputs supplies. This changes inputs prices through inverse demand equations (1). In turn, it induces 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.

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*

[13]The program solved by individuals of type $\mathbf{w}$ can be decomposed into two consecutive stages. In the first stage, taxpayers choose their inputs supplies $\mathbf{x}$ to earn a given taxable income $v = \sum_{i=1}^{n} x_i$:

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

In the second stage, taxpayers choose their taxable income:

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

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

[14]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 if 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.

the tax reform and I is an open interval containing 0 such that, for all $\mathbf{y} \in \mathbb{R}^n_+$, one has $\widetilde{\mathcal{T}}(\mathbf{y},0) = \mathcal{T}(\mathbf{y})$. Therefore, $\widetilde{\mathcal{T}}(\mathbf{y},0)$ 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 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:

$$\widetilde{U}(\mathbf{w},t) \overset{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,...,y_n)} \quad \mathscr{U}\left(\sum_{i=1}^{n} y_i - \widetilde{\mathcal{T}}(\mathbf{y},t), \frac{y_1}{\widetilde{p}_1(t)}, ..., \frac{y_n}{\widetilde{p}_n(t)}; \mathbf{w}\right) \tag{21}$$

In a similar way, we denote $\widetilde{\mathscr{B}}(t)$, $\widetilde{\mathscr{W}}(t)$ and $\widetilde{\mathscr{L}}(t) \overset{\text{def}}{\equiv} \widetilde{\mathscr{B}}(t) + \frac{1}{\lambda}\widetilde{\mathscr{W}}(t)$, 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 $(\mathbf{y},t) \mapsto \widetilde{\mathcal{T}}(\mathbf{y},t)$.[15]

IV.1.a Behavioral responses

We now explain how the economy adjusts to tax reforms. To do so, we present, for each type of taxpayers $\mathbf{w}$, the responses of each kind of income due to behavioral responses and endogenous prices. As in the case of exogenous prices and a single kind of income (Saez, 2001), any tax reform can imply income and compensated responses. All these responses are *total* responses, 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: When $\mathbf{w}$-taxpayers modify their income, it endogenously creates a change in the marginal tax rate they face so that they further adjust their income.

Income responses

We define the income responses as the behavioral responses to a small change in the tax liability of $\mathbf{w}$-taxpayers. Their tax schedule becomes:

$$\widetilde{\mathcal{T}}(\mathbf{y},\rho) = \mathcal{T}(\mathbf{y}) - \rho, \tag{22a}$$

where ρ measures the magnitude of this *lump-sum* perturbation. Let $\frac{\partial Y_i(\mathbf{w})}{\partial \rho}$ denote how $\mathbf{w}$-taxpayers modify their i^{th} income after this lump-sum tax perturbation. We call $\frac{\partial Y_i(\mathbf{w})}{\partial \rho}$ (where $i = 1, ..., n$) their income responses.

Compensated responses

We now study a tax reform that impacts the individual first-order conditions only through substitution effects. We let $\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}$ denote 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 $Y_j(\mathbf{w})$,

[15] Note that when we define the perturbed Lagrangian $\widetilde{\mathscr{L}}(t)$, we keep $1/\lambda$ at its value before the perturbation. This will appear convenient in Proposition 6 below.

while leaving unchanged the level of tax at the initial incomes $\mathbf{Y}(\mathbf{w})$. That is, after a compensated tax reform, the tax schedule becomes:

$$\widetilde{\mathcal{T}}(\mathbf{y},\tau_j) = \mathcal{T}(\mathbf{y}) - \tau_j\left(y_j - Y_j(\mathbf{w})\right). \tag{22b}$$

where τ_j measures the magnitude of this specific perturbation. The response and reform are said to be *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 $Y_i(\mathbf{w})$ $(i = 1, ..., n)$.

Uncompensated responses

We let $\frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_j}$ denote 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:

$$\widetilde{\mathcal{T}}(\mathbf{y},\tau_j) = \mathcal{T}(\mathbf{y}) - \tau_j\, y_j. \tag{22c}$$

An uncompensated tax reform of size τ_j is the combination of a compensated tax reform of size τ_j with a lump-sum tax perturbation of size $\tau_j\ Y_j(\mathbf{w})$. Therefore, 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:

$$\frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_j} = \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} + Y_j(\mathbf{w})\frac{\partial Y_i(\mathbf{w})}{\partial \rho}. \tag{22d}$$

Price responses

Finally, let $\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}$ denote the price response. It is defined as 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 the individual effort to generate this input, hence the level of aggregate income associated with 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.

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 outcome 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 make $\left.\frac{\partial \widetilde{Y}_i(\mathbf{w},t)}{\partial t}\right|_{t=0}$ explicit. As shown in Appendix D, this leads to the following expression:[16]

[16]We derive (23) using the implicit function theorem thanks to Assumption 3 in Appendix D.

$$\left.\frac{\partial \widetilde{Y}_i(\mathbf{w},t)}{\partial t}\right|_{t=0} = \underbrace{-\sum_{j=1}^{n}\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}}_{\text{Compensated responses}} \underbrace{-\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}}_{\text{Income responses}}$$
$$+\underbrace{\sum_{j=1}^{n}\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0}}_{\text{Price responses}} \tag{23}$$

A tax perturbation affects taxpayers' first-order conditions (6) 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 *income* responses. Third, price responses $\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}$ occur as soon as input prices respond to tax reforms, i.e. $\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0} \neq 0$.

Effects on tax liability

Following Saez (2001), 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:

$$\left.\frac{\mathrm{d}\widetilde{\mathcal{T}}\left(\widetilde{\mathbf{Y}}(\mathbf{w},t),t\right)}{\mathrm{d}t}\right|_{t=0} = \underbrace{\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0}}_{\text{Mechanical effects}} + \underbrace{\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\left.\frac{\partial \widetilde{Y}_i(\mathbf{w},t)}{\partial t}\right|_{t=0}}_{\text{Behavioral effects}}. \tag{24}$$

The first term in the right-hand side of (24) is the mechanical effect of the tax reform. It consists in the mechanical change in individual tax liability when one assumes inputs prices and individual decisions are constant. The second term captures the behavioral effects of the reform. Behavioral responses modify the levels of the different incomes. Each modification of a specific income implies a change in tax liability. This change is proportional to the marginal tax rate $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))$ that prevails for this income.

We plug Equation (23) into (24) in order to rewrite the impact of a tax perturbation as the effects induced by the changes in tax liabilities (i.e. mechanical effects and income effects), those induced by the changes in marginal tax rates (compensated effects) and the log changes in prices.

$$\left.\frac{\mathrm{d}\widetilde{\mathcal{T}}\left(\widetilde{\mathbf{Y}}(\mathbf{w},t),t\right)}{\mathrm{d}t}\right|_{t=0} = \left[1-\sum_{i=1}^{n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right]\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \tag{25}$$
$$-\sum_{1\le i,j\le n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} + \sum_{1\le i,j\le n}\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0}.$$

Effects on welfare

The marginal social welfare weight, expressed in monetary units, for $\mathbf{w}$-taxpayers, is defined as:

$$g(\mathbf{w}) \stackrel{\text{def}}{\equiv} \frac{\Phi_U(U(\mathbf{w});\mathbf{w})\,\mathscr{U}_c\left(C(\mathbf{w}),\mathbf{X}(\mathbf{w});\mathbf{w}\right)}{\lambda} \tag{26}$$

It is the social value of giving one extra unit of consumption to a **w**-taxpayer, assuming constant prices.

The following equation gives, in monetary terms, the effects of a tax reform on the social welfare of a **w**-taxpayer:

$$\frac{1}{\lambda} \left. \frac{\partial \Phi\left(\widetilde{U}(\mathbf{w},t);\mathbf{w}\right)}{\partial t} \right|_{t=0} = \left(- \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} + \sum_{j=1}^{n} \left(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w}) \left. \frac{\partial \log \widetilde{p}_j(t)}{\partial t} \right|_{t=0} \right) g(\mathbf{w}). \quad (27)$$

The proof, in which we apply the envelope theorem to the individual's maximization program (21), is relegated to Appendix D. Changes in utility are driven by the mechanical effect in tax liability $- \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0}$ and by the effects of reform-induced changes in prices on utility. Taxpayers' decisions conversely do not show up in (27). Indeed, the taxpayers' decisions are perturbed from their optimum and they are indifferent to small changes in their decisions to a first-order approximation. This envelope argument is well understood since Saez (2001). It however does not apply to prices changes since 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 consumption increase of $(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w})))Y_j(\mathbf{w})$. Multiplying the mechanical effect and the effects of prices changes on utility by the welfare weight $g(\mathbf{w})$ leads to the right-hand side of (27).

Effects on government's Lagrangian

We can now give the impact of a tax reform on the government's Lagrangian (12). We sum, across all types **w**, the impact on their tax liability (25) and on their welfare (27). This yields:

$$\begin{aligned} \frac{\partial \widetilde{\mathscr{L}}}{\partial t} = & \int_{\mathbf{w}\in W} \left\{ \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \right. \\ & - \sum_{1\le i,j\le n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left. \frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \\ & \left. + \sum_{j=1}^{n} \left[\left(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w})\, g(\mathbf{w}) + \sum_{i=1}^{n} \mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} \right] \left. \frac{\partial \log \widetilde{p}_j(t)}{\partial t} \right|_{t=0} \right\} f(\mathbf{w}) d\mathbf{w}. \end{aligned} \quad (28)$$

Equation (28) allows one to conclude whether a given tax reform is socially desirable. It is entirely expressed in terms of sufficient statistics and social welfare weights.

IV.2 General equilibrium

We now derive the impact of a tax reform on the equilibrium (see Definition 1).

Exogenous prices and micro responses

A tax reform impacts the general equilibrium because it impacts the decisions of taxpayers. When one ignores the effects of the reform on the prices of inputs, the taxpayers responses are called *micro responses*. Combining Equations (7) and (23) where one puts to zero the price responses term

$\sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} \left. \frac{\partial \log \widetilde{p}_j(t)}{\partial t} \right|_{t=0}$, the micro responses of the i^{th} aggregate income to a tax reform are defined by:

$$\left. \frac{\partial \widetilde{\mathcal{Y}}_i(t)}{\partial t} \right|_{t=0}^{Micro} = -\int_{\mathbf{w}\in W} \left\{ \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left. \frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} + \sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left. \frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t} \right|_{t=0} \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \quad (29)$$

Endogenous prices

With endogenous prices, the micro responses of aggregate incomes $\mathcal{Y}_i(t)$ modify all input levels, thereby the marginal product of each input, and eventually the inputs' prices, according to aggregate input demand equations (1). In turn, each taxpayer responds to these price changes according to (23). Therefore, all aggregate incomes $(\widetilde{\mathcal{Y}}_1(t), ..., \widetilde{\mathcal{Y}}_n(t))$ are impacted, which in turn feeds back into the prices, which further impacts the taxpayers' incomes, and so on. The key to incorporate this infinite sequence – hence general equilibrium effects– into the tax formulas consists of using the fixed-point conditions in the adjustment of prices that one combines with the matrices of inverse demand elasticities with respect to prices and the matrices of aggregate income elasticities with respect to prices.

Fixed-point conditions

At equilibrium, according to Definition 1, for each reform's magnitude t, prices $(\widetilde{p}_1(t), ..., \widetilde{p}_n(t))$ have to verify the following fixed-point conditions:

$$\forall t, \forall j \in \{1, ..., n\} \qquad \widetilde{p}_i(t) = \mathcal{F}_{\mathcal{X}_i} \left(\frac{\widetilde{\mathcal{Y}}_1(t)}{\widetilde{p}_1(t)}, ..., \frac{\widetilde{\mathcal{Y}}_n(t)}{\widetilde{p}_n(t)} \right). \quad (30)$$

Matrices of inverse demand elasticities and of aggregate income elasticities

Let Ξ denote the matrix of inverse demand elasticities. 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 $\mathcal{X}_j$:

$$\Xi_{i,j} \stackrel{\text{def}}{\equiv} \frac{\mathcal{X}_j \, \mathcal{F}_{\mathcal{X}_i \mathcal{X}_j}}{\mathcal{F}_{\mathcal{X}_i}}. \quad (31a)$$

Let Σ denote the matrix of the 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:

$$\Sigma_{i,j} \stackrel{\text{def}}{\equiv} \left. \frac{\partial \log \mathcal{Y}_i}{\partial \log p_j} \right|_{t=0} = \frac{1}{\mathcal{Y}_i} \int_{\mathbf{w}\in W} \frac{\partial Y_i(\mathbf{w})}{\partial \log p_j} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \quad (31b)$$

The percentage change in aggregate income i when price p_j changes is made of the sum, across taxpayers, of the percentage changes in the individual incomes i generated by all taxpayers when p_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-differentiates (30). Thanks to Assumption 1, Equation (30) is invertible and one can apply the Implicit Functions 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 inputs are sufficiently high.

Macroeconomic price spillovers

In Appendix D, we calculate how a tax reform impacts prices. We obtain the following equation:

$$\left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0} = \sum_{i=1}^{n} \Pi_{j,i} \left.\frac{\partial \widetilde{\mathcal{Y}}_i(t)}{\partial t}\right|_{t=0}^{Micro} \quad where: \; \Pi = (I_n + \Xi - \Xi \cdot \Sigma)^{-1} \cdot \Xi \cdot \begin{pmatrix} \frac{1}{\mathcal{Y}_1} & 0 & 0 \\ 0 & ... & 0 \\ 0 & 0 & \frac{1}{\mathcal{Y}_n} \end{pmatrix} \tag{32}$$

with Matrix $[A]^{-1}$ the inverse of matrix A. In (32), the micro responses $\frac{\partial \widetilde{\mathcal{Y}}_i(t)}{\partial t}\Big|_{t=0}^{Micro}$ multiply the matrix of price multipliers Π that describes how micro responses translate into a log-change of prices (see Appendix D). Each argument $\Pi_{j,i}$ of this matrix provides the relative change in the j^{th} price induced by an aggregate micro response of the i^{th} income to any tax reform. The general equilibrium loops from any reform are captured by Matrix Π. As one can see from the expression of Π in (32), it incorporates the general equilibrium interactions between input markets, thanks to the matrix of inverse inputs' demand elasticities Ξ and the matrix of aggregate income elasticities Σ.[17] Importantly, it does not depend on the tax reform that creates the micro responses. Thanks to the price multipliers, the incidence of any tax reform on welfare and on government's revenue, in general equilibrium, is straightforward to derive. It simply requires the estimate of a new sufficient statistic.

For each type $i \in \{1, ...n\}$ of income, we now define this new sufficient statistic, μ_i, that we call *macroeconomic price spillover* statistic. It indicates the impact on the Lagrangian (12) of a micro increase in the i^{th} income, through prices changes. The term *macroeconomic* emphasizes the general equilibrium mechanisms and, overall, that μ_i does depend neither on a particular tax reform nor on a specific type of taxpayers. The term *price spillover* stresses that firms and taxpayers' responses impact prices. We define for any $i \in \{1, ..., n\}$:

$$\mu_i \stackrel{\text{def}}{\equiv} \sum_{j=1}^{n} \Pi_{j,i} \int_{\mathbf{w} \in W} \left[\left(1 - \mathcal{T}_{y_j}(\mathbf{Y}(\mathbf{w}))\right) Y_j(\mathbf{w})\, g(\mathbf{w}) + \sum_{k=1}^{n} \mathcal{T}_{y_k}(\mathbf{Y}(\mathbf{w})) \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j} \right] f(\mathbf{w}) d\mathbf{w}. \tag{33}$$

From (32), the price multipliers $\Pi_{j,i}$ capture how any (micro) change in aggregate income i result in changes in price p_j, in general equilibrium. According to the last line of (28), the integral in (33) corresponds to the impact on the Lagrangian of a one-percent increase in the j^{th} price.

To calibrate the spillover statistics μ_i, one can use the values taken by the social welfare weights $g(\mathbf{w})$, by the behavioral responses of taxpayers to variations of prices as well as the values of the

[17] As a limit case, under the linear production function (2), matrix Ξ simplifies to the nil matrix according to (31a), in which case the price multipliers $\Pi_{j,i}$ are also nil and the process of prices' adjustments vanishes.

price multipliers. The latter can be obtained using the inverse demand elasticities and the matrix of aggregate responses to prices' variations. Moreover, as we detail in Appendix D, rather than estimating the behavioral responses of taxpayers to variations of prices, one can apply the Implicit Functions Theorem to the first-order conditions of the taxpayer's maximization program and easily calibrate a resulting expression, Equation (69c), which can be found in Appendix D. This is less demanding empirically. We follow this method in our numerical exercise.

Effects on government's Lagrangian in general equilibrium

Plugging Equations (29), (32) and (33) into (28) yields the impact of a tax reform on the government Lagrangian formulated as:

$$\begin{aligned}\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = & \int_{\mathbf{w}\in W} \left\{ \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} \left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) + \mu_i\right) \frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right] \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \right. \\ & \left. - \sum_{1\leq i,j\leq n} \left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) + \mu_i\right) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}_{y_j}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \right\} f(\mathbf{w})\mathrm{d}\mathbf{w}. \qquad (34)\end{aligned}$$

This tax formula can be used to evaluate the impact of any tax reform in terms of tax revenue and welfare. It is expressed as a function of behavioral responses, spillover statistic μ_i and other sufficient statistics. We detail its numerical implementation in Section VI. We now provide economic intuitions for each of its terms.

Absent any behavioral response, the tax reform mechanically impacts the government tax receipts and social welfare as reflected by $1 - g(\mathbf{w})$ in the first line of Equation (34). Then, for each $\mathbf{w}$-taxpayer and each type of income y_i, behavioral responses and price responses have to be taken into account. First behavioral (income 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 modify prices along the process described in Equation (32). The macro spillover statistics, defined in Equation (33), then indicate how an increase in aggregate income $\mathcal{Y}_i$ impacts the government Lagrangian through changes in prices. Incorporating these price spillover effects amounts to correcting (i.e. increasing or decreasing) the marginal tax rates $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))$. When $\mu_i > 0$, the government Lagrangian and aggregate income $\mathcal{Y}_i$ vary in the same direction, via the above-mentioned process. Ceteris paribus, when $\mu_i > 0$, one should decrease the marginal tax rate on income y_i. Indeed, a reduction of $\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w}))$ increases, in turn, aggregate income $\mathcal{Y}_i$ (through general equilibrium effects) hence, the Lagrangian. Similarly, when $\mu_i < 0$, the aggregate income $\mathcal{Y}_i$ and the Lagrangian vary in opposite direction due to endogenous prices. This pushes the corresponding marginal tax rate upwards in order to reduce the corresponding aggregate income (through general equilibrium effects).

More intuition on μ_i can be obtained in the classical general equilibrium context, with labor (indexed by 1) and capital (indexed by 2). For simplicity, consider tax revenue and social welfare are mainly fed by labor income y_1, which is empirically plausible. Lowering marginal tax rates on capital income encourages the supply of more capital. Assuming capital and labor are imperfect substitutes with $\mathcal{F}_{KL} > 0$, this raises the marginal productivity of labor hence, its price p_1. These spillover effects improve workers' well-being and increases tax revenue obtained from labor income. One therefore expects a positive macro spillover statistic for capital, $\mu_2 > 0$ in order to correct downwards the marginal tax rate on capital income $\mathcal{T}_{y_2}(\mathbf{Y}(\mathbf{w}))$.[18] Symmetrically, when one lowers marginal tax rates on labor income, it fosters incentives to work, hence aggregate labor income. In turn, it reduces the marginal productivity of labor, hence the price of labor p_1. This partially crowds out the impact of the initial reduction of labor income marginal tax rates on incentive to work and on workers' well-being. One thus 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, will be limited.

In an economy where labor is the main source of income, as in France for instance, the aggregate amount of labor income is larger than the aggregate amount of capital income. Therefore, a small change in aggregate capital income has a stronger impact on the prices of labor and capital, hence on the Lagrangian, than a small change in aggregate labor income. In this economy, consider any tax reform that slightly increases aggregate capital income $\mathcal{Y}_2$. In general equilibrium, p_2 strongly decreases and p_1 slightly increases hence, aggregate labor income $\mathcal{Y}_1$ rises. In turn, the decrease of p_2 reduces aggregate capital income, which only slightly countervails its initial increase. With a large share of labor income in the economy, the reform that increases $\mathcal{Y}_2$ is beneficial to social welfare and tax revenue. Therefore, one expects μ_2 to be larger than μ_1, in absolute values. One can also expect $\mu_2 > 0$ (which plays hand in hand with tax reforms that create an increase of aggregate capital income, that is, a reduction of capital taxation). Moreover, one can also expect that a lower elasticity of substitution between capital and labor increases these trickle-down effects and reduces capital taxation. Although these results are obtained under simplifying assumptions, the mechanisms we highlight are more general and will help us understand the values obtained in the numerical simulations.

Effects on Government's Lagrangian of balanced tax reforms

With Equation (34), policy advisers can determine the effects of any tax reform. However, a tax reform is not budget-balanced unless $\left.\frac{\partial \widetilde{\mathscr{B}}}{\partial t}\right|_{t=0} = 0$. It is very important 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). We now

[18] The rise in the supply of capital also reduces the marginal productivity of capital p_2. This is detrimental to tax revenues obtained from capital income. It also reduces the well-being of capital owners. However, provided that tax revenue and social welfare are mainly fed by labor income, the benefits obtained from the increase in p_1 more than offset these negative effects on the Lagrangian.

characterize the impact, in terms of welfare, of combining any tax reform with a lump-sum transfer such that their combination is budget-balanced.

We normalize the social value of public funds λ to ensure that the cost in terms of tax revenue implied by the lump-sum tax reform defined in (22a) is offset by the gains in terms of welfare. This implies that the lump-sum reform has no impact on the government's Lagrangian. Applying Equation (34) to the lump-sum reform (22a), the social value of public funds is pined down by:[19]

$$0 = \int_{\mathbf{w}\in W} \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} \left(\mathcal{T}_{y_i}(\mathbf{Y}(\mathbf{w})) + \mu_i \right) \frac{\partial Y_i(\mathbf{w})}{\partial \rho} \right] f(\mathbf{w}) \mathrm{d}\mathbf{w}, \tag{35}$$

with λ included into $g(\mathbf{w})$, see Equation (26). That amounts to the following proposition.

Proposition 4. *If the social value of public funds λ verifies (35) and if $\left.\frac{\partial \widetilde{\mathscr{L}}}{\partial t}\right|_{t=0}$ defined in (34) 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 4, 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 tax formula (34), 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.

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 (17) and (18). Reducing the n-dimensional tax schedule $\mathbf{y} \mapsto \mathcal{T}(\mathbf{y})$ to the sum of $n+1$ one-dimensional schedules $y_i \mapsto T_i(y_i)$ enables us to bypass the technical difficulties inherent to multidimensional screening, while keeping the overall tax system very flexible and realistic. 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 nonlinear and linear tax 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.

[19] We assume that:

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

i.e. that a positive lump-sum transfer to taxpayers reduces government's tax revenue, despite income responses. Otherwise, a lump-sum transfer would simultaneously increase taxpayers' well being and the government's revenue so that the initial economy would be Pareto-dominated.

V.1 Optimal mixed tax schedules

Effect of tax reforms on personal income

We first describe income responses, compensated responses, uncompensated responses and price responses (described in Subsection IV.1.a) of the personal income y_0 defined in (18), when the tax schedule takes the mixed form (Equation (17)). We combine the taxpayers' responses (22a)-(22c) and the changes in prices induced by general equilibrium described in (32) with the definition of personal income in (18). For any reform, the impact on personal income y_0 consists in the weighted sum of the induced changes in each specific income k, each income change being weighted by its marginal net-of-deduction rate $a_k'(y_k)$. Formally, the income response is:

$$\frac{\partial Y_0(\mathbf{w})}{\partial \rho} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k'(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}. \tag{36a}$$

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

$$\frac{\partial Y_0(\mathbf{w})}{\partial \tau_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k'(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}. \tag{36b}$$

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

$$\frac{\partial Y_0^u(\mathbf{w})}{\partial \tau_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k'(y_k) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_j}, \tag{36c}$$

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

$$\frac{\partial Y_0}{\partial \log p_j} \stackrel{\text{def}}{\equiv} \sum_{k=1}^{n} a_k'(y_k) \frac{\partial Y_k}{\partial \log p_j}. \tag{36d}$$

Note that, since personal income y_0 does not correspond to any input, and, each input $i = 1, ..., n$ is associated to a specific spillover statistics μ_i, we normalize $\mu_0 = 0$.

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:

$$\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_i(Y_i(\mathbf{w})). \tag{37}$$

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

$$\left.\frac{\partial \widetilde{\mathcal{T}}_{y_i}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_i'(Y_i(\mathbf{w})). \tag{38}$$

Note that it does not modify the other marginal tax rates. Substituting (37) and (38) into (34), we obtain the effect, on the Lagrangian, of reforming the tax schedule that prevails on the i^{th} income:

$$\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = \int_{\mathbf{w}\in W} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{k=0}^{n} \left(T'_k(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \right] R_i(Y_i(\mathbf{w})) + \left[\sum_{k=0}^{n} \left(T'_k(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i} \right] R'_i(Y_i(\mathbf{w})) \right\} f(\mathbf{w}) d\mathbf{w}. \tag{39}$$

The economic intuition behind Equation (39) is similar to the one we gave for Equation (34). However, Equation (39) is expressed in terms of the $n+1$ marginal tax rates $T'_k(Y_k(\mathbf{w}))$ associated with the $n+1$ one-dimensional schedules $y_k \mapsto T_k(Y_k(\mathbf{w}))$ for $k = 0, ..., n$ 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 income responses $\frac{\partial Y_k}{\partial \rho} R_i(Y_i(\mathbf{w}))$ for all incomes $k \in \{0, ..., n\}$. Compensated 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 prices through changes in every k^{th} input in the production process. The latter channel is taken into account by the macro spillover sufficient statistics μ_k. Aggregating these effects for all types leads to Equation (39). Importantly, Equation (39) also takes into account 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 $T_0(\cdot)$ and we show that it is also given by Equation (39) with $i = 0$. Consider a reform that 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:

$$\left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -R_0(Y_0(\mathbf{w})). \tag{40}$$

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

$$\left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), t)}{\partial t}\right|_{t=0} = -a'_j(Y_j(\mathbf{w}))\, R'_j(Y_0(\mathbf{w})) \tag{41}$$

In Equation (41), the marginal net-of-deduction rate that applies to the j^{th} income shows up when one reforms the personal income tax. This differs from (38), which was obtained from reforming a specific income tax schedule.

Now, according to (19), 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 times the marginal net-of-deduction 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 net-of-deduction rates a_j':

$$\forall i \in \{0, ..., n\} \qquad \frac{\partial Y_i}{\partial \tau_0} = \sum_{j=1}^{n} a_j'(Y_j(\mathbf{w})) \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}. \tag{42}$$

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 (39) with $i = 0$, as shown in Appendix H. Equation (39) therefore summarizes the first-order effects, on the government's Lagrangian, of a reform of both the personal income tax and of a specific income tax.

Optimal specific and personal income tax schedules

Under mixed taxation, the optimal tax formulas consist of the optimal marginal tax rates $T'(y_k)$, for each type of income $k = 0, ..., n$, expressed in terms of empirically meaningful sufficient statistics.

The tax schedule specific to the i^{th} income is optimal if its reform implies that the first-order effects on the Government's Lagrangian, described in Equation (39), are nil, 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 (39) 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:

$$\varepsilon_i(y_i) \stackrel{\text{def}}{\equiv} \frac{1 - T_i'(y_i)}{y_i} \overline{\frac{\partial Y_i}{\partial \tau_i}}\bigg|_{Y_i(\mathbf{w})=y_i}. \tag{43}$$

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 (36b) and (42) allows us to define this elasticity as:[20]

$$\varepsilon_0(y_0) = \frac{1 - T_0'(y_0)}{y_0} \overline{\frac{\partial Y_0}{\partial \tau_0}}\bigg|_{Y_0(\mathbf{w})=y_0} = \frac{1 - T_0'(y_0)}{y_0} \sum_{1 \le i,j \le n} \overline{a_i'(Y_i(\mathbf{w}))\, a_j'(Y_j(\mathbf{w}))\, \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}}\bigg|_{Y_0(\mathbf{w})=y_0}. \tag{44}$$

[20] 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 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 marginal net-of-deduction rates $a_i'(Y_i(\mathbf{w}))$ and $a_j'(Y_j(\mathbf{w}))$.

To write down the optimal income tax formulas, one needs the income densities. To define the latter, 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:

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

Assumption 2 implies that the mapping $\mathbf{y} \mapsto \mathbf{Y}(\mathbf{w})$ is globally invertible.[21]

Optimal linear and nonlinear tax schedules for each income

Proposition 5. *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 (39) is positive (negative).

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

$$\frac{T_i'(y_i) + \mu_i}{1 - T_i'(y_i)} \, \varepsilon_i(y_i) \, y_i \, h_i(y_i) + \sum_{0 \le k \le n, k \ne i} \overline{\left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_i}\Big|_{Y_i(\mathbf{w}) = y_i}} h_i(y_i) \tag{45}$$
$$= \int_{z=y_i}^{\infty} \left\{ 1 - \overline{g(\mathbf{w})}|_{Y_i(\mathbf{w})=z} - \sum_{k=0}^{n} \overline{\left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\Big|_{Y_i(\mathbf{w})=z}} \right\} h_i(z) \mathrm{d}z.$$

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

$$\frac{t_i + \mu_i}{1 - t_i} \int_{\mathbf{w} \in W} \varepsilon_i^u(\mathbf{w}) \, Y_i(\mathbf{w}) \, f(\mathbf{w}) \mathrm{d}\mathbf{w} + \int_{\mathbf{w} \in W} \sum_{k=0, k \ne i}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} f(\mathbf{w}) \mathrm{d}\mathbf{w}$$
$$= \int_{\mathbf{w} \in W} [1 - g(\mathbf{w})] \, Y_i(\mathbf{w}) \, f(\mathbf{w}) \mathrm{d}\mathbf{w}. \tag{46}$$

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] Assume,by contradiction, the existence of two types $\mathbf{w}, \mathbf{w}'$ such 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):

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

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

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 (45) 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. It 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 optimal nonlinear tax formula (45), 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 (45) describes the costs due to the rise in the marginal tax rate and its right-hand side details the gains due to the increase 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 (43)). This response is encapsulated into the first term in the left-hand side of Equation (45). On top of this response, which is already present in Saez (2001), there are 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 show up in the second term of the left-hand side of Equation (45). 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. Composition effects then 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 they also induce prices changes in general equilibrium. All *micro* compensated responses create prices changes which imply responses of taxpayers to these prices changes, and so on. The sufficient statistics that summarize the impact of these price spillover effects, on the Lagrangian, are the 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 all $z \geq y_i$ in the right-hand side of (45). It also creates income responses $\frac{\partial y_i(\mathbf{w})}{\partial \rho}$. Another key difference compared to the Mirrleesian framework is that these income responses not only have a direct impact on the

[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.

Lagrangian by modifying the tax revenue proportionally to marginal tax rates $T'_k(y_k)$, they also create macro price 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 the averaging of the mechanical losses and responses to wealth change across all taxpayers with the same i^{th} income.

We now discuss the determinants of optimal marginal tax rates. First, from (45), we see that 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 who earn an i^{th} income above y_i. These incomes are mechanically impacted by any change in the optimal marginal tax rate on the i^{th} income. This mechanically modifies the welfare weights assigned to taxpayers who earn an i^{th} income above y_i. The optimal marginal tax rate on the i^{th} income at income y_i 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.

The optimal marginal tax rate on the i^{th} income at income y_i is also, ceteris paribus, increasing when the mean compensated elasticity ε_i decreases. The inverse elasticity rule remains valid. From Equations (44) and (45), this implies that the optimal marginal tax rate on *personal* income $T'_0(y_0)$ increases when the incomes that are the most responsive to tax reforms are withdrawn from the definition of personal income. For instance, if the most responsive income is capital income, then, the mean compensated elasticity of personal income ε_0 is lower with a schedular tax on capital income than with a comprehensive tax system. This leads to a more progressive personal income tax schedule with a schedular tax system. This might explain why Scandinavian countries have implemented the dual tax system (Boadway, 2004, Sørensen, 2009) in the early nineties. With dual taxation, they keep the progressivity of their personal income tax schedule despite highly elastic capital incomes.

Equation (45) 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, given by $\Delta Y_k = -\Delta T'_{y_i} \frac{\partial Y_k}{\partial \tau_i}$.[24] 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 $\frac{\partial Y_k}{\partial \tau_i}$, the less costly or the more beneficial is the response of the k^{th} income for the government. From (45), one can then recommend a higher i^{th} optimal marginal tax rate. In particular, lower income-shifting leads to less positive or 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, since the cross-base responses they have in mind come mostly from income-shifting behaviors.

[24] 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.

Moreover, the macro spillover statistics μ_k magnify the compensated responses and the income responses. In particular, a larger macro spillover statistic on the i^{th} income μ_i (which appears in the first term, in the left-hand side of Equation (45)) tends, ceteris paribus, to reduce the optimal marginal tax rate $T_i'(y_i)$. Intuitively, consider a rise in $T_i'(y_i)$. It creates compensated responses that reduce the i^{th} income of the taxpayers concerned by this tax reform. These responses imply a detrimental reduction in tax liability (whenever $T'(y_i) > 0$) hence, a reduction of tax revenue. Moreover, these compensated responses, by decreasing the i^{th} aggregate income $\mathcal{Y}_i$ in turn, impact prices, create spillover effects and modify the government's Lagrangian. A larger μ_i increases the detrimental consequences of these price spillovers implied by the rise of $T_i'(y_i)$. Therefore, the optimal marginal tax rate $T_i'(y_i)$ decreases with μ_i, ceteris paribus.

From (46), we see that, 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. There are however several particularities. First, under a linear tax schedule, income effects and compensated effects can be combined and substituted with the uncompensated responses, as can been verified using the Slutsky Equation (22d). This implies fewer terms in the left-hand side of (46) compared to (45). Second, in the optimal linear tax formula (46), means of sufficient statistics over the whole population appear instead of means of sufficient statistics at a given income level. Last, as expected from the optimal linear tax formula (see e.g. Piketty and Saez (2013a)), 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 personal income?

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 marginal net-of-deduction rate of the i^{th} income $a_i(y_i)$, as follows:

$$\widetilde{\mathcal{T}}(\mathbf{y}, t) = T_0\left(\sum_{k=1}^{n} a_k(y_k) - t\, y_i\right) + \sum_{k=1}^{n} T_k(y_k) \tag{47}$$

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

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

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

$$\widetilde{\mathcal{T}}_{y_i}(\mathbf{y}, t) = \left(a_i'(y_i) - t\right)\, T_0'\left(\sum_{k=1}^{n} a_k(y_k) - t\, y_i\right) + T_i'(y_i). \tag{49}$$

When the tax system becomes more schedular, thanks to an increased deduction of the i^{th} income (i.e. a reduction in $a_i(y_i)$), the personal tax base $y_0(\mathbf{w})$ is reduced by $Y_i\Delta t$. Proposition 6 describes the impact on the Lagrangian and states when this budget-balanced reform is socially desirable.

Proposition 6. *(i) Under a mixed tax schedule, a small reduction of personal income, described by (47), modifies the government's Lagrangian as follows:*

$$\begin{aligned} \left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= \int_{\mathbf{w}\in W} \left\{ \left[(g(\mathbf{w})-1)\; Y_i(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} \right] T_0'(Y_0(\mathbf{w})) \right. \\ &\quad + \left. \sum_{k=0}^{n} (T_k'(Y_k(\mathbf{w})) + \mu_k) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w})\; T_0''(Y_0(\mathbf{w})) \right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \end{aligned} \tag{50}$$

ii) A reform that consists in combining a deduction of the i^{th} income from the taxable income according to (47) (with $t>0$) with a lump-sum transfer that makes the overall perturbation budget-balanced is socially desirable if Equation (50) is positive.

The proof is in Appendix J. A reduction of personal income y_0 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 (50) and other effects are propagated by way of $T_0''(\cdot)$ in the second line of Equation (50). 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,

$$\int_{\mathbf{w}\in W} [g(\mathbf{w})-1] T_0'(Y_0(\mathbf{w})) Y_i(\mathbf{w})\; \Delta t\; f(\mathbf{w}) \mathrm{d}\mathbf{w} \tag{51}$$

that are in the first line of Equation (50). This reduction in tax liability also creates income 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 income responses, that occur for each source of income, modify tax revenue and welfare (due to general equilibrium effects) as follows:

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

where general equilibrium effects are encapsulated into the sufficient statistics μ_k. Moreover, the withdrawal of some income y_i from personal income modifies the marginal tax rate on the i^{th} income $\mathcal{T}_{y_i}(\cdot)$. Indeed, the latter not only depends on $T_i'(\cdot)$ (which is not modified) but also on $T_0'(\cdot)$, as emphasized in Equation (49). 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:

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

Using the Slutsky Equation (22d), the impacts on the government's Lagrangian of the income effects (52) and compensated effects (53) are the same as the effects, on the government's Lagrangian, of any uncompensated tax reform of the i^{th} income with a size, for $\mathbf{w}$-taxpayers, equal to $T_0'(Y_0(\mathbf{w}))\,\Delta t$, i.e.:

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} \; T_0'(Y_0(\mathbf{w})) \, \Delta t \; f(\mathbf{w}) \mathrm{d}\mathbf{w}, \tag{54}$$

that one can see in the first line of Equation (50).

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 (48)) where the curvature of the personal income tax matters as emphasized by $T_0''(Y_0(\mathbf{w}))$. This change induces compensated responses from other sources of income that modify tax revenue and welfare by:

$$\int_{\mathbf{w}\in W} \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w}) \; T_0''(Y_0(\mathbf{w}))\Delta t f(\mathbf{w}) \mathrm{d}\mathbf{w}, \tag{55}$$

an expression that one can find in the last line of (50). Adding Equations (51), (54) and (55), we obtain the effect on the Lagrangian described in (50).

To investigate further the determinants of the optimal net-of-deduction rate, we now assume the i^{th} income is taxed linearly and the linear tax rate on the i^{th} income is optimal, so that it verifies (46). Then, the optimal deduction condition (50) and Proposition 6 simplify to:

Lemma 1. *(i) Under a mixed tax schedule with optimal linear tax rate on income i^{th}, a small reduction of personal income, described by (49), modifies the government's Lagrangian as follows:*

$$\begin{aligned} \left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} &= cov\left[\left(g(\mathbf{w}) - 1\right) Y_i(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i}; T_0'(Y_0(\mathbf{w}))\right] \\ &+ \int_{\mathbf{w}\in W} \left\{\sum_{k=0}^{n} (T_k'(Y_k(\mathbf{w})) + \mu_k) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_0} Y_i(\mathbf{w}) \; T_0''(Y_0(\mathbf{w}))\right\} f(\mathbf{w}) \mathrm{d}\mathbf{w}. \end{aligned} \tag{56}$$

ii) A reform that consists in combining a deduction of the i^{th} income from the taxable income according to (47) (with $t > 0$) with a lump-sum transfer that makes the overall perturbation budget-balanced is socially desirable if Equation (56) is positive.

The proof is in Appendix J. The first line of (56) is the covariance between (i) the marginal tax rate on personal income $T_0'(\cdot)$ and (ii) the contribution of a taxpayer of type $\mathbf{w}$ to the response of the government's Lagrangian when an uncompensated reform of the i^{th} income (22d) occurs. The second line of (56) contains all compensated responses from every source of income due to the fact that the marginal tax rate on every income is modified. With a U-shaped 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.

If all earners of the i^{th} income were facing the same marginal tax rate on personal income $T_0'(\cdot)$, for instance because the personal income tax schedule were piecewise linear and they all lied in the same income bracket, both lines of (56) would be equal to zero. In this very specific situation, further deducing the i^{th} income and compensating the induced revenue loss by a rise in the linear tax rate on the i^{th} income affects the economy only because of the nonlinearity of the personal income tax schedule $T_0(\cdot)$. Essentially, this reform consists in shifting the burden of the taxation of the i^{th} income from those facing a high marginal personal income tax rate $T_0'(\cdot)$ to those facing a low $T_0'(\cdot)$. The first line of Equation (56) states that such a reform is beneficial (detrimental) if taxpayers who face higher marginal personal income tax rates coincide with those whose contribution to an uncompensated tax cut of the i^{th} income on the government's Lagrangian is the more positive (negative). The second line states that such a reform is beneficial if, because of the nonlinearity of the tax schedule, the tax cut induces a rise (a decrease) in the marginal personal income tax rate $T_0'(\cdot)$ if $T_0''(\cdot) < 0$ (if $T_0''(\cdot) > 0$) that itself induces beneficial compensated responses.

VI Numerical illustrations with French taxation

In this section, we numerically implement the optimal mixed tax system and compare it to the optimal dual and optimal comprehensive tax systems. We then illustrate the impact general equilibrium effects and cross-base responses have on these optimal tax profiles. The numerical algorithm is described in Appendix L.

VI.1 Calibration

We first describe how we calibrate the model, i.e. how we select social preferences, individual preferences, the distribution of types and the production function.

To focus on the new responses that occur with different income sources, as highlighted in Section IV, we consider a maximin social objective. This objective shuts down the impact the heterogeneity in social welfare weights $g(\mathbf{w})$ could have on the optimal schedules. Another advantage is that the value it reaches at the optimum is equivalent to the amount of tax revenue raised per head in the economy (Boadway and Jacquet, 2008).

We calibrate our model on French 2017 *Enquête Revenus Fiscaux Sociaux* (ERFS) data. This database merges a part of the French Labor Force survey with a set of variables extracted from the respondents' tax records.[25] The tax unit is a single person with dependents if any or a married couple with dependents if any.

We assume two inputs and regroup the different sources of income that we observe in the administrative tax return data into two categories: labor income, denoted y_1, and capital income, denoted y_2. Capital income is the sum of the different sources of financial incomes included in the personal income tax base, financial incomes and, for the self-employed, $1/3$ of their declared income.[26] Importantly, ERFS data do not contain capital gains and losses. Moreover, we choose to exclude (public) pensions and social transfers since they are exogenous income levels.[27] Since we consider a two-factor production function, we also exclude rents and real estate income. In our data, reported taxable incomes are drastically different from primary incomes. For instance, the total reported capital income is much lower than the primary capital income in France.[28]. However, our aim is to focus on incomes that are currently reported on tax returns in France, and ERFS data perfectly fits this purpose. The sample consists of $27,804$ tax units with positive labor and capital income. Capital (labor) income represents 7.6% (92.4%) of total income in the sample.

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

$$\mathscr{U}(c, x_1, x_2; w_1, w_2) = c - \frac{e_1}{1+e_1} x_1^{\frac{1+e_1}{e_1}} w_1^{-\frac{1}{e_1}} - \frac{e_2}{1+e_2} x_2^{\frac{1+e_2}{e_2}} w_2^{-\frac{1}{e_2}} \tag{57}$$

For the direct elasticities of labor income and capital income, e_1 and e_2, we take the estimates of Lefebvre et al. (2021) based on French tax returns: $e_1 = 0.10$ and $e_2 = 0.65$.[29] This specification of preferences can be reinterpreted in the usual 2 two-period model using Equation (8).

We infer from each observation of labor and capital income the subjacent w_1 and w_2 by inverting taxpayers' first-order condition (6), using the marginal tax rates on labor and capital incomes one has in ERFS (see Appendix L). We then estimate the income density using a biweight kernel with a bandwidth of $89,028$€ .[30]

The production function is a CES:

$$\mathcal{F}(\mathcal{X}_1, \mathcal{X}_2) = \left[A_1\, \mathcal{X}_1^{1-\gamma} + A_2 \mathcal{X}_2^{1-\gamma}\right]^{\frac{1}{1-\gamma}} \tag{58}$$

[25] When we calibrate our model with U.S. Current Population Survey (CPS 2019) data, we find similar results to the ones we describe here.

[26] In the French tax records, incomes of self-employed are declared either as *Benefices Industriels et Commerciaux* (BIC), *Benefices Non Commerciaux* (BNC) or Benefices Agricoles (BA). We choose $1/3$ to approximate the share of these different sources of income that falls under the capital income tax since $1/3$ is the usual and relatively stable national income share for capital income vs labor income.

[27] Pensions in France are heavily reliant on a Pay-As-You-Go system combined with mandatory occupational pensions for employees. Voluntary schemes do exist; however, most companies are very reluctant to offer these in addition to compulsory pensions.

[28] In 2017, in France, declared capital income is only around 40 billion euros while the net operating surplus is 250 billion euros.

[29] An elasticity of capital income larger than the one of labor income is consistent with the findings of Kleven and Schultz (2014) in Denmark.

[30] 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 $x = -1, 0, 1$).

where A_1 and A_2 are the scale parameters of inputs and $1/\gamma$ is the elasticity of substitution between capital and labor. In order to pin down A_1 and A_2, without loss of generality, we normalize prices p_1 and p_2 to 1 in the actual economy (see Appendix L).

In our baseline scenario, we assume away general equilibrium effects (i.e. $\gamma = 0$) and cross-base responses.[31] In alternative scenarii, we relax these assumptions.

In France and in many OECD countries, labor income is included, after possible deductions, into the personal income tax base. We therefore tax y_1 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 consider a flat tax rate on capital income, $T_2(y_2) = t_2\, y_2$. The mixed tax schedule (17) becomes:

$$\mathcal{T}(y_1, y_2) = T_0(y_1 + a_2\, y_2) + t_2\, y_2 \qquad a_2 \in [0,1]. \tag{59}$$

On top of the nonlinear personal income tax schedule $T_0(\cdot)$, the government has two instruments: (i) the capital income net-of-deduction rate a_2, which gives the proportion of capital income one includes into personal income and (ii) the tax rate specific to capital income, t_2. This relatively simple framework is sufficiently rich to encompass, among its sub-cases, the comprehensive tax system when one put $a_2 = 1$ and $t_2 = 0$ in (59) and the schedular tax system when $a_2 = 0$. The schedular tax system is then the dual one since $T_2(y_2) = t_2\, y_2$.

VI.2 Baseline scenario

Figure 1 shows the optimal marginal tax rates on personal income (Equation (45)) under the comprehensive, dual and mixed tax systems. These marginal tax rates follow the usual U-shaped pattern taken by marginal tax rates on labor income (Diamond, 1998). In Figure 1, the optimal marginal tax rates on personal income are described by the red curve with dots under the comprehensive tax system (i.e. the optimal system under the constraints $a_2 = 1$ and $t_2 = 0$), by the blue curve with triangles under the dual tax system (i.e. the optimal system under the constraint $a_2 = 0$) and by the solid black curve under the mixed tax system.

Optimal comprehensive vs dual tax systems

With a comprehensive tax system, the personal income y_0 consists of the rather inelastic labor income y_1 and of the very elastic capital income y_2. Conversely, only labor income enters personal income under dual taxation, $y_0 = y_1$. The mean elasticity ε_0 of personal income y_0 is thus smaller under dual taxation (see (44)) than under the comprehensive tax system, so the optimal marginal tax rates $T_0'(\cdot)$ are higher (see (45)) under dual taxation. The difference is especially important for income levels above 100,000€ since the share of capital income in total income is larger for these

[31] Lefebvre et al. (2021) find zero income-shifting in France following 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.

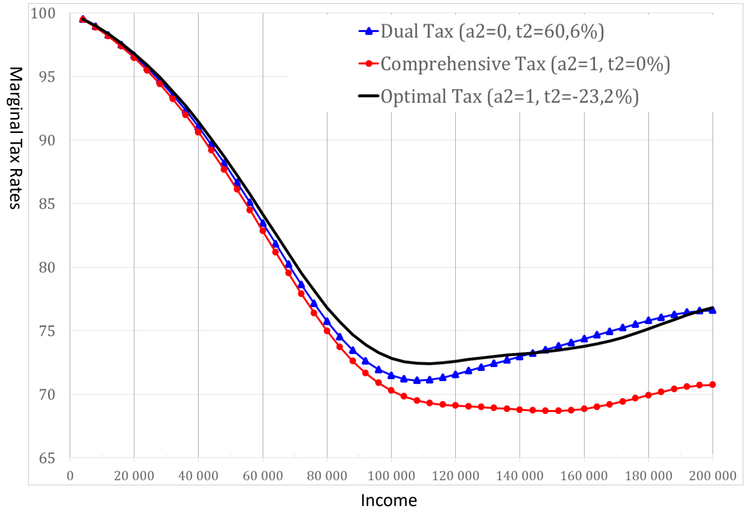

Figure 1: Optimal marginal tax rates $T_0'(Y_0(\mathbf{w}))$ under comprehensive taxation (where $a_2 = 1$ and $t_2 = 0$ by def.), under dual taxation (where $a_2 = 0$ by def., optimal $t_2 = 60.6\%$), and under mixed taxation (optimal $a_2 = 1$, optimal $t_2 = -23.2\%$)

levels of income.[32] The marginal personal income tax rates are therefore close to each other below € 100,000 and start to diverge above this level. In particular, at 200,000€ , the optimal marginal tax rate on personal income is equal to 70.7% under the comprehensive tax system, while it is 5.9 points larger, reaching 76.6%, under the dual tax system. Moreover, from Figure 1 and the fact that the optimal marginal tax rate on labor income is $\mathcal{T}_{y_1}(\mathbf{y}) = T_0'(\cdot)$ under both the comprehensive and dual tax systems, one knows that the comprehensive tax system taxes labor income more weakly than the dual tax system.

Going from the optimal comprehensive tax regime to the optimal dual one, the capital income net-of-deduction rate, a_2, decreases from 1 to 0. When a_2 decreases, all effective marginal tax rates on capital income, provided by Equation (19), decrease because they depend less on the marginal tax rate $T_0'(Y_0(\mathbf{w}))$. This implies a mechanical reduction in tax revenue combined with an increase in capital income. To countervail a negative net impact in terms of tax revenue, increasing t_2 is required. Indeed when one shifts from comprehensive taxation, where $\mathcal{T}_{y_2}(\mathbf{y}) = T_0'(y_1 + y_2)$ with $t_2 = 0$, to dual taxation where $\mathcal{T}_{y_2}(\mathbf{y}) = t_2$, the optimal tax rate on capital income reaches $t_2 = 60,6\%$.[33]

Moreover, in Figure 2, one can see that the U-shaped profile of the effective marginal tax rates on capital income $\mathcal{T}_{y_2}(\mathbf{Y}(\mathbf{w})) = T_0'(Y_0(\mathbf{w})) + a_2\, t_2$ under comprehensive taxation (the red curve with dots) lie, for all income levels, above $60,6\%$. The comprehensive tax system taxes capital income more strongly than does the dual tax system.

In terms of tax revenue, thanks to its intrinsic progressiveness on every source of income, the comprehensive tax regime is performing much better than the dual one, as illustrated in Table 1.

[32] 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 € $100,000$.

[33] In the baseline scenario, obtaining the optimal dual capital tax rate is straightforward. From (46), one obtains $t_2 = 1/(1+e_2) \simeq 60.6\%$.

Shifting from the dual tax system to the comprehensive one brings the non-negligible additional amount of 1108 € per capita.

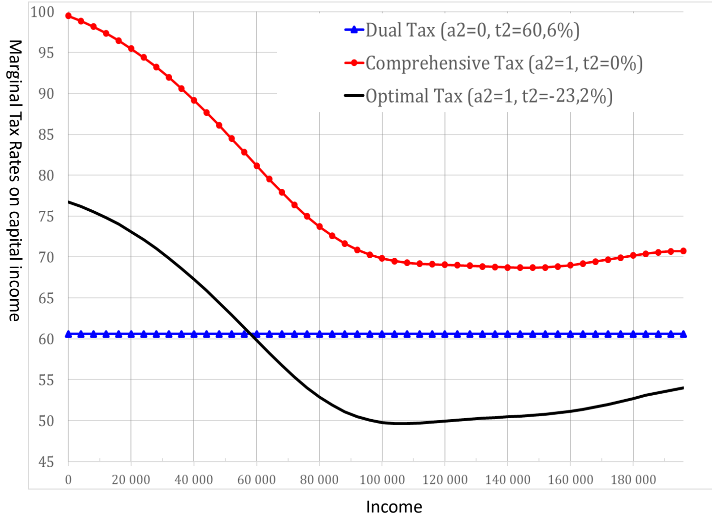

Figure 2: Optimal (effective) marginal tax rates $\mathcal{T}'_{y_2}(\mathbf{Y}(\mathbf{w}))$ on capital income under comprehensive taxation (where $\mathcal{T}'_{y_2}(\mathbf{Y}(\mathbf{w})) = T'_0(Y_1(\mathbf{w}) + Y_2(\mathbf{w}))$), under dual taxation (where $\mathcal{T}'_{y_2}(\mathbf{Y}(\mathbf{w})) = t_2$), and under mixed taxation (where $\mathcal{T}'_{y_2}(\mathbf{Y}(\mathbf{w})) = a_2 T'_0(Y_1(\mathbf{w}) + a_2\, Y_2(\mathbf{w})) + t_2$)

Optimal mixed tax system

All our simulations confirm the conjecture we had in Section III, according to which the optimal mixed tax system is neither dual nor comprehensive. In Figure 1, the three distinct curves of optimal $T'_0(y_0)$ illustrate this outcome. However, all our simulations point to the fact that the optimal mixed system borrows characteristics from both the comprehensive and dual tax systems.

First, the optimal U-shaped schedule under mixed taxation (solid black curve in Figure 1) is very close to the one obtained under dual taxation (blue curve with triangles). This similarity happens despite the fact that the composition of personal income is very different under both systems, $y_0 = y_1$ under dual taxation versus $y_0 = y_1 + a_2 y_2$ under mixed taxation. This means that relatively large effective marginal tax rates on high levels of labor income are made optimal under the optimal mixed tax system, with $\mathcal{T}_{y_1}(\mathbf{y}) = T'_0(y_1 + a_2 y_2)$, as is the case under optimal dual taxation with $\mathcal{T}_{y_1}(\mathbf{y}) = T'_0(y_1)$. In contrast, under comprehensive and mixed taxation, the optimal marginal tax rates $T'_0(\cdot)$ are quite distinct for relatively high income earners. As illustrated in Figure 1, they are up to 6 percentage points larger under mixed taxation (solid black curve) compared to the comprehensive one (red curve with dots). Therefore, for these top earners, the optimal marginal tax rates on labor income, $\mathcal{T}_{y_1}(\mathbf{y})$ are larger under the optimal mixed tax system than under the optimal comprehensive tax system, since $\mathcal{T}_{y_1}(\mathbf{y}) = T'_0(\cdot)$ under both systems.

Second, the optimal mixed regime includes all income sources in the personal tax base at the optimum, i.e. $a_2 = 1$, which replicates a crucial characteristic of comprehensive taxation. This optimal level $a_2 = 1$ is given in the second row of Table 2. From Figure 1, we know that high capital incomes are heavily taxed by the way of T_0 under optimal mixed taxation. This should create important distortions in capital income. However, these distortions are mitigated thanks to a subsidy for capital income with the optimal marginal tax rate on capital, t_2, equal to -23.2%, as shown in the fourth row of Table 2. Thanks to this negative tax rate on capital income, the optimal marginal tax rates on capital income $\mathcal{T}_{y_2}(\mathbf{y}) = a_2 T_0'(y_1 + y_2) + t_2$ have a similar U-shaped profile than under the comprehensive tax system but they are drastically lower, as one can see in Figure 2. On the same figure, one also sees that, compared to the optimal dual tax system, under the optimal mixed tax system, marginal tax rates are larger for levels of personal income below $58,000$€ and they are lower beyond. They are up to 17 percentage points larger at the very bottom of the personal income distribution and up to 11 percentage points lower around $100,000$€.

Last but not least, when the mixed tax in (59) is optimized, the social objective increases by an additional 811€ per head, compared to the comprehensive tax system (see Table 1). Compared to the optimal dual tax system, the gain is substantial with 1919€ per head.

Moreover, whatever the scenario – and even when allowing for general equilibrium effects or income-shifting, as we will in the next sections, we always find that the highest tax revenue per head is collected with $a_2 = 1$ and $t_2 < 0$, a striking outcome of optimal mixed tax profiles. The second highest tax revenue per head is always obtained under comprehensive taxation which always dominates dual taxation. This can be seen when comparing the columns of Table 1, row by row.

Calibration	Dual tax	Comprehensive tax	Mixed tax
Baseline scenario	31,345€	32,453€	33,264€
With G.E. effects $1/\gamma = 4$	31,262€	32,442€	33,312€
With G.E. effects $1/\gamma = 2$	31,203€	32,440€	33,448€
With G.E. effects $1/\gamma = 0.67$	31,125€	32,449€	33,449€
With income shifting $\theta = 0.1$	31,290€	32,453€	33,222€
With income shifting $\theta = 0.25$	31,215€	32,453€	33,163€
With income shifting $\theta = 0.5$	31,121€	32,453€	33,084€

Table 1: Value of tax revenue per head

VI.3 Scenario with general equilibrium effects

To investigate how our previous results are affected by the assumption of zero general equilibrium effects, we run simulations with values of $\gamma > 0$ in the production function (58). We here reproduce results with $\gamma = 0.25$, $\gamma = 0.5$ and $\gamma = 1.5$, i.e. with an elasticity of substitution between capital and labor $1/\gamma$ of 4, 2 and 0.67 respectively.

With general equilibrium effects, the optimal mixed tax system still consists in including all capital income in personal income, i.e. $a_2 = 1$, and in a subsidy on capital income $t_2 < 0$. As expected for the French economy, which has a large share of labor income (see Section 5), we find that the optimal

Calibration	Dual system	Optimal system	
	t_2	a_2	t_2
Baseline scenario	60.6%	100%	– 23.2%
With G.E. effects, elast. of substitution 4	59.9%	100%	– 23.9%
With G.E. effects, elast. of substitution 2	59.2%	100%	– 24.7%
With G.E. effects, elast. of substitution 0.67	56.4%	100%	– 27.8%
With income shifting $\theta = 0.1$	62.3%	100%	– 21.7%
With income shifting $\theta = 0.25$	64.5%	100%	– 19.6%
With income shifting $\theta = 0.5$	67.5%	100%	– 17.3%

Table 2: Optimal t_2 under dual taxation and optimal a_2 and t_2 under mixed taxation, w/o and w/ general equilibrium effects ($1/\gamma = 4$, 2 and 0.67) and w/o and w/ income shifting ($\theta = 0.1, 0.25$).

macro spillover statistic on capital income μ_2 is positive and larger than μ_1 in absolute value. For instance, when the elasticity of substitution $1/\gamma = 4$, one has $\mu_2 = 2.3\%$ and $\mu_1 = -0.2\%$. When $1/\gamma = 0.67$, one obtains $\mu_2 = 12.5\%$ and $\mu_1 = -1\%$. Compared to the baseline scenario, general equilibrium effects (and the positive μ_2) make it socially desirable to decrease the tax burden on capital income. As expected from our theoretical outcomes in Section 5, the lower the elasticity of substitution between capital and labor, the stronger the decrease. More precisely, general equilibrium effects increase the subsidy t_2 by 0.7 percentage points when $1/\gamma = 4$ and by 4.7 percentage points when $1/\gamma = 0.67$, as shown in the last column, second to fifth rows, in Table 2.

Under the optimal dual taxation, general equilibrium effects also decrease the tax burden on capital income, with the macro spillover statistics $\mu_2 = 1.79\%$ and $\mu_1 = -0.13\%$ when $1/\gamma = 4$ while $\mu_2 = 10.7\%$ and $\mu_1 = -0.9\%$ when $1/\gamma = 0.67$. The (positive) tax rate t_2 is reduced by 0.7 percentage points when $1/\gamma = 4$ and by 4.2 percentage points when $1/\gamma = 0.67$.

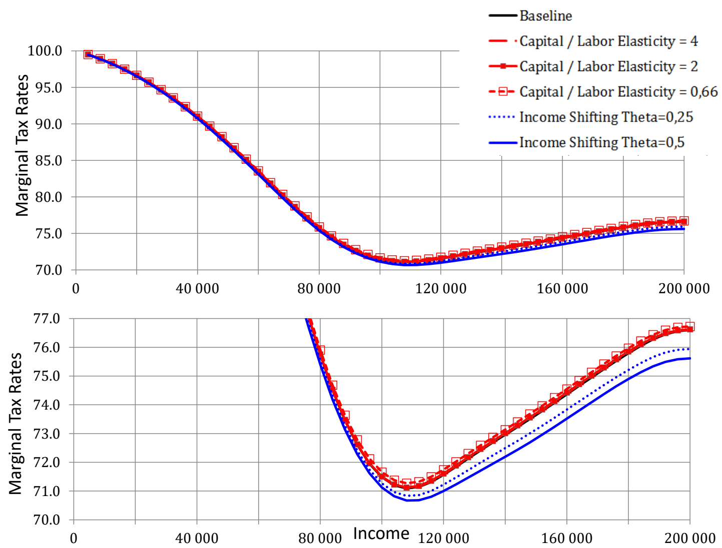

Figure 3: Optimal $T_0'(Y_0(\mathbf{w}))$ under mixed taxation, in the baseline, general equilibrium and income shifting scenarii. In the lower graph, we zoom in the upper part of the income distribution.

In all our simulations, under dual and mixed taxation, the personal income tax schedule T_0' is

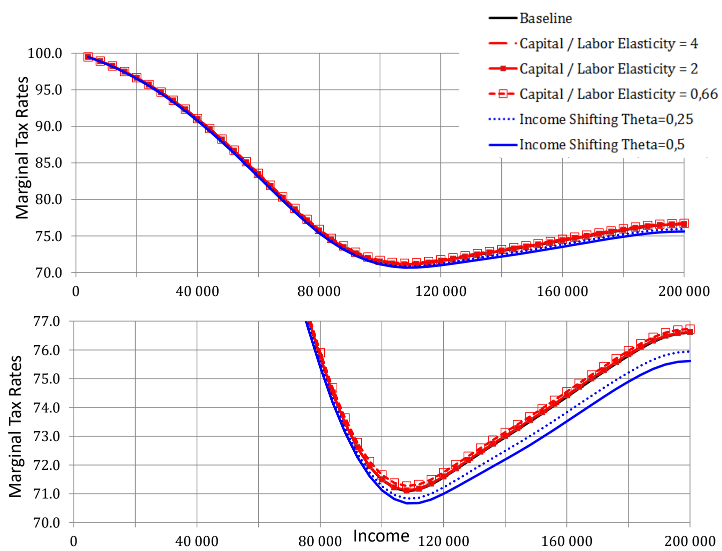

Figure 4: Optimal $T_0'(Y_0(\mathbf{w}))$ under dual taxation, in the baseline, general equilibrium and income shifting scenarii. In the lower graph, we zoom in the upper part of the income distribution.

nearly not impacted by general equilibrium effects. For the optimal mixed tax system, this is visible from the three red curves that represent T_0' with $1/\gamma = 4, 2$ and 0.66, in the upper and lower panels of Figure 3. They are confounded. For the optimal dual tax system, these three curves are also confounded as one can observe in the upper and lower panels of Figure 4. Under dual and mixed tax systems, it is the tax instrument that specifically targets capital income (t_2) which is impacted by general equilibrium effects, as shown in Table 2.

Conversely, under the comprehensive tax system, which has no specific tax rate t_2 on capital income, the marginal tax rates T_0' are modified with general equilibrium effects and so is $\mathcal{T}_{y_2}(\mathbf{y})$. Marginal tax rates slightly decrease for incomes beyond $150,000$€, as illustrated in Figure 5. At 200,000€ instead of 70.7% in the baseline scenario, the optimal marginal tax rate under the comprehensive tax system is 70.2% in the general equilibrium scenario with $1/\gamma = 4$ and it is 69% when $1/\gamma = 0.67$ i.e. 0.5 and 1.8 percentage points lower. When $1/\gamma = 4$, we obtain macro spillover statistics $\mu_2 = -4.9\%$ and $\mu_1 = 0.5\%$. When $1/\gamma = 0.66$, the macro spillover statistics are $\mu_2 = -17.6\%$ and $\mu_1 = 1.6\%$. In absolute values, μ_2 is always larger than μ_1, as expected from theory due to the larger share of labor income in the economy. We observe that, with comprehensive taxation, these statistics have the opposite signs to the ones obtained under the dual and mixed tax systems. The negative sign of μ_2 means that when one reduces capital taxation to stimulate aggregate capital income, general equilibrium effects negatively impact the Lagrangian. The induced reduction of capital price p_2 reduces tax revenue more than the implied increase of p_1. This pleads in favor of higher marginal tax rates on capital incomes. In contrast, the positive sign of μ_1 means that when one reduces taxation on labor income to boosts aggregate labor income, general equilibrium effects (concretely, the implied reduction of p_1 and increase of p_2 and their consequences) positively impact the Lagrangian. This pleads in favor of reducing marginal tax rates on labor income. Under comprehensive taxation, the only way to reduce the tax burden on capital income and the tax burden on labor income consists in modifying the personal income tax schedule $T_0(\cdot)$. When one modifies $T_0'(\cdot)$, one then modifies the corresponding marginal tax rates on both labor and capital income. Our simulations show that, because these opposite forces are at play, general equilibrium effects have only a moderate impact on the optimal comprehensive tax schedule. As shown in Figure 5, when general equilibrium effects are taken into account, only the optimal tax rates beyond 120,000€are reduced and the U-shape of the $T_0'(\cdot)$ schedule is more pronounced.[34]

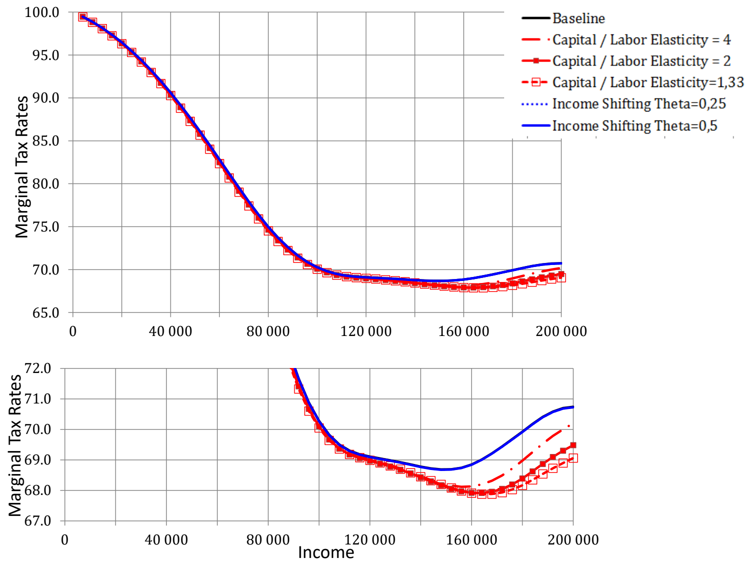

Figure 5: Optimal $T_0'(Y_0(\mathbf{w}))$ under comprehensive taxation, in the baseline, general equilibrium and income shifting scenarii. In the lower graph, we zoom in the upper part of the income distribution.

VI.4 Scenario with cross-base responses

Cross-base responses are a major issue in the policy debate, specifically the possibility of shifting income between the labor and capital bases, see e.g. Saez et al. (2012) and Saez and Zucman (2019). To model income-shifting, we add a quadratic cost of income-shifting to the utility function (57). Using (9) (introduced in the income-shifting model of Section II.4), the utility function is:

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

The amount of shifted income verifies $\sigma = \Gamma(w_1, w_2)(\mathcal{T}_{y_2} - \mathcal{T}_{y_1})$. We calibrate the scale parameter Γ to θ% of the minimum between labor income and capital income, in the no-tax economy. We assume away general equilibrium effects and use the linear technology described by Equation (2)

[34]These findings with a comprehensive income tax are in the vein of those obtained by Sachs et al. (2020) with a single source of income (labor) and nonlinear labor income taxation. In their exogenous-wage optimum the optimal marginal tax rates are U-shaped and they find that endogenous wages imply lower marginal tax rates on intermediate and large labor incomes (see their Figure 4, p.491).

Comprehensive taxation prevents income-shifting. Under comprehensive taxation, when one shifts from the baseline scenario to the income-shifting scenario, the optimal $T_0'(y_0)$ and the value of tax revenue per head remain identical, as one can confirm with Figure 5 and the third column of Table 1.

Under dual and mixed taxation, income-shifting takes place. To reduce it, it is optimal to reduce the discrepancy between the marginal tax rates on capital and on labor incomes, away from their optimal values with no shifting. This occurs via a relatively large increase of t_2 and a relatively small reduction of $T_0'(\cdot)$ for large income levels (where one observes larger proportions of capital income). Moving from the scenario without shifting to one with income-shifting, the optimal U-shaped schedule of $T_0'(\cdot)$ is slightly reduced for income levels beyond $36,000$€ . This reduction is a bit more pronounced beyond $88,000$€ as one can see in the lower graphs of Figures 3 and 4. For instance, under mixed taxation, with $\theta = 0.5$ (with $\theta = 0.25$), the reduction of $T_0'(y)$ is of 1.2 (0.7) percentage point for $y = 200,000$€ as illustrated in Figure 3. By contrast, the optimal t_2 increases, under dual and mixed taxation, in order to reduce cross-base responses. Under mixed taxation, the optimal t_2 increases by 1.5 percentage points (from -23.2% to -21.7%) with $\theta = 0.1$, by 3.4 percentage points (from -23.2% to -19.6%) with $\theta = 0.25$ and by 5.9 percentage points (from -23.2% to $-17,3\%$) with $\theta = 0.5$, as one can see in the second and sixth to eight rows in Table 2. Intuitively, when the intensity of income-shifting increases, negative cross-base responses are stronger so that one needs to increase t_2 to maintain the equal sign in (46). In the extreme case of an infinite income-shifting elasticity, the optimum is to set a comprehensive tax system so that $t_2 = 0$ and $a_2 = 1$. An infinite elasticity then provides a rationale for a tax based on total comprehensive income as advocated in Saez and Stantcheva (2018).

VII Conclusion

In this paper, we have derived the optimal tax formulas when taxpayers earn different incomes, without having to rely on specific micro-foundations. We have shown how general equilibrium effects and cross-base responses arise. We have described the optimal marginal tax rate on each source of income as a function of empirically meaningful sufficient statistics. Our optimal tax formulas have extended the one-dimensional optimal tax formula of Diamond (1998) and Saez (2001) by taking into account general equilibrium effects as well as the cross-effects, on other tax bases, of the marginal tax rates that apply on a given source of income.

We have also provided policy implications for the debate about comprehensive versus schedular (and dual) taxation. We have shown that conditions for obtaining one or the other system are deprived of empirical relevance. We have then provided an easily implementable formula that indicates whether a government should shift its tax system towards a more comprehensive or a more schedular one.

Last but not least, we have developed a new algorithm that we have used to implement our tax formulas on French data. In the process, we have thoroughly compared the optimal tax system, the optimal dual tax system and the optimal comprehensive one. We have found that the optimal system borrows characteristics from both the comprehensive and the dual tax systems. Our simulations point to a system in which the personal income should encompass all incomes without deduction (as in comprehensive taxation), with a distinct negative tax rate on the most elastic source of income (as could be the case under schedular taxation). Our simulations show that U-shaped marginal tax rates on personal income should prevail. As in schedular taxation, incentives to earn the most elastic source of income (capital, in our illustration) should also be provided. While this is usually done under dual taxation with a positive and relatively low linear tax rate, our simulations suggest to impose a negative tax rate on capital income. All our simulations emphasize that income-shifting reduces this subsidy and that a weaker elasticity of substitution between capital and labor increases the spillover statistic associated with capital, hence this subsidy. Similar to comprehensive taxation, all income should be included into personal income, without deduction, and U-shaped marginal tax rates on personal income should prevail. But we have also found that incentives to earn the most elastic source of income should be provided, as done with a relatively low positive linear tax rate under dual taxation. Differing from the dual tax however, we have found a negative tax rate on capital income. All our simulations have emphasized that income-shifting reduces this subsidy and that a weaker elasticity of substitution between capital and labor increases the spillover statistic associated with capital, hence this subsidy.

Our paper can be extended in many directions. In our application, we have focused on labor and (financial) capital income in France. Our numerical analysis can obviously be replicated for other sources of income and other countries. In our research agenda, we plan to extend the model to the optimal taxation of labor and housing. Another planned extension consists in relaxing the assumption of homogeneous financial capital. In particular, interests, dividends and capital gains are no longer equivalent when one adds financial frictions in the analysis. Extending the analysis of political feasibility of tax reforms by Bierbrauer et al. (2021) to our environment with general equilibrium effects and cross-base responses is also on our agenda.

References

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

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.

Armstrong, Mark, "Multiproduct Nonlinear Pricing," *Econometrica*, 1996, *64*, 51–76.

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

Bastani, Spencer and Daniel Waldenström, "How should capital be taxed," *Journal of Economic Surveys*, 2020, *34* (4), 812–846.

Benoteau, Isabelle and Olivier Meslin, "Les prélèvements obligatoires sur le capital des ménages : comparaisons internationales," Technical Report, Rapport particulier pour le Conseil des prélèvements obligatoires 2017.

Bierbrauer, Felix J., Pierre Boyer, and Andreas Peichl, "Politically Feasible Reforms of Nonlinear Tax Systems," *American Economic Review*, 2021, *111* (1), 153–91.

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.

_ **and Laurence Jacquet**, "Optimal marginal and average income taxation under maximin," *Journal of Economic Theory*, 2008, *143* (1), 425–441.

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.

Chamley, Christophe, "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives," *Econometrica*, 1986, *54* (3), 607–622.

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, "National Debt in a Neoclassical Growth Model," *American Economic Review*, 1965, *55* (5), 1126–1150.

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

_ **and Emmanuel Saez**, "The Case for a Progressive Tax: From Basic Research to Policy Recommendations," *Journal of Economic Perspectives*, 2011, *25* (4), 165–190.

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

Dixit, Avinash and Joseph Stiglitz, "Monopolistic Competition and Optimum Product Diversity," *American Economic Review*, 1977, *76* (3), 297–308.

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

Ferey, Antoine, Benjamin Lockwood, and Dmitry Taubinsky, "Sufficient statistics for nonlinear tax systems with preference heterogeneity," mimeo, 2021.

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

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

— , **Narayana Kocherlakota, and Aleh Tsyvinski**, "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 2003, *70* (3), 569–587.

Hendren, Nathaniel, "Measuring Economic Efficiency Using Inverse-Optimum Weights," *Journal of Public Economics*, 2020, *187*, 104–119.

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

Hungerbühler, Mathias, Parmentier Alexis Lehmann Etienne, and Bruno Van der Linden, "Optimal Redistributive Taxation in a Search Equilibrium Model," 2006, *73* (3), 743–67.

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), 1299–1341.

Judd, Kenneth L., "Redistributive taxation in a simple perfect foresight model," *Journal of Public Economics*, 1985, *28* (1), 59–83.

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

Kleven, Henrik and Esben Schultz, "Estimating Taxable Income Responses Using Danish Tax Reforms," 2014, *6(4)*, 271–301.

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.

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

Kristjánsson, Arnaldur Sölvi, "Optimal taxation with endogenous return to capital," Memorandum 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, "Capital incomes are more elastic than labor incomes," CRED Working paper 2021-01, 2021.

Mankiw, N. Gregory, Matthew Weinzierl, and Danny Yagan, "Optimal Taxation in Theory and Practice," *Journal of Economic Perspectives*, 2009, *23* (4), 147–174.

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.

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.

— **and** — , "A Theory of Optimal Inheritance Taxation," *Econometrica*, 2013, *81* (5), 1851–1886.

— **,** — **, and Stefanie Stantcheva**, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 2014, *6* (1), 230–71.

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

— , "A Mathematical Theory of Saving," *The Economic Journal*, 1928, *38* (152), 543–559.

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.

— , "Do Taxpayers Bunch at Kink Points?," *American Economic Journal: Economic Policy*, 2010, *2* (3), 180–212.

— **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.

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

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

— , "Inequality and Redistribution: The Need for New Perspectives," NHH Discussion Papers 2005.

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

Slemrod, Joel and Zeuthen Gillitzer, *Tax Systems*, Zeuthen Lectures Book Series (Cambridge, MA: MIT Press), 2014.

Spiritus, Kevin, Etienne Lehmann, Sander Renes, and Floris Zoutman, "Optimal Taxation with Multiple Incomes and Types," Unpublished manuscript, 2021.

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.

Straub, Ludwig and Iván Werning, "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," *American Economic Review*, 2020, *110* (1), 86–119.

A Convexity of the Indifference Set

Let $\mathscr{C}(\cdot, \mathbf{x}; \mathbf{w})$ denote the reciprocal of $\mathscr{U}(\cdot, \mathbf{x}; \mathbf{w})$. Taxpayers of type $\mathbf{w}$ who supply inputs $\mathbf{x}$ obtain consumption $c = \mathscr{C}(u, \mathbf{x}; \mathbf{w})$ to enjoy utility $u = \mathscr{U}(c, \mathbf{x}; \mathbf{w})$. Using (3), we obtain:

$$\mathscr{C}_u(u, \mathbf{x}; \mathbf{w}) = \frac{1}{\mathscr{U}_c(\mathscr{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w})} \qquad \mathscr{C}_{x_i}(u, \mathbf{x}; \mathbf{w}) = \mathcal{S}^i(\mathscr{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}) \tag{60}$$

For each type $\mathbf{w} \in W$ and each utility level u, we assume that the indifference set $\mathbf{y} \mapsto \mathscr{C}(u, \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w})$ is 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

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

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:

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

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 Y_i^{-1} defined on $[Y_i(\underline{w}), Y_i(\overline{w})]$. Under quasilinear and additively separable utility function (20), 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:

$$T_i'(y_i) = 1 - \frac{1}{p_i} v^i_{x_i}\left(\frac{y_i}{p_i}; Y_i^{-1}(y_i)\right) \tag{61}$$

To determine the separate tax schedule that decentralizes the optimal allocation, one simply needs to integrate (61). 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:

$$\mathcal{T}(\mathbf{y}) = \left(\sum_{i=1}^{n} Y_i(w^\star)\right) - C(w^\star) + \sum_{i=1}^{n} T_i(y_i) \tag{62}$$

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

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 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 as follows. One integrates the marginal tax rate in (61) between $Y_i(w^\star)$ and y_i. Otherwise, the tax liability is infinite.

We now show that the separate tax schedule (62) is sufficient to decentralize the allocation $w \mapsto (C(w), Y_1(w, ..., Y_n(w))$. As (62) 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 (62) that:

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

So, we have:

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

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 (62), a $\mathbf{w}$-taxpayer chooses $y_i = Y_i(\mathbf{w})$ which ends the proof of Proposition 1.

C Proof of Proposition 3

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).[35] Our proof is constructed on a similar reasoning but is valid with general tax instruments and multidimensional incomes.

[35] 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.

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:

$$\max_{\mathbf{y}} \quad \mathcal{U}\left(\sum_{i=1}^{n} y_i - \mathcal{T}(\mathbf{y}), \mathcal{V}\left(y_1, ..., y_n\right); \mathbf{w}\right). \tag{63}$$

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 $U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \mathscr{U}\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) = \mathcal{U}(C(\mathbf{w}), V(\mathbf{w}); \mathbf{w})$.

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, $\mathscr{U}\left(C(\mathbf{w}^\star), \mathbf{X}(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}^\star), V(\mathbf{w}^\star); \mathbf{w}'\right) > \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}^\star); \mathbf{w}'\right) = \mathcal{U}\left(C(\mathbf{w}'), V(\mathbf{w}'); \mathbf{w}'\right) = \mathscr{U}\left(C(\mathbf{w}'), \mathbf{X}(\mathbf{w}'); \mathbf{w}'\right)$. This would contradict that $\mathbf{y} = \mathbf{Y}(\mathbf{w}')$ solves (63) 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 (63) for individuals of type $\mathbf{w}'$.

For individuals of type $\mathbf{w}$ solving (63) amounts to solve:

$$\max_{v} \quad \mathcal{U}(\mathcal{R}(v), v; \mathbf{w}). \tag{64}$$

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

$$V(g) \stackrel{\text{def}}{\equiv} \min_{\mathbf{y}} \quad \mathcal{V}\left(y_1, ..., y_n\right) \quad s.t: \quad \sum_{i=1}^{n} y_i = g \tag{65}$$

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 (65), 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:

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

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 (65):

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

for a given level of subdisutility v, since $\mathcal{R}$ and V are increasing mappings. Second, one solves Program (64). 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, tax revenue is lower under the initial tax schedule, since $\mathbf{Y}(\mathbf{w})$ is solving

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

instead of solving:

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

The latter program has the same solution as:

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

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.[36]

Parts (*i*) and (*ii*) 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.[37]

[36] In practice, most of real world tax schedules are piecewise linear. In theory, bunching should occur at convex kink points. 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. $\mathbf{y} \mapsto \int \mathcal{T}(\mathbf{y} + \mathbf{u})\, \mathrm{d}\Psi(\mathbf{u})$ where $\mathbf{u}$ is an n-dimensional random shock on incomes with joint CDF Ψ, which does verify part *i*) of Assumption 3.

[37] 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.

We derive optimality conditions and verify ex-post whether this presumption is validated by the obtained solution.

Derivation of Equation (23) with exogenous and endogenous prices and of Equations (22d), (27) and (32)

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. It depends on the magnitude t of the tax perturbation and on the price vector $\mathbf{p}$:

$$\widehat{U}(\mathbf{w}; t, \mathbf{p}) \stackrel{\text{def}}{\equiv} \max_{\mathbf{y}=(y_1,...,y_n)} \quad \mathscr{U}\left(\sum_{i=1}^{n} y_i - \widetilde{\mathcal{T}}(\mathbf{y}, t), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right). \tag{66}$$

The first-order conditions are:

$$\forall i \in \{1, ..., n\}: \qquad \frac{1}{p_i}\mathcal{S}^i\left(\sum_{k=1}^{n} y_k - \widetilde{\mathcal{T}}(\mathbf{y}, t), \frac{y_1}{p_1}, ..., \frac{y_n}{p_n}; \mathbf{w}\right) = 1 - \widetilde{\mathcal{T}}_{y_i}(\mathbf{y}, t). \tag{67}$$

Let $\widehat{\mathbf{Y}}(\mathbf{w}, t, \mathbf{p}) = (\widehat{Y}_1(\mathbf{w}, t, \mathbf{p}), ..., \widehat{Y}_n(\mathbf{w}, t, \mathbf{p}))$ denote the solution. 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 (66) is differentiable with respect to t and to $\mathbf{p}$. Moreover, its partial derivatives, at $\mathbf{p} = (\widetilde{p}_1(0), ..., \widetilde{p}_n(0))$ and $t = 0$, can be obtained by differentiating Equations (67) 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:

$$\begin{aligned}\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c\mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j} \cdot \mathrm{d}\mathbf{y}^T \;&=\; \left\{-\left[\frac{\partial \widetilde{\mathcal{T}}_{y_1}}{\partial t}, ..., \frac{\partial \widetilde{\mathcal{T}}_{y_n}}{\partial t}\right]^T_i + \left[\frac{\mathcal{S}^1_c}{p_1}, ..., \frac{\mathcal{S}^n_c}{p_n}\right]^T_i \frac{\partial \widetilde{\mathcal{T}}}{\partial t}\right\} \mathrm{d}t \\ &\quad + \left[\mathbb{1}_{i=j}(1 - \mathcal{T}_{y_i}) + \frac{\mathcal{S}^i_{x_j}\, y_j}{p_i\, p_j}\right]_{i,j} \cdot \left(\frac{\mathrm{d}p_1}{p_1}, ..., \frac{\mathrm{d}p_n}{p_n}\right)^T.\end{aligned} \tag{68}$$

where superscript T denotes the transpose operator $[A_{i,j}]^T_{i,j} = [A_{j,i}]_{i,j}$ 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 (22b), 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 (68), the matrix of compensated responses is given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} = \left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c\mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \tag{69a}$$

and is symmetric since the Hessian matrix is symmetric. Under the lump-sum tax reform defined in (22a), one has $\frac{\partial \widetilde{\mathcal{T}}}{\partial t} = -1$ and $\frac{\partial \widetilde{\mathcal{T}}_{y_k}}{\partial t} = 0$. Hence, according to (68), the vector of income responses is given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \rho}\right]^T_i = -\left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c\mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \cdot (\mathcal{S}^1_c, ..., \mathcal{S}^n_c)^T. \tag{69b}$$

Finally, according to (68), the responses to changes in log prices are given by:

$$\left[\frac{\partial Y_i(\mathbf{w})}{\partial \log p_j}\right]_{i,j} = \left(\left[\frac{\mathcal{S}^i_{x_j} + \mathcal{S}^i_c\mathcal{S}^j}{p_i\, p_j} + \mathcal{T}_{y_i y_j}\right]_{i,j}\right)^{-1} \cdot \left[\mathbb{1}_{i=j}(1 - \mathcal{T}_{y_i}) + \frac{\mathcal{S}^i_{x_j}\, y_j}{p_i\, p_j}\right]_{i,j}. \tag{69c}$$

Consider a tax perturbation as defined in Definition 2, plugging (69a) and (69b) into (68) yields:

$$\frac{\partial \widehat{Y}_i(\mathbf{w}, t=0, \mathbf{p})}{\partial t} = \underbrace{-\frac{\partial Y_i(\mathbf{w})}{\partial \rho} \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}), 0)}{\partial t}\right|_{t=0}}_{\text{Income responses}} \underbrace{- \sum_{j=1}^{n} \frac{\partial Y_i(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}), 0)}{\partial t}\right|_{t=0}}_{\text{Compensated responses}} \tag{70}$$

which, with exogenous prices, leads to (23).

Under an uncompensated tax reform of the j^{th} marginal tax rate as defined in (22c), 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 (70) leads to the Slutsky Equation (22d).

Finally, applying the envelope theorem to (66) leads to:

$$\frac{\partial \widehat{U}(\mathbf{w}, t=0, \mathbf{p})}{\partial t} = -\mathscr{U}_c\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \left.\frac{\partial \widetilde{\mathcal{T}}\left(\mathbf{Y}(\mathbf{w}), t\right)}{\partial t}\right|_{t=0} \tag{71a}$$

$$\begin{aligned} \frac{\partial \widehat{U}(\mathbf{w}, t=0, \mathbf{p})}{\partial \log p_j} &= -\mathscr{U}_{x_j}\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \frac{Y_j(\mathbf{w})}{p_j} \\ &= \mathscr{U}_c\left(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\right) \Big(1 - T_{y_j}(\mathbf{Y}(\mathbf{w}))\Big) \; Y_j(\mathbf{w}) \end{aligned} \tag{71b}$$

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:

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

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

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

Log-differentiating the latter equation leads to:

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

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 (32) holds. Adding these price responses to Equation (70) and using (69c) leads to Equation (23). Combining Equations (26), (71a) and (71b) yields (27).

E Proof of Proposition 4

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 $\left.\frac{\partial A}{\partial t}\right|_{t=0}^{\star}$ 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 $\left.\frac{\partial \widetilde{\mathscr{B}}}{\partial t}\right|_{t=0}^{\star} = 0$ and so:

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

Let $\left.\frac{\partial A^{\rho}}{\partial t}\right|_{t=0}$ be the partial derivative of A along the lump-sum perturbation (22a). According to (34), we get:

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

Equation (35) implies:

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

Combing these three equations leads to:

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

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

F Responses of taxable income under a mixed tax schedule

According to (18) and (23), we get:

$$\begin{aligned}
\left.\frac{\partial \widetilde{Y}_0(\mathbf{w},t)}{\partial t}\right|_{t=0} &= \sum_{k=1}^{n} a_k'(y_k) \left.\frac{\partial \widetilde{Y}_k(\mathbf{w},t)}{\partial t}\right|_{t=0} \\
= \; - \sum_{1\le j,k\le n} a_k'(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} \left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} &- \sum_{k=1}^{n} a_k'(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \rho} \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \\
- \sum_{1\le j,k\le n} a_k'(y_k) \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j} \left.\frac{\partial \log \widetilde{p}_j(t)}{\partial t}\right|_{t=0}
\end{aligned}$$

Equation (23) 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 (36a), (36b) and (36d). 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 (18) and (19) and by inverting subscripts i and k. The second equality is obtained using Equations (36a), (36b) and (36d). The last equality holds because we have normalized $\mu_0 = 0$. Equation (34) then becomes:

$$\begin{aligned}
\left.\frac{\partial \widetilde{\mathscr{L}}(t)}{\partial t}\right|_{t=0} = \int_{\mathbf{w}\in W} \Bigg\{ &\left[1 - g(\mathbf{w}) - \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\right] \left.\frac{\partial \widetilde{\mathcal{T}}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \\
&- \sum_{j=1}^{n} \left(\sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w})) + \mu_k\right) \frac{\partial Y_k(\mathbf{w})}{\partial \tau_j}\right) \left.\frac{\partial \widetilde{\mathcal{T}}_{y_j}(\mathbf{Y}(\mathbf{w}),t)}{\partial t}\right|_{t=0} \Bigg\} f(\mathbf{w}) d\mathbf{w}.
\end{aligned} \tag{72}$$

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 (37) and (38). Plugging these equations into (72) leads to Equation (39). When one applies our proof of Proposition 4 (Appendix E), it is then straightforward to proof part *i*) of Proposition 5.

Using the law of iterated expectations to condition types $\mathbf{w}$ on $Y_i(\mathbf{w}) = y_i$ and using (43), we obtain:

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

Integrating the latter equation by parts and using (35) leads to:

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

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 (45) and thereby, to part *ii*) of Proposition 5.

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 (40) and (41). Using (23), one obtains the impact of this type of reform of the personal income tax on all types of income, $\forall k \in \{1,...,n\}$:

$$\left.\frac{\partial \widetilde{Y}_k(\mathbf{w},t)}{\partial t}\right|_{t=0} = \sum_{j=1}^{n} a_j'(Y_j(\mathbf{w}))\frac{\partial Y_k(\mathbf{w})}{\partial \tau_j} R_0'(Y_0(\mathbf{w})) + \frac{\partial Y_k(\mathbf{w})}{\partial \rho}\, R_0(Y_0(\mathbf{w})) + \sum_{j=1}^{n} \frac{\partial Y_k(\mathbf{w})}{\partial \log p_j}\left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0}. \tag{73}$$

We combine $\widetilde{\mathcal{T}}(\mathbf{y},t) = \mathcal{T}(\mathbf{y}) - t\, R_i(y_i)$ with a compensated tax reform of the personal income described in Equation (22b), so that, in (73), one has $R_0(\cdot) = 0$, $R_0'(\cdot) = -1$ and $\left.\frac{\partial \log \widetilde{p}_j}{\partial t}\right|_{t=0} = 0$. This yields (42). Given (42), for $k \in \{1,..,n\}$, Equation (73) 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 (18), (36a), (36b) and (42) for $i = k = 0$ leads to:

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

We can conclude that (23) also holds for $j = 0$, i.e. with reforms of the personal income tax. According to Equation (72), 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 (42) has been used for the second equality. We thus obtain (39) with $i = 0$. Part (*i*) of Proposition 5 is therefore also valid for $i = 0$, thereby also its Part (*ii*).

I Optimal linear tax schedule

We rewrite Equation (39) with the uncompensated tax perturbation of the i^{th} income defined in (22c) (i.e. taking $R_i(Y_i(\mathbf{w})) = Y_i(\mathbf{w})$ and $R'(Y_i(\mathbf{w})) = 1$) and we use the Slutsky equations (22d) to obtain:

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

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:

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

We equalize to zero the latter expression and substitute in it the uncompensated elasticity $\varepsilon_i^u(\mathbf{w}) \stackrel{\text{def}}{\equiv} \frac{1-t_i}{Y_i(\mathbf{w})} \frac{\partial Y_i^u(\mathbf{w})}{\partial \tau_i}$. We then rearrange terms and we obtain (46).

J Proof of Proposition 6 and Lemma 1

The reform of the i^{th} net-of-deduction rate defined in (47) 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 (49) and (48) have been used for the second and third equation, respectively. Combining these expressions with (72) leads to:

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

Using the Slutsky equation (22d) and Equation (42), the preceding equation simplifies to (50), which, combined with Proposition 4, ends the proof of Proposition 6.

We define $\Psi_i(\mathbf{w}) \stackrel{\text{def}}{\equiv} (g(\mathbf{w}) - 1) Y_i(\mathbf{w}) + \sum_{k=0}^{n} \left(T_k'(Y_k(\mathbf{w}) + \mu_k) \frac{\partial Y_k^u(\mathbf{w})}{\partial \tau_i} \right)$. The first line of (50) is equal to

$$\int_{\mathbf{w}\in W} \Psi_i(\mathbf{w})\ T_0'(Y_0(\mathbf{w}))\ f(\mathbf{w}) d\mathbf{w} = cov\left(\Psi_i(\mathbf{w}), T_0'(Y_0(\mathbf{w}))\right) + \mathbb{E}(\Psi_i(\mathbf{w}))\ \mathbb{E}(T_0'(Y_0(\mathbf{w}))).$$

As the optimal linear tax formula (46) writes $\mathbb{E}(\Psi_i(\mathbf{w})) = 0$, one obtains Equation (56) whenever the i^{th} income is taxed at a linear rate given by the optimal linear tax rate formula (46). Combined with Proposition 4, this concludes the proof of Lemma 1.

K Input Taxation is superfluous

In this appendix, we show that the taxation of inputs can be replicated by an adequate re-scaling of the income tax function $\mathcal{T}(\cdot)$. Assuming no taxation of inputs is therefore without loss of generality.

Assume input i is taxed at rate $\alpha_i < 1$. For each $i \in \{1,...,n\}$, p_i denotes the price of input i for the firm and $q_i = p_i(1-\alpha_i)$ denotes the price after taxation faced by the suppliers of input i. The i^{th} market income is $y_i = p_i\ x_i$ while the i^{th} taxable income is equal to $q_i\ x_i = p_i(1-\alpha_i)x_i$. The tax schedule is a function of the vector of taxable income $(q_1\ x_1, ..., q_n\ x_n) = ((1-\alpha_1)y_1, ..., (1-\alpha_n)y_n)$. After-tax income c therefore verifies:

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

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

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

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 tax rates $(\alpha_1, ..., \alpha_n)$ and income tax $\mathcal{T}(\cdot)$ or without 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 algorithm

The initial dataset –hereafter, real dataset– contains 27,804 observations from ERFS.[38] For each observation, we have the levels of labor and capital income, the ERFS weight of the observation and an approximation of the marginal tax rates on labor income and on capital income. We use taxpayers' first-order conditions (6) to assign a type (w_1, w_2) to each observation of the real dataset. We calibrate A_1 and A_2 in the production function (58) to ensure that prices are at their normalized values of 1 in the real economy.

We have three distinct algorithms to find the optimal comprehensive, dual and mixed tax systems. Each algorithm iterates different operations on the real data set and on a virtual dataset that contains personal marginal income tax rates at each stage of the iteration process. More precisely, each of the 50 observations of this virtual dataset consists of a personal income y_0 and a value of the marginal tax rate on personal income $T_0'(y_0)$. The initial marginal tax rates are set from the French tax code. Personal income levels in the virtual dataset are evenly distributed between 4,000 € and 200,000 €.

The algorithm for the optimal mixed system is made of three nested loops. The inner loop computes the personal income tax schedule $T_0(\cdot)$ for a given linear tax rate and a given deduction rate on capital income, t_2 and a_2, using Equation (45). Given a_2, the middle loop optimizes the linear tax rate t_2 on capital income, using (46). The outer loop optimizes over a_2, between 0 and 1, using (50).

The inner loop is made of the following successive steps:

1. For each observation of the real dataset, we compute the solutions $Y_1(\mathbf{w})$ and $Y_2(\mathbf{w})$ of taxpayers' program (5). In this optimization step, we use a linear interpolation of marginal personal income tax rates in the virtual dataset to approximate $y_0 \mapsto T_0'(y_0)$.[39]

2. We compute the intercept of the personal income tax schedule that binds the budget constraint (10). We calculate, for each observation of the real dataset, the utility level. We then calculate λ using (35). We further compute, for each observation of the real dataset, the welfare weight (using (26)), total compensated responses (using (69a)) and total responses to prices changes (using (69c)).[40] Finally, we calculate macro spillover statistics using (32) and (33).

3. For each observation of the virtual dataset, to determine the personal income density $h_0(y_0)$, we use a kernel density estimation on the personal incomes in the real dataset. We use kernel density regressions to get the average compensated responses at y_0. In these kernel procedures, we use a biweight kernel $K(x) = (15/16)(1-x^2)^2$ with a bandwidth of 89,028 €.[41] We follow similar routines, respectively for the CDF $H_0(y_0)$ and for the mean of welfare weights above income y_0, using a primitive of the kernel $K(\cdot)$. Hence, we have all the sufficient statistics that show up in (45).

[38] These data are available from https://quetelet.progedo.fr

[39] We artificially extend the virtual dataset as follows. We add an observation at $y_0 = 0$ with the same marginal tax rate as the one for $y_0 = 4000$. We also add an observation at $y_0 = 10,000,000$ with the same marginal tax rate as the one obtained at $y_0 = 200,000$. This is useful to obtain the $T_0'(y_0)$ value for any y_0 outside $[4,000; 200,000]$.

The program only considers the minimum of (effective) marginal tax rates on capital income $T_0'(y_0) + a_2 \, t_2$ and of 0.99 to avoid a negative value of effective marginal net-of-tax rate on capital income. We make sure this correction does not happen in the last stage.

[40] As we assume income effect away, there is no need to compute income responses.

[41] The biweight kernel eases the computation since it is a 4$^{\text{th}}$-degree polynomial. It provides differentiable estimated densities since $K(\cdot)$ is differentiable with zero derivatives at $x = -1, 0, 1$.

4. For each observation y_0 of the virtual dataset, we update the marginal personal income tax schedule $T_0'(y_0)$, using (45).

5. We go back to step 1, unless the marginal tax rates are updated in the last step by less than 0.1 percentage point in absolute value.

The middle loop converges faster if instead of using (46), the linear tax rate on capital income t_2 is updated using the average between its value in the preceding middle loop and the value obtained with (46). The middle loop stops when two consecutive values of t_2 differ by less than 0.1 percentage points in absolute values.

At each iteration of the outer loop, there are a minimum, a maximum and an effective value of a_2. The effective value is the mean between the minimum and the maximum values. The initial triplet is $(0, 1, 0.5)$. It is the effective value that is used during the inner and middle loops. At the end of the iteration of the outer loop, Equation (50) is computed. If the obtained value of (50) is negative, the current effective value of a_2 becomes the new maximum value. Otherwise, it becomes the new minimum value. The outer loop stops after 10 iterations, so a_2 is obtained with a precision of $2^{-10} \simeq 0,000976$.

The algorithm for the dual system is the same with no outer loop and a_2 fixed at 0. The algorithm for the comprehensive system contains only the above inner loop with a_2 fixed at 1 and t_2 fixed at zero.