TI 2022-018/VI
Tinbergen Institute Discussion Paper

Optimal Commodity Taxation Under Non-linear Income Taxation

Kevin Spiritus[1]

1 Erasmus University Rotterdam

Optimal Commodity Taxation Under Non-linear Income Taxation

Kevin Spiritus*

21st February 2022

Abstract

I characterize the optimal linear commodity taxes when households differ in multiple characteristics, in presence of an optimal non-linear tax schedule on the households' labour incomes. The optimal distortions caused by a linear commodity tax are larger if, conditional on labour income, more deserving individuals consume more of that commodity. This is the case for merit goods, or if the government otherwise seeks to compensate individuals who consume larger quantities of that commodity. Furthermore, the government wishes to tax commodities at different rates to the extent that doing so reduces the distortions caused by the labour income tax. This is the case when individuals with different incomes have different preferences, or when individuals who supply different quantities of labour have different consumption patterns. I extend these findings to the case where households earn multiple incomes.

Keywords: optimal commodity taxation, multidimensional taxation
JEL Classification: H21, H24

*Department of Economics, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands, e-mail: spiritus@ese.eur.nl. I am grateful to Aart Gerritsen and Bas Jacobs for valuable comments.

1 Introduction

The optimal taxation of commodities remains a contentious issue. International organizations and academic reviews recommend that VAT rates are uniform (e.g., OECD, 2020, p.42 and Mirrlees et al., 2011, ch.9). Yet most governments do apply reduced VAT rates. By doing so, they aim to increase the purchasing power of low-income groups, to stimulate the consumption of merit goods such as culture, and to redistribute towards specific groups, for example, through reduced rates for female hygiene products.[1]

The recommendation that commodities are taxed at uniform rates is rooted in a seminal paper by Atkinson and Stiglitz [1976]. They show that when individual preferences are weakly separable between leisure and consumption and individuals differ only in their labour productivities, commodities should indeed be taxed at uniform rates. The reason is that with such a one-dimensional population, if all individuals have equal preferences, individual consumption patterns contain no information about the individuals' types beyond what is already revealed by the individuals' labour incomes. Given then that any redistribution attained by differentiated commodity taxes can also be attained by a tax on labour income with equal distortions of labour supply but without distortions of consumption decisions, it is best to not use differentiated commodity taxes for redistribution. Only when there are clear complementarities between labour supply and the consumption of specific goods, most notably childcare and transportation to work, can differentiated commodity taxes help reduce the distortions caused by a progressive tax on labour income (Gordon and Kopczuk, 2014 and Pirttilä and Suoniemi, 2014).

A notable shortcoming of the model of Atkinson and Stiglitz [1976] is that individuals differ only in their labour productivities. Saez [2002] studies the desirability of differentiated linear commodity taxes when individuals also differ in their tastes. Studying linear commodity taxes besides a non-linear income tax is useful due to the possibility of arbitrage between taxpayers. Saez [2002] shows that starting from a situation without commodity taxes, introducing a small linear tax on one commodity increases social welfare if, on average, higher-income individuals have a relatively higher taste for consuming this commodity. Remarkably, to the best of my know-

[1] See, for example, OECD [2020] and Terricabras [2016]. According to OECD [2020], other goods enjoy reduced VAT rates because they are locally supplied and their production is labour-intensive. In this paper, I do not consider this motive.

ledge, no optimal tax formulas exist when the assumptions leading to the Atkinson and Stiglitz [1976] result are not fulfilled. Indeed, Saez [2002, p.229] leaves this as an "extremely useful" task for future research.

In this paper, I derive formulas for the optimal linear taxation of commodities in a broad generalization of the model of Atkinson and Stiglitz [1976]. By doing so, I identify the sufficient statistics that, when quantified, allow determining the optimal commodity tax rates. Extending Saez [2002], I allow for taxpayers to differ in multiple characteristics such as tastes, productivities, gender, etc. Furthermore, I allow for taxpayers to earn multiple labour incomes. Allowing multiple labour incomes in the household allows studying, for example, the effect of different complementarities of primary and secondary labour supply with the consumption of certain goods (e.g., childcare). I briefly characterize the optimal non-linear schedule for the multidimensional labour income tax, and then focus on the optimal linear tax rates on commodities.[2]

Main Results My results are as follows. If the tax schedule on the labour incomes is optimal, then one should only deviate from uniform commodity taxes if doing so either contributes to the distributional objectives of the government in ways that cannot be accomplished through the tax on labour income, or if it reduces the distortions caused by the taxation of labour income. We know from Saez [2001] and Werquin et al. [2015] that the optimal tax schedule on the labour incomes balances the economic costs of perturbing the tax schedule against the distributional benefits of doing so, taking into account all the information contained in the labour incomes. The sufficient statistics that are then needed to characterize the optimal commodity taxes, are conditional on labour income.

A first reason to deviate from uniform commodity taxes is when, on average and conditional on labour income, the covariance between the welfare weights and the consumption of a particular good differs from zero. This can be the case when a strong preference for that good is considered to be socially desirable (merit goods), or when taxpayers who, conditional on income, consume larger amounts of that good are otherwise more deserving (e.g. to compensate for other injustices).[3] The larger is the covariance conditional on income between the welfare weights and the consumption

[2]See Mirrlees [1976], Werquin et al. [2015] and Spiritus et al. [2022] for a treatment of optimal non-linear tax schedules in a multidimensional setting.

[3]In this paper, a good is a merit good if the government considers a strong taste for it socially desirable. I disregard positive externalities.

of that particular good, the larger is the distortion that the government should accept from taxing that good. This motive for deviating from uniform commodity taxes is absent if taxpayers differ only in their labour productivities.

A second reason to deviate from uniform commodity taxes is when, on average and conditional on labour income, there is a difference between the cross-sectional variation and the individual variation in the consumption of some good as one changes one of the labour incomes. This can be the case for two reasons. First, it is possible that around the income under consideration, individuals with different incomes have different tastes for that good. Second, it is possible that taxpayers who supply different quantities of labour, consume different quantities of the good under consideration. In both cases, the consumption patterns reveal information about the types of the individuals, allowing for more efficient redistribution than through the labour tax alone. The acceptable size of the distortion from taxing the good under consideration is then larger if the difference between the cross-sectional variation and the individual variation in the consumption of that good is larger as one varies a particular labour income, and if the marginal excess burden of taxing that labour income is larger.

A third reason to deviate from uniform commodity taxes is more technical. If on average and conditional on labour income, households for whom the individual variation of the consumption of some good with one of the labour incomes is larger also face a larger marginal excess burden of taxing that labour income, then shifting the burden of taxation from that labour income to the good under consideration increases the efficiency of the tax system. As noted by Saez [2002], there is little reason why such covariance would exist, and it would in any case be difficult to measure empirically. This third term is absent in models where individuals differ only in their labour productivities.

Relationship to Existing Literature This paper makes several contributions to the literature. Saez [2002] studies the linear taxation of commodities when individuals differ in their labour productivities and their tastes, but they earn only one labour income. He only finds desirability conditions for non-uniform commodity taxes, and does not characterize the optimal tax rates. Jacobs and Boadway [2014] do derive optimal linear commodity taxes, for the case with one labour income and where individuals differ only in their labour productivities. Because they consider a population that differs in one dimension only, they find less reasons than Saez [2002] to levy differentiated commodity taxes. I extend the literature by deriving the optimal linear commodity tax rates for the general case, allowing households to differ in characteristics beyond their labour productivities and preferences, and to earn multiple labour incomes.

Road Map We proceed as follows. I introduce the model in Section 2. In Section 3, I derive properties of individual behaviour, and I introduce sufficient statistics conditional on labour income. I derive the optimal tax schedules in Section 4, before concluding in Section 5.

2 The model

2.1 Households

Households differ in labour productivities $\boldsymbol{w} \equiv (w^1, \ldots, w^L) \in \mathscr{W}$, where $L \geq 1$ and $\mathscr{W} \subset \mathbb{R}^L_+$ is closed and convex. For each productivity dimension $l = 1, \ldots, L$, households supply labour ℓ^l, yielding labour income $z^l = w^l \ell^l$. Households pay labour income taxes $T(\boldsymbol{z})$, where the tax schedule can depend on all labour incomes $\boldsymbol{z} \equiv (z^1, \ldots, z^L)$ simultaneously in complicated ways.

Households spend their disposable income on *numéraire* good c and commodities $\boldsymbol{x} \equiv (x^1, \ldots, x^J)$. Producer prices are normalized to one, and commodities are taxed at linear rates $\boldsymbol{t} \equiv (t^1, \ldots, t^J)$. The budget constraint for a household is:

$$c + \sum_{j=1}^{J} x^j (1 + t^j) = \sum_{l=1}^{L} z^l - T(\boldsymbol{z}). \tag{1}$$

Besides their labour productivities, households can differ in other dimensions. We parametrise the additional characteristics using the vector $\boldsymbol{\theta} \equiv (\theta^1, \ldots, \theta^K) \in \Theta \subset \mathbb{R}^K_+$. We denote the type vectors as $\boldsymbol{\omega} \equiv (\boldsymbol{w}, \boldsymbol{\theta}) \in \Omega \equiv \mathscr{W} \times \Theta$. We denote the joint cumulative distribution of the types as $G^{\boldsymbol{\omega}}(\boldsymbol{\omega})$ with corresponding twice continuously differentiable density function $g^{\boldsymbol{\omega}}(\boldsymbol{\omega})$. We do not make assumptions about the joint distribution of the types.

Each household makes its decisions as one unit. They choose consumption c and $\boldsymbol{x}$, and labour incomes $\boldsymbol{z}$, taking into account their budget constraint (1), to maximize a thrice continuously differentiable, weakly concave utility function:

$$\max_{c,\boldsymbol{x},\boldsymbol{z}} u\left(c, \boldsymbol{x}, \frac{z^1}{w^1}, \ldots, \frac{z^L}{w^L}, \boldsymbol{\omega}\right), \tag{2}$$

with partial derivatives $u_c > 0$, $u_{x^j} > 0$ and $u_{\ell^l} < 0$, and with indifference sets that are strictly convex in $(c, \boldsymbol{x}, \boldsymbol{z})$ for all utility levels and types. We assume that for all households, the second-order conditions hold for optimization problem (1)–(2), and that the household's optimization problem admits a single global maximum. These assumptions, together with the assumption that the tax schedule $\boldsymbol{z} \mapsto T(\boldsymbol{z})$ is twice continuously differentiable, prevent that small tax perturbations lead to "jumps" in the choices of the households.[4] For a household of type $\boldsymbol{\omega}$ and for tax policies T and $\boldsymbol{t}$, we denote indirect utility as $v(\boldsymbol{\omega}, T, \boldsymbol{t})$.

2.2 Government

The government sets its tax policies to maximize an additive social welfare function subject to a budget constraint. The social welfare function is:[5]

$$\max_{T,\boldsymbol{t}} \iint_{\Omega} W\left(v(\boldsymbol{\omega}, T, \boldsymbol{t}), \boldsymbol{\omega}\right) \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega}), \tag{3}$$

where $(v, \omega) \mapsto W(v, \omega)$ is a twice continuously differentiable, increasing, and weakly concave transformation of household utility that captures the government's normative preferences. Note that the weight given to a household can depend on its type $\boldsymbol{\omega}$.[6] The government's budget constraint is:

$$\iint_{\Omega} \left(T(\boldsymbol{z}) + \sum_{j=1}^{J} t^j x^j\right) \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega}) \geq E, \tag{4}$$

with E an exogenously given revenue requirement.

[4] See Spiritus et al. [2022] for a proof.

[5] $\iint_{\Omega}$ denotes a multidimensional integral over the set Ω.

[6] I remain ambiguous about how the government should compare individuals with different preferences. See Fleurbaey and Maniquet [2018] for a review of principles of fairness in optimal taxation theory.

3 Individual behaviour

The first-order conditions for household optimization problem (1)–(2) for labour incomes $l = 1, \dots, L$ and commodities $j = 1, \dots, J$ are:

$$\frac{u_{\ell^l}}{u_c} = w^l(1 - T_{z^l}) \text{ and } \frac{u_{x^j}}{u_c} = 1 + t^j. \tag{5}$$

For each household type $\boldsymbol{\omega}$, we denote the optimal labour incomes as $\boldsymbol{Z}(\boldsymbol{\omega}) \equiv (Z^1(\boldsymbol{\omega}), \dots, Z^L(\boldsymbol{\omega}))$, the optimal consumption of the numéraire good as $C(\boldsymbol{\omega})$, and the optimal consumption of the other commodities as $\boldsymbol{X}(\boldsymbol{\omega}) \equiv (X^1(\boldsymbol{\omega}), \dots, X^J(\boldsymbol{\omega}))$. We denote the income space as the set $\mathscr{Z} \equiv \{\boldsymbol{z} \in \mathbb{R}_+^L | \exists \boldsymbol{\omega} \in \mathscr{W} : \boldsymbol{Z}(\boldsymbol{\omega}) = \boldsymbol{z}\}$.

With the assumptions made in Subsection 2.1, the households' responses to small perturbations of the tax policies are well defined. Let ρ denote a lump sum perturbation, and let τ^l denote a compensated perturbation of the l-th marginal tax rate, such that the perturbed tax schedule is:

$$\boldsymbol{z} \mapsto T(\boldsymbol{z}) - \rho + \sum_{l=1}^{L} \tau^l \cdot (z^l - Z^l(\boldsymbol{\omega})). \tag{6}$$

We denote the behavioural responses to an increase in the lump sum income for any commodity j as $\partial X^j(\boldsymbol{\omega})/\partial\rho$, and for any labour income l as $\partial Z^l(\boldsymbol{\omega})/\partial\rho$. We denote the behavioural responses to a compensated perturbation of the m-th marginal income tax rate as $\partial X^{j*}(\boldsymbol{\omega})/\partial\tau^m$ and $\partial Z^{l*}(\boldsymbol{\omega})/\partial\tau^m$. Lastly, we denote the behavioural responses to a change of the linear tax rate on the i-th commodity as $\partial X^j(\boldsymbol{\omega})/\partial t^i$ and $\partial Z^l(\boldsymbol{\omega})/\partial t^i$.

Standard derivations yield Slutsky decompositions (omitting function arguments, and with asterisks denoting compensated effects):

$$\forall j, m : \frac{\partial X^m}{\partial t^j} = \frac{\partial X^{m*}}{\partial t^j} - \frac{\partial X^m}{\partial \rho} X^j, \tag{7}$$

and:

$$\forall j, l : \frac{\partial Z^l}{\partial t^j} = \frac{\partial Z^{l*}}{\partial t^j} - \frac{\partial Z^l}{\partial \rho} X^j, \tag{8}$$

and Slutsky symmetries:

$$\forall l, m, j, k : \frac{\partial Z^{l*}}{\partial \tau^m} = \frac{\partial Z^{m*}}{\partial \tau^l}, \quad \frac{\partial X^{j*}}{\partial \tau^l} = \frac{\partial Z^{l*}}{\partial t^j} \quad \text{and} \quad \frac{\partial X^{k*}}{\partial t^j} = \frac{\partial X^{j*}}{\partial t^k}. \tag{9}$$

Following Christiansen [1984], Saez [2002] and Jacobs and Boadway [2014], we split the decision process of the individuals into two phases. In the second phase, individuals take their labour incomes and thus their disposable incomes as given, and they decide how to spend their money on consumption goods. In the first phase, individuals choose their labour incomes, taking the ensuing outcome of the second phase as given. Conditional on the combination of labour incomes $\boldsymbol{z}$, we denote the corresponding optimal consumption of the j-th good as $X^{jc}(\boldsymbol{\omega}, \boldsymbol{z})$. The compensated response of the consumption of the j-th consumption good to a reform in the k-th commodity tax can then be decomposed as follows:

$$\frac{\partial X^{j*}(\boldsymbol{\omega})}{\partial t^k} = \frac{\partial X^{jc*}(\boldsymbol{\omega}, \boldsymbol{Z}(\boldsymbol{\omega}))}{\partial t^k} + \sum_{m=1}^{L} \frac{\partial X^{jc}(\boldsymbol{\omega}, \boldsymbol{Z}(\boldsymbol{\omega}))}{\partial z^m} \frac{\partial Z^{m*}(\boldsymbol{\omega})}{\partial t^k}. \tag{10}$$

Keeping constant the chosen combination of labour incomes, a perturbation of the tax on labour income does not directly affect the consumption choices. We thus find:

$$\frac{\partial X^{j*}(\boldsymbol{\omega})}{\partial \tau^l} = \sum_{m=1}^{L} \frac{\partial X^{jc}(\boldsymbol{\omega}, \boldsymbol{Z}(\boldsymbol{\omega}))}{\partial z^m} \frac{\partial Z^{m*}(\boldsymbol{\omega})}{\partial \tau^l}. \tag{11}$$

4 Optimal policies

Saez [2001] was the first to characterize the optimal non-linear tax schedule on a single labour income in terms of sufficient statistics. Werquin et al. [2015] did the same for the case with multiple labour incomes. These issues go far beyond the scope of the present paper: I simply state the optimality conditions for the tax on the labour incomes in Subsection 4.1, and take these optimality conditions as given in Subsection 4.2 to characterize the optimal linear commodity taxes.

4.1 Optimal tax on labour incomes

The derivations and the interpretation of the optimal tax schedule on the labour incomes are beyond the scope of this paper. Werquin et al. [2015] and Spiritus et al.

[2022] show that, given our assumptions, the optimal tax schedule is characterized by:[7]

$$\forall z : \sum_{l=1}^{L} \frac{\partial}{\partial z^l} \left(\overline{\mathcal{W}^l}(\boldsymbol{z}) g^{\mathrm{z}}(\boldsymbol{z}) \right) = -(1 - \overline{\beta}(z)) g^{\mathrm{z}}(\boldsymbol{z}), \tag{12}$$

where for any function $\boldsymbol{\omega} \mapsto h(\boldsymbol{\omega})$, the notation $\boldsymbol{z} \mapsto \overline{h}(\boldsymbol{z})$ denotes the average of $h(\boldsymbol{\omega})$ over all types $\boldsymbol{\omega}$ who choose labour income $\boldsymbol{Z}(\boldsymbol{\omega}) = \boldsymbol{z}$; $\boldsymbol{z} \mapsto g^{\mathrm{z}}(\boldsymbol{z})$ denotes the probability density function for the labour incomes and $\boldsymbol{z} \mapsto G^{\mathrm{z}}(\boldsymbol{z})$ denotes the cumulative distribution function; we denote the marginal excess burden of a tax on the i-th labour income as follows:

$$\forall \boldsymbol{\omega} : \mathcal{W}^i(\boldsymbol{\omega}) \equiv -\sum_{l=1}^{L} T_{z_l} \frac{\partial Z^{l*}(\boldsymbol{\omega})}{\partial \tau^i} - \sum_{j=1}^{J} t^j \frac{\partial X^{j*}(\boldsymbol{\omega})}{\partial \tau^i}; \tag{13}$$

and we denote the marginal social welfare weights net of income effects as follows (with λ denoting the government's budget Lagrange multiplier):

$$\forall \boldsymbol{\omega} : \beta(\boldsymbol{\omega}) \equiv \frac{W_v u_c}{\lambda} + \sum_{l=1}^{L} T_{z^l} \frac{\partial Z^l(\boldsymbol{\omega})}{\partial \rho} + \sum_{j=1}^{J} t^j \frac{\partial X^j(\boldsymbol{\omega})}{\partial \rho}. \tag{14}$$

At each point $\boldsymbol{z}$ on the boundary of the income space $\mathcal{Z}$, the following boundary condition must hold:[8]

$$\sum_{l=1}^{L} \overline{\mathcal{W}^l}(\boldsymbol{z}) e^l(\boldsymbol{z}) = 0, \tag{15}$$

where $e^l(\boldsymbol{z})$ is the l-th component of the unit vector normal to the boundary of the income space at the point $\boldsymbol{z}$. Applying the divergence theorem to condition (12) and substituting the boundary conditions (15) shows that the welfare weights should average to one:

$$\iint_{\Omega} \beta(\boldsymbol{\omega}) \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega}) = 1. \tag{16}$$

[7]Integrating (12) for $L = 1$ shows that it is equivalent to Saez' [2001] *ABC*-style characterization for the optimal one-dimensional tax on labour income.

[8]See Werquin et al. [2015] and Spiritus et al. [2022] for proofs. For $L = 1$, condition (15) is equivalent to demanding that the marginal tax rates equal zero at the bottom and at the top of the income distribution.

4.2 Optimal commodity taxes

The linear tax rate on the j-th commodity is optimal if perturbing it leaves social welfare unaffected. Demanding that the sum of all effects of a perturbation of t^j on social welfare equals zero, the government's first-order condition for the optimal linear tax on the j-th commodity is:

$$\iint_\Omega \left\{ \left(1 - \frac{W' u_c}{\lambda}\right) X^j + \sum_{l=1}^{L} T_{z^l} \frac{\partial Z^l}{\partial t^j} + \sum_{k=1}^{J} t^k \frac{\partial X^k}{\partial t^j} \right\} \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega}) = 0, \tag{17}$$

where we omit the function arguments in the integrand to simplify notations.

Substitute Slutsky decompositions (7)–(8) and definition (14) of the social welfare weights:

$$\iint_\Omega \left\{ (1 - \beta)\, X^j + \sum_{l=1}^{L} T_{z^l} \frac{\partial Z^{l*}}{\partial t^j} + \sum_{k=1}^{J} t^k \frac{\partial X^{k*}}{\partial t^j} \right\} \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega}) = 0. \tag{18}$$

Condition (18) characterizes the optimal linear taxes on the commodities, even when the tax schedule on the labour incomes is not optimal. We are interested in the optimal tax rates on the commodities when the tax schedule on the labour incomes is optimal, thus optimal tax condition (12) and boundary conditions (15) are fulfilled.[9]

I present the results in two steps. First I present the optimality conditions in the presence of a single labour income, so we remain close to the model of Saez [2002]. This allows comparing our optimality conditions to his desirability conditions. Next I present the results for an arbitrary number of labour incomes. The proof for all results is in Appendix A.

Proposition 1. *If households earn a single labour income z and the tax schedule on labour income is optimal, then the optimal linear commodity tax rates are character-*

[9] If the marginal tax rates on labour income are not optimal, one can use the derivations of Atkinson and Stiglitz [1980] to study the properties of condition (18). For example, assuming the population average of the welfare weights equals one, so (16) is satisfied, we find:

$$\frac{\operatorname{cov}\left(\beta, X^j\right)}{\iint_\Omega X^{j*} \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega})} = \frac{\iint_\Omega \left\{ \sum_{l=1}^{L} T_{z^l} \frac{\partial Z^{l*}}{\partial t^j} + \sum_{k=1}^{J} t^k \frac{\partial X^{k*}}{\partial t^j} \right\} \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega})}{\iint_\Omega X^{j*} \mathrm{d}G^{\boldsymbol{\omega}}(\boldsymbol{\omega})}. \tag{19}$$

We thus find that in the optimum, the distributional characteristic of each good should be equal to the normalized efficiency losses caused by taxing that good.

ized by the following conditions, for all commodities j:

$$\sum_{k=1}^{J} t^k \int_{\mathcal{Z}} \overline{\frac{\partial X^{kc*}}{\partial t^j}} \mathrm{d}G^z(z) = \int_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|z) \mathrm{d}G^z(z) \qquad (20)$$
$$- \int_{\mathcal{Z}} \overline{\mathcal{W}} \left(\frac{\mathrm{d}\overline{X^j}}{\mathrm{d}z} - \frac{\partial \overline{X^{jc}}}{\partial z} \right) \mathrm{d}G^z(z)$$
$$+ \int_{\mathcal{Z}} \mathrm{cov}\left(\mathcal{W}, \left. \frac{\partial X^{jc}}{\partial z} \right| z \right) \mathrm{d}G^z(z).$$

The interpretation of the terms in Proposition 1 corresponds directly to the desirability conditions of Saez [2002]. The first term on the right-hand side of (20) indicates the covariance of the welfare weights with the consumption of the commodity under consideration, conditional on the labour incomes. It indicates whether conditional on the labour incomes, individuals who consume more of the good under consideration are more deserving of additional resources, or the government considers a strong taste for the goods under consideration to be desirable from a normative perspective. This first term is zero in models with one-dimensional types, such as the models of Atkinson and Stiglitz [1976] and of Jacobs and Boadway [2014].

The second term on the right-hand side of (20) captures the difference between the cross-sectional variation and the individual variation in the consumption of the good under consideration. This term was also identified by Jacobs and Boadway [2014] for the case where individuals differ only in a single labour productivity. This term can differ from zero for two reasons, as discussed by Saez [2002]. The first possible reason is that around the income level under consideration, individuals with different incomes have relatively different tastes for the consumption good under consideration. The second reason why this term can differ from zero, is that households who supply different quantities of labour, consume different amounts of the good under consideration. If there is a difference between the cross-sectional variation and the individual variation in the consumption of the good under consideration, the consumption pattern reveals information about the individual types beyond what is revealed by the labour incomes alone. Shifting from a tax on labour income to a tax on that commodity then reduces the distortions caused by the tax system, by an amount proportional to the marginal excess burden of the labour income tax.

The third term on the right-hand side of (20) captures the possibility that conditional on labour income, household for whom a higher labour income leads to a higher consumption of the good under consideration, have higher compensated responses to taxes on labour income. By then shifting the burden of taxation from that labour income to the good under consideration, the government can attain equal benefits of redistribution, with less efficiency losses. As noted by Saez [2002], this third term is difficult to measure empirically, and there is intuitively little reason why it would differ from zero. This third term is absent in models with a single dimension of heterogeneity.

If the right-hand side of (20) differs from zero, then it becomes desirable to levy differentiated taxes on the commodities. The left-hand side of (20) then indicates that the optimal size of the marginal excess burden caused by taxing a good under consideration, conditional on labour income, is larger in absolute value as the absolute value of the right-hand side of (20) increases.

We now turn to the optimal taxes in presence of multiple labour incomes. Recall from Subsection 4.1 that when turning to a multidimensional context, studying the optimal taxation of the labour incomes becomes considerably more complicated. We show in the following Theorem that this complexity does not transfer to the optimal linear taxation of commodities.

Theorem 1. *If the tax schedule on the labour incomes of the households is optimal, then the optimal linear commodity tax rates are characterized by the following conditions, for all commodities j:*

$$\begin{aligned}\iint_{\mathcal{Z}} \sum_{k=1}^{J} t^k \overline{\frac{\partial X^{kc*}}{\partial t^j}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) = &\iint_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &- \sum_{l=1}^{L} \iint_{\mathcal{Z}} \overline{\mathcal{W}^l} \left(\frac{\mathrm{d}\overline{X^j}}{\mathrm{d}z^l} - \frac{\partial \overline{X^{jc}}}{\partial z^l} \right) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &+ \sum_{l=1}^{L} \iint_{\mathcal{Z}} \mathrm{cov}\left(\mathcal{W}^l, \left.\frac{\partial X^{jc}}{\partial z^l} \right| z \right) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}).\end{aligned} \tag{21}$$

The optimality condition (21) is very similar to the optimality condition (20) in presence of a single labour income. The left-hand sides and the first term on the right-hand sides are equal for both conditions, with equal interpretations.

The second term on the right-hand side of (21) now contains a sum over all labour incomes. This term still captures, for each labour income, the difference between the cross-sectional variation and the individual variation in the consumption of the good under consideration. The interpretation remains the same as in the case with a single labour income, only now the effect is summed over the different labour incomes. For example, there can now be multiple Corlett and Hague [1953] effects, one for each of the household members.

The third term on the right-hand side of (21) is also similar to the third term on the right-hand side of condition (20), except that it is again summed over all the labour incomes. This term again captures the technical possibility that conditional on the labour incomes, household for whom a higher labour income leads to a higher consumption of the good under consideration, have higher compensated responses to taxes on that labour income. The interpretation of this term remains the same as in the case with a single labour income.

5 Conclusions

There are several possible reasons to deviate from uniform tax rates on commodities. A first possible reason is that conditional on labour income, individuals who consume more of a particular good are more deserving in the eyes of the government. This can be the case when a strong taste for a particular good is considered socially desirable (merit goods) or if the government wishes to compensate types who typically consume certain goods (e.g., to compensate for other injustices). The acceptable size of the distortion from taxing a particular good, conditional on labour income, is then larger if the conditional covariance between the welfare weights and the consumption of that good is larger.

A second possible reason to differentiate commodity taxes is that on average, individuals with different incomes have different consumption patterns. A third, related reason is that conditional on income, individuals who supply different quantities of labour have different consumption patterns. In each case, shifting the burden of taxation from the labour incomes to differentiated commodity taxes increases the efficiency of the tax system. The acceptable size of the distortion of taxing a particular good then depends on the size of the difference between the cross-sectional variation and the individual variation of the consumption of that good along each labour income, and the size of the marginal excess burdens of the taxes on those labour incomes.

The formulas in this paper identify which statistics are sufficient to characterize the optimal commodity taxes. Most of these statistics are not yet available in the empirical literature. Moreover, it remains unclear how the social welfare function can capture common ethical intuitions that may lead to differentiated commodity taxes in the optimum. For these reasons, numerical simulations of the optimal commodity tax rates would currently be based on arbitrary parameters. It would be very useful if further research could quantify the needed sufficient statistics for a numerical exercise. Further useful extensions would include moving beyond the unitary decision framework for the households. Such an extension, however, would first require advances in our knowledge on the optimal taxation of labour income when households do not take their decisions as one unit.

References

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

Anthony B. Atkinson and Joseph E. Stiglitz. *Lectures on public economics*. McGraw-Hill College, 1980.

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

Wilfred J Corlett and Douglas C Hague. Complementarity and the excess burden of taxation. *The Review of Economic Studies*, 21(1):21–30, 1953.

Marc Fleurbaey and François Maniquet. Optimal income taxation theory and principles of fairness. *Journal of Economic Literature*, 56(3):1029–79, 2018.

Roger H Gordon and Wojciech Kopczuk. The choice of the personal income tax base. *Journal of Public Economics*, 118:97–110, 2014.

Bas Jacobs and Robin Boadway. Optimal linear commodity taxation under optimal non-linear income taxation. *Journal of Public Economics*, 117:201–210, 2014.

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

James Mirrlees, Stuart Adam, Tim Besley, Richard Blundell, Steve Bond, Robert

Chote, Malcolm Gammie, Paul Johnson, Gareth Myles, and James Poterba. *Tax By Design: The Mirrlees Review*. Oxford University Press, Sep 2011.

OECD. *Consumption Tax Trends*. 2020. doi: https://doi.org/https://doi.org/10.1787/152def2d-en.

Jukka Pirttilä and Ilpo Suoniemi. Public provision, commodity demand, and hours of work: An empirical analysis. *The Scandinavian Journal of Economics*, 116(4): 1044–1067, 2014.

Emmanuel Saez. Using elasticities to derive optimal income tax rates. *Review of Economic Studies*, 68(1):205–29, January 2001.

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

Kevin Spiritus, Etienne Lehmann, Sander Renes, and Floris Zoutman. Optimal taxation with multiple incomes and types. *CEPR Discussion Paper*, (DP16797), 2022.

Josep-Maria Terricabras. Question for written answer E-006958-16 to the Commission. Rule 130. European Parliament parliamentary questions, September 2016. URL `https://www.europarl.europa.eu/doceo/document/E-8-2016-006958_EN.html`.

Nicolas Werquin, Aleh Tsyvinski, and Mikhail Golosov. A variational approach to the analysis of tax systems. Technical report, Society for Economic Dynamics, 2015.

A Proof of Proposition 1 and Theorem 1

To account for the conditions for the optimal tax schedule on the labour incomes, we reformulate optimality condition (18) in the income space. Let $G^{\boldsymbol{\omega}|\mathrm{z}}(\boldsymbol{\omega}|\boldsymbol{z})$ denote the cumulative distribution function of the characteristics $\boldsymbol{\omega}$ conditional on labour incomes $\boldsymbol{z}$, with corresponding density function $g^{\boldsymbol{\omega}|\mathrm{z}}(\boldsymbol{\omega}|\boldsymbol{z})$. We can then change the integration domain of (18), to find for any commodity j:

$$\iint_{\mathcal{Z}} \iint_{\{\boldsymbol{\omega}|\boldsymbol{Z}(\boldsymbol{\omega})=\boldsymbol{z}\}} \left\{(1-\beta)\, X^j + \sum_{l=1}^{L} T_{z^l} \frac{\partial Z^{l*}}{\partial t^j} + \sum_{k=1}^{J} t^k \frac{\partial X^{k*}}{\partial t^j}\right\} \mathrm{d}G^{\boldsymbol{\omega}|\mathrm{z}}(\boldsymbol{\omega}|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) = 0, \tag{22}$$

or shorter:

$$\iint_{\mathcal{Z}} \overline{(1-\beta)X^j} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) + \iint_{\mathcal{Z}} \left\{ \sum_{l=1}^{L} T_{z^l} \frac{\overline{\partial Z^{l*}}}{\partial t^j} + \sum_{k=1}^{J} t^k \frac{\overline{\partial X^{k*}}}{\partial t^j} \right\} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) = 0. \tag{23}$$

We now rewrite the first integral in (23) as a covariance conditional on labour income:

$$\iint_{\mathcal{Z}} \overline{(1-\beta)X^j} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) = -\iint_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) + \iint_{\mathcal{Z}} (1-\overline{\beta})\overline{X^j} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}). \tag{24}$$

Substitute optimality condition (12) and then apply partial integration taking into account boundary conditions (15):

$$\begin{aligned} &\iint_{\mathcal{Z}} \overline{(1-\beta)X^j} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &\quad = -\iint_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) - \iint_{\mathcal{Z}} \sum_{l=1}^{L} \frac{\partial}{\partial Z^l} \left(\overline{\mathcal{W}^l} g^{\mathrm{z}}(\boldsymbol{z}) \right) \overline{X^j} \mathrm{d}\boldsymbol{z} \qquad (25) \\ &\quad = -\iint_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) + \iint_{\mathcal{Z}} \sum_{l=1}^{L} \left(\overline{\mathcal{W}^l} \frac{\mathrm{d}\overline{X^j}}{\mathrm{d}z^l} \right) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}). \qquad (26) \end{aligned}$$

Substitute the latter result into optimality condition (23) and rearrange:[10]

$$\begin{aligned} &\iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} \left(\mathcal{W}^l \frac{\partial X^{jc}}{\partial z^l} + T_{z^l} \frac{\partial Z^{l*}}{\partial t^j} \right) + \sum_{k=1}^{J} t^k \frac{\partial X^{k*}}{\partial t^j} \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &\quad = \sum_{l=1}^{L} \iint_{\mathcal{Z}} \mathrm{cov}\left(\mathcal{W}_l, \left. \frac{\partial X^{jc}}{\partial z^l} \right| \boldsymbol{z} \right) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &\quad - \sum_{l=1}^{L} \iint_{\mathcal{Z}} \overline{\mathcal{W}^l} \left(\frac{\mathrm{d}\overline{X^j}}{\mathrm{d}z^l} - \frac{\partial \overline{X^{jc}}}{\partial z^l} \right) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\ &\quad + \iint_{\mathcal{Z}} \mathrm{cov}(\beta, X^j|\boldsymbol{z}) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}). \qquad (27) \end{aligned}$$

Finally, we rewrite the left-hand side of the latter equation. Substitute definition

[10]To find (27), first add $\sum_{l=1}^{L} \iint_{\mathcal{Z}} \overline{\mathcal{W}^l} (\partial \overline{X^{jc}} / \partial z^l) \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z})$ on both sides of (23).

(13) of the excess burden and rearrange:

$$
\begin{aligned}
&\iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} \left(\mathcal{W}^{l} \frac{\partial X^{jc}}{\partial z^{l}} + T_{z^{l}} \frac{\partial Z^{l*}}{\partial t^{j}} \right) + \sum_{k=1}^{J} t^{k} \frac{\partial X^{k*}}{\partial t^{j}} \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\
&= \iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} T_{z^{l}} \left(\frac{\partial Z^{l*}}{\partial t^{j}} - \sum_{m=1}^{L} \frac{\partial Z^{l*}(\boldsymbol{\omega})}{\partial \tau^{m}} \frac{\partial X^{jc}}{\partial z^{m}} \right) \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\
&\quad + \iint_{\mathcal{Z}} \overline{\left\{ \sum_{k=1}^{J} t^{k} \left(\frac{\partial X^{k*}}{\partial t^{j}} - \sum_{l=1}^{L} \frac{\partial X^{k*}(\boldsymbol{\omega})}{\partial \tau^{l}} \frac{\partial X^{jc}}{\partial z^{l}} \right) \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}).
\end{aligned}
\tag{28}
$$

Apply Slutsky symmetries (9):

$$
\begin{aligned}
&\iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} \left(\mathcal{W}^{l} \frac{\partial X^{jc}}{\partial z^{l}} + T_{z^{l}} \frac{\partial Z^{l*}}{\partial t^{j}} \right) + \sum_{k=1}^{J} t^{k} \frac{\partial X^{k*}}{\partial t^{j}} \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\
&= \iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} T_{z^{l}} \left(\frac{\partial X^{j*}}{\partial \tau^{l}} - \sum_{m=1}^{L} \frac{\partial Z^{m*}}{\partial \tau^{l}} \frac{\partial X^{jc}}{\partial z^{m}} \right) \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\
&\quad + \iint_{\mathcal{Z}} \overline{\left\{ \sum_{k=1}^{J} t^{k} \left(\frac{\partial X^{j*}}{\partial t^{k}} - \sum_{l=1}^{L} \frac{\partial Z^{l*}}{\partial t^{k}} \frac{\partial X^{jc}}{\partial z^{l}} \right) \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}).
\end{aligned}
\tag{29}
$$

Substitute properties (10)–(11) on the right-hand side to simplify:

$$
\begin{aligned}
&\iint_{\mathcal{Z}} \overline{\left\{ \sum_{l=1}^{L} \left(\mathcal{W}^{l} \frac{\partial X^{jc}}{\partial z^{l}} + T_{z^{l}} \frac{\partial Z^{l*}}{\partial t^{j}} \right) + \sum_{k=1}^{J} t^{k} \frac{\partial X^{k*}}{\partial t^{j}} \right\}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}) \\
&= \iint_{\mathcal{Z}} \sum_{k=1}^{J} t^{k} \frac{\overline{\partial X^{kc*}}}{\partial t^{j}} \mathrm{d}G^{\mathrm{z}}(\boldsymbol{z}).
\end{aligned}
\tag{30}
$$

Substitute the latter result into condition (27) to find the characterization of the optimal tax rates in Theorem 1. Proposition 1 follows by setting $L = 1$ in Theorem 1.