Theoretical Economics 20 (2025), 543–581 1555-7561/20250543

Optimal taxation with multiple incomes and types

Kevin Spiritus
Erasmus School of Economics, Erasmus University Rotterdam

Étienne Lehmann
CRED, Université Paris-Panthéon-Assas, CEPR, CESifo, IZA, and TEPP

Sander Renes
Faculty of Technology Policy and Management, Delft University of Technology

Floris T. Zoutman
Department of Business and Management Science, NHH Norwegian School of Economics

We analyze the optimal nonlinear income tax schedule for taxpayers with multiple incomes and multiple unobserved characteristics. We identify smoothness assumptions and extensions of the single crossing conditions that enable the characterization of the optimum through variational calculus. Both the tax perturbation and mechanism design approaches yield identical results when the number of incomes equals the number of unobserved characteristics. Notably, the mechanism design approach requires slightly less stringent assumptions than the tax perturbation approach. Additionally, we introduce a numerical method to determine the optimal tax schedule. Applied to couples, the optimal isotax curves are nearly linear and parallel. Additional contributions include a Pareto efficiency test and a condition on primitives ensuring the sufficiency of the government's necessary conditions, thereby guaranteeing the uniqueness of the solution.

Keywords. Nonlinear optimal taxation, multidimensional screening, household income taxation.

JEL classification. D82, H21, H23, H24.

Kevin Spiritus: spiritus@ese.eur.nl
Étienne Lehmann: elehmann@u-paris2.fr
Sander Renes: s.renes-1@tudelft.nl
Floris T. Zoutman: floris.zoutman@nhh.no

This paper merges earlier projects of Renes and Zoutman (2017) and Spiritus (2017). We are thankful for valuable suggestions from Spencer Bastani, Felix Bierbrauer, Pierre Boyer, Robin Boadway, Katherine Cuff, André Decoster, Eva Gavrilova, Aart Gerritsen, Tom Gresik, Nathan Hendren, Yasusi Iwamoto, Bas Jacobs, Laurence Jacquet, Roman Kozlof, Jonas Loebbing, Luca Micheletto, John Morgan, Nicola Pavoni, Emmanuel Saez, Dominik Sachs, Leif Sandal, Dirk Schindler, Guttorm Schjelderup, Erik Schokkaert, Matti Tuomala, Casper de Vries, Bauke Visser, Hendrik Vrijburg, and Nicolas Werquin. This paper benefited from suggestions by participants at numerous conferences. Kevin Spiritus acknowledges financial support of BELSPO via BRAIN.be project BR/121/A5/CRESUS. Étienne Lehmann acknowledges financial supports from CY Initiative, a program supported by the French National Research Agency (ANR) under the French government grant: "Investissements d'avenir" #France2030 (ANR-16-IDEX-0008).

© 2025 The Authors. Licensed under the Creative Commons Attribution-NonCommercial License 4.0. Available at https://econtheory.org. https://doi.org/10.3982/TE5479

1. Introduction

Studying optimal tax problems involving multiple income sources and dimensions of unobserved heterogeneity presents a significant challenge in the field of public finance. It is crucial to consider this dual-layered multidimensionality when exploring topics such as the optimal taxation of a couple's incomes, the optimal taxation of income from labor and capital, and the optimal means testing of benefits. However, the investigation of such issues introduces considerable theoretical challenges, complicating the formulation of policy recommendations.

Mirrlees (1976) pioneered a Mechanism Design (MD) approach to characterize the optimal incentive-compatible allocation of multiple taxable incomes. This approach results in a partial differential equation that proves challenging to interpret. Golosov, Tsyvinski, and Werquin (2014) develop a Tax Perturbation (TP) approach to characterize the optimal tax schedule. They, too, derive a partial differential equation describing the optimal tax function. Their formulation, however, offers the benefit of being expressed in terms of observable sufficient statistics.

The extent to which the MD approach by Mirrlees (1976) and the TP approach by Golosov, Tsyvinski, and Werquin (2014) are equivalent has not been examined yet. At first glance, it might appear that both approaches must be equivalent, since the taxation principle (Hammond (1979)) proves that selecting a tax function is equivalent to choosing an incentive-compatible allocation. However, this principle holds true only when there are no additional constraints on the tax schedule or the allocation of taxable incomes. Both Mirrlees' MD approach and Golosov, Tsyvinski, and Werquin's TP approach introduce additional smoothness assumptions—the former on the allocation and the latter on the tax schedule—to facilitate the application of variational calculus. These smoothness assumptions address the possibility that distinct types may be allocated the same income bundles, a situation referred to as "bunching," or instances where the mapping between types and income is discontinuous, a situation referred to as "jumping." In this paper, we derive optimal-tax formulas using both approaches. We study a model where taxpayers differ in multiple unobservable characteristics and in multiple incomes, and assume that individuals respond along the intensive margins. We assume that multidimensional versions of the single crossing conditions hold and make standard assumptions on the smoothness of the optimal-tax function and the optimal allocation. We show that under these assumptions the optimal-tax formulas derived through both approaches are equivalent when the number of unobservable characteristics equals the number of taxable incomes. The assumptions required to apply the TP approach are slightly more demanding than those required to apply the MD approach. With the TP approach, we need to assume the tax-function is thrice differentiable, whereas the MD approach only requires twice-differentiability of the tax function.

We also investigate the cases where the numbers of taxable incomes and characteristics are not equal. The TP approach solves the optimal-tax problem in the income space, whereas the MD approach solves the same problem in the type space. Therefore, if the number of characteristics exceeds the number of incomes, solving the optimal-tax

problem with the TP approach reduces the dimensionality of the problem, and thus its complexity. Conversely, the MD approach reduces the dimensionality when the number of incomes exceeds the number of characteristics.

Determining the optimal tax schedule involves solving a partial differential equation, a task that is significantly more complex than solving the ordinary differential equation implied by the optimal tax formula for a single tax base. To illustrate the complexity, consider that in a one-dimensional scenario, one can examine the effects of perturbing the marginal tax rate at one income level. The optimal marginal tax rate at that income level is then determined by the ratio of mechanical and income effects at all incomes above, to compensated effects at the income level under consideration. However, in a multidimensional scenario, it is not possible to examine the effects of a change in the tax gradient at one combination of incomes without inducing additional changes in the tax gradients at other combinations of incomes.

We develop a numerical algorithm that tackles this geometric complexity and can solve the optimal multidimensional tax problem in its general form. We apply our algorithm to the taxation of couples. In our application, we adopt simplifying assumptions akin to those made by Kleven, Kreiner, and Saez (2007). We presume quasilinear and additively separable household preferences. Furthermore, consistent with the empirical literature, we assume that the labor supply of wives is more elastic (0.43) than that of husbands (0.11) (Bargain and Peichl (2016)). Lastly, we calibrate the joint distribution of skills nonparametrically, starting from the joint distribution of incomes in the Current Population Survey (CPS) of the US census.

To facilitate our exposition, we introduce the notion of an "isotax curve," which refers to a set of income bundles that incur the same tax liability. Our findings indicate that the optimal isotax curves are nearly linear and parallel, with both spouses subjected to positive marginal tax rates. A joint income tax that discounts female income by approximately 53% closely approximates the fully optimized schedule in terms of social welfare. Additionally, we explore the concept of negative jointness, which stipulates that the optimal marginal tax rates of males should decrease with an increase in female income (and vice versa). Kleven, Kreiner, and Saez (2007) analytically demonstrate that negative jointness is desirable when the productivities of both spouses are assumed to be uncorrelated. However, our numerical findings suggest that this result does not hold up under a more realistic joint distribution of productivities.

In addition to our comparison between the MD approach and the TP approach, and our numerical algorithm, we make several theoretical contributions. First, we formulate a test to verify the Pareto efficiency of a given tax schedule. If the welfare weights revealed by the optimal tax formula are negative for certain income bundles, then reducing tax liabilities at these income bundles results in a self-financed Pareto improvement. This extends the revealed social preference approach of Werning (2007), Bourguignon and Spadaro (2012), Bargain, Dolls, Neumann, Peichl, and Siegloch (2014), Jacobs, Jongen, and Zoutman (2017), Scheuer and Werning (2017), Hendren (2020), and Bierbrauer, Boyer, and Hansen (2023) to a multidimensional context.

Second, we employ the MD approach to establish conditions under which the first-order conditions are sufficient to characterize the optimal allocation. This holds true

when the government's Lagrangian is concave both with respect to the taxpayers' utilities, and with respect to the gradient of the mapping between the taxpayers' types and utilities. We analytically confirm that the specification used in our numerical exercise complies with these sufficiency conditions. Therefore, once we have obtained a numerical solution that satisfies the government's first-order conditions, we know that it is the unique solution. Consequently, there is no need to perform sensitivity analyses concerning the initial conditions of our algorithm.

Third, we address a concern in the TP approach, where it is assumed by both Saez (2001) and Golosov, Tsyvinski, and Werquin (2014) that incomes respond smoothly to tax perturbations. We contribute by explicitly outlining the assumptions about the tax schedule that ensure smooth responses of taxpayers to tax perturbations. Our assumptions rule out kinks in the tax schedule and the presence of multiple global optima, thereby ensuring that incremental tax perturbations do not lead to jumps in taxpayers' behavior.

Fourth, we introduce a new method to derive the optimal mechanism. When Mirrlees (1976) and Kleven, Kreiner, and Saez (2007) derive necessary conditions for the optimal incentive-compatible allocation, they use taxpayers' utilities and taxable incomes as controls. The challenge here is that there are numerous income allocations that satisfy their necessary conditions for the optimum. We show how for each such income allocation, the first-order incentive constraints imply the partial derivatives of the attained utilities with respect to the types. At this point, nothing guarantees that the obtained partial derivatives of the achieved utilities are mutually consistent, meaning they imply symmetric second-order partial derivatives. Mirrlees (1976, p. 343) and Kleven, Kreiner, and Saez (2007, p. 18) recognize this issue by stating that among the different solutions of the partial differential equation, only the one implying symmetric second-order cross- derivatives should be considered. We circumvent these challenges by dividing the government's problem into two stages. In the first stage, the government chooses the optimal type-to-utility mapping from the set of possible mappings. In the second stage, the taxable incomes are determined as functions of the utility profile and its partial derivatives.

Finally, we investigate the cases where the number of characteristics p differs from the number of incomes n. If the number of characteristics exceeds the number of incomes, opting for the TP method and using average sufficient statistics among taxpayers with identical income bundles reduces the complexity of the problem. This extends the findings of Saez (2001), Scheuer and Werning (2016) and Jacquet and Lehmann (2021) to situations where taxpayers have multiple income sources. Conversely, applying the MD approach in this setting is only feasible under strong restrictions on preferences (see, e.g., Choné and Laroque (2010), Rothschild and Scheuer (2013, 2014, 2016), Scheuer (2014), and Jacquet and Lehmann (2023) for the $n = 1 < p$ case).

When the number of incomes n exceeds the number of unobservable characteristics p, it generally makes more sense to use the MD approach. Indeed, by working within the type space rather than the income space, one reduces the problem's dimensionality. In this case, the government's problem involves the two-step process described above. The second step then is a subprogram that determines the most efficient distribution

of income choices to produce the type-to-utility mapping that was selected in the first step. The solution to this subprogram is independent of government preferences and solely depends on the resource costs of providing these utility levels. Our findings shed light on similar subprograms implicitly present in the works of, among others, Atkinson and Stiglitz (1976), Golosov, Kocherlakota, and Tsyvinski (2003), Gerritsen, Jacobs, Rusu, and Spiritus (2025), and Ferey, Lockwood, and Taubinsky (2022).

Related literature

Our derivations crucially rely on assumptions on the smoothness of the allocations and of the tax schedule. Such assumptions are also found in the MD approaches proposed by Mirrlees (1976) and further developed by Kleven, Kreiner, and Saez (2007), and the TP approach of Golosov, Tsyvinski, and Werquin (2014). When comparing the two approaches, we rule out the possibilities of jumping and bunching.

While bunching can occur in one-dimensional models, it is more likely when taxpayers have multiple income sources and multidimensional unobserved characteristics. To see this, note that our paper is related to the multidimensional screening problem, which has been studied in the context of monopoly pricing by Armstrong (1996), Rochet and Choné (1998), and Basov (2005). Rochet and Choné (1998) demonstrate that bunching is a significant concern due to the interplay between the participation constraint and the second-order incentive constraints. However, our model does not include participation constraints, making the argument of Rochet and Choné (1998) not directly applicable to our model.

Still, Dodds (2023) shows that bunching is optimal in the optimal tax problem if social preferences are sufficiently close to maximin. The intuition is that, following Boadway and Jacquet (2008), the dual of the optimal tax problem with maximin social preferences consists in maximizing tax revenue subject to incentive constraints and a lower bound at the lowest utility level. The latter constraint is mathematically equivalent to the participation constraint in the monopoly model of Rochet and Choné (1998). Conversely, Kleven, Kreiner, and Saez (2007) argue that a range of moderate inequality aversions exists where bunching does not occur in the optimum. Therefore, the approach we adopt in this paper is valid when social preferences remain sufficiently far from maximin.

Most closely related to our work is a concurrent working paper by Golosov and Krasikov (2023). They study the optimal taxation of couples with a general tax function that depends on both spouses' incomes and allows for different earnings abilities for both spouses. Using a mechanism-design approach, they derive the optimum and focus on obtaining new theoretical results by applying the Coarea formula to the optimal tax expression. This method enables them to find closed-form solutions for various conditional moments of the optimal tax formula, such as linking optimal tax rates to the correlation in spousal earnings. Our study, on the other hand, combines the mechanism-design and tax-perturbation approaches. We demonstrate when these two methods yield the same optimal tax formula and discuss the pros and cons of each approach.

Another important paper on taxation with multiple dimensions of labor through MD tools is Boerma, Tsyvinski, and Zimin (2022). Their paper differs from ours in multiple aspects. First, they solve the government's problem using Legendre transformations. With this method, individual utility functions must be additively separable with isoelastic cost of effort. Conversely, neither our numerical algorithm nor our analytical results rely on such restrictions on individual preferences. Second, Boerma, Tsyvinski, and Zimin (2022) assume that production requires sorting between manual and cognitive labor. In this case, bunching is a robust property throughout the income distribution, for reasons similar to those set forth by Rochet and Choné (1998).

Our paper also relates to the literature, which studies multidimensional heterogeneity in the context where the government can only observe and tax a single income (e.g., Choné and Laroque (2010), Rothschild and Scheuer (2013, 2014, 2016), Lockwood and Weinzierl (2015), Jacquet and Lehmann (2021), Bergstrom and Dodds (2021)). We rely on the insights in this literature to formulate our expressions in terms of sufficient statistics. Specifically, in the context of multidimensional heterogeneity, sufficient statistics can be strongly endogenous to the tax schedule. We use the approach of Jacquet and Lehmann (2021) to overcome this issue by expressing our optimal tax formulas in terms of *total elasticities* that incorporate this endogeneity. We expand on this literature by allowing for multidimensional incomes in addition to multidimensional heterogeneity in unobserved characteristics.

Scheuer (2014) and Gomes, Lozachmeur, and Pavan (2018) study a setting with multidimensional heterogeneity in which agents choose to earn income in one of two different sectors, and the government can tax the income of each sector according to a separate tax schedule. The main difference with our approach is that in our model agents can earn multiple incomes at the same time.

Like in our application, Frankel (2014) studies the optimal taxation of couples in a setting with multidimensional heterogeneity and taxation of both male and female income. The main contrast between the approaches is that we allow for a continuous type distribution, whereas Frankel (2014) studies a discrete 2×2 distribution of married couples. Cremer, Pestieau, and Rochet (2001, 2003) also consider multidimensional settings. However, they only allow labor income to be taxed nonlinearly, whereas taxes on commodity/capital are constrained to be linear.

The paper is organized as follows. We describe the problem of multidimensional optimal taxation in Section 2. Section 3 is devoted to the TP approach, and Section 4 is devoted to the MD approach. We compare both approaches in Section 5. We present our numerical algorithm and results in Section 6.

2. The model

2.1 *Taxpayers*

The economy consists of a unit mass of taxpayers who differ in a p-dimensional vector of characteristics denoted $\mathbf{w} \stackrel{\text{def}}{\equiv} (w_1, \dots, w_p)$, that we call type. The type space is denoted $\mathcal{W} \subset \mathbb{R}^p$ and is assumed to be closed and convex. Types are distributed according to a twice continuously differentiable density denoted by $f(\cdot)$, which is positive over $\mathcal{W}$.

Taxpayers make n choices. The n observable tax bases are denoted $\mathbf{x} \overset{\text{def}}{\equiv} (x_1, \dots, x_n) \in \mathbb{R}^n_+$. We call these tax bases incomes for brevity.[1] Taxpayers pay a tax $T(\mathbf{x})$ that can depend on all incomes in a nonlinear way. Taxpayers who earn incomes $\mathbf{x}$ consume after-tax income $c = \sum_{i=1}^n x_i - T(x_1, \dots, x_n)$.

The preferences of taxpayers of type $\mathbf{w}$ over consumption c and income choices $\mathbf{x}$ are described by a thrice continuously differentiable utility function $\mathcal{U}(c, \mathbf{x}; \mathbf{w})$ defined over $\mathbb{R}^{n+1}_+ \times \mathcal{W}$. Taxpayers enjoy utility from consumption but endure disutility to obtain income, so $\mathcal{U}_c > 0$ and $\mathcal{U}_{x_i} < 0$. Let $\mathcal{C}(\cdot, \mathbf{x}; \mathbf{w})$ be the inverse of $\mathcal{U}(\cdot, \mathbf{x}; \mathbf{w})$. That is, a taxpayer of type $\mathbf{w}$ earning incomes $\mathbf{x}$ should consume $\mathcal{C}(u, \mathbf{x}; \mathbf{w})$ to enjoy utility level u. It follows that $\mathcal{C}_u = 1/\mathcal{U}_c$ and $\mathcal{C}_{x_i} = -\mathcal{U}_{x_i}/\mathcal{U}_c$. We assume the utility function $\mathcal{U}(\cdot, \cdot; \mathbf{w})$ is weakly concave in $(c, \mathbf{x})$ and indifference sets defined by $c = \mathcal{C}(u, \mathbf{x}; \mathbf{w})$ are strictly convex in $\mathbf{x}$.

We assume taxpayers maximize utility subject to their budget constraints. Therefore, a taxpayer of type $\mathbf{w}$ solves

$$U(\mathbf{w}) \overset{\text{def}}{\equiv} \max_{x_1, \dots, x_n} \mathcal{U}\left(\sum_{i=1}^n x_i - T(x_1, \dots, x_n), x_1, \dots, x_n; \mathbf{w}\right). \tag{1}$$

Let $\mathbf{X}(\mathbf{w}) \overset{\text{def}}{\equiv} (X_1(\mathbf{w}), \dots, X_n(\mathbf{w}))$ denote the solution to this program and let $C(\mathbf{w}) \overset{\text{def}}{\equiv} \sum_{i=1}^n X_i(\mathbf{w}) - T(\mathbf{X}(\mathbf{w}))$ denote the corresponding consumption. In addition, we denote the marginal rate of substitution between the ith income and consumption as

$$\mathcal{S}^i(c, \mathbf{x}; \mathbf{w}) \overset{\text{def}}{\equiv} -\frac{\mathcal{U}_{x_i}(c, \mathbf{x}; \mathbf{w})}{\mathcal{U}_c(c, \mathbf{x}; \mathbf{w})} = \mathcal{C}_{x_i}\big(U(c, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big) > 0. \tag{2}$$

The first-order conditions for taxpayers of type $\mathbf{w}$ are

$$\forall j \in \{1, \dots, n\}: \quad \mathcal{S}^j\big(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big) = 1 - T_{x_j}\big(\mathbf{X}(\mathbf{w})\big). \tag{3}$$

2.2 *Government*

The government's budget constraint is given by

$$\int_{\mathcal{W}} T\big(\mathbf{X}(\mathbf{w})\big) f(\mathbf{w}) \, \mathrm{d}\mathbf{w} - E \geq 0, \tag{4}$$

where $E \geq 0$ is an exogenous amount of public expenditure. The government's objective aggregates the utility of the households in the economy:

$$\int_{\mathcal{W}} \Phi\big(U(\mathbf{w}); \mathbf{w}\big) f(\mathbf{w}) \, \mathrm{d}\mathbf{w}, \tag{5}$$

where the transformation $(u; \mathbf{w}) \mapsto \Phi(u; \mathbf{w})$ is twice continuously differentiable in $(u, \mathbf{w})$, increasing and weakly concave in u and potentially type-dependent. The government's problem consists of finding the tax function $T(\cdot)$ that maximizes the social

[1] Our model could be extended to include observable actions like private expenditures in education, which correspond to negative cash-flows for the households. This extension would not affect the validity of our results.

welfare function (5) subject to revenue constraint (4), taking into account taxpayers' behaviors defined by (1).

Following Diamond (1975) and Saez (2001), we define the welfare weights of taxpayers of type $\mathbf{w}$ as the social marginal utility of consumption expressed in monetary terms:

$$g(\mathbf{w}) \stackrel{\text{def}}{\equiv} \frac{\Phi_u\big(U(\mathbf{w});\mathbf{w}\big)\,\mathcal{U}_c\big(C(\mathbf{w}),\mathbf{X}(\mathbf{w});\mathbf{w}\big)}{\lambda} \geq 0, \tag{6}$$

where $\lambda > 0$ is the Lagrange multiplier associated to the budget constraint.

We compare two strategies to solve the government's problem.

- In the tax perturbation (TP) approach, we consider the effects of marginal reforms of the tax schedule $\mathbf{x} \mapsto T(\mathbf{x})$, taking into account taxpayer's behavioral responses to tax reforms. The tax schedule is optimal only if any tax reforms induce no first-order effect on the government's objective.

- In the mechanism design (MD) approach, we acknowledge the taxation principle according to which selecting a tax schedule $\mathbf{x} \mapsto T(\mathbf{x})$ taking into account taxpayer's behavioral responses is equivalent to choosing an incentive-compatible allocation $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$. At the second-best optimum, no incentive-compatible perturbation of the allocation should induce a first-order effect on the government's objective.

For tractability reasons, most authors make smoothness assumptions to pursue either of these two approaches. In the following sections, we clarify the relations between the assumptions. We proceed by first introducing the TP approach and the MD approach separately, in Sections 3 and 4, and comparing both approaches in Section 5.

3. The tax perturbation approach

In this section, we derive the optimal tax formula using the TP approach, which was previously derived by Golosov, Tsyvinski, and Werquin (2014). A necessary condition for a tax schedule to be optimal is that small perturbations of the schedule do not change social welfare. Golosov, Tsyvinski, and Werquin (2014) assume that individuals respond smoothly to such perturbations. We contribute by revealing underlying assumptions that ensure that the responses of the taxpayers to tax reforms are smooth. We also identify additional assumptions that allow the characterization of the optimum in a partial differential equation. Identifying these assumptions allows us to compare the TP approach to the MD approach.

We first formally introduce the perturbations to the tax schedule. Perturbing the tax schedule $\mathbf{x} \mapsto T(\mathbf{x})$ in the *direction* $R(\cdot)$ by *magnitude* $t \lesseqgtr 0$ leads to the perturbed tax schedule $\mathbf{x} \mapsto T(\mathbf{x}) - t\,R(\mathbf{x})$. The utility of taxpayers of type $\mathbf{w}$ then becomes a function of the magnitude t through

$$\widetilde{U}^R(\mathbf{w}, t) \stackrel{\text{def}}{\equiv} \max_{x_1,\dots,x_n} \mathcal{U}\left(\sum_{i=1}^n x_i - T(x_1,\dots,x_n) + t\,R(x_1,\dots,x_n), x_1,\dots,x_n;\mathbf{w}\right). \tag{7}$$

By definition, we know that $\widetilde{U}^R(\mathbf{w}, 0) = U(\mathbf{w})$. The first-order conditions associated with (7) are

$$\forall j \in \{1, \dots, n\} : \mathcal{S}^j\left(\sum_{i=1}^n x_i - T(\mathbf{x}) + t\, R(\mathbf{x}), \mathbf{x}; \mathbf{w}\right) = 1 - T_{x_j}(\mathbf{x}) + t\, R_{x_j}(\mathbf{x}). \tag{8}$$

If we perturb the tax schedule or any of the characteristics of the households, then the households will update their choices $\widetilde{\mathbf{X}}^R(\mathbf{w}, t)$ such that first-order conditions (8) remain satisfied. We now introduce assumptions on the unperturbed tax schedule that allow applying the implicit function theorem to (8) to derive these behavioral responses.

Assumption 1. *The tax schedule $T(\cdot)$ verifies the following assumptions:*

(i) The tax schedule $\mathbf{x} \mapsto T(\mathbf{x})$ is thrice continuously differentiable.

(ii) For each type $\mathbf{w} \in \mathcal{W}$, the second-order conditions associated with (1) are strictly verified, that is, the matrix $[\mathcal{S}^i_{x_j} + \mathcal{S}^j\mathcal{S}^i_c + T_{x_i x_j}]_{i,j}$ is positive definite at $c = C(\mathbf{w})$ and $\mathbf{x} = \mathbf{X}(\mathbf{w})$.[2]

(iii) For each type $\mathbf{w} \in \mathcal{W}$, the function $\mathbf{x} \mapsto \mathcal{U}(\sum_{i=1}^n x_i - T(\mathbf{x}), \mathbf{x}; \mathbf{w})$ admits a single global maximum.

Assumption 1(i) rules out kinks like those in piecewise linear tax schedules. Moreover, it ensures that the first-order conditions (8) are twice continuously differentiable in t, $\mathbf{w}$, and $\mathbf{x}$, provided that the direction $R(\cdot)$ is thrice continuously differentiable. We require thrice differentiability to derive Proposition 1 and 3, as we explain in more detail below. Assumption 1(ii) ensures that the first-order conditions (8) are associated with a local maximum of the taxpayers' program (7). Parts (i) and (ii) of Assumption 1 together enable one to apply the implicit function theorem to determine how a local maximum of (7) is affected by a small tax perturbation or a small change in types. Assumption 1(iii) rules out the existence of multiple global maxima. This prevents an incremental tax perturbation from causing a "jump" in the taxpayers' choices from one maximum to another. At such jumps, the derivative of $\widetilde{\mathbf{X}}^R(\mathbf{w}, t)$ with respect to the size t of the perturbation tends to infinity.[3] If Assumption 1 is satisfied, then the function $\widetilde{\mathbf{X}}^R(\mathbf{w}, t)$ that solves (7) is continuously differentiable for t close to 0, that is, the behavioral responses to tax reforms are smooth.

Geometrically, Assumption 1 implies that for each type $\mathbf{w}$, the indifference set defined by $c = \mathcal{C}(U(\mathbf{w}), \mathbf{x}; \mathbf{w})$ admits a single tangency point with the budget set defined by $c = \sum_{i=1}^n x_i - T(\mathbf{x})$ and lies strictly above the budget set elsewhere. Given that we assume

[2]We let $[a(k)]_k$ denote a column vector whose k^{th} row is $a(k)$, $[A(k, \ell)]_{k,\ell}$ denotes a rectangular matrix whose k^{th} row and ℓ^{th} column is $A(k, \ell)$, and $\cdot$ stands for the matrix product. The transpose operator is denoted with superscript T, and the inverse operator is denoted with superscript -1.

[3]Note that in the $n = 1$ case, Bergstrom and Dodds (2023) relax part (iii) of our Assumption 1 and allow for some individuals' optimization problems to admit multiple global maxima, leading to jumping responses. Hence, Assumption 1 is only sufficient to apply the TP, but not necessary.

that the indifference sets defined by $c = \mathcal{C}(u, \mathbf{x}; \mathbf{w})$ are strictly convex, Assumption 1 is automatically verified if the tax schedule is linear (see Appendix A.1).

We characterize the optimal tax schedule under the presumption that Assumption 1 holds, which then needs to be verified ex post in applications. This approach is analogous with the standard first-order MD approach, which presumes that the second-order incentive constraints do not bind in the optimum, and verifies ex post that this is indeed the case (Mirrlees (1971, p. 188)).

A variation in t affects the first-order conditions (8) through the changes in the marginal tax rates on the right-hand side and through the changes in the tax liabilities that determine the marginal rates of substitution on the left-hand side. Thanks to Assumption 1, one can differentiate Equations (8) with respect to t and $\mathbf{x}$, which leads to (see Appendix A.2):

$$\frac{\partial \widetilde{X}_i^R(\mathbf{w}, t)}{\partial t} = \underbrace{\frac{\partial X_i(\mathbf{w})}{\partial \rho} R\big(\mathbf{X}(\mathbf{w})\big)}_{\text{Income responses}} + \underbrace{\sum_{j=1}^{n} \frac{\partial X_i(\mathbf{w})}{\partial \tau_j} R_{x_j}\big(\mathbf{X}(\mathbf{w})\big)}_{\text{Compensated responses}}. \tag{9}$$

Here, $\partial X_i(\mathbf{w})/\partial \tau_j$ denotes type-$\mathbf{w}$ taxpayers' i^{th} income response to a *compensated*[4] *perturbation of the jth marginal tax rate* in the direction $R(\mathbf{x}) = x_j - X_j(\mathbf{w})$. Moreover, $\partial X_i(\mathbf{w})/\partial \rho$ denotes their i^{th} income response to a *lump sum* perturbation in the direction $R(\mathbf{x}) = 1$. Note that we do not explicitly assume that responses to tax perturbations are smooth. Instead, we show that if the unperturbed tax schedule verifies Assumption 1, the function $t \mapsto \widetilde{\mathbf{X}}^R(\mathbf{w}, t)$ is continuously differentiable at $t = 0$, and that Equation (9) holds in that case.[5]

We now investigate whether, starting from a tax schedule $T(\cdot)$, a perturbation in a direction $R(\cdot)$ is socially desirable by investigating its effects on the government's perturbed Lagrangian:

$$\widetilde{\mathcal{L}}^R(t) \stackrel{\text{def}}{\equiv} \int_{\mathcal{W}} \left\{ T\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big) - tR\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big) + \frac{\Phi\big(\widetilde{U}^R(\mathbf{w}, t); \mathbf{w}\big)}{\lambda} \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w}, \tag{10}$$

which is written in monetary terms. To compute the partial derivative of the Lagrangian with respect to the magnitude t of the tax reform, we also need the partial derivative of social welfare. Apply the envelope theorem to (7) and use (6) and (9) to find (See

[4]These perturbations are said to be "compensated for taxpayers of type $\mathbf{w}$" because they change the marginal tax rate of type $\mathbf{w}$ but leave the tax liability at incomes $\mathbf{x} = \mathbf{X}(\mathbf{w})$ unchanged.

[5]Strictly speaking, these responses do not just depend on the type $\mathbf{w}$, but also on the Hessian of the tax function. When the tax function is nonlinear, the responses to a tax reform generate changes in the marginal tax rates, which further induce compensated responses to these changes in marginal tax rates, etc. (Jacquet and Lehmann (2021)). By applying the implicit function theorem, the behavioral responses $\partial X_i(\mathbf{w})/\partial \rho$ and $\partial X_i(\mathbf{w})/\partial \tau_j$ encapsulate this "circular process" through the endogeneity of the marginal tax rates. We refer to these responses as *total* responses. We discuss the relation between direct and total responses in the working paper version of this article (Appendix A.3 of Spiritus, Lehmann, Renes, and Zoutman (2022)). Finally, throughout the paper, we evaluate the partial derivatives with respect to t only at $t = 0$.

Appendix A.3):[6]

$$\frac{\partial \widetilde{\mathcal{L}}^R(t)}{\partial t} = \int_{\mathcal{W}} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big) \frac{\partial X_i(\mathbf{w})}{\partial \rho} \right] R\big(\mathbf{X}(\mathbf{w})\big) \right.$$
$$\left. + \sum_{1 \le i,j \le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big) \frac{\partial X_i(\mathbf{w})}{\partial \tau_j} R_{x_j}\big(\mathbf{X}(\mathbf{w})\big) \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w}. \tag{11}$$

At the optimum, there should not exist an infinitesimal perturbation of the tax schedule that would induce a first-order effect on the government's objective. Therefore, the right-hand side of (11) should be equal to zero for any direction $R(\cdot)$. We derive an optimal tax formula from this requirement in Appendix A.4. To do so, we first rewrite (11) in the income space. For this purpose, let $\mathcal{X} \overset{\text{def}}{\equiv} \{\mathbf{x} | \exists \mathbf{w} \in \mathcal{W} \ : \mathbf{x} = \mathbf{X}(\mathbf{w})\}$ denote the range of the type set $\mathcal{W}$ under the allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$. Let $h(\mathbf{x})$ denote the joint density of incomes $\mathbf{x}$, which is defined over $\mathcal{X}$. Furthermore, for each combination of incomes $\mathbf{x} \in \mathcal{X}$, let $\partial \overline{X_i}(\mathbf{x})/\partial \tau_j$, $\partial \overline{X_i}(\mathbf{x})/\partial \rho$, and $\overline{g}(\mathbf{x})$, respectively, denote the means of $\partial X_i(\mathbf{w})/\partial \tau_j$, $\partial X_i(\mathbf{w})/\partial \rho$ and $g(\mathbf{w})$ among taxpayers that earn the combination of incomes $\mathbf{X}(\mathbf{w}) = \mathbf{x}$. Second, we use the divergence theorem to rewrite the second line of (11). For this purpose, we need income densities and compensated responses to be continuously differentiable. Furthermore, we can only apply the divergence theorem on the income space $\mathcal{X}$ if it is of dimension n. We thus make the following assumption.

Assumption 2. *The number of characteristics is greater than or equal to the number of incomes ($p \ge n$), the income space $\mathcal{X}$ is of dimension n, and the sufficient statistics $h(\mathbf{x})$, $\partial \overline{X_i}(\mathbf{x})/\partial \tau_j$, $\partial \overline{X_i}(\mathbf{x})/\partial \rho$, and $\overline{g}(\mathbf{x})$ are continuously differentiable functions of $\mathbf{x}$.*

At the end of this subsection, we provide sufficient microfoundations to illustrate the plausibility of Assumption 2. The following proposition then characterizes the optimal tax schedule (see the proof in Appendix A.4).

Proposition 1. *If the optimal tax schedule satisfies Assumptions 1 and 2, the optimum verifies the Euler–Lagrange equation:*

$$\left[1 - \overline{g}(\mathbf{x}) - \sum_{i=1}^{n} T_{x_i}(\mathbf{x}) \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \rho} \right] h(\mathbf{x}) = - \sum_{1 \le i,j \le n} \frac{\partial \left[T_{x_i}(\mathbf{x}) \dfrac{\partial \overline{X_i}(\mathbf{x})}{\partial \tau_j} h(\mathbf{x}) \right]}{\partial x_j}, \tag{12a}$$

for all $\mathbf{x}$ in $\mathcal{X}$, and it verifies the boundary conditions:

$$\forall \mathbf{x} \in \partial \mathcal{X} : \quad \sum_{1 \le i,j \le n} T_{x_i}(\mathbf{x}) \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \tau_j} h(\mathbf{x})\, e_j(\mathbf{x}) = 0, \tag{12b}$$

[6] In Spiritus et al. (2022), we show that the effect of any perturbation on the government's Lagrangian has the same sign as the effect on the social objective of that perturbation combined with a lump-sum transfer that keeps the budget restriction satisfied. This result also holds outside of the optimum, as long as the weight λ put on government's revenue verifies (15d).

where $\partial\mathcal{X}$ denotes the boundary of $\mathcal{X}$, and $\mathbf{e}(\mathbf{x}) = (e_1(\mathbf{x}), \ldots, e_n(\mathbf{x}))$ denotes the outward unit vector normal to the boundary at $\mathbf{x}$. Finally, when conditions (12a) and (12b) hold, the Lagrange multiplier λ is implicitly determined by

$$0 = \int_{\mathcal{W}} \left[1 - \overline{g}(\mathbf{x}) - \sum_{i=1}^{n} T_{x_i}(\mathbf{x}) \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \rho} \right] h(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{12c}$$

Proposition 1 provides necessary conditions for the government's optimum in the income space. It is consistent with Proposition 3 in Golosov, Tsyvinski, and Werquin (2014). The Euler–Lagrange equation (12a) provides a divergence equation that should hold for any income $\mathbf{x} \in \mathcal{X}$. Note that we use thrice differentiability of the tax schedule (Assumption 1(i) to derive the proposition. Equation (26b) in Appendix A.2 shows that the sufficient statistics $\partial X_i(\mathbf{w})/\partial \tau_j$ depend on the second-order partial derivatives of the tax schedule. Since the right-hand side of (12a) once again contains partial derivatives of these sufficient statistics, we require the tax system to be at least thrice differentiable. Equations (12b) are boundary conditions that should hold at any income $\mathbf{x} \in \partial\mathcal{X}$ in the boundary of $\mathcal{X}$. Finally, Equation (12c) states that, starting from the optimum, a lump sum perturbation implies no first-order effect on the Lagrangian. Using (6), the latter condition determines the Lagrange multiplier λ. To provide more intuition, the working paper version of this article contains a heuristic proof of Proposition 1 based on a reform that uniformly changes tax liability within a closed convex subset of $\mathcal{X}$ and changes marginal tax rates around that subset (Section III.3 of Spiritus et al. (2022)).

We use the Euler–Lagrange equation (12a) to derive a test to verify whether a given tax schedule is Pareto efficient, and, if not, what reform can lead to a Pareto improvement. We thus extend results by Werning (2007), Lorenz and Sachs (2016), Scheuer and Werning (2017), Hendren (2020), and Bierbrauer, Boyer, and Hansen (2023) to a setting with multiple incomes. Solving (12a) for $g(\mathbf{x})$, the revealed marginal welfare weights are defined as

$$\widehat{g}(\mathbf{x}) \stackrel{\text{def}}{\equiv} 1 - \sum_{i=1}^{n} T_{x_i}(\mathbf{x}) \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \rho} + \frac{1}{h(\mathbf{x})} \sum_{1 \le i, j \le n} \frac{\partial \left[T_{x_i}(\mathbf{x}) \dfrac{\partial \overline{X_i}(\mathbf{X}(\mathbf{w}))}{\partial \tau_j} h(\mathbf{x}) \right]}{\partial x_j}. \tag{13}$$

This formula extends the inverse-optimum approach of Bourguignon and Spadaro (2012), Bargain et al. (2014), and Jacobs, Jongen, and Zoutman (2017) to a setting with multidimensional incomes. If for some income $\mathbf{x}$ these revealed marginal welfare weights are negative, then there exists a Pareto improvement to the current tax schedule (see Appendix A.5 for the proof).

Proposition 2. *Under Assumptions 1 and 2:*

(i) An incremental tax perturbation that decreases tax liabilities where $\widehat{g}(\mathbf{x}) < 0$, and that does not change tax liabilities elsewhere, is Pareto improving.

(ii) A Pareto efficient tax schedule must lead to $\widehat{g}(\mathbf{x}) \ge 0$ for all $\mathbf{x} \in \mathcal{X}$.

Part (ii) of Proposition 2 provides a necessary condition in terms of observable statistics to test whether the current tax system is Pareto efficient. If the test fails, part (i) of Proposition 2 provides a Pareto improving tax reform. The Pareto improving reform we provide is different than the one provided by Lorenz and Sachs (2016) and Bierbrauer, Boyer, and Hansen (2023). In a unidimensional setting, one can decrease the marginal tax rate in a small income interval and decrease the tax liability above. The current situation is Pareto dominated if such a reform generates extra revenue for the government. In the multidimensional setting, such a reform is not feasible because it is geometrically not possible to change the gradient of the tax function at one income bundle without affecting this gradient for at least some other income bundles. This is the reason we consider reforms that decrease tax liability at income bundles where revealed welfare weights are negative instead of reforms changing the tax gradient. This is generalization of the Pareto-improving reforms studied by Scheuer and Werning (2017) and Hendren (2020) to a multidimensional context.[7]

We have shown that the optimality conditions in Proposition 1 and the condition for a Pareto improvement in Proposition 2 are valid if Assumptions 1 and 2 hold. As Assumption 2 may appear overly demanding, we now discuss a microfoundation to demonstrate its plausibility.

Assumption 2′. *The utility function $\mathcal{U}$ satisfies:*

(i) The number of incomes is equal to the number of unobserved characteristics $n = p$.

(ii) The matrix $[\mathcal{S}^i_{w_j}]_{i,j}$ is invertible.

(iii) The mapping $\mathbf{w} \mapsto (\mathcal{S}^1(c, \mathbf{x}; \mathbf{w}), \dots, \mathcal{S}^n(c, \mathbf{x}; \mathbf{w}))$ defined on $\mathcal{W}$ is injective.

Part (iii) of Assumption 2′ guarantees that at each income bundle $\mathbf{x} \in \mathcal{X}$, each vector of marginal rates of substitution is at most assigned to a single type. For $n = p = 1$, Parts (ii) and (iii) of Assumption 2′ are both equivalent to the standard single crossing condition. For $n = p > 1$, part (iii) of Assumption 2′ is stronger than part (ii), as the latter only demands local invertibility between the types and the marginal rates of substitution. Assumption 2′ is then a natural extension of the unidimensional single crossing condition, which corresponds to Assumption 1 in Dodds (2023). One case in which Assumption 2′ holds is when the utility function is additively separable:

$$\mathcal{U}(c, \mathbf{x}; \mathbf{w}) = \gamma(c) - \sum_{i=1}^{n} \upsilon^i(x_i, w_i) \quad \text{where } \gamma', \upsilon^i_{x_i}, \upsilon^i_{x_i, x_i} > 0 \neq \upsilon^i_{x_i, w_i}. \tag{14}$$

Both parts (ii) and (iii) of Assumption 2′ then become equivalent to $\upsilon^i_{x_i, w_i} \neq 0$.[8]

[7]The Pareto improving reforms of Scheuer and Werning (2017) and Hendren (2020) decentralize the Pareto improvements studied by Werning (2007) in a one-dimensional MD context. Bergstrom and Dodds (2023) extend our Proposition 2 for the presence of bunching.

[8]When the utility function takes the form (14), we get $\mathcal{S}^i(c, \mathbf{x}; \mathbf{w}) = \upsilon^i_{x_i}(x_i, w_i)/\gamma'(c)$. Assumption 2′ then amounts to demanding that the n one-dimensional mappings $w_i \mapsto \upsilon^i_{x_i}(x_i, w_i)/\gamma'(c)$ are injective, which is guaranteed by $\upsilon^i_{x_i, w_i}$ being either everywhere positive or everywhere negative.

Together, Assumptions 1 and 2′ guarantee that the tax schedule effectively separates taxpayers by type, so no two types choose the same bundle of incomes. We thus obtain the following lemma, which we prove in Appendix A.6.

Lemma 1. *Under Assumptions 1 and 2′, the mapping* $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ *is a continuously differentiable bijection from* $\mathcal{W}$ *into* $\mathcal{X}$*, and Assumption 2 holds.*

Lemma 1 allows us to rewrite the necessary conditions for the optimal tax schedule in the type space. This is important because the type space $\mathcal{W}$ is exogenous to the tax schedule whereas the income space $\mathcal{X}$ is not. In the numerical computations, this enables us to solve the Euler–Lagrange partial differential equation over a fixed space. Additionally, it is useful because we will also be able to retrieve this optimal tax formula in the type space using the MD approach, proving the consistency of the two approaches. We derive the following proposition in Appendix A.7.[9]

Proposition 3. *Under Assumption 2′, if the optimal tax schedule satisfies Assumption 1, the optimum verifies the Euler–Lagrange equation in the type space:*

$$\left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\frac{\partial X_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w}) = \sum_{1\le i,j\le n} \frac{\partial\big[T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\,\mathcal{A}_{j,i}(\mathbf{w})\,f(\mathbf{w})\big]}{\partial w_j}, \tag{15a}$$

for all $\mathbf{w}$ *in* $\mathcal{W}$*, while the boundary conditions become*

$$\sum_{1\le i,j\le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\,\mathcal{A}_{j,i}(\mathbf{w})\,e_j(\mathbf{w}) = 0, \tag{15b}$$

for all $\mathbf{w}$ *in* $\mathcal{W}$*, where the matrix* $\mathcal{A}$ *is defined by*

$$[\mathcal{A}_{i,j}]_{i,j} \stackrel{\text{def}}{\equiv} \big[\mathcal{S}^{i}_{w_j}\big]^{-1}_{i,j} = -\left[\frac{\partial X_i(\mathbf{w})}{\partial w_j}\right]^{-1}_{i,j} \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j}. \tag{15c}$$

Finally, the Lagrange multiplier λ *is implicitly determined by*

$$0 = \iint_{\mathcal{W}} \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\,\frac{\partial X_i(\mathbf{w})}{\partial \rho}\right] f(\mathbf{w})\,\mathrm{d}\mathbf{w}. \tag{15d}$$

4. The mechanism design approach

In this section, we rederive the optimal tax system using the mechanism-design approach instead of the tax-perturbation approach. This exercise serves two purposes. First, it allows us to verify under what conditions the two approaches result in the same

[9] Note that thrice differentiability of the tax schedule, as we assume in Assumption 1, remains necessary for the derivation of (15a), as it presumes knowledge of the Jacobian $[\partial X_i/\partial w_j]_{i,j}$ of the allocation. This Jacobian depends on the Hessian of the tax schedule, as we show in Appendix A.2.

optimal-tax function. Second, we use the mechanism-design approach to verify under what conditions the solution to the government's first-order conditions uniquely describes the social maximum.

The MD approach consists in optimizing over the set of allocations $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$ that verify the self-selection (or incentive) constraints:

$$\forall \mathbf{w}, \hat{\mathbf{w}} \in \mathcal{W}: \quad U(\mathbf{w}) \stackrel{\text{def}}{\equiv} \mathcal{U}\big(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big) \geq \mathcal{U}\big(C(\hat{\mathbf{w}}), \mathbf{X}(\hat{\mathbf{w}}); \mathbf{w}\big). \tag{16}$$

Instead of dealing with the double continuum of inequalities in (16), we follow Mirrlees (1976) by adopting a First-Order MD approach (henceforth the FOMD approach). The FOMD amounts to finding a continuously differentiable allocation $\mathbf{w} \mapsto (U(\mathbf{w}), \mathbf{X}(\mathbf{w}), C(\mathbf{w}))$ that verifies only the first-order incentive constraints:

$$\forall \mathbf{w} \in \mathcal{W}, \forall i \in \{1, \dots, p\}: \quad U_{w_i}(\mathbf{w}) = \mathcal{U}_{w_i}\big(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big), \tag{17}$$

and maximizes the government's Lagrangian:

$$\int_{\mathcal{W}} \left\{ \sum_{i=1}^{n} X_i(\mathbf{w}) - C(\mathbf{w}) + \frac{\Phi\big(U(\mathbf{w}); \mathbf{w}\big)}{\lambda} \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w} - E. \tag{18}$$

We restrict our attention to allocations that are continuously differentiable and satisfy the incentive constraint (16). This is formalized in Assumption 3.

Assumption 3. *The allocation $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$ is continuously differentiable and incentive-compatible, that is, it verifies (16).*

We divide the optimization problem into two stages. In the first stage, the government chooses the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$. In the second stage, the government chooses the incentive compatible allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ to maximize the resources extracted from taxpayers conditional on the utility profile chosen in the first stage, thus guaranteeing that a Pareto efficient allocation is chosen. Formally, the government chooses the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$ to maximize

$$\int_{\mathcal{W}} L\big(U(\mathbf{w}), U_{w_1}(\mathbf{w}), \dots, U_{w_p}(\mathbf{w}); \mathbf{w}, \lambda\big)\, \mathrm{d}\mathbf{w}, \tag{19}$$

where function $L(\cdot)$ is defined as

$$L(u, \mathbf{z}; \mathbf{w}, \lambda) \stackrel{\text{def}}{\equiv} \left(\mathcal{R}(u, \mathbf{z}; \mathbf{w}) + \frac{\Phi(u; \mathbf{w})}{\lambda} \right) f(\mathbf{w}) - E, \tag{20}$$

and the function $\mathcal{R}(\cdot)$ is defined via the subprogram

$$\begin{aligned} \mathcal{R}(u, \mathbf{z}; \mathbf{w}) &\stackrel{\text{def}}{\equiv} \max_{x_1, \dots, x_n} \sum_{i=1}^{n} x_i - \mathcal{C}(u; \mathbf{x}; \mathbf{w}) \\ \text{s.t. } \forall i \in \{1, \dots, p\}: &\quad z_i = \mathcal{U}_{w_i}\big(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big). \end{aligned} \tag{21}$$

Our approach differs from the traditional approach in Mirrlees (1976), Kleven, Kreiner, and Saez (2007), and Renes and Zoutman (2017), who directly maximize Lagrangian (18) subject to the incentive constraint (17) with respect to both the utility profile and the allocation. As noted by Mirrlees (1976, p. 343) and Kleven, Kreiner, and Saez (2007, p. 18), the traditional approach hides a conceptual problem in the multidimensional context. To see this, consider an example in which utility is additively separable as in (14). In that case, for any given candidate allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$, the first-order incentive constraints (17) form a system of partial differential equations in $\mathbf{w} \mapsto U(\mathbf{w})$. If there is only one type, $p = 1$, the system simplifies to an ordinary differential equation, which can be integrated to provide the corresponding mapping $\mathbf{w} \mapsto U(\mathbf{w})$, up to a constant. Conversely, when $p \geq 2$, the system of partial differential equations (17) for a given *candidate* mapping $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ yields a *candidate* for the gradient of $\mathbf{w} \mapsto U(\mathbf{w})$ with components $\mathbf{w} \mapsto Z_i(\mathbf{w}) \stackrel{\text{def}}{=} -v^i_{w_i}(X_i(\mathbf{w}), w_i)$ for all $i \in \{1, \dots, p\}$. However, not every combination of mappings $\mathbf{w} \mapsto Z_i(\mathbf{w})$ can be the gradient of a mapping $\mathbf{w} \mapsto U(\mathbf{w})$. The utility profile $\mathbf{w} \mapsto U(\mathbf{w})$ must exhibit symmetric second-order cross-derivatives, that is, $U_{w_j,w_k}(\mathbf{w}) = U_{w_k,w_j}(\mathbf{w})$ for all j, k, and all $\mathbf{w}$. Hence, only candidate mappings $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ that imply a utility profile that verifies $\partial Z_k(\mathbf{w})/\partial w_j = \partial Z_j(\mathbf{w})/\partial w_k$ for all j, k, and for all $\mathbf{w}$, are implementable. These additional *implementability* constraints are irrelevant in one-dimensional optimal tax problems but cannot be ignored in the multidimensional case. Our approach overcomes this challenge by explicitly choosing the utility profile $U(\mathbf{w})$ in the first stage, and choosing $C(\mathbf{w})$ and $\mathbf{X}(\mathbf{w})$ from the incentive-compatible allocations that implement that utility profile in the second stage. Therefore, the solution automatically satisfies the implementability condition $U_{w_i,w_j}(\mathbf{w}) = U_{w_j,w_i}(\mathbf{w})$.

To apply methods from variational calculus to the government's problem (19)–(21), we make regularity assumptions about subprogram (21) in Assumption 4. First, we rule out the possibility that two allocations that yield the same utility profile also extract an identical amount of resources. Second, we make differentiability assumptions about the unique solution to subprogram (21). Together, these assumptions ensure the differentiability of the function L, defined in Equation (20), with respect to all of its arguments. We will provide a plausible microfoundation in Assumption 4′.

Assumption 4. *Subprogram (21) admits a single solution for each* $(u, \mathbf{z}; \mathbf{w})$. *We denote this solution by* $\mathbb{X}_1(u, \mathbf{z}; \mathbf{w}), \dots, \mathbb{X}_n(u, \mathbf{z}; \mathbf{w})$ *and assume that it is twice continuously differentiable in* $(u, \mathbf{z}; \mathbf{w})$.

Note that subprogram (21) selects n incomes subject to p constraints. Assumption 4 thus implies that the MD approach is generally restricted to cases with at least as many incomes as types, $n \geq p$, although some exceptions exist, as we discuss in Section 5. Assumptions 3 and 4 allow us to derive necessary conditions for the FOMD problem (19) by considering continuously differentiable perturbations in the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$, and deducing the resulting perturbed allocations $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$ from subprogram (21). Assumption 4 ensures a unique perturbed allocation exists for every perturbed $\mathbf{w} \mapsto U(\mathbf{w})$. Mirrlees (1976, p. 342) implicitly makes a similar assumption to ensure his system of equations (63) admits a single solution. This leads to the following proposition, which we prove in Appendix B.1.

PROPOSITION 4. *Under Assumption 4, if among the allocations that verify the first-order incentive constraints (17), the optimal one verifies Assumption 3, then the optimal utility profile* $\mathbf{w} \mapsto U(\mathbf{w})$ *must verify for all* w *in* $\mathcal{W}$*:*

$$\big(1-\mathcal{S}^i\big)f(\mathbf{w}) = \mathcal{U}_c \sum_{j=1}^{p} \theta_j(\mathbf{w})\mathcal{S}^i_{w_j}, \quad \forall i \in \{1,\dots,n\}, \tag{22a}$$

$$\sum_{j=1}^{p} \frac{\partial \theta_j}{\partial w_j}(\mathbf{w}) = \left(\frac{1}{\mathcal{U}_c} - \frac{\Phi_u\big(U(\mathbf{w});\mathbf{w}\big)}{\lambda}\right) f(\mathbf{w}) - \sum_{j=1}^{p} \theta_j(\mathbf{w}) \frac{\mathcal{U}_{cw_j}}{\mathcal{U}_c}, \tag{22b}$$

$$0 = \sum_{j=1}^{p} \theta_j(\mathbf{w})\, e_j(\mathbf{w}), \quad \forall \mathbf{w} \in \partial\mathcal{W}, \tag{22c}$$

where we define

$$\theta_j(\mathbf{w}) \stackrel{\text{def}}{\equiv} -L_{z_j}\big(U(\mathbf{w}), U_{w_1}(\mathbf{w}), \dots, U_{w_p}(\mathbf{w}); \mathbf{w}, \lambda\big). \tag{22d}$$

Equation (22a) characterizes the optimal incomes $\mathbf{X}(\mathbf{w})$. Equation (22b) is the Euler–Lagrange equation characterizing the cost $\theta(w_j)$ of distorting the jth component of the gradient of $\mathbf{w} \mapsto U(\mathbf{w})$ (see (22d)). Equation (22c) corresponds to the boundary conditions that must hold along the boundary of the type space $\partial\mathcal{W}$. Equations (22a), (22b), and (22c), respectively, correspond to equations (60), (61), and (62) in Mirrlees (1976). Note that $\theta_j(\cdot)$ corresponds to the multiplier of the incentive constraints in Mirrlees (1976), as well as to the multiplier of the incentive constraints in the resource maximization subprogram (21). Our approach of perturbing $\mathbf{w} \mapsto U(\mathbf{w})$ and deducing the implied perturbation of the allocation $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$ from the first-order incentive constraints, thus shows that the shadow cost on the incentive constraint can be interpreted as the resource cost of changing $U(\mathbf{w})$.

In a setting where the number of incomes equals the number of characteristics, $n = p$, there is usually only one incentive-compatible allocation that can implement the same utility profile, because the number of free variables in the system of equations (17) is equal to the number of equations. In a setting with more incomes than types, $n > p$, the same utility profile can typically be offered through multiple incentive compatible allocations. In that case, through subprogram (21), the n first-order conditions (22a) can be decomposed into p conditions characterizing the optimal $\mathbf{w} \mapsto U(\mathbf{w})$ profile and $n-p$ supplementary conditions describing how to decentralize the mapping $\mathbf{w} \mapsto U(\mathbf{w})$ at the lowest cost.

Several results in the literature that are derived in settings where $n > p$ correspond to these supplementary conditions. A famous example is the Atkinson and Stiglitz (1976) theorem, which states for $p = 1$ that when preferences are weakly separable between leisure and consumption, commodity taxes should be uniform. This result remains valid regardless of social preferences for redistribution and can be seen as a way to realize a desired distribution of utilities with the least distortions (see also Jacobs and Boadway (2014)). In the same vein, Boadway and Keen (1993), Gauthier and Laroque (2009), and

Jacobs and de Mooij (2015) retrieve first-best principles such as the Samuelson rule for the provision of public good or the Pigouvian tax rule in case of externalities in models with weakly separable preferences and one-dimensional unobserved heterogeneity. Another strand of literature considers capital income taxation in settings with endogenous labor supply and savings and one dimension of unobserved heterogeneity. Assuming that preferences, inherited wealth or returns to capital vary along the ability distribution, the Atkinson and Stiglitz (1976)'s theorem no longer applies and the optimal capital tax is nonzero.[10] These authors show how to split the deadweight losses of redistribution between labor and capital income taxation, relying only on efficiency considerations without reference to social preferences for redistribution. Finally, the "new dynamic public finance" literature (Golosov, Kocherlakota, and Tsyvinski (2003)) considers models, where at each period, there is a new productivity drawn (a new dimension of unobserved heterogeneity) and agents make a labor supply and a saving decision at each period, such that $n = 2p$ where p is equal to the number of periods. The inverse Euler equation then describes how the planner should allocate consumption between the present period and each state of nature of the following period at the lowest cost. The finding that such supplementary efficiency conditions arise when $n > p$ is summarized in Corollary 1.

Corollary 1. *When $n > p$, subprogram (21) implies $n - p$ supplementary efficiency conditions describing how to decentralize a given mapping $\mathbf{w} \mapsto U(\mathbf{w})$ at the lowest cost.*

An additional advantage of our approach is that it becomes straightforward to provide conditions under which the government's necessary conditions (22) are also unique and sufficient. We do so in the following proposition.

Proposition 5. *Under Assumption 4, if for each type $\mathbf{w} \in \mathcal{W}$ and each $\lambda \in \mathbb{R}_+$ the mapping $(U, \mathbf{z}) \mapsto L(U, \mathbf{z}; \mathbf{w}, \lambda)$ is concave and if an allocation $\mathbf{w} \mapsto (U^\star(\mathbf{w}), C^\star(\mathbf{w}), \mathbf{X}^\star(\mathbf{w}))$ verifies Assumption 3 and Equations (22), then it is the unique solution to the government's problem.*

This result is especially important for our numerical simulations. We will demonstrate the use of Proposition 5 in Section 6, where we prove that in our simulations the mapping $(U, \mathbf{z}) \mapsto L(U, \mathbf{z}; \mathbf{w}, \lambda)$ is concave. As we find in the simulations an allocation that verifies the necessary conditions, Proposition 5 then ensures that this allocation is the unique solution to the government's problem.

We now provide a microfoundation to show the plausibility of Assumption 4. This additional assumption allows us to retrieve the optimal tax formula in the type space provided by Proposition 3 using the FOMD approach, instead of the TP approach.

[10]See, for example, Saez (2002), Cremer, Pestieau, and Rochet (2003), Diamond and Spinnewijn (2011), Gahvari and Micheletto (2016), Kristjánsson (2016), Saez and Stantcheva (2018), Gerritsen et al. (2025), Ferey, Lockwood, and Taubinsky (2022), Boadway and Spiritus (2024), and Zanoutene (2023).

Assumption 4′. *The number of incomes equals the number of unobserved characteristics, that is, $n = p$, and the mapping*

$$(u, \mathbf{x}; \mathbf{w}) \mapsto \big(\mathcal{U}_{w_1}\big(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big), \dots, \mathcal{U}_{w_p}\big(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big)\big)$$

is twice continuously differentiable in $(u, \mathbf{x}, \mathbf{w})$, and bijective in $\mathbf{x}$ with an invertible Jacobian.

When the utility function is of the additively separable form described in (14), Assumption 4′ is equivalent to $v^i_{x_i, w_i} \neq 0$. Hence, Assumption 4′ is a way to extend the single crossing condition to a multidimensional context. Recall that in subprogram (21), for given type $\mathbf{w}$, utility level u and utility gradient $\mathbf{z}$, the government chooses the incentive compatible allocation $\mathbf{x}$ that maximizes the government's revenues. Because of the incentive constraints of subprogram (21), Assumption 4′ implies that for each type $\mathbf{w}$, the value of $(u, \mathbf{z})$ uniquely determines the allocation $\mathbf{x}$. The subprogram thus admits a single solution and Assumption 4 is verified. When the number of incomes is equal to the number of unobserved characteristics ($n = p$), the incentive constraints of subprogram (21) imply that Assumption 4 also leads to Assumption 4′, so both assumptions are equivalent.

The fact that Assumptions 4 and 4′ are equivalent when $n = p$ enables us to show that the formulation of the optimality conditions in Proposition 3 can also be derived using the MD approach. We prove the following proposition in Appendix B.2.

Proposition 6. *Under Assumption 4′, if among the allocations that verify the first-order incentive constraints (17), the optimal one verifies Assumption 3, then the optimum verifies the Euler–Lagrange equation (15a) in the type space with Boundary conditions (15b), and the Lagrange multiplier λ is determined by (15d).*

We have thus shown that with the right assumptions, the same optimal-tax equations can be derived in the type space using either the FOMD approach or the TP approach. We elaborate further on the correspondence between the two approaches in the next section.

5. Comparing the TP and MD approaches

We now compare TP and MD approaches. In Section 5.1, we focus on the case when the numbers of incomes and characteristics are equal ($n = p$), before turning our attention to the settings where incomes outnumber characteristics ($n > p$, Section 5.2) and where characteristics outnumber incomes ($p > n$, Section 5.3).

5.1 *Equal numbers of incomes and characteristics ($n = p$)*

When the number of incomes n is equal to the number of characteristics p, we show in Propositions 3 and 6 that the same optimal tax formulas (15a)–(15c) can be obtained

using either the TP or the MD approach. We thus show the consistency of both approaches for $n = p > 1$, as Saez (2001) does for $n = p = 1$. This result may not seem surprising in light of the taxation principle (Hammond (1979)), which states that choosing an incentive-compatible allocation is equivalent to choosing a tax function. However, neither the TP nor the MD approaches in the literature solve the fully general case, as they both adopt smoothness assumptions.

For tractability reasons, we conduct both approaches under extensions of the single crossing condition to the multidimensional $n = p > 1$ case, namely Assumption 2′ in the TP approach, versus Assumption 4′ in the MD approach. Assumption 2′ states that for each $(c, \mathbf{x})$, the mapping $\mathbf{w} \mapsto (\mathcal{S}^1(c, \mathbf{x}; \mathbf{w}), \dots, \mathcal{S}^n(c, \mathbf{x}; \mathbf{w}))$ is globally invertible, that is, each bundle of marginal rates of substitution corresponds to at most one type. Assumption 4′ states that for each $(u, \mathbf{w})$, the mapping $\mathbf{x} \mapsto (\mathcal{U}_{w_1}(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}), \dots, \mathcal{U}_{w_p}(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}))$ is globally invertible, that is, each gradient of the utility profile corresponds to at most one vector of incomes. Note that local invertibility of both mappings are equivalent since the marginal utility of consumption times the Jacobian of $\mathbf{w} \mapsto (\mathcal{S}^1(c, \mathbf{x}; \mathbf{w}), \dots, \mathcal{S}^n(c, \mathbf{x}; \mathbf{w}))$ is equal to minus the transpose of the Jacobian of $\mathbf{x} \mapsto (\mathcal{U}_{w_1}(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}), \dots, \mathcal{U}_{w_p}(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}))$. Furthermore, if preferences are additively separable as in Equation (14), Assumptions 2′ and 4′ are equivalent to the single crossing assumptions $v^i_{x_i, w_i} \neq 0$ for $i = 1, \dots, n$. Otherwise, when preferences are not additively separable, the two assumptions differ only by assuming global invertibility of different mappings. From here on, we compare the TP and MD approaches when $n = p$ assuming preferences satisfy both Assumptions 2′ and 4′.

When the tax schedule is smooth in the sense of Assumption 1, we show in Appendix A.2 that the allocation is smooth in the sense of Assumption 3. Conversely, combining Assumption 3 with the left-hand side of individual first-order condition (3) implies that the marginal tax rates are only once-differentiable functions of type $\mathbf{w}$, while Assumption 1 requires they are twice-differentiable in incomes $\mathbf{x}$. Therefore, provided that the FOMD approach is valid, the TP approach is slightly more demanding than the MD approach as it requires stronger differentiability assumptions on the tax function.

A fundamental difference between the TP and MD approaches is the way they deal with bunching and jumping. Bunching and jumping are sometimes confused in the literature, but are actually polar issues. For instance, consider the case $n = p = 1$ with a single crossing condition in place. It follows that $w \mapsto X(w)$ is nondecreasing.[11] In that case, bunching arises when the optimal mapping $w \mapsto X(w)$ becomes constant. Conversely, jumping arises when $w \mapsto X(w)$ is upward discontinuous. From a tax function perspective, bunching implies the presence of a kink, and hence a nondifferentiability in the tax function. Jumping can arise even if the tax function is smooth, for instance, if the tax function is so concave in some region that taxpayers' second-order conditions cannot be met and they prefer to locate elsewhere. It is difficult to address jumping in an MD approach, since this requires optimization over non-continuous allocations.[12]

[11] See, for instance, Salanié (2003) for a formal proof.

[12] To our knowledge, only Hellwig (2010) addresses jumping with an MD approach, and he focuses only on the case $n = p = 1$.

Conversely, it is feasible to address bunching in the MD approach by adopting a second-order MD approach (see, e.g., Lollivier and Rochet (1983), Ebert (1992); and Boerma, Tsyvinski, and Zimin (2022)). On the other hand, the TP approach is better equipped to handle jumping, since jumping does not require a nondifferentiability in the tax function (see, e.g., Bergstrom and Dodds (2021)). Equivalence between the two approaches only arises when (i) one makes no smoothness assumptions whatsoever as in the taxation principle, or (ii) one makes sufficient assumptions to rule out both bunching and jumping as we do here.

A famous result in the multidimensional screening literature states that bunching is generic in the multidimensional nonlinear monopoly model of Armstrong (1996) and Rochet and Choné (1998). The latter state that bunching occurs "because of a strong conflict between participation constraints and second- order incentive compatible conditions." However, as there are no participation constraints in our optimal tax problem, the argument of Rochet and Choné (1998) for bunching does not apply to our model.

The absence of a participation constraint does not necessarily imply the absence of bunching. Boadway and Jacquet (2008) show that when the government's objective is maximin, the dual of the optimal tax problem consists in maximizing tax revenue subject to incentive constraints and a lower bound at the lowest utility level. The latter constraint is mathematically equivalent to the participation constraint in the monopoly model of Rochet and Choné (1998). In addition, Proposition 4 in Dodds (2023) shows that bunching is optimal in the optimal tax problem if preferences are sufficiently close to maximin, using a continuity argument. Conversely, under an additive social welfare function and quasilinear preferences in consumption, no bunching occurs. It follows, as argued by continuity by Kleven, Kreiner, and Saez (2007), that a range of moderate inequality aversions exists for which no bunching occurs. In our simulations, we address this by considering relatively light preferences for redistribution.

5.2 *Incomes outnumber characteristics ($n > p$)*

We now consider cases where incomes outnumber unobserved characteristics. The MD approach extends to the case where $n > p$. The n first-order conditions (22a) given in Proposition 4 can be decomposed into p conditions characterizing the optimal $\mathbf{w} \mapsto U(\mathbf{w})$ profile and $n - p$ supplementary efficiency conditions describing how to decentralize the mapping $\mathbf{w} \mapsto U(\mathbf{w})$ at the lowest cost. These supplementary conditions can be reinterpreted in a TP approach as describing how to minimize tax distortions while keeping the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$ unchanged.

It is much more difficult to apply the TP approach in this setting. This is because the range $\mathcal{X}$ of the type set under the allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ has a lower dimension than its containing space. Our definitions of the boundary $\partial\mathcal{X}$ and of the unit vector $\mathbf{e}(\mathbf{x}) = (e_1(\mathbf{x}), \dots, e_n(\mathbf{x}))$ normal to it in Proposition 1 are no longer meaningful and Proposition 1 loses its validity. More specifically, the divergence theorem used in the proof to Proposition 1 no longer applies.

Even if the TP approach as introduced by Golosov, Tsyvinski, and Werquin (2014), that we investigate in this paper, can no longer directly be applied in the context

where $n > p$, imposing additional assumptions may still enable the use of some version of the TP approach. Doing so generally requires assumptions that project the n-dimensional income space on the p-dimensional range $\mathcal{X}$ of the type set under the allocation. For instance, in the context $n = 2 > p = 1$, Gerritsen et al. (2025) assume that all incomes are increasing in ability, allowing them to project n-dimensional income to one-dimensional type. Ferey, Lockwood, and Taubinsky (2022) make a similar assumption in their Theorem 1 (see their Condition 2-UD) when they assume $n = 2 > p = 1$. Under their assumptions, one can retrieve the supplementary efficiency conditions by considering tax reforms that increase marginal tax rates on labor income and decrease marginal tax rates on capital income (or vice versa) for some taxpayers without changing tax liabilities for the others. Finding restrictions that allow for the projection of the n-dimensional income space to the p-dimensional type space is significantly more complicated when $n > p > 1$. Contrarily, the MD-approach naturally applies in the p-dimensional type space and, therefore, does not require such a projection when $n > p$.

5.3 *Characteristics outnumber incomes ($p > n$)*

When the number of characteristics is larger than the number of incomes, $p > n$, the TP approach continues to apply if one averages sufficient statistics among the different types assigned to the same income bundles. This averaging procedure enables projecting the p-dimensional type space into the n-dimensional income space. Assumption 2 ensures this projection can be done is a smooth way. Hence, we here generalize findings of Saez (2001), Scheuer and Werning (2016), and Jacquet and Lehmann (2021) to the case with multiple incomes $n \geq 2$.

Conversely, the MD approach does not immediately apply since subprogram (21) admits more constraints than free variables. This implies that starting from a utility profile $\mathbf{w} \mapsto U(\mathbf{w})$, which satisfies the incentive constraints, there exist perturbations to $\mathbf{w} \mapsto U(\mathbf{w})$ that cannot be made incentive compatible by changing the allocation $\mathbf{w} \mapsto (C(\mathbf{w}), \mathbf{X}(\mathbf{w}))$. This imposes additional constraints on the first stage (19) of the optimization problem (18). In the case where $p > n = 1$, there are different approaches to making these constraints tractable. First, Choné and Laroque (2010), Rothschild and Scheuer (2013, 2014, 2016) assume that labor supply decisions depend only on a one-dimensional function of type. Second, Jacquet and Lehmann (2023) assume preferences are additively separable between consumption and pre-tax income. Moreover, they assume that types matter only for the utility cost of earning pre-tax income. In all of these papers, restrictions on preferences imply that the type space can be projected onto a single dimension, thereby making the MD approach tractable. To the best of our knowledge, the literature has not yet been able to derive optimal tax formulas using the MD approach without such further assumptions on preferences when $p > n$. Moreover, this hurdle is even more difficult to overcome in the case where the types are multidimensional. While we do not exclude that ways can be found to apply the MD approach more generally to the cases where $p > n$, the TP approach deals with this case most naturally and Proposition 1 can be readily applied.

6. Numerical simulations

If both the type space and the income space are multidimensional, the optimal-tax formulas do not take the form of ordinary differential equations as in Mirrlees (1971), Diamond (1998), and Saez (2001), but they take the form of a second-order partial differential equation, as in Mirrlees (1976) and Golosov, Tsyvinski, and Werquin (2014). This significantly complicates the process of solving the optimal tax equations. To understand this, it helps to consider the effects of a tax perturbation from a geometric perspective. In the one-dimensional case, the change in the marginal tax rate at a given income level is directly connected to changes in tax liabilities at all higher incomes. In the multidimensional case, the relation is more complicated. To change the gradient of the tax function at a given point, one must change the tax liabilities near that point, causing changes in the tax gradient elsewhere; see, for instance, the graphical proof in the working paper version of this article, Section III.3 of Spiritus et al. (2022). To deal with this complexity, we rely on numerical simulations.

We develop a new numerical algorithm and apply it to the optimal taxation of couples. We consider an economy where couples differ in the productivity of females (w_f) and males (w_m), so unobserved heterogeneity is bi-dimensional ($p = 2$). Each couple chooses the labor supply of both spouses, so there are two incomes (i.e., $n = p = 2$). Preferences over the couple's consumption c, female income x_f and male income x_m are quasilinear in consumption, additively separable, and isoelastic in each income:

$$\mathcal{U}(c, x_f, x_m; w_f, w_m) = c - \frac{\varepsilon_f}{1+\varepsilon_f} x_f^{\frac{1+\varepsilon_f}{\varepsilon_f}} w_f^{-\frac{1}{\varepsilon_f}} - \frac{\varepsilon_m}{1+\varepsilon_m} x_m^{\frac{1+\varepsilon_m}{\varepsilon_m}} w_m^{-\frac{1}{\varepsilon_m}}. \tag{23}$$

The quasilinearity of taxpayers' preferences implies that there are no income effects (i.e., $\partial X_f(\mathbf{w})/\partial\rho = \partial X_m(\mathbf{w})/\partial\rho = 0$). Moreover, one can verify that if the tax schedule is additively separable, the cross-responses are equal to zero (i.e., $\partial X_f(\mathbf{w})/\partial\tau_m = \partial X_m(\mathbf{w})/\partial\tau_f = 0$). Finally, ε_f and ε_m, respectively, denote the direct elasticities of male and female incomes with respect to their own net-of-marginal-tax rates. Our baseline values are $\varepsilon_f = 0.43$ and $\varepsilon_m = 0.11$, which correspond to the mean labor supply elasticity for married women and for men in the meta-analysis of Bargain and Peichl (2016, Figure 1).

We calibrate the skill density $f(\cdot)$ using the Current Population Survey (CPS) of the US census of March 2016. We focus on married, mixed-gender couples that live together. We only consider income from labor. We drop couples in which either partner earns less than \$1000 per year or in which either of the partners' incomes is top-coded. We drop same-sex couples because in our simulations we attach labor elasticities based on gender in each couple. From each observed couple, we recover their type (w_f, w_m) from their labor earnings (x_f, x_m) by inverting the first-order conditions (3). For this purpose, we use a rough approximation of the current tax schedule in the US by assuming a constant marginal tax rate of 37%, a figure which is consistent with Barro and Redlick (2011, Table 1). Next, we estimate the type density through a bi-dimensional kernel. We specify the social welfare function to be CARA with $\Phi(u, w_1, w_2) = 1 - \exp(-\beta\, u)/\beta$, where $\beta > 0$ stands for the degree of inequality aversion. For our baseline simulation, we

select β such that the assumed 37% tax rate coincides with the optimal linear tax rate. This leads to $\beta = 0.0061$. Throughout the simulations, we assume that the government's revenue requirement equals 15% of GDP, which is close to the observed share of public spending in GDP for the US.

With our functional specifications, the government's Lagrangian (20) becomes

$$L(u, \mathbf{z}; \mathbf{w}, \lambda) = \left[\sum_{i=f,m} \left((1+\varepsilon_i)^{\frac{\varepsilon_i}{1+\varepsilon_i}} w_i \, z_i^{\frac{\varepsilon_i}{1+\varepsilon_i}} - \varepsilon_i \, z_i \, w_i \right) - u - \frac{\exp(-\beta \, u)}{\lambda} \right] f(\mathbf{w}),$$

which is concave in (u, z_f, z_m). Since the Lagrangian is concave, Proposition 5 applies, meaning that our optimal tax formulas are both necessary and sufficient for the unique optimum.

We first give an overview of the simulation algorithm, in Section 6.1. Next, in Section 6.2, we report the results of the simulations for the baseline calibration.

6.1 *Simulation algorithm*

The idea of our numerical algorithm is to first solve an optimal tax formula for given values of sufficient statistics, then to update the sufficient statistics using the tax schedule derived from the optimal tax formula, and to repeat this procedure until it converges to the optimal tax schedule. To do so, we can a priori use three optimal tax formulas, namely (12), (22), and (15). Let us explain why we choose (15). The optimal formula in (22) takes the form of a second-order *nonlinear* partial differential equation in the type space, which is numerically much more challenging than solving a linear second-order partial differential equation. Conversely, the optimal formula in Equations (12) is a *linear* second-order partial differential equation. However, it is defined in the income set $\mathcal{X}$. Hence, if one solves the optimal tax formula (12a) using the same income set from one iteration to the next, which is required given the boundary conditions (12b), then the corresponding typeset is changing from one iteration to the next. This is problematic when, for instance, comparing the values obtained for the tax revenue or for the social objective from one iteration to the next. Finally, the partial differential equation described in (15) is linear, provided that the sufficient statistics $g(\mathbf{w})$ and $\mathcal{A}(\mathbf{w})$ are taken as given. In addition, it is defined over the fixed type set $\mathcal{W}$.

Here again, there is a difficulty. Equations (15a)–(15b) are defined in the type space, while $(T_{x_1}, \dots, T_{x_n})$ stands for the gradient of tax liability with respect to incomes. However, one can rewrite (15a)–(15b) in terms of the gradient of tax liability in the skill space by scaling matrix $\mathcal{A}$ by the matrix $[\partial X_j(\mathbf{w})/\partial w_i]_{i,j}^{-1}$. We then iterate by (i) finding the mapping $\mathbf{w} \mapsto T(\mathbf{X}(\mathbf{w}))$ that solves Equations (15a)–(15b) for given matrix $\mathcal{A}$, Jacobian $[\partial X_j(\mathbf{w})/\partial w_i]_{i,j}$ and type density $f(\mathbf{w})$ and getting a tax schedule $\mathbf{x} \mapsto T(\mathbf{x})$ from this solution, and (ii) updating the matrix $\mathcal{A}$ and the Jacobian $[\partial X_j(\mathbf{w})/\partial w_i]_{i,j}$ given the new tax schedule. This hybrid approach thus combines the strength of the MD approach (a fixed typeset over which to integrate), with the strength of the TP approach (a linear PDE).

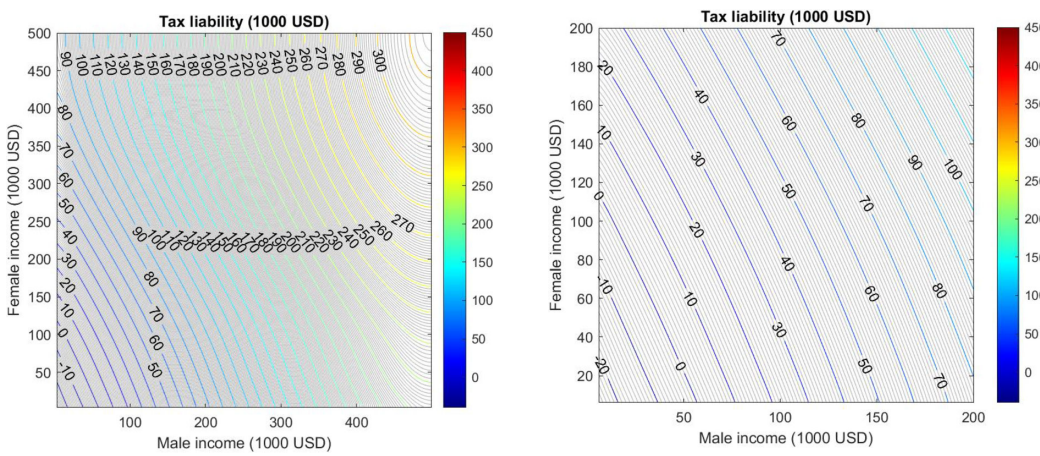
We describe the algorithm in more detail in the Supplementary Materials (available at http://econtheory.org/supp/5479/supplement.pdf).

6.2 *Results*

Figure 1 displays the solution to the optimal tax problem using our baseline calibration. The optimal tax schedule is represented by the isotax curves, which are the loci of incomes for which the tax liability is constant at a given value. Male income is shown on the horizontal axis, while female income is shown on the vertical axis. The left panel displays the whole domain of the simulations running up to \$500,000, while the right panel zooms in at incomes below \$200,000, where we find most taxpayers, roughly 97% of males and 99% of females.

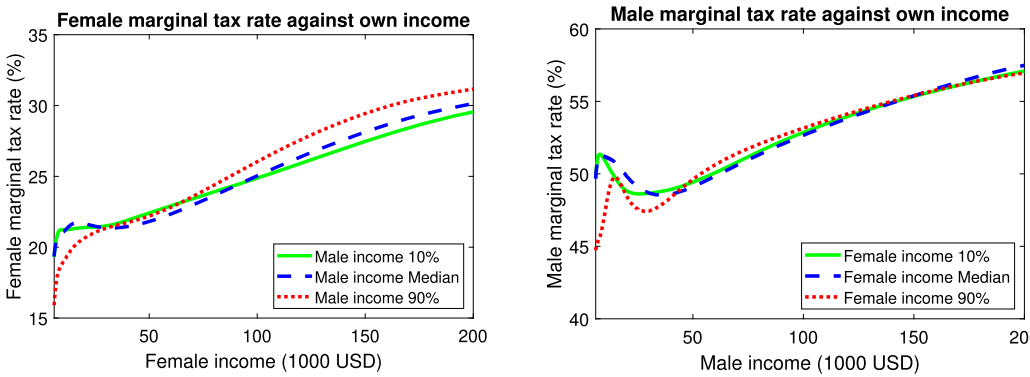
Strikingly, isotax curves are almost linear and parallel, except close to the boundaries. There, isotax curves are curved to satisfy boundary constraints (12b). This curvature pattern is most notable at high income levels where there are very few taxpayers. For lower incomes, the curvature only affects isotax curves very close to the lower bound.

Compared with the current economy, which is approximated by a linear tax rate of 37%, the optimal tax schedule leads to an improvement of the social objective equivalent to 0.82% of GDP in monetary terms. To understand which forces drive this gain, we decompose the welfare gain in different steps. Going from our approximation of the current economy (where we assume linear tax rates) to the optimal *joint* tax $(x_f, x_m) \mapsto T(x_f + x_m)$ captures the welfare gain of allowing the joint income tax schedule to be nonlinear. We find this welfare gain to be only 0.03%. If we now maintain the requirement that the isotax curves are linear and parallel but remove the requirement that both marginal tax rates are equal, so $(x_f, x_m) \mapsto T(x_f + \alpha\, x_m)$ where α is optimized, we obtain a welfare gain from the current economy equal to 0.81%. The optimal value of α is 2.13, which implies that female income is discounted by 53%. Hence, while the gain



(a) Full simulation set (b) Zoom on $x_f, x_m \leq \$200,000$

FIGURE 1. Isotax curves in the baseline case.



(a) Optimal marginal tax rates of female income against own income

(b) Optimal marginal tax rates of male income against own income

FIGURE 2. Optimal jointness.

of optimizing the slope of the isotax curves (optimizing α) is economically significant, the welfare gain of relaxing the constraint that isotax curves must be linear and parallel appears to be small.

Kleven, Kreiner, and Saez (2007) show that under our individual and social preferences, when the abilities of both spouses are not correlated, the optimal marginal tax rates of each partner decrease in the income of the other partner. This is the so-called *negative jointness* of the optimal tax system. In a separate simulation with a population that replicates the moments of male and female incomes, but removes any correlation between the two, we confirm the optimality of the negative jointness of the tax system. In reality, however, the assumption that the skills of both partners are not correlated does not hold. We show in Figure 2 that the optimality of negative jointness is not robust to using more realistic type densities with positive assortative matching. Figure 2a (resp., Figure 2b) displays the marginal tax rate for females (males) as a function of their own income. Each curve graphs this marginal income while fixing male (female) income at the 10th, 50th, and 90th percentile of the male (female) income distribution. In case of negative jointness, the curve corresponding to male (female) income at the 10th percentile should be everywhere above the curve corresponding to male (female) income at 50th and 90th percentiles of the distribution. Figures 2a and 2b contradict this prediction, thereby rejecting the idea that negative jointness holds at the optimum. We rather find that, except at the very bottom of the income distribution, marginal tax rates exhibit minimal jointness, since in Figures 2a and 2b the three lines are close.[13]

[13]Golosov and Krasikov (2023) show that when ability is correlated between spouses, optimal jointness depends on a complex interplay between redistributive and efficiency motives. Hence, except in the tails of the distribution (Lemma 7 and 8 of Golosov and Krasikov (2023)), there is little theoretical guidance on whether jointness should be positive or negative. Since we estimate the type density using a bi-dimensional kernel instead of a parametric distribution with a Pareto tail, we are unable to verify the asymptotic properties presented by Golosov and Krasikov (2023).

7. Conclusion

We study the optimal tax problem with multiple incomes and multiple dimensions of unobserved heterogeneity. We identify assumptions on the smoothness of the allocation and of the tax schedule, and multidimensional extensions of the single crossing assumptions, that enable the use of variational calculus to characterize the optimum. When comparing the MD approach to the TP approach, we demonstrate that when the numbers of types and of incomes are equal, the latter implies slightly more demanding restrictions on the smoothness of the tax schedule. When there are more unobserved characteristics than incomes, the TP approach is more suitable than the MD approach. Conversely, when there are more incomes than unobserved characteristics, the TP approach as we apply it cannot be used to solve the problem. We show that in terms of rigor, the TP method is on par with the MD method.

We propose a numerical algorithm that addresses the difficulties inherent to the multidimensional tax problem. We apply this algorithm to the optimal taxation of couples. Our findings indicate that the optimal isotax curves are nearly linear and parallel. We show that the optimal negative jointness of the tax schedules when skills are uncorrelated does not hold up when a more realistic distribution is introduced.

In addition to our primary findings, we obtain several theoretical results. First, we identify a necessary condition for the tax schedule to be Pareto efficient. If this condition is not met, we describe a Pareto-improving tax reform. Second, we identify conditions that ensure the necessary conditions of the optimal tax problem are unique and sufficient. Third, we contribute to the TP approach by proposing conditions under which income bundles respond smoothly to small tax reforms. Fourth, we introduce a MD approach that encapsulates not only incentive constraints, but also the implementability constraints embedded in the multidimensional optimal tax problem. Lastly, we examine the cases where the number of incomes differs from the number of characteristics.

Appendix A: TP approach

A.1 *Convexity of the indifference sets*

We verify that assuming convex indifference sets implies that the second-order conditions of the taxpayers' program strictly hold when the tax schedule is linear. On the one hand, the indifference sets are defined by $c = \mathcal{C}(u, \mathbf{x}; \mathbf{w})$. Applying the implicit function theorem to the definition of $\mathcal{C}(u, \mathbf{x}; \mathbf{w})$, we find the gradient of the indifference sets:

$$\mathcal{C}_{x_i}(u, \mathbf{x}; \mathbf{w}) = -\frac{\mathcal{U}_{x_i}\big(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big)}{\mathcal{U}_c\big(\mathcal{C}(u, \mathbf{x}; \mathbf{w}), \mathbf{x}; \mathbf{w}\big)}.$$

The Hessian is therefore a matrix with i^{th} row and j^{th} column:

$$\mathcal{C}_{x_i,x_j} = -\frac{\mathcal{U}_{x_i,x_j}\mathcal{U}_c - \mathcal{U}_{c,x_i}\dfrac{\mathcal{U}_{x_j}}{\mathcal{U}_c}\mathcal{U}_c - \mathcal{U}_{c,x_j}\mathcal{U}_{x_i} + \mathcal{U}_{c,c}\dfrac{\mathcal{U}_{x_j}}{\mathcal{U}_c}\mathcal{U}_{x_i}}{\mathcal{U}_c^2}.$$

On the other hand, from (2), we get

$$\mathcal{S}^i_{x_j} + \mathcal{S}^j \mathcal{S}^i_c = -\frac{\mathcal{U}_{x_i,x_j}\mathcal{U}_c - \mathcal{U}_{c,x_j}\mathcal{U}_{x_i}}{\mathcal{U}_c^2} + \frac{\mathcal{U}_{x_j}}{\mathcal{U}_c}\frac{\mathcal{U}_{c,x_i}\mathcal{U}_c - \mathcal{U}_{c,c}\mathcal{U}_{x_i}}{\mathcal{U}_c^2} = \mathcal{C}_{x_i,x_j}. \tag{24}$$

The assumption that indifference sets are convex thus implies that the matrix $[\mathcal{S}^i_{x_j} + \mathcal{S}^j\mathcal{S}^i_c]_{i,j}$ is symmetric and positive definite. If then taxes are linear, so $T_{x_ix_j} = 0$, Assumption 1 is fulfilled.

A.2 *Behavioral responses*

Assumption 1 enables differentiating (8) with respect to t, $\mathbf{x}$ and $\mathbf{w}$ to get

$$\begin{aligned}[\mathcal{C}_{x_jx_i} + T_{x_jx_i}]_{j,i} \cdot [\mathrm{d}x_i]_i = \big[R_{x_j}\big(\mathbf{X}(\mathbf{w})\big)\big]_j \,\mathrm{d}t - \big[\mathcal{S}^j_c\big]_j R\big(\mathbf{X}(\mathbf{w})\big)\,\mathrm{d}t \\ - \big[\mathcal{S}^j_{w_k}\big]_{j,k} \cdot [\mathrm{d}w_k]_k,\end{aligned} \tag{25}$$

where the expressions are evaluated at $t = 0$, $\mathbf{x} = \mathbf{X}(\mathbf{w})$, and $c = C(\mathbf{w})$, and we use (3) and (24).

A compensated reform of the jth marginal tax rate for taxpayers of type $\mathbf{w}$ is defined by

$$R(\mathbf{x}) = x_j - X_j(\mathbf{w}), \tag{26a}$$

where we use τ_j to denote the magnitude of these specific perturbations. It implies $R(\mathbf{X}(\mathbf{w})) = 0$, $R_{x_j}(\mathbf{X}(\mathbf{w})) = 1$ and $R_{x_k}(\mathbf{X}(\mathbf{w})) = 0$ for $k \neq j$. Using (25), the matrix of compensated responses for type $\mathbf{w}$ is

$$\left[\frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} = \left[\frac{\partial X_j(\mathbf{w})}{\partial \tau_i}\right]_{i,j} = [\mathcal{C}_{x_jx_i} + T_{x_jx_i}]^{-1}_{j,i}. \tag{26b}$$

Since the matrix of compensated responses is the inverse of the symmetric and positive definite matrix $[\mathcal{C}_{x_jx_i} + T_{x_jx_i}]_{j,i}$, it is also symmetric and positive definite.

A lump-sum perturbation of the tax function is defined by

$$R\big(\mathbf{X}(\mathbf{w})\big) = 1, \tag{26c}$$

where we use ρ to denote the magnitude of this specific perturbation. It is characterized by $R_{x_j}(\mathbf{X}(\mathbf{w})) = 0$. Using (25), the vector of income responses of type $\mathbf{w}$ is therefore given by

$$\left[\frac{\partial X_i(\mathbf{w})}{\partial \rho}\right]_i = -[\mathcal{C}_{x_jx_i} + T_{x_jx_i}]^{-1}_{j,i} \cdot \big[\mathcal{S}^j_c\big]_j = -\left[\frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} \cdot \big[\mathcal{S}^j_c\big]_j. \tag{26d}$$

Multiplying both sides of (25) by the matrix $[\mathcal{C}_{x_jx_i} + T_{x_jx_i}]^{-1}_{j,i}$ and using (26b)–(26d) leads to (9).

Finally, the implicit function theorem ensures that the mapping $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is differentiable for all $\mathbf{w} \in \mathcal{W}$ with a Jacobian given by

$$\left[\frac{\partial X_i(\mathbf{w})}{\partial w_k}\right]_{i,k} = -[\mathcal{C}_{x_j x_i} + T_{x_j x_i}]^{-1}_{j,i} \cdot \left[\mathcal{S}^j_{w_k}\right]_{j,k} = -\left[\frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} \cdot \left[\mathcal{S}^j_{w_k}\right]_{j,k}. \tag{26e}$$

Equation (26e) shows that when the tax schedule verifies Assumption 1 and individual preferences verify Assumption 2′, the ensuing allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ verifies Assumption 3.

A.3 *The derivative of the perturbed Lagrangian*

We first compute the response of tax liabilities $T(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)) - t\, R(\widetilde{\mathbf{X}}^R(\mathbf{w}, t))$ to a change in the magnitude t of the tax perturbation. Using (9) yields at $t = 0$:

$$\begin{aligned}\frac{\partial T\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big) - t\, R\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big)}{\partial t} &= \left[-1 + \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, \frac{\partial X_i(\mathbf{w})}{\partial \rho}\right] R\big(\mathbf{X}(\mathbf{w})\big) \\ &\quad + \sum_{1 \le i,j \le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, \frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\, R_{x_j}\big(\mathbf{X}(\mathbf{w})\big).\end{aligned} \tag{27}$$

Next, we evaluate the effect of the tax perturbation on the social objective. Apply the envelope theorem to (7) and use (6) to find at $t = 0$:

$$\frac{1}{\lambda} \frac{\partial \Phi\big(\widetilde{U}^R(\mathbf{w}, t); \mathbf{w}\big)}{\partial t} = g(\mathbf{w})\, R\big(\mathbf{X}(\mathbf{w})\big). \tag{28}$$

To find the derivative of (10) with respect to t, we add (27) to (28). We integrate the result over all types $\mathbf{w}$ to obtain (11). The above derivation is valid if Assumption 1 holds, regardless of whether $n \lesseqgtr p$.

A.4 *Proof of Proposition 1*

We rewrite (11) in terms of the income density $h(\cdot)$ (which is well-defined under Assumption 2):

$$\begin{aligned}\frac{\partial \widetilde{\mathcal{L}}^R(t)}{\partial t} = \int_{\mathcal{X}} \Bigg\{ &\left[\overline{g}\big(\mathbf{X}(\mathbf{w})\big) - 1 + \sum_{i=1}^{n} T_{x_i}(\mathbf{x})\, \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \rho}\right] R(\mathbf{x}) \\ &+ \sum_{1 \le i,j \le n} T_{x_i}(\mathbf{x})\, \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \tau_j}\, R_{x_j}(\mathbf{x}) \Bigg\} h(\mathbf{x})\, \mathrm{d}\mathbf{x}.\end{aligned} \tag{29}$$

When $n \le p$, the set $\mathcal{X}$ has the same dimensions as its containing space, and we can use the divergence theorem to integrate the term on the second line of this equation by

parts. Rearranging terms yields

$$\frac{\partial \widetilde{\mathcal{L}}^R(t)}{\partial t} = \oint_{\partial\mathcal{X}} \sum_{1\le i,j\le n} T_{x_i}(\mathbf{x})\, \frac{\partial \overline{X_i}\big(\mathbf{X}(\mathbf{w})\big)}{\partial \tau_j}\, h(\mathbf{x})\, e_j(\mathbf{x})\, R(\mathbf{x})\, \mathrm{d}\Sigma(\mathbf{x})$$
$$- \int_{\mathcal{X}} \left\{ \left[1 - \overline{g}\big(\mathbf{X}(\mathbf{w})\big) - \sum_{i=1}^{n} T_{x_i}(\mathbf{x})\, \frac{\partial \overline{X_i}(\mathbf{x})}{\partial \rho} \right] h(\mathbf{x}) \right.$$
$$\left. + \sum_{1\le i,j\le n} \frac{\partial \left[T_{x_i}(\mathbf{x})\, \frac{\partial \overline{X_i}\big(\mathbf{X}(\mathbf{w})\big)}{\partial \tau_j}\, h(\mathbf{x}) \right]}{\partial x_j} \right\} R(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{30}$$

If the tax schedule $T(\cdot)$ is optimal, (30) must equal 0 for all possible directions $R(\cdot)$. This is only possible if the Euler–Lagrange partial differential equation (12a) and the boundary conditions (12b) are both satisfied. Using the divergence theorem, Equations (12a) and (12b) implies (12c). Alternatively, we can use the lump-sum perturbation in (11), that is, we set $R(\mathbf{X}(\mathbf{w})) = 1$ and $R_{x_j}(\mathbf{X}(\mathbf{w})) = 0$ in (11) to retrieve (12c).

A.5 *Proof of Proposition 2*

Under Assumptions 1 and 2, the revealed welfare weights $\widehat{g}(\cdot)$ are defined such that $0 = \partial\mathcal{L}^R/\partial t$ for any perturbation R. However, using (28), we get that

$$\frac{\partial \mathcal{L}^R}{\partial t} = \int_{\mathcal{X}} \widehat{g}(\mathbf{x})\, R(\mathbf{x})\, h(\mathbf{x})\, \mathrm{d}\mathbf{x} + \int_{\mathcal{X}} \frac{\partial \big(T\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big) - t\, R\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big)\big)}{\partial t}\, h(\mathbf{x})\, \mathrm{d}\mathbf{x}.$$

Therefore, for any perturbation R, its effects on tax revenue are simply given by

$$\int_{\mathcal{X}} \frac{\partial T\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big) - t\, R\big(\widetilde{\mathbf{X}}^R(\mathbf{w}, t)\big)}{\partial t}\, h(\mathbf{x})\, \mathrm{d}\mathbf{x} = -\int_{\mathcal{X}} \widehat{g}(\mathbf{x})\, R(\mathbf{x})\, h(\mathbf{x})\, \mathrm{d}\mathbf{x}. \tag{31}$$

Note that the right-hand side of (13), and thus also $\widehat{g}(\cdot)$, is continuous with respect to $\mathbf{x}$. Consider a tax perturbation $\mathbf{x} \mapsto T(\mathbf{x}) - t\, R(\mathbf{x})$ where $R(\cdot)$ is twice continuously differentiable, positive where $\widehat{g}(\mathbf{x}) < 0$ and nil otherwise. Implementing such a perturbation with a small positive t increases taxpayers' welfare for those earning income bundles such that $\widehat{g}(\mathbf{x}) < 0$, and leave the other welfare unchanged, according to (28). Moreover, such perturbation is self-financed according to (31). It is therefore Pareto improving.

A.6 *Proof of Lemma 1*

Given that $\mathcal{X}$ is defined as the range of the typeset $\mathcal{W}$ under the allocation $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$, it is sufficient to show that the mapping $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is injective to establish that it is a bijection. Assume there exists $\mathbf{x} \in \mathcal{X}$ and $\mathbf{w}, \widehat{\mathbf{w}} \in \mathcal{W}$ such that $\mathbf{X}(\mathbf{w}) = \mathbf{X}(\widehat{\mathbf{w}}) = \mathbf{x}$. From Assumption 1, the first-order conditions (3) must be verified both at $(c, \mathbf{x}; \mathbf{w})$ and at $(c, \mathbf{x}; \widehat{\mathbf{w}})$, so we get $\mathcal{S}^i(c, \mathbf{x}, \mathbf{w}) = \mathcal{S}^i(c, \mathbf{x}, \widehat{\mathbf{w}})$ for all $i \in \{1, \dots, n\}$. According to part (iii)

of Assumption 2′, these n equalities imply that $\mathbf{w} = \widehat{\mathbf{w}}$. Differentiability of $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is ensured under Assumption 1 by the implicit function theorem applied to (3). Part (ii) of Assumption 2′ then ensures the Jacobian of $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is invertible (see (26e) in Appendix A.2).

Because the mapping $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is injective, we get that $\overline{g}(\mathbf{X}(\mathbf{w})) = g(\mathbf{w})$, $\partial\overline{X_i}(\mathbf{X}(\mathbf{w}))/\partial\tau_j = \partial X_i(\mathbf{w})/\partial\tau_j$ and $\partial\overline{X_i}(\mathbf{X}(\mathbf{w}))/\partial\rho = \partial X_i(\mathbf{w})/\partial\rho$. According to Equations (6), (26b), and (26d), $g(\mathbf{w})$, $\partial X_i(\mathbf{w})/\partial\tau_j$ and $\partial X_i(\mathbf{w})/\partial\rho$ are continuously differentiable functions of c, $\mathbf{x}$, $\mathbf{w}$, and for the latter two of the terms $T_{x_i x_j}$ in the Hessian of the tax schedule. Hence, because the mapping $\mathbf{w} \mapsto \mathbf{X}(\mathbf{w})$ is continuously differentiable and invertible, and because of part *(iv)* of Assumption 2′, $\partial\overline{X_i}(\mathbf{x})/\partial\tau_j$, $\partial\overline{X_i}(\mathbf{x})/\partial\rho$ and $\overline{g}(\mathbf{x})$ are continuously differentiable in $\mathbf{x}$. Finally, the income density is given by

$$h\big(\mathbf{X}(\mathbf{w})\big) = f(\mathbf{w})\left|\det\left[\frac{\partial X_i(\mathbf{w})}{\partial w_j}\right]_{i,j}\right|^{-1}, \tag{32}$$

which ensures the income density is also continuously differentiable in income. Hence, Assumption 2 holds.

A.7 *Proof of Proposition 3*

To get an optimal tax formula in the type space, we need to rewrite the derivative of the perturbed Lagrangian, (11), in the type space rather than in the income space. To reparametrize the direction of a tax perturbation as a function of types, define $\widehat{R}(\mathbf{w}) \stackrel{\text{def}}{=} R(\mathbf{X}(\mathbf{w}))$. Differentiating both sides with respect to w_j yields

$$\widehat{R}_{w_j}(\mathbf{w}) = \sum_{i=1}^{n}\big(\partial X_i(\mathbf{w})/\partial w_j\big)\, R_{x_i}\big(\mathbf{X}(\mathbf{w})\big).$$

In matrix notation, the latter equality becomes

$$\begin{aligned}
&\big[\widehat{R}_{w_j}(\mathbf{w})\big]_j^T = \big[R_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\big]_i^T \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial w_j}\right]_{i,j} \\
\Leftrightarrow\quad &\big[R_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\big]_i^T = \big[\widehat{R}_{w_j}(\mathbf{w})\big]_j^T \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial w_j}\right]_{i,j}^{-1},
\end{aligned}$$

where we use parts (i) and (ii) of Assumption 2′ and Equation (26e) to ensure that matrix $[\partial X_i(\mathbf{w})/\partial w_j]_{i,j}$ is invertible. Using the symmetry of the matrix of compensated effects $[\partial X_i(\mathbf{w})/\partial\tau_j]_{i,j}$, we can rewrite the last term of (11):

$$\begin{aligned}
&\sum_{1\le i,j\le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\frac{\partial X_i(\mathbf{w})}{\partial\tau_j}\, R_{x_j}\big(\mathbf{X}(\mathbf{w})\big) \\
&\quad = \big[R_{x_j}(\mathbf{X}(\mathbf{w})\big]_j^T \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial\tau_j}\right]_{i,j} \cdot \big[T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\big]_i
\end{aligned}$$

$$= \left[\widehat{R}_{w_j}(\mathbf{w})\right]_j^T \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial w_j}\right]_{i,j}^{-1} \cdot \left[\frac{\partial X_i(\mathbf{w})}{\partial \tau_j}\right]_{i,j} \cdot \left[T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\right]_i$$

$$= -\left[\widehat{R}_{w_j}(\mathbf{w})\right]_j^T \cdot \left[\mathcal{S}_{w_j}^i\right]_{i,j}^{-1} \cdot \left[T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\right]_i,$$

where the last equality follows from (26e). Using the definition of matrix $\mathcal{A}_{i,j}(\mathbf{w})$ in (15c), Equation (11) can be rewritten as

$$\frac{\partial \widetilde{\mathcal{L}}^R(t)}{\partial t} = \int_{\mathcal{W}} \left\{ \left[g(\mathbf{w}) - 1 + \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big) \frac{\partial X_i(\mathbf{w})}{\partial \rho} \right] \widehat{R}(\mathbf{w}) \right.$$
$$\left. - \sum_{1 \le i,j \le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, \mathcal{A}_{j,i}(\mathbf{w})\, \widehat{R}_{w_j}(\mathbf{w}) \right\} f(\mathbf{w})\, \mathrm{d}\mathbf{w}.$$

Using the divergence theorem to perform integration by parts, we get

$$\frac{\partial \widetilde{\mathcal{L}}(t)}{\partial t} = -\oint_{\partial \mathcal{W}} \sum_{1 \le i,j \le n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, \mathcal{A}_{j,i}(\mathbf{w})\, e_j(\mathbf{w})\, f(\mathbf{w})\, \widehat{R}(\mathbf{w})\, \mathrm{d}\Sigma(\mathbf{w})$$
$$- \int_{\mathcal{W}} \left\{ \left[1 - g(\mathbf{w}) - \sum_{i=1}^{n} T_{x_i}\big(\mathbf{X}(\mathbf{w})\big) \frac{\partial X_i(\mathbf{w})}{\partial \rho} \right] f(\mathbf{w}) \right.$$
$$\left. + \sum_{1 \le i,j \le n} \frac{\partial \big(T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, \mathcal{A}_{j,i}(\mathbf{w})\, f(\mathbf{w})\big)}{\partial w_j} \right\} \widehat{R}(\mathbf{w})\, \mathrm{d}\mathbf{w}.$$

This partial derivative equals zero for any direction of tax perturbation $\widehat{R}(\cdot)$ if and only if Euler–Lagrange Equation (15a) and boundary conditions (15b) are verified.

Appendix B: FOMD approach

B.1 *Proof of Proposition 4*

Let R be a twice differentiable function defined over $\mathcal{W}$ into $\mathbb{R}$. We consider the effects of perturbing the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$ in the direction R. Using (20), define

$$\widetilde{L}^R(t) \stackrel{\text{def}}{\equiv} \int_{\mathcal{W}} L\big(U(\mathbf{w}) + t\, R(\mathbf{w}), U_{w_1}(\mathbf{w}) + t\, R_{w_1}(\mathbf{w}), \ldots,$$
$$\ldots, U_{w_p}(\mathbf{w}) + t\, R_{w_p}(\mathbf{w}); \mathbf{w}, \lambda\big)\, \mathrm{d}\mathbf{w}. \tag{33}$$

Applying the chain rule and denoting $\langle \mathbf{w} \rangle$ as a shortcut to denote that a function is evaluated at $(U(\mathbf{w}), U_{w_1}(\mathbf{w}), \ldots, U_{w_p}(\mathbf{w}); \mathbf{w}, \lambda)$, we obtain at $t = 0$:

$$\frac{\partial \widetilde{L}^R(t)}{\partial t} = \int_{\mathcal{W}} \left\{ L_u \langle \mathbf{w} \rangle R(\mathbf{w}) + \sum_{j=1}^{p} L_{z_j} \langle \mathbf{w} \rangle\, R_{w_j}(\mathbf{w}) \right\} \mathrm{d}\mathbf{w}.$$

Applying integration by parts using the divergence theorem leads to

$$\frac{\partial \widetilde{L}^R(t)}{\partial t} = \int_{\mathcal{W}} \left\{ L_u\langle \mathbf{w}\rangle - \sum_{j=1}^{p} \frac{\partial L_{z_j}\langle \mathbf{w}\rangle}{\partial w_j} \right\} R(\mathbf{w})\,\mathrm{d}\mathbf{w} + \oint_{\partial\mathcal{W}} \sum_{j=1}^{p} L_{z_j}\langle \mathbf{w}\rangle e_j(\mathbf{w}) R(\mathbf{w})\,\mathrm{d}\Sigma(\mathbf{w}).$$

At the optimal allocation, the latter expression is nil for any perturbation R.[14] Using (22d), we find boundary conditions (22c), and the Euler–Lagrange equation

$$\forall \mathbf{w} \in \mathcal{W}: \quad \sum_{j=1}^{p} \frac{\partial \theta_j(\mathbf{w})}{\partial w_j} = -L_u\langle \mathbf{w}\rangle. \tag{34}$$

Using incentive compatibility constraint (17), we can rewrite (20) as

$$\begin{aligned}\left[\sum_{i=1}^{p} X_i(\mathbf{w}) - \mathcal{C}\big(U(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big) + \frac{\Phi\big(U(\mathbf{w}); \mathbf{w}\big)}{\lambda}\right] f(\mathbf{w}) \\ = L\big(U(\mathbf{w}), \mathcal{U}_{w_1}\big(\mathcal{C}\big(U(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big), \mathbf{X}(\mathbf{w}); \mathbf{w}\big), \ldots, \\ \ldots, \mathcal{U}_{w_n}\big(\mathcal{C}\big(U(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w}\big), \mathbf{X}(\mathbf{w}); \mathbf{w}\big); \mathbf{w}, \lambda\big).\end{aligned} \tag{35}$$

Differentiating both sides of (35) with respect to $X_i(\mathbf{w})$ and using (2) and (22d):

$$\big(1 - \mathcal{S}^i\langle \mathbf{w}\rangle\big) f(\mathbf{w}) = -\sum_{j=1}^{p} \theta_j(\mathbf{w})\big[\mathcal{U}_{x_i w_j}\langle \mathbf{w}\rangle + \mathcal{S}^i\langle \mathbf{w}\rangle \mathcal{U}_{c w_j}\langle \mathbf{w}\rangle\big],$$

which leads to (22a) given that $\mathcal{S}^i_{w_j} = (\mathcal{U}_{c w_j}\, \mathcal{U}_{x_i} - \mathcal{U}_{x_i w_j}\, \mathcal{U}_c)/\mathcal{U}_c^2 = -[\mathcal{U}_{x_i w_j} + \mathcal{S}^i\, \mathcal{U}_{c w_j}]/\mathcal{U}_c$. Differentiating (35) with respect to $U(\mathbf{w})$ and using $\mathcal{C}_u = 1/\mathcal{U}_c$ and (22d) leads to:

$$\left(-\frac{1}{\mathcal{U}_c\langle \mathbf{w}\rangle} + \frac{\Phi_u\big(U(\mathbf{w}); \mathbf{w}\big)}{\lambda}\right) f(\mathbf{w}) = L_u\langle \mathbf{w}\rangle - \sum_{j=1}^{p} \theta_j(\mathbf{w}) \frac{\mathcal{U}_{c,w_j}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle}. \tag{36}$$

Substituting (34) into (36) yields (22b).

B.2 *Derivation of the optimal tax formula in the type space*

Using (3), Equation (22a) leads to

$$T_{x_i}\big(\mathbf{X}(\mathbf{w})\big)\, f(\mathbf{w}) = \sum_{j=1}^{p} \mu_j(\mathbf{w})\, \mathcal{S}^i_{w_j}\langle \mathbf{w}\rangle, \tag{37}$$

where we denote $\mu_j(\mathbf{w}) \overset{\text{def}}{\equiv} \theta_j(\mathbf{w})\ \mathcal{U}_c(C(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w})$. This can be rewritten $[T_{x_i}(\mathbf{X}(\mathbf{w}))]_i f(\mathbf{w}) = [\mathcal{S}^i_{w_j}]_{i,j} \cdot [\mu_j(\mathbf{w})]_j$ in matrix notation, which leads to $[\mu_j(\mathbf{w})]_j =$

[14] Depending on the specification of the utility function, the image of the functions U and the U_{w_i} may be a finite interval, implying that the domain of L may be restricted. We assume the optimum exists and is interior. Note that this is impossible if there are more types than incomes ($p > n$). Investigating cases where the optimum binds feasibility constraints is beyond the scope of the present paper.

$[\mathcal{S}^i_{w_j}]^{-1}_{i,j} \cdot [T_{x_i}(\mathbf{X}(\mathbf{w}))]_i f(\mathbf{w})$. Using (15c), we therefore get

$$\forall \mathbf{w} \in \mathcal{W}, \forall i \in \{1, \dots, p\}: \quad \mu_i(\mathbf{w}) = \sum_{j=1}^{n} \mathcal{A}_{i,j}(\mathbf{w})\, T_{x_j}\big(\mathbf{X}(\mathbf{w})\big)\, f(\mathbf{w}). \tag{38}$$

Combining (22c) with (38) thus leads to (15b). Using (6), Equation (22b) implies that

$$\begin{aligned}\sum_{j=1}^{p} \frac{\partial \mu_j(\mathbf{w})}{\partial w_j} &= \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) - \sum_{j=1}^{p} \theta_j(\mathbf{w})\, \mathcal{U}_{cw_j}\langle \mathbf{w}\rangle \\ &\quad + \sum_{j=1}^{p} \theta_j(\mathbf{w}) \left[\mathcal{U}_{cc}\langle \mathbf{w}\rangle \frac{\partial C(\mathbf{w})}{\partial w_j} + \sum_{i=1}^{n} \mathcal{U}_{cx_i}\langle \mathbf{w}\rangle \frac{\partial X_i(\mathbf{w})}{\partial w_j} + \mathcal{U}_{cw_j}\langle \mathbf{w}\rangle \right] \\ &= \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) + \sum_{j=1}^{p} \theta_j(\mathbf{w}) \left[\mathcal{U}_{cc}\langle \mathbf{w}\rangle \frac{\partial C(\mathbf{w})}{\partial w_j} + \sum_{i=1}^{n} \mathcal{U}_{cx_i}\langle \mathbf{w}\rangle \frac{\partial X_i(\mathbf{w})}{\partial w_j} \right]. \end{aligned} \tag{39}$$

Differentiating $C(\mathbf{w}) = \mathcal{C}(U(\mathbf{w}), \mathbf{X}(\mathbf{w}); \mathbf{w})$ with respect to w_j and using $\mathcal{C}_u = 1/\mathcal{U}_c$, $\mathcal{C}_{x_i} = -\mathcal{U}_{x_i}/\mathcal{U}_c$, $\mathcal{C}_{w_j} = -\mathcal{U}_{w_j}/\mathcal{U}_c$ and (17) leads to

$$\frac{\partial C(\mathbf{w})}{\partial w_j} = \frac{\mathcal{U}_{w_j}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle} - \sum_{i=1}^{n} \frac{\mathcal{U}_{x_i}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle} \frac{\partial X_i(\mathbf{w})}{\partial w_j} - \frac{\mathcal{U}_{w_j}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle} = -\sum_{i=1}^{n} \frac{\mathcal{U}_{x_i}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle} \frac{\partial X_i(\mathbf{w})}{\partial w_j}.$$

Plugging this equality into (39) leads to

$$\begin{aligned}\sum_{j=1}^{p} \frac{\partial \mu_j(\mathbf{w})}{\partial w_j} &= \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) + \sum_{j=1}^{p} \sum_{i=1}^{n} \theta_j(\mathbf{w}) \left[\mathcal{U}_{cx_i}\langle \mathbf{w}\rangle - \frac{\mathcal{U}_{x_i}\langle \mathbf{w}\rangle}{\mathcal{U}_c\langle \mathbf{w}\rangle} \mathcal{U}_{cc}\langle \mathbf{w}\rangle \right] \frac{\partial X_i(\mathbf{w})}{\partial w_j} \\ &= \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) - \sum_{j=1}^{p} \sum_{i=1}^{n} \mu_j(\mathbf{w})\, \mathcal{S}^i_c\langle \mathbf{w}\rangle\, \frac{\partial X_i(\mathbf{w})}{\partial w_j}. \end{aligned} \tag{40}$$

Substituting (26e) into (40) yields

$$\sum_{j=1}^{p} \frac{\partial \mu_j(\mathbf{w})}{\partial w_j} = \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) + \sum_{j=1}^{p} \sum_{i=1}^{n} \sum_{k=1}^{n} \mu_j(\mathbf{w})\, \mathcal{S}^i_c\langle \mathbf{w}\rangle\, \frac{\partial X_i(\mathbf{w})}{\partial \tau_k}\, \mathcal{S}^k_{w_j}(\mathbf{w}). \tag{41}$$

Substituting (26d) into (41) and using $\frac{\partial X_i(\mathbf{w})}{\partial \tau_k} = \frac{\partial X_k(\mathbf{w})}{\partial \tau_i}$ yields

$$\sum_{j=1}^{p} \frac{\partial \mu_j(\mathbf{w})}{\partial w_j} = \big(1 - g(\mathbf{w})\big) f(\mathbf{w}) - \sum_{j=1}^{p} \sum_{k=1}^{n} \mu_j(\mathbf{w})\, \mathcal{S}^k_{w_j}(\mathbf{w})\, \frac{\partial X_k(\mathbf{w})}{\partial \rho}. \tag{42}$$

Plugging (37) into (42) leads to

$$\sum_{j=1}^{p} \frac{\partial \mu_j(\mathbf{w})}{\partial w_j} = \left(1 - g(\mathbf{w}) - \sum_{k=1}^{n} T_{x_k}\big(\mathbf{X}(\mathbf{w})\big)\, \frac{\partial X_k(\mathbf{w})}{\partial \rho} \right) f(\mathbf{w}). \tag{43}$$

Plugging (38) into (43) leads to (15a). The last equality in (15c) follows from (26e).

B.3 *Proof of Proposition 5*

If $(u, \mathbf{z}) \mapsto L(u, \mathbf{z}; \mathbf{w}, \lambda)$ is concave, then for any perturbation p, the function $t \mapsto \widetilde{L}^R(t)$ defined in (33) is concave. Let $\mathbf{w} \mapsto U(\mathbf{w})$ be another utility profile that verifies (22a) and take the perturbation $R(\mathbf{w}) = U(\mathbf{w}) - U^\star(\mathbf{w})$. As the utility profile $\mathbf{w} \mapsto U(\mathbf{w})$ verifies (22), we get that function $t \mapsto \widetilde{L}^R(t)$ admits a zero derivative at $t = 0$ and is concave. So, $\widetilde{L}^R(0) > \widetilde{L}^R(1)$ and $U^\star(\cdot)$ provides a strictly higher welfare than $U(\cdot)$.

If two distinct allocations $\mathbf{w} \mapsto U^\star(\mathbf{w})$ and $\mathbf{w} \mapsto U(\mathbf{w})$ verify (22), then following the reasoning above, $U(\cdot)$ strictly dominates $U^\star(\cdot)$ and $U^\star(\cdot)$ strictly dominates $U(\cdot)$, a contradiction. So, at most one allocation can verify (22).

References

Armstrong, Mark (1996), "Multiproduct nonlinear pricing." *Econometrica*, 64, 51–75. [0547, 0563]

Atkinson, Anthony and Joseph Stiglitz (1976), "The design of tax structure: Direct versus indirect taxation." *Journal of Public Economics*, 6, 55–75. [0547, 0559, 0560]

Bargain, Olivier, Mathias Dolls, Dirk Neumann, Andreas Peichl, and Sebastian Siegloch (2014), "Comparing inequality aversion across countries when labor supply responses differ." *International Tax and Public Finance*, 21, 845–873. [0545, 0554]

Bargain, Olivier and Andreas Peichl (2016), "Own-wage labor supply elasticities: Variation across time and estimation methods." *IZA Journal of Labor Economics*, 5, 1–31. [0545, 0565]

Barro, Robert and Charles Redlick (2011), "Macroeconomic effects from government purchases and taxes." *Quarterly Journal of Economics*, 126, 51–102. [0565]

Basov, Suren (2005), *Multidimensional Screening Studies in Economic Theory*. Springer-Verlag Berlin Heidelberg. [0547]

Bergstrom, Katy and William Dodds (2021), "Optimal taxation with multiple dimensions of heterogeneity." *Journal of Public Economics*, 200, 104442. [0548, 0563]

Bergstrom, Katy and William Dodds (2023), "A general theory of inverse welfare functions." Tulane University Working Papers 2308. [0551, 0555]

Bierbrauer, Felix, Pierre Boyer, and Emanuel Hansen (2023), "Pareto-improving tax reforms and the earned income tax credit." *Econometrica*, 91, 1077–1103. [0545, 0554, 0555]

Boadway, Robin and Laurence Jacquet (2008), "Optimal marginal and average income taxation under maximin." *Journal of Economic Theory*, 143, 425–441. [0547, 0563]

Boadway, Robin and Michael Keen (1993), "Public goods, self-selection and optimal income taxation." *International Economic Review*, 34, 463–478. [0559]

Boadway, Robin and Kevin Spiritus (2024), "Optimal taxation of normal and excess returns to risky assets." *The Scandinavian Journal of Economics*. [0560]

Boerma, Job, Aleh Tsyvinski, and Alexander Zimin (2022), "Bunching and taxing multidimensional skills." NBER Working Paper 30 015. [0548, 0563]

Bourguignon, François and Amadeo Spadaro (2012), "Tax-benefit revealed social preferences." *Journal of Economic Inequality*, 10, 75–108. [0545, 0554]

Choné, Philippe and Guy Laroque (2010), "Negative marginal tax rates and heterogeneity." *American Economic Review*, 100, 2532–2547. [0546, 0548, 0564]

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet (2001), "Direct versus indirect taxation: The design of the tax structure revisited." *International Economic Review*, 42, 781–799. [0548]

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet (2003), "Capital income taxation when inherited wealth is not observable." *Journal of Public Economics*, 87, 2475–2490. [0548, 0560]

Diamond, Peter (1998), "Optimal income taxation: An example with U-shaped pattern of optimal marginal tax rates." *American Economic Review*, 88, 83–95. [0565]

Diamond, Peter and Johannes Spinnewijn (2011), "Capital income taxes with heterogeneous discount rates." *American Economic Journal: Economic Policy*, 3, 52–76. [0560]

Diamond, Peter A. (1975), "A many-person Ramsey tax rule." *Journal of Public Economics*, 4, 335–342. [0550]

Dodds, William (2023), "Solving multidimensional screening problems using a generalized single crossing property." *Economic Theory*, 77, 1025–1084. [0547, 0555, 0563]

Ebert, Udo (1992), "A reexamination of the optimal nonlinear income tax." *Journal of Public Economics*, 49, 47–73. [0563]

Ferey, Antoine, Benjamin Lockwood, and Dmitry Taubinsky (2022), "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity." CESifo Working Paper 9958. [0547, 0560, 0564]

Frankel, Alexander (2014), "Taxation of couples under assortative mating." *American Economic Journal: Economic Policy*, 6, 155–177. [0548]

Gahvari, Firouz and Luca Micheletto (2016), "Capital income taxation and the Atkinson-Stiglitz theorem." *Economics Letters*, 147, 86–89. [0560]

Gauthier, Stéphane and Guy Laroque (2009), "Separability and public finance." *Journal of Public Economics*, 93, 1168–1174. [0559]

Gerritsen, Aart, Bas Jacobs, Alexandra Rusu, and Kevin Spiritus (2025), "Optimal taxation of capital income when rates of return are heterogeneous." *The Economic Journal*, 135, 180–211. [0547, 0560, 0564]

Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003), "Optimal indirect and capital taxation." *Review of Economic Studies*, 70, 569–587. [0547, 0560]

Golosov, Mikhail and Ilia Krasikov (2023), "The optimal taxation of couples." NBER Working Papers 31140. [0547, 0568]

Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin (2014), "A variational approach to the analysis of tax systems." NBER Working Papers 20780. [0544, 0546, 0547, 0550, 0554, 0563, 0565]

Gomes, Renato, Jean-Marie Lozachmeur, and Alessandro Pavan (2018), "Differential taxation and occupational choice." *The Review of Economic Studies*, 85, 511–557. [0548]

Hammond, Peter (1979), "Straightforward individual incentive compatibility in large economies." *The Review of Economic Studies*, 46, 263–282. [0544, 0562]

Hellwig, Martin (2010), "Incentive problems with unidimensional hidden characteristics: A unified approach." *Econometrica*, 78, 1201–1237. [0562]

Hendren, Nathaniel (2020), "Measuring economic efficiency using inverse-optimum weights." *Journal of Public Economics*, 187, 104198. [0545, 0554, 0555]

Jacobs, Bas and Robin Boadway (2014), "Optimal linear commodity taxation under optimal non-linear income taxation." *Journal of Public Economics*, 117, 201–210. [0559]

Jacobs, Bas and Ruud A. de Mooij (2015), "Pigou meets Mirrlees: On the irrelevance of tax distortions for the second-best Pigouvian tax." *Journal of Environmental Economics and Management*, 71, 90–108. [0560]

Jacobs, Bas, Egbert Jongen, and Floris Zoutman (2017), "Revealed social preferences of Dutch political parties." *Journal of Public Economics*, 156, 81–100. [0545, 0554]

Jacquet, Laurence and Etienne Lehmann (2021), "Optimal income taxation with composition effects." *Journal of the European Economic Association*, 19, 1299–1341. [0546, 0548, 0552, 0564]

Jacquet, Laurence and Etienne Lehmann (2023), "Optimal tax problems with multidimensional heterogeneity: A mechanism design approach." *Social Choice and Welfare*, 60, 135–164. [0546, 0564]

Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez (2007), "The optimal income taxation of couples as a multi-dimensional screening problem." CESifo Working Paper 2092, CESifo. [0545, 0546, 0547, 0558, 0563, 0568]

Kristjánsson, Arnaldur Sölvi (2016), "Optimal taxation with endogenous return to capital." Memorandum 06/2016, Oslo University, Department of Economics. [0560]

Lockwood, Benjamin and Matthew Weinzierl (2015), "De Gustibus non est Taxandum: Heterogeneity in preferences and optimal redistribution." *Journal of Public Economics*, 124, 74–80. [0548]

Lollivier, Stefan and Jean-Charles Rochet (1983), "Bunching and second-order conditions: A note on optimal tax theory." *Journal of Economic Theory*, 31, 392–400. [0563]

Lorenz, Normann and Dominik Sachs (2016), "Identifying Laffer bounds: A sufficient-statistics approach with an application to Germany." *Scandinavian Journal of Economics*, 118, 646–665. [0554, 0555]

Mirrlees, James (1971), "An exploration in the theory of optimum income taxation." *Review of Economic Studies*, 38, 175–208. [0552, 0565]

Mirrlees, James (1976), "Optimal tax theory: A synthesis." *Journal of Public Economics*, 6, 327–358. [0544, 0546, 0547, 0557, 0558, 0559, 0565]

Renes, Sander and Floris Zoutman (2017), "As easy as ABC? Multidimensional screening in public finance." CESifo Working Paper 6301. [0543, 0558]

Rochet, Jean-Charles and Philippe Choné (1998), "Ironing, sweeping, and multidimensional screening." *Econometrica*, 66, 783–826. [0547, 0548, 0563]

Rothschild, Casey and Florian Scheuer (2013), "Redistributive taxation in the Roy model." *The Quarterly Journal of Economics*, 128, 623–668. [0546, 0548, 0564]

Rothschild, Casey and Florian Scheuer (2014), "A theory of income taxation under multidimensional skill heterogeneity." Technical report, National Bureau of Economic Research. [0546, 0548, 0564]

Rothschild, Casey and Florian Scheuer (2016), "Optimal taxation with rent-seeking." *The Review of Economic Studies*, 83, 1225–1262. [0546, 0548, 0564]

Saez, Emmanuel (2001), "Using elasticities to derive optimal income tax rates." *Review of Economic Studies*, 68, 205–229. [0546, 0550, 0562, 0564, 0565]

Saez, Emmanuel (2002), "Optimal income transfer programs: Intensive versus extensive labor supply responses." *Quarterly Journal of Economics*, 117, 1039–1073. [0560]

Saez, Emmanuel and Stefanie Stantcheva (2018), "A simpler theory of optimal capital taxation." *Journal of Public Economics*, 162, 120–142. [0560]

Salanié, Bernard (2003), *The Economics of Taxation*. The MIT Press. [0562]

Scheuer, Florian (2014), "Entrepreneurial taxation with endogenous entry." *American Economic Journal: Economic Policy*, 6, 126–163. [0546, 0548]

Scheuer, Florian and Iván Werning (2016), "Mirrlees meets Diamond-Mirrlees." NBER Working Papers 22076. [0546, 0564]

Scheuer, Florian and Iván Werning (2017), "The taxation of superstars." *The Quarterly Journal of Economics*, 132, 211–270. [0545, 0554, 0555]

Spiritus, Kevin (2017), "Characterizations for optimal multidimensional taxation with multidimensional heterogeneity of agents." Report. [0543]

Spiritus, Kevin, Etienne Lehmann, Sander Renes, and Floris Zoutman (2022), "Optimal taxation with multiple incomes and types." CESifo Working Paper Series 9534. [0552, 0553, 0554, 0565]

Werning, Iván (2007), "Pareto efficient taxation." Report. [0545, 0554, 0555]

Zanoutene, Eddy (2023), "Scale-dependent and risky returns to savings: Consequences for optimal capital taxation." *Journal of Public Economic Theory*, 25, 532–569. [0560]

Co-editor Pierre-Olivier Weill handled this manuscript.

Manuscript received 21 October, 2022; final version accepted 23 September, 2024; available online 23 October, 2024.