

Optimal Tax and Benefit Policies with Multiple Observables

Etienne Lehmann, Sander Renes, Kevin Spiritus and Floris Zoutman *

February 1, 2018

PRELIMINARY VERSION

This note was initially written individually by Kevin Spiritus. It is now a part of a project with Etienne Lehmann, Sander Renes, and Floris Zoutman. It is intended to be combined with Renes and Zoutman: "As Easy as ABC? Multi-dimensional Screening in Public Finance" (2017) and Lehmann: "Optimal multidimensional and nonlinear taxation" (2017).

We study the optimum for tax problems with multidimensional tax bases and multidimensional heterogeneity of agents. We use the Euler-Lagrange formalism to show how the optimal tax function balances efficiency versus equity considerations. The equity considerations are captured in a localized distributional characteristic, a generalization of the distributional characteristic first introduced by Feldstein (1972a,b). We apply these findings to the optimal joint taxation of couples, and to the optimal mixed taxation of capital and labour income. We show robustness for pooling, bunching and restrictions to the tax base.

Keywords: optimal mixed taxation, heterogeneous agents

JEL Classification: H21, H23, H24

1. Introduction

The problem of optimal mixed taxation with multidimensional types, as stated by Mirrlees (1976), has long eluded an exact solution. Most optimal-tax models taking into account distributional concerns, in the tradition of Mirrlees (1971) and Atkinson and

*Lehmann: CRED (TEPP) University Panthéon-Assas Paris 2; Renes: Erasmus School of Economics at University Rotterdam; Zoutman: NHH Norwegian School of Economics; Spiritus: Erasmus University Rotterdam, KU Leuven and Tinbergen Institute.

Stiglitz (1976), assume that a single parameter, often an ability parameter, drives differences between agents. This makes it difficult to study problems with multiple tax bases where more extensive type heterogeneity is deemed relevant. The literature offers no exact solution for example for the optimal mixed tax problem for labour and capital income when individuals differ simultaneously in a number of dimensions, e.g. in their abilities and their preferences for labour supply and for investment, or in the inheritances they receive. Atkinson and Stiglitz (2015, p.xxi) list this question as one of the central current challenges in public finance. My aim in this paper is to provide a solution.

This paper makes two contributions. The first is that it is the first to characterize the full optimum in terms of sufficient statistics and distributional characteristics. The second is that it introduces a solution method which avoids the labour intensity and technical complications inherent to existing approaches. I apply my results to a number of previously unsolved problems.

In optimal-tax theory it is customary to search for a tax system that maximizes a *social welfare function*. It maps the “social state”, e.g. the distribution of attained utility levels or disposable incomes, on a scalar indicating a social judgment about that social state. It takes into account normative criteria, for example social preferences for redistribution, efficiency and responsibility. The maximization of the social welfare function occurs subject to constraints such as the government budget, incentive compatibility and restrictions to tax and benefit policies. Boadway (2012) gives an overview of the recent literature.

The case with a one-dimensional tax base has been well studied. Saez (2001) and Jacquet and Lehmann (2016) offer exact solutions. The optimal tax rate at each income level is formulated as a product of three terms, traditionally denoted as A , B and C ,¹ where the A -term relates negatively to the excess burden associated with the tax, the B -term concerns the distributional advantage of levying an additional tax from all individuals at higher income levels, and the C -term takes into account the thickness of the income distribution at higher levels.²

There are two broad approaches in the literature to study the problem with multidimensional tax bases, both of which lead to technical complications. The *dual* approach to optimal taxation, which directly uses the tax function as an instrument, is gaining popularity (see e.g. Christiansen (1984), Saez (2002), Werquin, Tsyvinski, and Golosov (2015)). It is customary to first construct perturbations of the tax function, and then to find conditions for such perturbations to have zero effect on social welfare. Some authors construct combined perturbations such that the total effect on the government budget is zero (e.g. Christiansen (1984), Saez (2002)). Others use Gateaux derivatives in the direction of some reform function (Werquin, Tsyvinski, and Golosov (2015)). Although the perturbation approach yields important insights, it becomes cumbersome when more complex tax systems are studied.

Saez (2002), one of the founders of the perturbation approach, makes an important

¹See for example Diamond (1998, p.86) for an early use of this notation.

²Gerritsen (2016) and Scheuer and Werning (2016) use perturbations to study the problem with a one-dimensional tax base from different angles.

advancement by finding conditions for optimal commodity taxes to be non-uniform. His model is robust for multidimensional heterogeneity of the agents. He leaves finding optimal tax formulas as an “extremely useful” and “important task” for future research (2002, p.229). Werquin, Tsyvinski, and Golosov (2015) go beyond desirability conditions, finding a *second-order partial differential equation* for the tax function, characterizing the optimum for a quite general multidimensional problem. They stop short of solving this equation.

The second, more traditional approach to study the optimal-tax problem is referred to as the *primal* or the *mechanism design* approach. Instead of using variations to the tax functions to figure out the optimum, one looks for an optimal allocation using individual decision variables as controls, for example the consumption and labour supply variables of the different individuals. The social welfare function is optimized using a Hamiltonian or a Lagrangian. This leads to optimality conditions for the tax wedges in a fairly straightforward way. The greater challenge is to find a tax function which implements the optimum. An optimal tax system will generally be prohibitively complicated. In a setting with multiple periods, for example, tax liabilities will generally depend in non-separable ways on labour income and capital incomes in all periods at the same time. Often it appears impossible to characterize this function analytically. Even in cases where an analytical solution can be found, the task of reformulating it in terms of observable statistics can be labour intensive. A further limitation of the approach is that it has difficulties to handle multiple dimensions of heterogeneity of the agents, or important restrictions to the tax function.

Diamond and Spinnewijn (2011) use this mechanism design approach to study the problem of optimal mixed taxation of capital and labour income with heterogeneous abilities and discount rates. They find desirability conditions in terms of the fundamentals of the model. They do not characterize the optimum.

Mirrlees (1976) states the full optimization problem with a multidimensional tax base and multidimensional heterogeneity of the agents, but solves it only for the case with one-dimensional heterogeneity of the agents. For the case with multidimensional heterogeneity he finds a second-order partial differential equation for the utility profile as a necessary condition for the optimum. Kleven, Kreiner, and Saez (2007) study the optimal joint taxation of couples, finding a necessary condition of the exact same form.

Interestingly, the second-order partial differential equation found by Mirrlees (1976) has the same form as the previously mentioned equation found by Werquin, Tsyvinski, and Golosov (2015) – though it is expressed in terms of fundamentals of the model rather than in terms of observable variables. Since second-order partial differential equations of similar form are encountered in both approaches, the key to solving the optimal-tax problem appears to lie in finding solutions for these equations.

Renes and Zoutman (2016a) make important headway by treating the second-order differential equation found by Mirrlees (1976) as a first-order differential equation in the multipliers of the incentive compatibility constraint. They show that these multipliers form a conservative vector field, and use this fact to convert the problem into a new second-order partial differential equation for which solution methods are more readily available. They apply the *Green functions* approach, well-studied in physics and en-

gineering but less familiar in economics, to characterize the optimum. Their method still has a number of limitations. It cannot handle the case with more types than tax bases – e.g. the cases treated by Saez (2001, 2002) and Jacquet and Lehmann (2016). It also requires perfect screening, meaning that at most one type pools at each tax-base level. The solutions are formulated in terms of fundamentals of the problem rather than observable variables, lacking a straightforward economic interpretation. And the procedure to convert the problem to a new second-order partial differential equation is complicated.

The present paper suggests an alternative. The central methodological insight is that the problem of optimal multidimensional taxation is essentially a field-theoretical problem: the tax function is a scalar field over the tax base space, which is shaped in order to optimize an objective, a social welfare functional subject to a resource constraint. Both the level of the tax function at each value of the tax base (the tax liability) as its gradient (the marginal tax rates) have an impact and need to be taken into account.

A central difference with traditional field theory is that the objective function is not defined over the tax base space, but over a set of individuals, members of a multidimensional type space. These individuals are free to choose any value in the tax base space, as long as they can afford it. It is even possible for multiple types to choose the same value of the tax base. This difference leads to a number of complications, but these are not insurmountable. With some modifications it is still possible to characterize the optimal tax function using well-established techniques.

I show how this optimum is characterized by an Euler-Lagrange equation, first graphically and then formally. This is a second-order partial differential equation which generalizes the equation found by Werquin, Tsyvinski, and Golosov (2015). Next I show how standard methods can be used to characterize solutions to the Euler-Lagrange equation. I treat it as a first-order partial differential equation, and introduce multidimensional Green functions to characterize the optimum directly in terms of observable quantities such as population densities, average behavioural responses and average welfare weights at each value of the tax base.

An advantage of using the tax function as an instrument, rather than the allocation, is that it immediately leads to a characterization in terms of sufficient statistics, similar to the findings of Saez (2001) for a one-dimensional tax base. This makes the results much easier to interpret. Using the tax function as an instrument also makes that no incentive compatibility constraints need to be taken into account, avoiding the usual implementation issues that surface in multidimensional problems. Moreover, it allows to remain agnostic about the process determining individual behaviour, as I do in most of the paper, and it allows to distinguish between the wellbeing measures taken into account by the planner on one hand, and the preferences driving individual behaviour on the other.

The resulting characterization of the optimum is very similar to that for the problem with linear taxes as stated by Atkinson and Stiglitz (1980, p.386-390). At any given value of the multidimensional tax base, the efficiency effects of a marginal tax reform should be balanced against its distributional effects. The proportional reduction of any component of aggregate demand, along compensated demand curves, should be equal

to a local distributional characteristic of that component. The distributional side of this equation is new. It is a localized version of the distributional characteristic first introduced by Feldstein (1972a,b). If we associate a marginal social welfare weight with each potential value of the tax base, then the *local* covariance of these welfare weights with the tax-base component under consideration determines the distributional impact on social welfare. These findings apply even when the dimensionality of the type space is higher than that of the tax base space.

Although these are interesting insights, they still do not teach us much about the underlying economics. I go on to repeat the same techniques in the type space, still using the tax function as an instrument. This again leads to a characterization of the optimum where efficiency effects of a tax reform are balanced against equity effects. The equity effects now consist of two parts. First there are again local distributional characteristics of the different traits of the individuals. If for example individuals who receive a larger inheritance have much lower marginal social welfare weights, this will lead to a large distributional characteristic for this trait. Second, there are additional terms indicating how well a given component of the tax base is suited to target a given trait of the individuals. These two elements, how much we care about particular traits (a normative part) and how well these traits can be targeted (an informational part), determine the distributional impact of a given tax base component. These results extend findings by Mirrlees (1976) and Renes and Zoutman (2016a).

I illustrate my results using a number of examples, applying either to the joint taxation of couples or to the joint taxation of labour and capital incomes. My aim is not to give a thorough discussion of these topics. Indeed, important normative issues such as whether the government should tax individuals or families, or how families should enter in social preferences, are not taken into account. Neither do I discuss what social welfare weights should look like, e.g. whether the government should support couples where one individual stays at home, or rather whether it should stimulate equal participation in the labour market. The point I want to illustrate is that once the tax base, the social objective and a behavioural model have been chosen, then the techniques in this paper can be used to gain insight into the optimal tax system.

Still these examples yield some interesting results. For the problem of the optimal joint taxation of couples, as posed by Kleven, Kreiner, and Saez (2007), I am able to characterize for the first time the solution in terms of local distributional characteristics of the households. For the optimal tax mix between capital income and labour income, I study the case where discount rates are heterogeneous, a case that is excluded e.g. by Werquin, Tsyvinski, and Golosov (2015). I confirm a number of classical reasons for marginal tax rates on capital income to be non-zero: when discount rates are directly driven by labour abilities (Saez, 2002), when labour supply is a complement or a substitute to consumption (Corlett and Hague, 1953), and when investment technology is affected by an ability parameter. The latter complements the findings of (Gerritsen et al., 2015). Even when neither of these reasons applies, an accidental correlation between the labour abilities and the discount rates suffices for the optimal marginal tax rate on capital income to be non-negative (Diamond and Spinnewijn, 2011). I show that the latter correlation remains sufficient even when social welfare weights are constructed

such that they do not depend on the discount rates, holding individuals responsible for their preferences. I show that when rates of returns depend on the individuals' discount rates, then the optimal marginal tax rates on capital income again differ from zero. I state for the first time the optimality condition, extending on the desirability conditions formulated in the past literature. It is straightforward to extend this formulation e.g. to heterogeneous wealth endowments.

There are a number of possible complications while dealing with multidimensional tax bases and multidimensional heterogeneity of the agents. The most important is the possibility of bunching. Although Kleven, Kreiner, and Saez (2007) show that bunching may not be a problem for a large range of social welfare objectives, it cannot simply be ignored. I do not attempt to characterize when and where bunching will occur in the optimum. I rather give necessary conditions to check whether a given tax function, with or without bunching, is optimal given information about population densities and behavioural responses. I show that the approach described in this paper remains valid even when bunching occurs. Similarly, I show how to deal with potential issues such as pooling, double deviations, and restrictions to the tax base space.

To the best of my knowledge, the possible use of the Euler-Lagrange formalism is mentioned just once in the optimal-tax literature. Boháček and Kejak (2016) construct an Euler-Lagrangian to solve a dynamic problem with heterogeneous types and a one-dimensional tax base. They prove the correctness of the approach for their specific problem. They provide an analytical solution only for the end-points, resorting to simulations for the rest of the type distribution. The above-mentioned partial differential equation identified by Werquin, Tsyvinski, and Golosov (2015) is a specialized version of the Euler-Lagrange equation. It was derived for a specific optimal tax problem, and excludes for example heterogeneous discount rates.

The present paper shows the general applicability of the Euler-Lagrange framework, so that it is no longer necessary to construct new perturbations and list their effects each time a new problem is encountered. Furthermore, as far as I know, it is the first to identify the localized distributional characteristics that characterize the optimum.

In this paper I do not address the optimal choice of the tax base. I take a tax base as given, and I determine the optimal tax function over this base. Since the problem is not so difficult to solve with a separable tax function (see e.g. Kleven, Kreiner, and Saez (2007)), I consider non-separable tax bases only – although I will encounter an example where the optimal tax system turns out to be separable. The methods in this paper remain valid even when the tax system is incomplete, e.g. in the cases of tax avoidance or income shifting as studied by Christiansen and Tuomala (2008). The approach also allows solving more general problems, e.g. with a government aiming to gain sufficient votes to remain in power, or a formulation incorporating alternative normative convictions, e.g. using generalized Pareto weights as discussed for a one-dimensional tax base by Saez and Stantcheva (2016).

I start in section 2 by introducing the model and identifying the behavioural responses to tax reforms. I introduce the Euler-Lagrange formalism in section 3. I show intuitively and formally how the Euler-Lagrange equation is a necessary condition for the optimum. In section 4 I then use the Euler-Lagrange equation to find a more intuitive characteri-

zation of the optimum, using sufficient statistics in the tax base space, and in section 5 I reformulate this characterization in terms of economic fundamentals. Finally in section 6, I list a number of potential complications, and I show how they can be overcome.

2. Model

We study a population of individuals who are distinguished by a K -dimensional column vector of characteristics $\boldsymbol{\theta} \equiv (\theta^1, \dots, \theta^K)^\top$, with superscript $\cdot^\top$ denoting a transpose. This vector is also called the *type* of the individual. The set of all such vectors is denoted Θ , which is a convex subset of the real vector space $\mathbb{R}^K$. The type of an individual can include characteristics such as ability, gender or preferences. The types have a cumulative distribution function $F^\Theta(\boldsymbol{\theta})$, with corresponding density function $f^\Theta(\boldsymbol{\theta})$.

There is a government which does not observe the type of the individuals, but it does observe for each individual an L -dimensional column vector $\mathbf{x} \equiv (x^1, \dots, x^L)^\top$ which is called the *tax base*. The values of the tax base are elements of the *tax base space* $\mathcal{X}$, which is a convex subset of the real vector space $\mathbb{R}^L$. I denote the cumulative distribution function as $F^\mathcal{X}(\mathbf{x})$, with corresponding density function $f^\mathcal{X}(\mathbf{x})$. The government uses the value $\mathbf{x}$ of the tax base for each individual to determine their tax liability $T(\mathbf{x})$, which is a scalar-valued, potentially nonlinear function.

Each individual makes a number of decisions, given the tax function and his type. These decisions may include for example labour supply, consumption of different goods, savings, and so on. They may also include reporting decisions such as how much taxes to evade, or how much labour income to declare as capital income. The government does not observe all of these decisions. Together with the type, these decisions do determine the tax base, which is observed by the government. For example, although the government may not observe an individual's labour effort and his ability, it does observe the gross labour income declared by the individual. To facilitate my exposition, I will act as if the individual chooses the value of his tax base directly, given his type and the tax function. This leads to the vector-valued tax base function $\mathbf{x}(\boldsymbol{\theta}, T)$.

The government chooses the function T in order to maximize social welfare, which is the sum of individual wellbeing measures $v(\boldsymbol{\theta}, T)$:

$$\max_T \int_{\Theta} v(\boldsymbol{\theta}, T) dF^\Theta(\boldsymbol{\theta}). \quad (1)$$

The individual wellbeing measure is determined by the individual type and by the tax function. It may simply be defined as the indirect utility of an individual, or as some other measure of individual wellbeing, taking into account social normative judgments. There is not necessarily a direct link between the wellbeing measure taken into account by the government, and individual behaviour.

The government aims to levy an exogenously determined revenue, which without loss of generality is normalized to zero. This leads to the government budget constraint:

$$\int_{\Theta} T(\mathbf{x}(\boldsymbol{\theta}, T)) dF^\Theta(\boldsymbol{\theta}) = 0. \quad (2)$$

The simultaneity in this formulation requires careful attention. The government budget constraint depends on the tax liability owed by each individual, while the chosen values of the tax base by the individuals depend again on the tax function. One should distinguish conceptually between the function T , and its value $T(\mathbf{x})$ at a particular value of the tax base.

Denote the partial derivatives of the tax function using a subscript: $T_l \equiv \partial T / \partial x^l$, and denote the gradient of the tax function as follows:

$$\nabla_{\mathbf{x}} T(\mathbf{x}) \equiv (T_1(\mathbf{x}), \dots, T_L(\mathbf{x})). \quad (3)$$

I will refer to it in short as the *tax gradient*.

To find necessary conditions for a tax function T to be optimal, I use the fact that a small tax reform should leave social welfare unchanged. I assume that for an individual who chooses value $\mathbf{x}(\boldsymbol{\theta}, T)$ of the tax base, his response to a small tax reform will be will be along the intensive margin, excluding discrete jumps.³ I assume that his response will be determined only by the change in the tax liability at $\mathbf{x}$, and by the change in the tax gradient at $\mathbf{x}$. Locally, individual behaviour thus remains unchanged if the tax function is replaced by a linearized version, since this leaves the tax liability and the tax gradient unchanged. For example in the neoclassical labour supply model, where an individual chooses his labour supply such that the indifference curve at his chosen bundle is tangent to his budget set, the individual would not change his behaviour if the budget set were linearized.⁴ Furthermore, I assume that individual wellbeing is affected only by a change in the tax liability, and not by a change in the gradient or higher-order derivatives of the tax function. If the chosen wellbeing measure is the indirect utility function, then this assumption is automatically fulfilled because of the envelope property.

I use a subscript T to denote the effect of a local change in the tax liability. For example, $\mathbf{x}_T(\boldsymbol{\theta}, T)$ denotes the column vector containing the effects of a change in the tax liability on the different components of the tax base chosen by a type- $\boldsymbol{\theta}$ individual in presence of tax function T :

$$\mathbf{x}_T(\boldsymbol{\theta}, T) \equiv \begin{pmatrix} x_T^1(\boldsymbol{\theta}, T) \\ \vdots \\ x_T^L(\boldsymbol{\theta}, T) \end{pmatrix}.$$

Similarly, the marginal social value of an extra unit of income for a type- $\boldsymbol{\theta}$ individual equals $-v_T(\boldsymbol{\theta}, T)$.

I use a subscript T_l to denote the effect of a local change in the l -th component of the tax gradient, holding constant the tax liability at the original value of the tax base. The

³More precisely, I assume that the impact of discrete jumps on social welfare is of second order, and that it can be ignored when studying marginal tax reforms. This excludes the empirically relevant cases where individuals face discrete choice sets, or where they do not re-optimize continuously.

⁴This assumption is not innocuous. For example in presence of uncertainty, the local curvature of the tax function will affect the variance of after-tax income, which will affect individual behaviour. The derivations in this paper can be extended to include higher-order derivatives of the tax function, analogous to the methods used in the calculus of variations with higher-order derivatives. For an overview, see for example Courant and Hilbert (1953, p.190).

effects of a change in the tax gradient on the tax base chosen by a type- θ individual are summarized in the following matrix:

$$\mathbf{x}_{\nabla T}(\theta, T) \equiv \begin{pmatrix} x_{T_1}^1(\theta, T) & \cdots & x_{T_L}^1(\theta, T) \\ \vdots & \ddots & \vdots \\ x_{T_1}^L(\theta, T) & \cdots & x_{T_L}^L(\theta, T) \end{pmatrix}. \quad (4)$$

Note that even if I assume that small local reforms to the curvature of the tax function do not affect the chosen values of the tax base, the responses to a marginal reform of the tax liability or the tax gradient are affected by the curvature. The reason is that when the tax function is nonlinear, the response of the value of the tax base to a reform triggers an additional change in the gradient of the tax function, which causes second-round behavioural effects.⁵

3. Euler-Lagrange Formalism

The problem facing the government is to maximize social welfare (1), subject to government budget constraint (2), using the tax function T as an instrument. A necessary condition for the function T to be optimal, is that a marginal tax reform leaves social welfare unchanged.

Any tax reform triggers a number of effects. There is an impact on individual well-being, and as such a direct effect on social welfare. There is also a mechanical effect on government revenue, controlling for changes in individual behaviour, and there are additional effects on government revenue, attributable to the changes in individual behaviour. Requiring that these effects sum to zero then yields a necessary condition for the optimum.

A difficulty is that it is impossible to study all admissible reforms and to list their effects. Luckily we can use the *fundamental lemma of the calculus of variations*, which states that it suffices to study every extremely local reform to the tax function, changing the liability at one particular value of the tax base and taking into account the induced changes in the tax gradient around it.

In the present section I will construct such local tax reforms, and show intuitively how they lead to a partial differential equation, a necessary condition for the optimum, which is similar to the *Euler-Lagrange equation* often used in the calculus of variations. I will first show in subsection 3.1 how for one-dimensional tax bases, this equation corresponds to the *ABC*-style optimal-tax equation that is found in the literature. Next I will argue how it is straightforward to construct similar local tax reforms for multidimensional bases, and I show how a similar Euler-Lagrange equation characterizes the optimum. In subsection 3.2 I do this graphically for a two-dimensional tax base, and in subsection 3.3 I formally prove the multidimensional optimal-tax condition. I will treat the boundary conditions for the optimum in subsection 3.4.

⁵These additional effects are similar to those discussed by Jacquet and Lehmann (2016), and are in contrast with the linearized counterparts in Saez (2001).

I exclude for now the possibility of bunching. Bunching occurs when there is a mass point, and thus a discontinuity, in the population density $f^{\mathcal{X}}$ in the tax base space. This situation could occur e.g. at an edge of the tax base space, because a range of types ends up in a corner solution, or as Ebert (1992) illustrates it could also occur in the interior of the tax base space, leading to a kink in the optimal tax function. I will treat this possibility in section 6. Throughout the present section I assume that the tax function and the objective function are sufficiently smooth, excluding the possibility of kinks or bunching.

Note that I follow Jacquet and Lehmann (2016) in distinguishing between bunching and pooling: the latter situation occurs when different types choose the same value of the tax space without forming a mass point. This situation occurs for example when the dimension of the type space is higher than the dimension of the tax base space. Throughout this section I allow for any number of dimensions in the type space, and thus for pooling.

3.1. One-Dimensional Tax Base

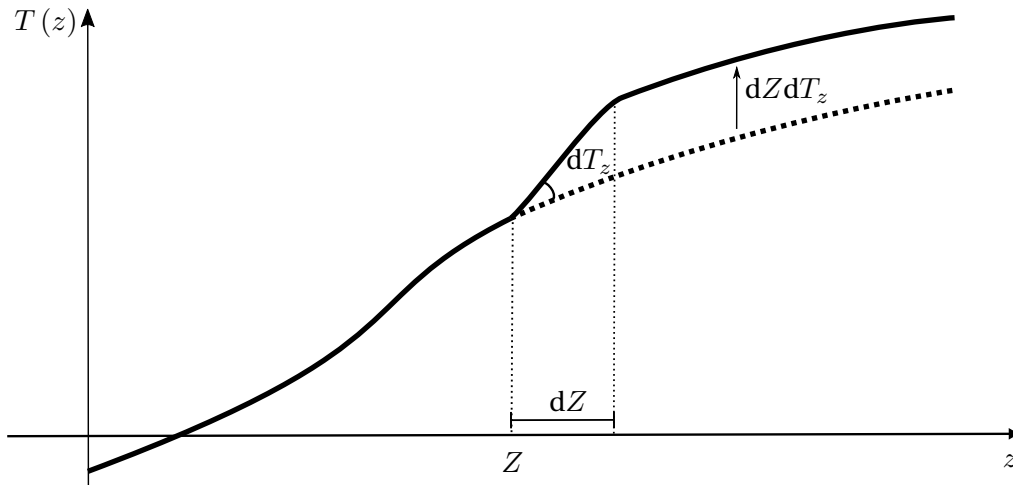
The case with a one-dimensional tax base is well understood. It was first solved for one-dimensional types by Mirrlees (1971). Its intuition for multidimensional types is discussed by Saez (2001), and formalized by Jacquet and Lehmann (2016). I will treat it here in a different way, that can be extended more readily to the case of a multidimensional tax base.

To focus our thoughts, I will treat the case of an optimal labour income tax. I denote gross labour income as z , the corresponding tax liability is $T(z)$ and the corresponding marginal tax rate is $T_z(z)$. I denote the cumulative distribution function of gross labour income as $F^{\mathcal{X}}(z)$, with corresponding density function $f^{\mathcal{X}}(z)$. I denote the support of the latter function as $[\underline{z}, \bar{z}]$. Saez (2001) finds the optimal marginal tax rate at an arbitrary gross income level Z by increasing the marginal tax rate by a small quantity dT_z over an interval of small width dZ , as illustrated in figure 1. Individuals at income level Z experience an increase in their marginal tax rate, which leads to compensated effects, and individuals at higher income levels experience an increase in their tax liability, which leads to income effects. Saez (2001) adds together all mechanical, behavioural and welfare effects of this reform, and requires that the total effect on social welfare should be zero. This, together with a transversality condition, leads directly to a necessary condition for the tax optimum.

The difficulty for us is that it is not clear how to extend this reform to the case with a multidimensional tax base. I will introduce a different kind of reform which leads to the same characterization of the optimum as the one found by Saez (2001), but which can be readily extended to higher dimensions. The trick lies in still performing the reform introduced by Saez (2001), namely increasing the marginal tax rate at gross income Z by a quantity dT_z over an interval of width dZ , and adding a reverse reform at a small distance $\delta \gg dZ$. In other words, at gross income $Z + \delta$, there is a decrease in the marginal tax rate by the same quantity dT_z over an interval of the same width dZ . The resulting tax reform is shown in figure 2.

Figure 1: The tax reform as introduced by Saez (2001). At gross income level Z , over an income range of width dZ , the marginal tax rate is increased by dT_z . At income levels above Z , the tax liability is increased by $dZdT_z$.

(a) The original tax function and the reformed tax function.



(b) The size of the tax reform.

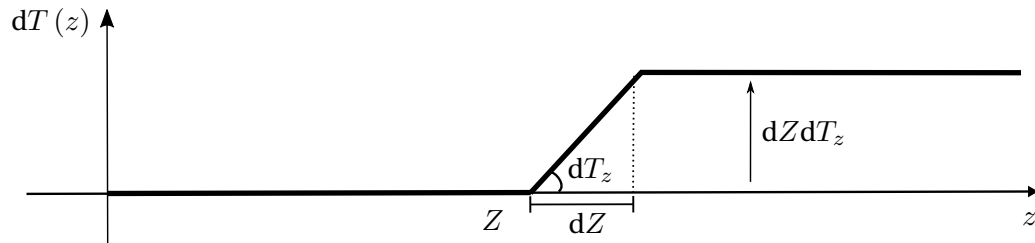
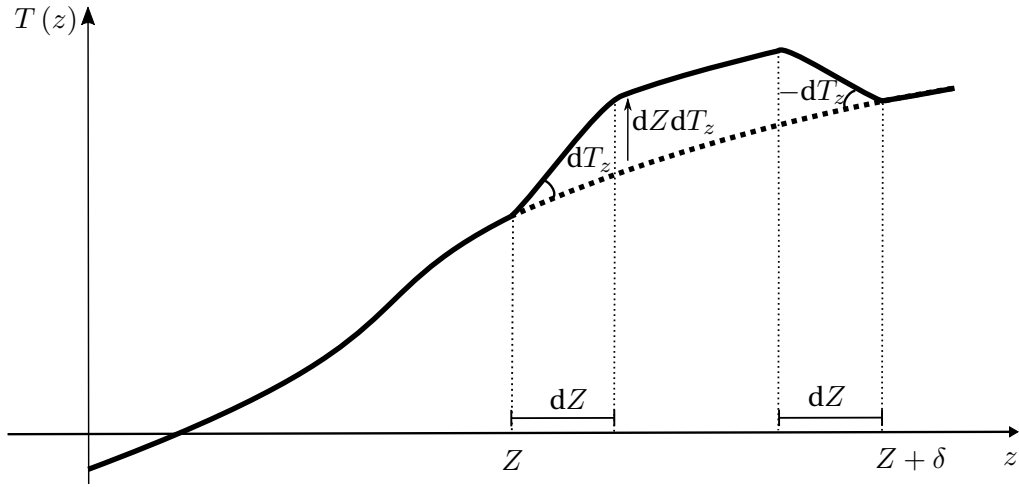
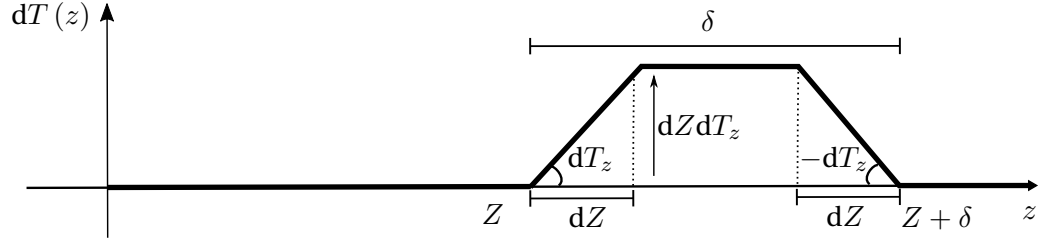


Figure 2: The tax reform for a one-dimensional tax base as introduced in this paper. The tax reform introduced by Saez (2001) is maintained at gross income level Z , but it is reversed at gross income level $Z + \delta$, with $\delta \gg dZ$ small.

(a) The original tax function and the reformed tax function.



(b) The size of the tax reform.



The social-welfare effect of this marginal reform should of course be zero in the optimum, as is the case for any marginal reform. A reason for choosing this particular reform, is that any broader reform could be decomposed into smaller reforms of the kind introduced here. This intuition is captured more formally in the *fundamental lemma of the calculus of variations*, which I will apply in subsection 3.3.

This reform has a number of effects. First, there are behavioural effects at Z , where individuals experience an increase dT_z in their marginal tax rate. The number of individuals experiencing this change equals $f^x(Z) dZ$. A difficulty is that individuals of different types might pool at the same gross income level Z . Denote the average response of all individuals pooling at Z to a reform of the marginal tax rate as $\overline{z_{T_z}}(Z)$.⁶ The total resulting change in government revenue will then be $[\overline{z_{T_z}} T_z f^x](Z) dT_z dZ$.⁷ Second, and similarly, the decrease in the marginal tax rate at $Z + \delta$ causes a decrease in government revenue by $[\overline{z_{T_z}} T_z f^x](Z + \delta) dT_z dZ$. Third, individuals on the interval $[Z, Z + \delta]$ experience an increase of their tax liability by $dT_z dZ$. This leads to a mechanical effect on tax revenue equal to $dT_z dZ \int_Z^{Z+\delta} f^x(z) dz$. Fourth, this increased tax liability leads to welfare effects, which equal $dT_z dZ \int_Z^{Z+\delta} (\overline{v_T}(z) / \lambda) f^x(z) dz$, where λ is the government budget multiplier which converts the individual wellbeing measure to monetary units. Fifth, the increased tax liability also leads to behavioural effects, which have revenue effect $dT_z dZ \int_Z^{Z+\delta} [\overline{z_T} T_z f^x](z) dz$. Note that we do not need to account for the change in the density function $f^x(z)$: since we multiply the average per-individual effect with the number of individuals experiencing the reform, we need to work with the original distribution before the reform takes place.⁸

A necessary condition for the tax function T to be optimal, is that the above reform does not affect social welfare. Adding all effects, dropping the common factor $dT_z dZ$, dividing by δ , and requiring that the terms sum to zero leads to the following condition:

$$\frac{1}{\delta} \int_Z^{Z+\delta} (1 - \overline{\alpha}(z)) f^x(z) dz = \frac{[\overline{z_{T_z}} T_z f^x](Z + \delta) - [\overline{z_{T_z}} T_z f^x](Z)}{\delta}, \quad (5)$$

where $\overline{\alpha}(z) \equiv -[\overline{v_T}/\lambda + \overline{z_T} T_z](z)$ denotes the average net marginal effect on social welfare, in monetary units, of an exogenous income increase for individuals at income level z , as first introduced by Diamond (1975). This equation tells us that the social-welfare effect of a change in the tax liability at a small interval $[Z, Z + \delta]$ in the income distribution, should be compensated by the effects of the ensuing changes in the tax gradient around this interval.

If we now take the limit for very small values of δ , keeping $dz \ll \delta$, we are studying the effect of an extremely local reform to the tax function. The left-hand side of equation (5) will converge to $(1 - \overline{\alpha}(Z)) f^x(Z)$, while the right-hand side by definition converges

⁶Throughout this paper, for any function $g : \Theta \rightarrow \mathbb{R}^D$, with D some dimensionality, I will denote using a bar $\overline{g}(x)$ the average of the function g for all individuals pooling at tax base x .

⁷Where necessary, I will denote the product of the values of any functions $g(x)$ and $h(x)$ as $[gh](x)$.

⁸I ignore the impact of individuals making a discrete jump due to the reform, e.g. because they were almost indifferent between two bundles before the reform. It is of second order compared to the total impact.

to the derivative of $\bar{z}T_z f^{\mathcal{X}}$ with respect to z at Z :

$$(1 - \bar{\alpha}(Z)) f^{\mathcal{X}}(Z) = \frac{d}{dz} [\bar{z}T_z f^{\mathcal{X}}](Z). \quad (6)$$

This differential equation is similar to the one-dimensional Euler-Lagrange equation found in the calculus of variations. The main difference is that the objective of our optimization is defined as an integral over the type space, while the tax instrument is defined in the tax base space. In the traditional version of the Euler-Lagrange equation, both the objective and the instrument would be defined in the same vector space. Averaging though at each value of the tax base takes care of this difference. In subsection 3.3 I derive this equation more rigorously.

Since I exclude for now the possibility of bunching, the marginal tax rates should be zero at the edges of the tax base space:

$$T_z(\underline{z}) = T_z(\bar{z}) = 0.$$

There is no point in distorting labour supply at the lowest income level, as there are no individuals at lower income levels to whom the proceeds can be redistributed. Similarly, there is no point distorting labour supply at the highest income level, as there is nobody at higher income levels to levy additional taxes from. This point was first made for the top by Sadka (1976), and for the bottom by Seade (1977).⁹

Integrating (6) and using the latter boundary conditions leads to the necessary optimal-tax condition:

$$\forall z : \int_z^{\bar{z}} (1 - \bar{\alpha}(z')) f^{\mathcal{X}}(z') dz' = -\bar{z}T_z(z) T_z(z) f^{\mathcal{X}}(z), \quad (7)$$

with transversality condition:

$$\int_{\underline{z}}^{\bar{z}} \bar{\alpha}(z') f^{\mathcal{X}}(z') dz' = 1.$$

The transversality condition conforms to the notion, discussed by Jacobs (2013), that the *marginal cost of public funds* equals one when the tax system is optimal.

By performing this integration, the local reform that I constructed at one specific value of the tax base, is repeated over the interval $[z, \bar{z}]$, recovering the tax reform constructed by Saez (2001). Defining the elasticity $\bar{e}(z) \equiv -\bar{z}T_z(z) (1 - T_z(z)) / z$, holding constant the tax liability $T(z)$ at z , we can reformulate this in a more traditional *ABC*-form:

$$\forall z : \frac{T_z(z)}{1 - T_z(z)} = A(z) B(z) C(z), \quad (8)$$

⁹I will derive these boundary conditions more formally for the formalism of this paper in subsection 3.3.

with:

$$\begin{aligned} A(z) &\equiv \frac{1}{\bar{e}(z)}, \\ B(z) &\equiv \frac{\int_z^\infty (1 - \bar{\alpha}(z')) f^{\mathcal{X}}(z') dz'}{1 - F^{\mathcal{X}}(z)}, \\ C(z) &\equiv \frac{1 - F^{\mathcal{X}}(z)}{z f^{\mathcal{X}}(z)}. \end{aligned}$$

The A -term contains the inverse of the average elasticity of gross income with respect to local changes in the marginal tax rate, keeping constant the tax liability at income level z . The B -term contains the average redistributational effect of taking one euro from each individual above income level z and lowering the tax liability elsewhere in the income distribution. The C -term is a hazard rate, which relates the number of individuals who are at the income level where the reform to the marginal tax rate takes place – whose labour supply decisions are being distorted, to the number of individuals who are at higher income levels – from whom additional funds are levied for redistribution.

3.2. Two-Dimensional Tax Base

The tax reform which I introduced for a one-dimensional tax base in the previous subsection, can be extended to a two-dimensional tax base. Suppose the government observes a tax base (z, y) for all individuals. Its components z and y could be thought of for example as labour income and capital income, or the incomes of two individuals in a couple. The government aims to maximize social welfare by setting a joint, non-separable tax function $T(z, y)$, taking into account the multidimensional heterogeneity of the agents. A necessary condition for the tax function to be optimal is that a small tax reform leaves social welfare unchanged. To fix thoughts, I will refer to the tax base components as labour income and capital income.

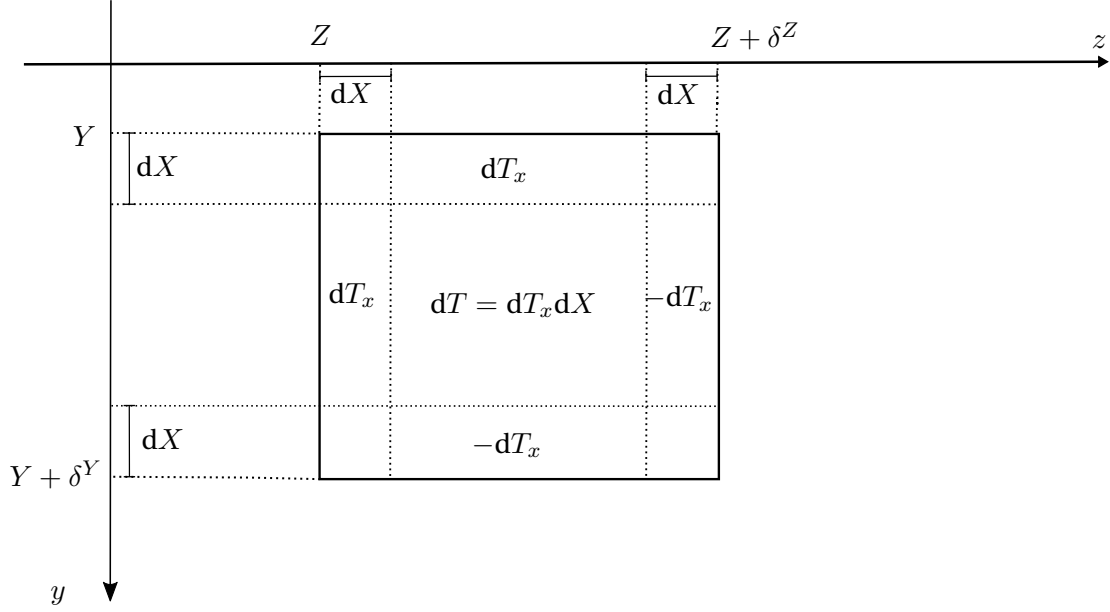
Consider a small rectangular area in the tax base space delineated by an interval $[Z, Z + \delta^Z]$ in the labour income distribution, and an interval $[Y, Y + \delta^Y]$ in the capital income distribution. I introduce a number of tax reforms within this rectangle $[Z, Z + \delta^Z] \times [Y, Y + \delta^Y]$. These reforms are illustrated in figure 3. Around labour income Z , over an infinitesimal width dX , increase the marginal labour income tax by dT_x . Conversely, decrease it by dT_x over width dX around labour income $Z + \delta^Z$. Similarly, increase the marginal capital income tax by the same quantity dT_x over the same width dX around capital income Y , and decrease it by dT_x around capital income $Y + \delta^Y$.

There are again a number of effects from these reforms. Let us first consider the increase in the marginal tax rate on labour income in the rectangle delineated by intervals $[Z, Z + dX]$ and $[Y, Y + \delta^Y]$. Reasoning similar to that of the previous subsection shows that for each level of capital income y within this rectangle, the total effect on the government budget equals:

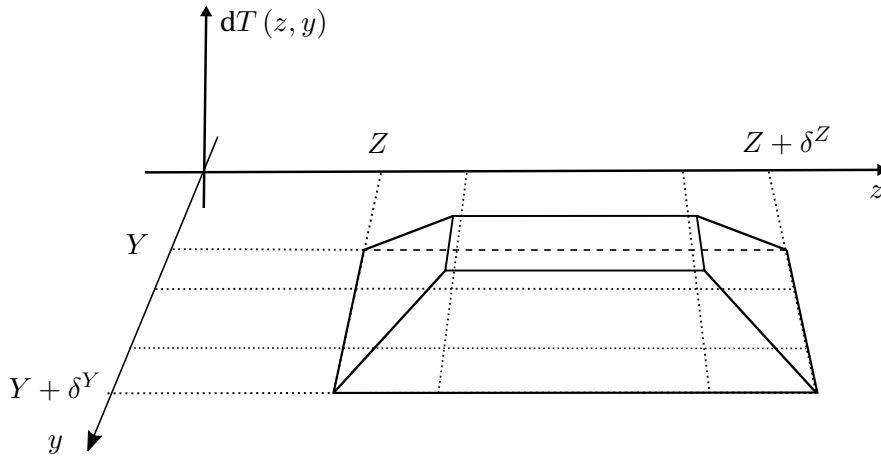
$$dT_x dX \left[\{\bar{z}_{T_z} T_z + \bar{y}_{T_z} T_y\} f^{\mathcal{X}} \right] (Z, y).$$

Figure 3: The tax reforms in two dimensions. At level Z of the first component of the tax base, over an income range of width dX , the marginal tax rate is increased by dT_x . At income level $Z + \delta^Z$, with $\delta^Z \gg dX$ small, this reform is reversed by a decrease in the marginal tax rate by dT_x over an income range of width dX . Similar reforms are implemented for the second component of the tax base. Within the rectangle at which the reforms take place, the tax liability is increased by $dT_x dX$.

(a) The tax base values where the reform takes place, form a small rectangle in the tax base space.



(b) The size of the tax reforms.



The total effect over the interval $[Y, Y + \delta^Y]$ follows by integrating the latter equation:

$$dT_x dX \int_Y^{Y+\delta^Y} [\{\bar{z}_{T_z} T_z + \bar{y}_{T_z} T_y\} f^{\mathcal{X}}] (Z, y) dy.$$

Similar effects occur at the other three edges of the rectangle on which the reforms take place.¹⁰

Within the edges of the rectangle the tax liability increases by $dT_x dX$. Again following reasoning similar to that of the previous subsection, we find the following effect on social welfare:

$$dT_x dX \int_Z^{Z+\delta^Z} \int_Y^{Y+\delta^Y} (1 - \bar{\alpha}(z, y)) f^{\mathcal{X}}(z, y) dy dz,$$

with $\bar{\alpha}(z, y) \equiv -[\bar{v}_T/\lambda + \bar{z}_T T_z + \bar{y}_T T_y](z, y)$ the average net marginal social welfare weight of individuals with labour income z and capital income y .

Setting the sum of all effects to zero, dividing by $dT_x dX$ and by $\delta^Y \delta^Z$, and grouping similar terms on the same line yields:

$$\begin{aligned} & \frac{1}{\delta^Y \delta^Z} \int_Z^{Z+\delta^Z} \int_Y^{Y+\delta^Y} (1 - \bar{\alpha}(z, y)) f^{\mathcal{X}}(z, y) dy dz \\ &= \frac{1}{\delta^Y} \int_Y^{Y+\delta^Y} \frac{[\{\bar{z}_{T_y} T_z + \bar{y}_{T_y} T_y\} f^{\mathcal{X}}] (Z + \delta^Z, y) - [\{\bar{z}_{T_y} T_z + \bar{y}_{T_y} T_y\} f^{\mathcal{X}}] (Z, y)}{\delta^Z} dy \\ & \quad + \frac{1}{\delta^Z} \int_Z^{Z+\delta^Z} \frac{[\{\bar{z}_{T_y} T_z + \bar{y}_{T_y} T_y\} f^{\mathcal{X}}] (z, Y + \delta^Y) - [\{\bar{z}_{T_y} T_z + \bar{y}_{T_y} T_y\} f^{\mathcal{X}}] (z, Y)}{\delta^Y} dz. \end{aligned} \tag{9}$$

The intuition here is similar to the one-dimensional case. The income effects of a local tax increase should be exactly cancelled out by the ensuing compensated effects in the surrounding region.

Now take the limits $\delta^Z \rightarrow 0$ and $\delta^Y \rightarrow 0$, keeping $dX \ll \delta^Z$ and $dX \ll \delta^Y$. Recognize on the right-hand side the definition of partial derivatives with respect to z and y . We can rewrite:

$$\begin{aligned} (1 - \bar{\alpha}(Z, Y)) f^{\mathcal{X}}(Z, Y) &= \frac{\partial}{\partial z} [\{\bar{z}_{T_z} T_z + \bar{y}_{T_z} T_y\} f^{\mathcal{X}}] (Z, Y) \\ & \quad + \frac{d}{dy} [\{\bar{z}_{T_y} T_z + \bar{y}_{T_y} T_y\} f^{\mathcal{X}}] (Z, Y). \end{aligned}$$

Denoting the tax-base vector $\mathbf{x} \equiv (z, y)$ and writing $\mathbf{X} \equiv (Z, Y)$, this can be shortened in vector notation:

$$(1 - \bar{\alpha}(\mathbf{X})) f^{\mathcal{X}}(\mathbf{X}) = \sum_{l=1}^2 \frac{\partial}{\partial x^l} [(\nabla_{\mathbf{x}} T \cdot \bar{\mathbf{x}}_{T_l}) f^{\mathcal{X}}] (\mathbf{X}), \tag{10}$$

¹⁰Note that with dX sufficiently small, the small inaccuracies in my exposition at the four corners of the square become negligible. A more formal prove is provided in the next subsection.

with gradient $\nabla_{\mathbf{x}}T$ introduced in equation (3), and the matrix containing behavioural responses $\mathbf{x}_{T_i}$ in equation (4). This partial differential equation again is similar to the Euler-Lagrange equation found in multidimensional variational analysis, with the added complication that we need to average behavioural effects and welfare weights at each income level. This partial differential equation is a necessary condition for the tax function T to be optimal. As in the one-dimensional case, it is amended with a boundary condition for the gradient $\nabla_{\mathbf{x}}T(\mathbf{X})$, which I will discuss in subsection 3.4, and with the government revenue constraint. Finding a fixed-point equation that characterizes the optimum is not as straightforward as in the one-dimensional case, and is postponed until section 4.

Before doing this, we might gain some further intuition by integrating the Euler-Lagrange equation. Let $V \subseteq \mathcal{X}$ be a compact area in the tax base space, with a piecewise smooth boundary $\Gamma(V)$.¹¹ Integrate Euler-Lagrange equation (10) over this area:

$$\int_V (1 - \bar{\alpha}(\mathbf{x})) dF^{\mathcal{X}}(\mathbf{x}) = \int_V \sum_{l=1}^2 \frac{\partial}{\partial x^l} [\{\nabla_{\mathbf{x}}T(\mathbf{x}) \cdot \overline{\mathbf{x}_{T_l}}(\mathbf{x})\} f^{\mathcal{X}}(\mathbf{x})] d\mathbf{x}.$$

The left-hand side of this equation is the effect on social welfare of a unit increase in the tax liability in the interior of the integration area V , as illustrated in figure 4a. To interpret the right-hand side, I rewrite it using a theorem from vector calculus, named the *divergence theorem* or *Gauss's theorem*. For any point $\mathbf{x}$ at the boundary $\Gamma(V)$ of our integration area, let $\hat{\mathbf{x}}$ be an outward-pointing unit vector that is perpendicular to the boundary, as illustrated in figure 4a. The divergence theorem then allows rewriting our integrated Euler-Lagrange equation:

$$\int_V (1 - \bar{\alpha}(\mathbf{x})) dF^{\mathcal{X}}(\mathbf{x}) = \int_{\Gamma(V)} [\{\nabla_{\mathbf{x}}T \cdot \overline{\mathbf{x}_{\nabla T}} \cdot \hat{\mathbf{x}}\} f^{\mathcal{X}}](\mathbf{x}) d\Gamma. \quad (11)$$

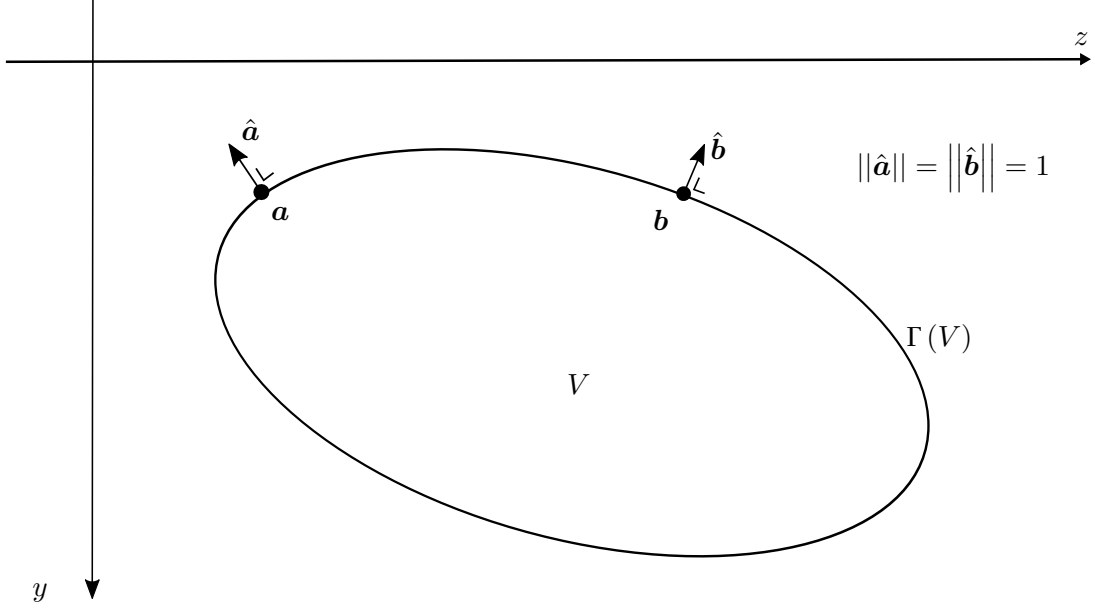
The right-hand side of this equation is a surface integral. The term within square brackets, which is a scalar, is integrated over the boundary $\Gamma(V)$.

Note in figure 4b that due to the increased tax liability in the interior of V , the tax gradient changes on the edge $\Gamma(V)$. The entity $\overline{\mathbf{x}_{\nabla T}}(\mathbf{x}) \cdot \hat{\mathbf{x}}$ is equal to the average behavioural effect caused by a unit change of the tax gradient that is perpendicular to the boundary $\Gamma(V)$ at $\mathbf{x}$. The term within curly brackets is the budgetary effect of this behavioural response for one individual at $\mathbf{x}$, and the term $f^{\mathcal{X}}(\mathbf{x})$ indicates how many individuals reside at that value of the tax base. The surface integral then adds these effects over the entire boundary $\Gamma(V)$. Equation (11) thus states that in the optimum, the effects caused by a unit increase in the tax liability at the interior of the area V should be exactly compensated by the behavioural effects due to the induced changes of the tax gradient at the edges of the area V .

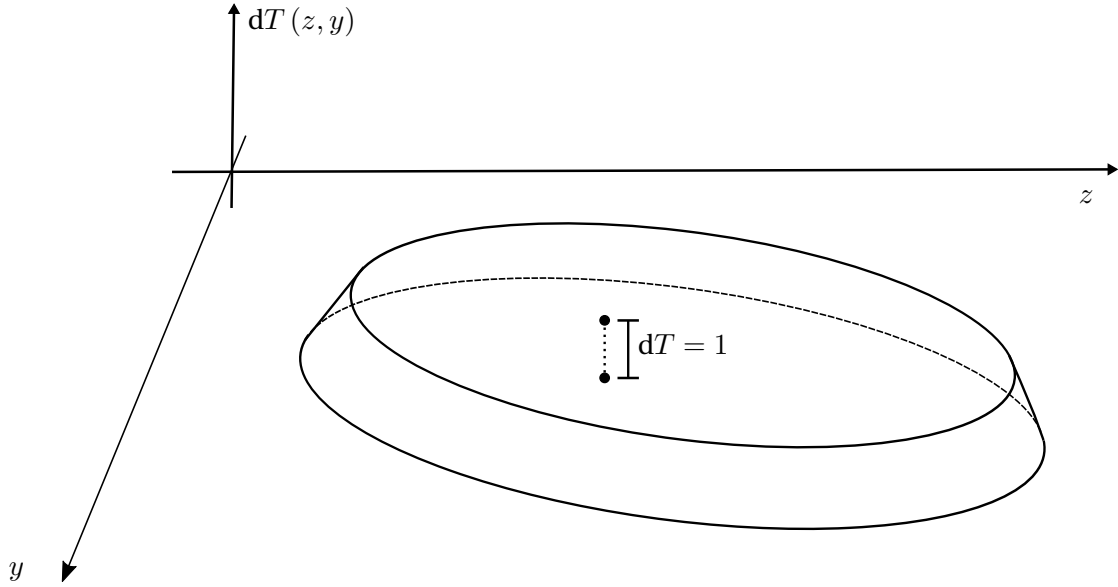
¹¹I maintain this notation throughout this paper, also in higher dimensions: for any compact subset of $\mathbb{R}^L$, let $\Gamma(\cdot)$ denote its boundary surface.

Figure 4: The tax reform in two dimensions, integrated over an area V .

- (a) The area V over which the integrated reform takes place. The vectors $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$ are unit normal vectors, at arbitrary points $\mathbf{a}$ and $\mathbf{b}$ on the edge $\Gamma(V)$ of the integration area, i.e. they are perpendicular to the edge and they have Euclidean norm $\|\hat{\mathbf{a}}\| = \|\hat{\mathbf{b}}\| = 1$.



- (b) The size of the integrated tax reform.



3.3. Higher-Dimensional Tax Bases

In the previous subsection I have shown that the reform that I have proposed for a one-dimensional tax base, can be readily extended to a two-dimensional tax base. I found a partial differential equation, the Euler-Lagrange equation, amended with boundary conditions and the government budget constraint, to be a necessary condition for the tax optimum.

This reasoning can now be further extended to tax bases with an arbitrary number of dimensions. The following theorem shows how Euler-Lagrange equation (10) can be extended to an L -dimensional tax base.¹²

Theorem 1. *The tax optimum with an L -dimensional tax base with multidimensional heterogeneity of the agents, in absence of bunching, complies to the following partial differential equation, referred to as the Euler-Lagrange condition:*

$$\forall \mathbf{x} \in \mathcal{X} : (1 - \bar{\alpha}(\mathbf{x})) f^{\mathcal{X}}(\mathbf{x}) = \sum_{l=1}^L \frac{\partial}{\partial x^l} \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}}_{T_l}) f^{\mathcal{X}} \right](\mathbf{x}), \quad (12)$$

subject to the boundary conditions:

$$\forall \mathbf{x} \in \Gamma(\mathcal{X}) : \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}}_{\nabla_{\mathbf{x}} T} \cdot \hat{\mathbf{x}}) f^{\mathcal{X}} \right](\mathbf{x}) = 0, \quad (13)$$

and the government budget constraint:

$$\int_{\mathbb{R}^L} T(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) d\mathbf{x} = 0. \quad (14)$$

Proof. See appendix A. □

Equation (12) is a multidimensional and extremely local version of equation (9). The intuition is the same: if there is a local increase of the tax liability at tax base value $\mathbf{x}$, then the income effects of this change should be cancelled out by the ensuing compensated effects in the infinitesimal region surrounding it.

The intuition perhaps becomes a bit clearer by integrating the equation. I extend in the following corollary the integration procedure from the two-dimensional case, set forth in the previous subsection, to the multidimensional case. It follows again that when the tax function in place is optimal, then if the tax liability is marginally increased on a subset of the tax base space, the effects on social welfare should be exactly compensated by the effects of the ensuing change in the tax gradient at the boundary of this subset.

¹²More generally, for an additively separable objective functional $\int_{\Theta} \mathcal{L}(\boldsymbol{\theta}, T) dF^{\Theta}(\boldsymbol{\theta})$ with non-linear instrument T , where reforms of T only have local effects on $\mathcal{L}(\boldsymbol{\theta}, T)$ and the curvature of T does not affect the value of $\mathcal{L}(\boldsymbol{\theta}, T)$, an analogous proof shows that $\forall \mathbf{x} : \overline{\mathcal{L}}_T(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) = \sum_j \frac{\partial}{\partial z^j} \left[\overline{\mathcal{L}}_{T_j} f^{\mathcal{X}} \right](\mathbf{x})$ is a necessary condition for the optimum, subject to boundary condition $\left[\left\{ \overline{\mathcal{L}}_{\nabla T} \cdot \hat{\mathbf{x}} \right\} f^{\mathcal{X}} \right](\mathbf{x}) = 0$ and the government budget constraint. This finding allows applying the techniques in this paper in a more general context.

Corollary 1. *The tax optimum with an L -dimensional tax base with multidimensional heterogeneity of the agents, in absence of bunching, complies to the following condition, for any compact volume $V \subseteq \mathcal{X}$ with piecewise smooth boundary $\Gamma(V)$, and with $\hat{\mathbf{x}}$ the unit vector normal to that boundary at point $\mathbf{x} \in \Gamma(V)$:*

$$\int_V (1 - \bar{\alpha}(\mathbf{x})) \, dF^{\mathcal{X}}(\mathbf{x}) = \int_{\Gamma(V)} \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}_{\nabla_{\mathbf{x}} T}} \cdot \hat{\mathbf{x}}) f^{\mathcal{X}} \right](\mathbf{x}) \, d\Gamma. \quad (15)$$

Proof. Integrate Euler-Lagrange equation (12) over volume V and apply the *divergence theorem*. $\square$

Even if this corollary has an intuitive interpretation, the fact that it has to be valid for any compact volume $V \subseteq \mathcal{X}$ makes it unpractical to use. Moreover, it does not give us the kind of insight into the properties of the gradient of the tax function that we obtain for the one-dimensional problem from *ABC*-equation (8). I provide a more intuitive characterization of the optimum in section 4. Before doing so, I will discuss the boundary condition for the optimum.

3.4. Boundary Conditions: Application to Joint Taxation of Couples

Euler-Lagrange condition (12) is a partial differential equation for the tax function T . If it has solutions, further restrictions need to be imposed to pin down the one that we are interested in. Like in the one-dimensional case, one needs to impose government budget constraint (14) on the one hand, and the set of boundary conditions (13) on the other. To gain intuition, I will describe here the boundary conditions for the joint taxation of couples.

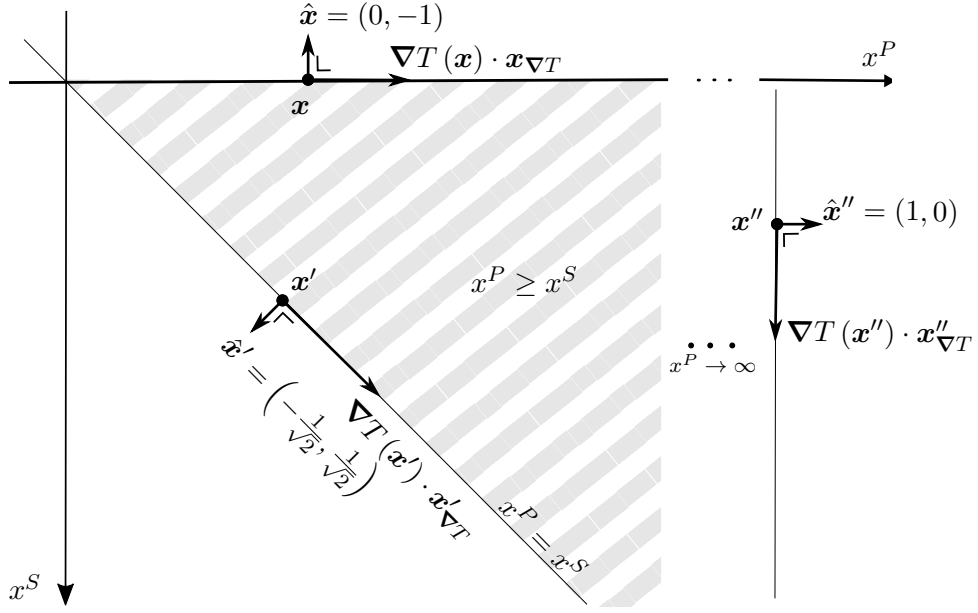
Let the tax base be defined by the gross earnings of a couple: $\mathbf{x} \equiv (x^P, x^S)$, where x^P are the earnings of the primary earner, and x^S are the earnings of the secondary earner. The primary earner is the individual with the highest income: $x^P \geq x^S$. The boundaries for this problem are illustrated in figure 5. The fact that the secondary earner, by definition, cannot earn more than the primary earner, makes that the line defined by $x^P = x^S$ is a boundary of the tax base space. Requiring that both incomes are non-negative: $x^S, x^P \geq 0$, yields another boundary at $x^S = 0$. Finally, there will be a boundary at the top income of the primary earner, $x^P \equiv \bar{x}^P$. If I do not impose an upper bound to the incomes, then the third “boundary” occurs as the income of the primary earner approaches infinity: $x^P \rightarrow \infty$. This last possibility is the one depicted in figure 5.

Theorem 1 tells us that the condition for the boundary of the tax base space $\Gamma(\mathcal{X})$ looks as follows:

$$\forall \underline{\mathbf{x}} \in \Gamma(\mathcal{X}) : f^{\mathcal{X}}(\underline{\mathbf{x}}) \left\{ \left[T_{x^P}(\underline{\mathbf{x}}) \overline{x_{\nabla T}^P} + T_{x^S}(\underline{\mathbf{x}}) \overline{x_{\nabla T}^S} \right] \cdot \hat{\mathbf{x}} \right\} = 0,$$

where $\hat{\mathbf{x}}$ is the unit normal vector perpendicular to the boundary at any point $\underline{\mathbf{x}}$ on that boundary. The term between square brackets is a row vector, with its l -th element $T_{x^P}(\underline{\mathbf{x}}) \overline{x_{T_l}^P} + T_{x^S}(\underline{\mathbf{x}}) \overline{x_{T_l}^S}$ indicating the average change in government revenue at $\underline{\mathbf{x}}$ that

Figure 5: The boundaries for the two-dimensional problem of the taxation of couples, with tax base $\mathbf{x} \equiv (x^P, x^S)$. The boundaries are determined by the fact that the primary earner has the higher income: $x^P \geq x^S$, by the fact that incomes are restricted to be positive: $x^S \geq 0$, and by the top income distribution for primary earners. The figure depicts the case where the income distribution is unbounded at the top: $x^P \rightarrow \infty$. The boundary condition states that at each point $\mathbf{x}$ at the boundary, the vector $\nabla T(\mathbf{x}) \cdot \mathbf{x}_{\nabla T}$ containing the effects of a change in the tax gradient on government revenue, should be tangent to the boundary. Put differently, changes in the tax gradient perpendicular to the boundary should have no effects on government revenues.



is caused by a small change in the marginal tax rate T_l . The left-hand side of the boundary condition thus indicates the change in government revenue that is caused by a unit change in the tax gradient, perpendicular to the surface of the tax base space. This change should be zero in the optimum.

Suppose that $f^{\mathbf{x}}(\mathbf{x}) \neq 0$. Since the term within square brackets is a vector, and since its dot product with the unit normal vector $\hat{\mathbf{x}}$ should be zero, it follows the vector described by this term within square brackets should be parallel to the boundary of the tax base space.

Let us look specifically at the point $\mathbf{x}$ in the figure, which lies on the boundary $x^S = 0$. The coordinates of the unit normal vector at this point are $(\hat{x}^P, \hat{x}^S) = (0, -1)$. Assuming that there are at least some individuals choosing this point, $f^{\mathbf{x}}(\mathbf{x}) \neq 0$, the boundary condition states:

$$-T_{x^P}(\mathbf{x}) \overline{x_{T^S}^P} - T_{x^S}(\mathbf{x}) \overline{x_{T^S}^S} = 0.$$

It immediately follows that, contrary to the one-dimensional case, the marginal tax rate

at the boundary is no longer necessarily zero. Assuming that the behavioural response of the primary earner is positive (working more as the marginal tax rate for his partner increases), and that of the secondary earner is negative (working less as his marginal tax rate increases), then all that one can infer from the boundary condition is that the marginal tax rates T_{x^P} and T_{x^S} have the same sign, and that their proportion is as follows:

$$\frac{T_{x^S}(\mathbf{x})}{T_{x^P}(\mathbf{x})} = -\frac{\overline{x_{T^S}^P}}{\overline{x_{T^S}^S}}.$$

In other words, suppose that the secondary earner has no income, and the primary earner has a positive income, facing a positive marginal tax rate. It follows in this situation that the secondary earner will also face a positive marginal tax rate, even if this person does not work and even if there is no bunching at the bottom. All that is required is that the total tax wedge for the family, the term within square brackets, is zero.

If we look at the point $\mathbf{x}'$ on the boundary where $x^S = x^P$, the coordinates of the unit normal vector are $\hat{\mathbf{x}}' = (-1/\sqrt{2}, 1/\sqrt{2})$.¹³ Assuming that at least some individuals are in this situation, so $f^{\mathcal{X}}(\mathbf{x}') \neq 0$, then the boundary condition at this point states:

$$\frac{T_{x^P}(\mathbf{x}')}{T_{x^S}(\mathbf{x}')} = -\frac{\overline{x_{T^S}^{'S}} - \overline{x_{T^P}^{'S}}}{\overline{x_{T^S}^{'P}} - \overline{x_{T^P}^{'P}}}.$$

An advantage of my approach is now that anonymity can be imposed through the sufficient statistics. If both individuals have equal incomes, $x^S = x^P$, the government has no way of telling them apart. Imposing anonymity on the sufficient statistics implies that the own-tax elasticities are equal, $\overline{x_{T^S}^{'S}} = \overline{x_{T^P}^{'P}}$, and that the cross-elasticities are equal: $\overline{x_{T^S}^{'P}} = \overline{x_{T^P}^{'S}}$. It follows that when both individuals have equal incomes, $x^S = x^P$, they should face equal marginal tax rates:

$$T_{x^P}(\mathbf{x}') = T_{x^S}(\mathbf{x}').$$

Finally let us study the situation for a tax base value $\mathbf{x}''$, where the earnings of the primary earner either reach a top value $x^P = \bar{x}^P$, or they approach infinity: $x^P \rightarrow \infty$. In this case the coordinates of the unit normal vector are $(1, 0)$, and the boundary condition becomes:

$$f^{\mathcal{X}}(\mathbf{x}'') \left[T_{x^P}(\mathbf{x}'') \overline{x_{T^P}^{'P}} + T_{x^S}(\mathbf{x}'') \overline{x_{T^P}^{'S}} \right] \rightarrow 0.$$

If the income distribution is bounded from above, with income density $f^{\mathcal{X}}(\mathbf{x}'')$ strictly larger than zero, then the term within square brackets must become zero for the top individual. Note that this does not imply that the top marginal tax rates become zero: if the behavioural responses have opposite signs, both individuals might face positive marginal tax rates at the top, even if the income distribution is bounded from above. It is only the total tax wedge for the couple, the term within square brackets, that should

¹³One can verify that the norm of this vector equals one: $\|\hat{\mathbf{x}}'\| = \sqrt{(-1/\sqrt{2})^2 + (1/\sqrt{2})^2}$.

become zero. With an unbounded distribution, what happens depends on the asymptotic behaviour of the income distribution and the elasticities. Like in the one-dimensional case, the term within square brackets no longer necessarily decreases at the top.

4. Characterizing the Optimum

The right-hand side of the Euler-Lagrange equation (12) is a first-order derivative of a term that itself already contains first-order derivatives $\nabla_x T$ of the tax function. The Euler-Lagrange equation is thus a second-order partial differential equation in T . Generally, this equation is hard to solve analytically. Finding an analytical solution though is not necessary to find an intuitive economic interpretation.

I will argue in this section that the optimum is determined by a balancing exercise between efficiency and equity, that is a localized version of the balancing exercise described for linear taxes by Atkinson and Stiglitz (1980, p.386-390). At any given value of the tax base, the proportionate reduction in a component of aggregate demand, along compensated demand curves, should be equal to a localized distributional characteristic for that tax base component similar to the one introduced by Feldstein (1972a,b).

Before arriving at this conclusion I need to introduce some methodology. For the one-dimensional problem, in subsection 3.1, we found the *ABC*-style formulation (8) by taking the integral of the Euler-Lagrange equation (6). This integration method treats the problem as if it were a first-order partial differential equation in $\overline{z}_{T_z} T_z f^x$, rather than a second-order partial differential equation in T .

Indeed, the *ABC*-equation (8) does not provide an analytical solution to the Euler-Lagrange equation. The three terms on the right-hand side still depend on the tax function. It is a necessary fixed-point equation for the optimum, a reformulation of the second-order Euler-Lagrange equation which has the more elegant economic interpretation described in subsection 3.1. I extend this reasoning to higher-dimensional tax bases, treating the multidimensional Euler-Lagrange equation (12) as a first-order rather than as a second-order partial differential equation.

We saw in corollary 1 that for multidimensional problems, simply integrating the Euler-Lagrange equation no longer yields straightforward insight into the optimal gradient of the tax function. I introduce a different solution method. In subsection 4.1 I illustrate this method for the one-dimensional case, where it coincides with the method of *Green functions* that is often used in the exact sciences, and that was introduced for mechanism design problems by Renes and Zoutman (2016a).¹⁴ I then extend this method to multidimensional tax bases in the subsequent subsections. For now, I still assume that the tax function is smooth, postponing the inclusion of kinks and bunching to a later section. I still allow for multidimensional heterogeneity of the agents and for pooling without bunching.

¹⁴An introduction and some examples of the use in the exact sciences can be found in Arfken and Weber (2005).

4.1. One-Dimensional Tax Base

I will now derive the *ABC*-style optimal-tax equation (8) from the one-dimensional Euler-Lagrange equation (6) using the method of *Green functions*. This may at first seem like a detour, but it will become clear in the following subsections how this allows reformulating the multidimensional Euler-Lagrange equation (12) in a more intuitive way. I will treat the one-dimensional Euler-Lagrange equation (6) as a first-order differential equation of the following form:

$$\forall z \in [\underline{z}, \bar{z}] : A(z) = \frac{dB(z)}{dz}, \quad (16)$$

with “unknown” function:

$$B(z) \equiv \left[\overline{z_{T_z}} T_z f^{\mathcal{X}} \right] (z),$$

with “known” function:

$$A(z) \equiv (1 - \bar{\alpha}(z)) f^{\mathcal{X}}(z),$$

and with initial condition $B(\underline{z}) = 0$.

First I introduce the concept of a *Dirac delta function*. It is a generalized function $z \mapsto \delta(z)$ which has value zero everywhere except at $z = 0$, and which integrates to one over the whole real line: $\int_{\mathbb{R}} \delta(z) dz = 1$. It is “generalized” in the sense that in order to satisfy these properties, it must take an infinite value in $z = 0$ (if not, it would integrate to zero).¹⁵ One way to represent the Dirac delta function is to see it as the limit of a series of functions. For example, define the following rectangular function:

$$\delta_{\varepsilon}(x) \equiv \begin{cases} \frac{1}{\varepsilon} & \text{if } 0 < x < \varepsilon, \\ 0 & \text{elsewhere.} \end{cases} \quad (17)$$

This function is depicted in figure 6. One can see that for any ε , $\int_{\mathbb{R}} \delta_{\varepsilon}(z) dz = 1$. Take the limit $\varepsilon \rightarrow 0$ to find that $\delta_{\varepsilon} \rightarrow \delta$.

The reason why the Dirac delta function is used, is that it has a desirable property that will help us. Intuitively it can be seen as follows. For any function $g : \mathbb{R} \rightarrow \mathbb{R}$, use definition (17) to find the following integral:

$$\int_{\mathbb{R}} g(z') \delta_{\varepsilon}(z') dz' = \frac{1}{\varepsilon} \int_0^{\varepsilon} g(z') dz'.$$

Take the limit $\varepsilon \rightarrow 0$ to find the property:

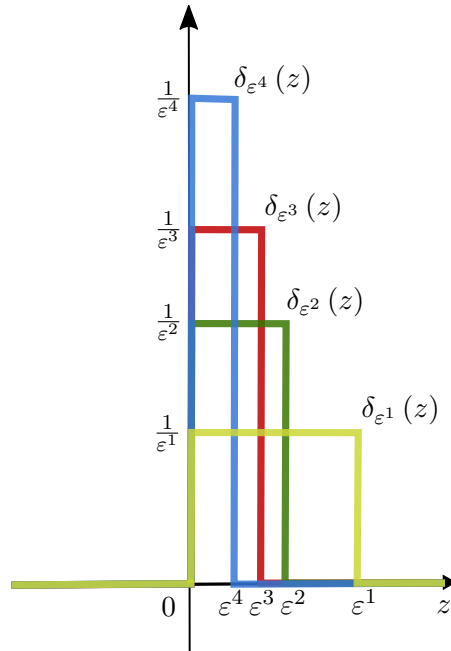
$$\int_{\mathbb{R}} g(z') \delta(z') dz' = g(0),$$

or more generally:

$$\forall z \in [\underline{z}, \bar{z}] : \int_{\mathbb{R}} g(z') \delta(z - z') dz' = g(z). \quad (18)$$

¹⁵The concept of a Dirac delta function is described more extensively e.g. by Arfken and Weber (2005).

Figure 6: The function δ_ε which takes value $1/\varepsilon$ on the interval $[0, \varepsilon]$, and which is zero elsewhere. It is depicted for four values $\varepsilon^4 < \varepsilon^3 < \varepsilon^2 < \varepsilon^1$. Taking the limit $\varepsilon \rightarrow 0$ leads to the one-dimensional Dirac delta function δ which is zero for any $z \neq 0$, and which approaches infinity at $z = 0$.



Integrating a function g with $\delta(z - z') \, dz'$ as measure thus selects its value $g(z)$.¹⁶

Now suppose that we can find a function $G : \mathbb{R}^2 \rightarrow \mathbb{R}$ whose derivative is the Dirac delta:

$$\forall z, z' \in [\underline{z}, \bar{z}] : \frac{dG(z, z')}{dz} \equiv \delta(z - z'), \quad (19)$$

and which complies to the following initial condition:

$$\forall z' \in [\underline{z}, \bar{z}] : G(\underline{z}, z') \equiv 0. \quad (20)$$

Then the following function solves differential equation (16):

$$B(z) \equiv \int_{\underline{z}}^{\bar{z}} A(z') G(z, z') \, dz'. \quad (21)$$

To see this, take its derivative:

$$\begin{aligned} \forall z \in [\underline{z}, \bar{z}] : \frac{dB(z)}{dz} &= \frac{d}{dz} \left[\int_{\underline{z}}^{\bar{z}} A(z') G(z, z') \, dz' \right] \\ &= \int_{\underline{z}}^{\bar{z}} A(z') \frac{dG(z, z')}{dz} \, dz' \\ &= \int_{\underline{z}}^{\bar{z}} A(z') \delta(z - z') \, dz' \\ &= A(z), \end{aligned}$$

where I used the fact that the derivative can be brought inside the integral, and I used definition (19) and property (18). Furthermore, verify that indeed the initial condition for B is fulfilled:

$$B(\underline{z}) = \int_{\underline{z}}^{\bar{z}} A(z') G(\underline{z}, z') \, dz' = 0,$$

because of condition (20).

It follows that finding a solution to a differential equation of form (16) boils down to finding a function G which solves equation (19) and which complies to initial condition (20). Such a function is called a *Green function* of the problem.

I will now show that the Green function for the one-dimensional problem is the *unit step function*.¹⁷ This function is defined as follows:

$$\forall z, z' : H(z - z') = \begin{cases} 1 & \text{if } z > z', \\ 0 & \text{if } z \leq z'. \end{cases} \quad (22)$$

It immediately follows that it complies to initial condition (20):

$$\forall z \in [\underline{z}, \bar{z}] : H(\underline{z} - z) = 0.$$

¹⁶This property is so fundamental that it is sometimes used as a defining property of the Dirac delta function.

¹⁷Again, see Arfken and Weber (2005) for a more traditional formulation of this property.

Furthermore, note that if we interpret $\delta(z) dz$ as a probability measure, then its cumulative distribution function is $H(z)$:

$$\int_{-\infty}^z \delta(z'' - z') dz'' = H(z - z'),$$

from the definition of the Dirac delta function. Take derivatives on both sides:

$$\frac{dH(z - z')}{dz} = \delta(z - z'). \quad (23)$$

We have thus found that the unit step function is the Green function for our problem. Substitute it into equation (21) to find the solution to differential equation (16):

$$B(z) = \int_{\underline{z}}^{\bar{z}} A(z') H(z - z') dz' = \int_{\underline{z}}^z A(z') dz'.$$

This is exactly the solution that we would have found by simply integrating the differential equation. Applying it to the one-dimensional Euler-Lagrange equation (6) leads to *ABC*-equation (8). I will extend this method to higher-dimensional problems in the following subsections.

4.2. Two-Dimensional Tax Base

The two-dimensional Euler-Lagrange equation (10) is a second-order partial differential equation, subject to the boundary condition that the entity $\{\nabla_{\mathbf{x}} T(\mathbf{x}) \cdot \overline{\mathbf{x}_{T_l}}(\mathbf{x}) \cdot \hat{\mathbf{x}}\} f^{\mathcal{X}}(\mathbf{x})$ equals zero on the boundary $\Gamma(\mathcal{X})$ of the tax base space. Following the procedures for the one-dimensional case, set forth in the previous subsection, I will treat this equation as if it were a first-order partial differential equation. In other words, I will solve a partial differential equation of the following form:

$$\forall \mathbf{x} \in \mathcal{X} : A(\mathbf{x}) = \sum_{l=1}^2 \frac{\partial B^l(\mathbf{x})}{\partial x^l}, \quad (24)$$

with:

$$A(\mathbf{x}) \equiv (1 - \bar{\alpha}(\mathbf{x})) f^{\mathcal{X}}(\mathbf{x}), \quad (25)$$

$$\forall l : B^l(\mathbf{x}) \equiv \left[\left(\sum_j T_j \overline{x_{T_l}^j} \right) f^{\mathcal{X}} \right](\mathbf{x}), \quad (26)$$

subject to boundary conditions:

$$\forall \underline{\mathbf{x}} \in \Gamma(\mathcal{X}) : \mathbf{B}(\underline{\mathbf{x}}) \cdot \hat{\mathbf{x}} = 0. \quad (27)$$

Note that this formulation immediately leads to a transversality condition. Use the divergence theorem to find:

$$\begin{aligned}\int_{\mathcal{X}} A(\mathbf{x}) \, d\mathbf{x} &= \int_{\mathcal{X}} \sum_{l=1}^2 \frac{\partial B^l(\mathbf{x})}{\partial x^l} d\mathbf{x} \\ &= \int_{\Gamma(\mathcal{X})} \mathbf{B}(\mathbf{x}) \cdot \hat{\mathbf{x}} d\Gamma \\ &= 0,\end{aligned}$$

or substituting (25) for $A(\mathbf{x})$:

$$\int_{\mathcal{X}} \bar{\alpha}(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) \, d\mathbf{x} = 1. \quad (28)$$

This condition extends the notion, discussed by Jacobs (2013), that the *marginal cost of public funds* equals one when the tax system is optimal.

To solve first-order partial differential equation (24), first introduce the two-dimensional Dirac delta function as the product of two one-dimensional Dirac delta functions:

$$\forall x, y \in \mathbb{R}^2 : \delta^2(x, y) \equiv \delta(x) \delta(y). \quad (29)$$

This function can again be constructed as the limit of a series of functions. I again use the rectangular function for this purpose:

$$\delta_{\varepsilon}^2(x^1, x^2) \equiv \delta_{\varepsilon}(x^1) \delta_{\varepsilon}(x^2) = \begin{cases} \frac{1}{\varepsilon^2} & \text{if } 0 < x^1 < \varepsilon \text{ and } 0 < x^2 < \varepsilon, \\ 0 & \text{elsewhere.} \end{cases}$$

This function is depicted in figure 7. It integrates to one for any ε :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta_{\varepsilon}^2(x^1, x^2) \, dx^1 dx^2 = 1. \quad (30)$$

Take the limit $\varepsilon \rightarrow 0$ to find that $\delta_{\varepsilon}^2 \rightarrow \delta^2$.

It is straightforward to check that property (18) extends to two dimensions. Integrating a function $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ with measure $\delta^2(\mathbf{x} - \mathbf{x}') \, d\mathbf{x}'$ selects its value $g(\mathbf{x})$:

$$\forall \mathbf{x} \in \mathbb{R}^2 : \int_{\mathbb{R}^2} g(\mathbf{x}') \delta^2(\mathbf{x} - \mathbf{x}') \, d\mathbf{x}' = g(\mathbf{x}).$$

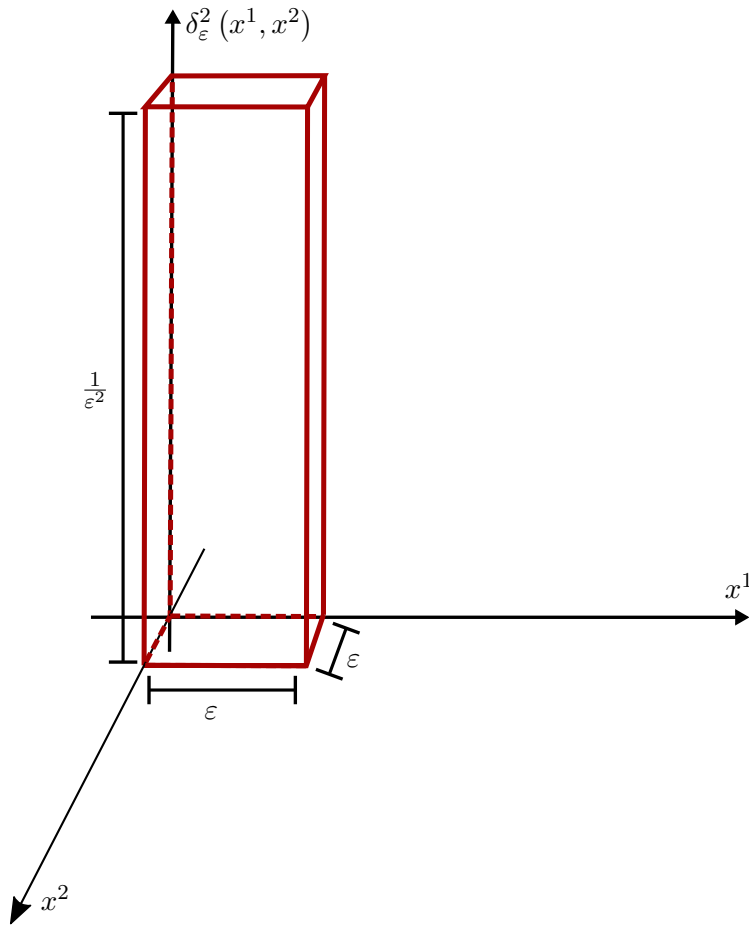
Like in the one-dimensional case described in the previous subsection, I will show that finding a solution to first-order partial differential equation (24) boils down to finding a Green function $\mathbf{G} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ which solves the following partial differential equation:

$$\forall \mathbf{x} \in \mathbb{R}^2 : \delta^2(\mathbf{x} - \mathbf{x}') = \sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l}, \quad (31)$$

complying to the following boundary condition:

$$\forall \underline{\mathbf{x}} \in \Gamma(\mathcal{X}), \forall \mathbf{x}' \in \mathcal{X} : \mathbf{G}(\underline{\mathbf{x}}, \mathbf{x}') \cdot \hat{\mathbf{x}} = 0. \quad (32)$$

Figure 7: The function δ_ε^2 which takes value $1/\varepsilon^2$ on the rectangle $[0, \varepsilon] \times [0, \varepsilon]$, and which is zero elsewhere. Taking the limit $\varepsilon \rightarrow 0$ leads to the two-dimensional Dirac delta function δ^2 which is zero for all points besides the origin, and which approaches infinity at the origin.



Note that the Green function $\mathbf{G}$ now is a vector-valued function, rather than a scalar-valued function as in the one-dimensional case.

If we find such a function $\mathbf{G}$, then first-order partial differential equation (24) has the following solution:

$$\forall \mathbf{x}, l : B^l(\mathbf{x}) \equiv \int_{\mathcal{X}} A(\mathbf{x}') G^l(\mathbf{x}, \mathbf{x}') d\mathbf{x}'. \quad (33)$$

This can be seen by substituting it into the right-hand side of partial differential equation (24), noting that the partial derivative can be brought inside the integral and using property (30):

$$\begin{aligned} \forall \mathbf{x} : \sum_{l=1}^2 \frac{\partial B^l(\mathbf{x})}{\partial x^l} &= \sum_{l=1}^2 \frac{\partial}{\partial x^l} \left[\int_{\mathcal{X}} A(\mathbf{x}') G^l(\mathbf{x}, \mathbf{x}') d\mathbf{x}' \right] \\ &= \int_{\mathcal{X}} A(\mathbf{x}') \sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} d\mathbf{x}' \\ &= \int_{\mathcal{X}} A(\mathbf{x}') \delta^2(\mathbf{x} - \mathbf{x}') d\mathbf{x}' \\ &= A(\mathbf{x}). \end{aligned}$$

Verify that indeed the boundary conditions (27) are fulfilled, using condition (32):

$$\forall \underline{\mathbf{x}} \in \Gamma(\mathcal{X}), \forall \mathbf{x} \in \mathcal{X} : \mathbf{B}(\underline{\mathbf{x}}) \cdot \hat{\mathbf{x}} = \int_{\mathcal{X}} A(\mathbf{x}') \mathbf{G}(\underline{\mathbf{x}}, \mathbf{x}') \cdot \hat{\mathbf{x}} d\mathbf{x}' = 0.$$

The question now is how to find the functions G^l that comply to partial differential equation (31) with boundary conditions (32). I will state a solution here for the case where the tax base space equals the entire two-dimensional real vector space, $\mathcal{X} = \mathbb{R}^2$. I will treat cases where the tax base set is a strict subset of the real vector space, $\mathcal{X} \subsetneq \mathbb{R}^2$, e.g. excluding negative values, in section 6.

I will show that the Green function for this problem looks as follows:

$$\forall \mathbf{x}, \mathbf{x}' : \mathbf{G}(\mathbf{x}, \mathbf{x}') \equiv \frac{\mathbf{x} - \mathbf{x}'}{D^2(\mathbf{x} - \mathbf{x}')}, \quad (34)$$

where I introduce the two-dimensional distance function for any vector $\mathbf{v}$ in $\mathbb{R}^2$:

$$D^2(\mathbf{v}) \equiv 2V^2(\|\mathbf{v}\|),$$

with $V^2(r)$ a function which maps a real number r on the surface area πr^2 of a circle with radius r , and with $\|\mathbf{v}\| \equiv \sqrt{(v^1)^2 + (v^2)^2}$ the Euclidean norm, the “length” of the vector $\mathbf{v}$. After showing that this function indeed is a Green function, I will show how it leads to intuitive optimal-tax results.

First I need to show that the function (34) solves partial differential equation (31). I show in appendix B that it complies to the following properties:

$$\forall \mathbf{x} \neq \mathbf{x}' : \sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} = 0, \quad (35)$$

$$\forall \mathbf{x}' : \int_{\mathbb{R}^2} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x} = 1. \quad (36)$$

These are the defining properties for the Dirac delta function. We conclude that the function $\mathbf{G}$ indeed solves partial differential equation (31).¹⁸

Next, note that the boundary conditions (32) are fulfilled:

$$\forall \mathbf{x}' \in \mathbb{R}^2 : \lim_{\|\underline{\mathbf{x}}\| \rightarrow \infty} \mathbf{G}(\underline{\mathbf{x}}, \mathbf{x}') \cdot \hat{\underline{\mathbf{x}}} = \lim_{\|\underline{\mathbf{x}}\| \rightarrow \infty} \frac{(\underline{\mathbf{x}} - \mathbf{x}') \cdot \hat{\underline{\mathbf{x}}}}{2\pi \|\underline{\mathbf{x}} - \mathbf{x}'\|^2} = 0. \quad (37)$$

It thus follows from (33) that the solution to partial differential equation (24) is as follows:

$$\forall \mathbf{x}, l : B^l(\mathbf{x}) = \int_{\mathbb{R}^2} A(\mathbf{x}') \frac{x^l - x'^l}{D^2(\mathbf{x} - \mathbf{x}')} d\mathbf{x}'.$$

Substituting the definitions (25) and (26) for the functions A and B^l leads to optimal-tax condition:

$$\forall \mathbf{x}, l : \sum_j T_j(\mathbf{x}) \overline{x_{T_l}^j}(\mathbf{x}) f^{\mathbf{x}}(\mathbf{x}) = \int_{\mathbb{R}^2} (1 - \bar{\alpha}(\mathbf{x}')) \frac{x^l - x'^l}{D^2(\mathbf{x} - \mathbf{x}')} dF^{\mathbf{x}}(\mathbf{x}'). \quad (38)$$

Using transversality condition (28), this equation can be simplified as follows:

$$\forall \mathbf{x}, l : \sum_j T_j(\mathbf{x}) X_{T_l}^j(\mathbf{x}) = \text{cov} \left(\bar{\alpha}(\mathbf{x}'), \frac{x'^l - x^l}{D^2(\mathbf{x}' - \mathbf{x})} \right), \quad (39)$$

where I introduce aggregate demand $X^j(\mathbf{x}) \equiv \overline{x^j}(\mathbf{x}) f^{\mathbf{x}}(\mathbf{x})$ for tax base component j at value $\mathbf{x}$, and where $X_{T_l}^j(\mathbf{x}) \equiv \overline{x_{T_l}^j}(\mathbf{x}) f^{\mathbf{x}}(\mathbf{x})$ equals the sum of the compensated responses for the individuals pooling at tax base value $\mathbf{x}$.¹⁹ The left-hand side in this equation equals the compensated change in public revenues which is induced by a reform to the marginal tax rate T_l . It is the sum of the compensated effects on the different components of aggregate demand, multiplied by the respective marginal tax rates that apply to these components. This side captures the efficiency effects of the reform.

The right-hand side captures the equity effects. It is a distance-weighted covariance between the average net marginal social welfare weight of the individuals pooling at tax base value $\mathbf{x}$ with the value of the l -th component of the tax base. If the welfare

¹⁸Partial differential equation (31) has multiple solutions. I argue in appendix B why solution 34 is the one that we are interested in.

¹⁹The entities $X^j(\mathbf{x})$ and $X_{T_l}^j(\mathbf{x})$ are density functions.

weights decrease strongly with the value of the tax base component, e.g. if individuals with higher gross labour income are weighted much less in the social welfare objective, then this covariance will be strongly negative. Balancing efficiency considerations against equity considerations, the government will accept larger efficiency losses due to taxation, thus allowing the absolute value of the left-hand side of the equation to become larger. For given behavioural responses, high marginal tax rates will be warranted. However, if the welfare weights decrease less strongly, the absolute value of the covariance will be smaller, and optimal marginal tax rates will be lower.

The covariance term which captures the equity effects is distance-weighted. What matters most is the covariance between the welfare weights and the tax base in the immediate environment of the value of the tax base $\mathbf{x}$ under consideration. Figure 8a shows graphically the values of the distance weights: they become infinitely large in the immediate environment of $\mathbf{x}$ and they decrease rapidly at slightly larger distances. Although the weights continue decreasing gradually at further distances, they only converge to zero at infinity. Figure 8b shows how the welfare weights are symmetric, in the sense that they depend only on the absolute value of the distance from the tax base level under consideration, independent of the direction.

We can reformulate equation (39) in a more familiar form. Denote the population average of the net marginal social welfare weights as $\mathcal{A} \equiv \int_{\mathcal{X}} \bar{\alpha}(\mathbf{x}') f^{\mathcal{X}}(\mathbf{x}') d\mathbf{x}'$. This quantity equals one in the optimum, by transversality condition (28). The optimal-tax condition can then be rewritten as follows:

$$\forall \mathbf{x}, l : \frac{\sum_j T_j(\mathbf{x}) X_{T_j}^l(\mathbf{x})}{X^l(\mathbf{x})} = \frac{\text{cov}\left(\bar{\alpha}(\mathbf{x}'), \frac{x'^l - x^l}{D^2(\mathbf{x}' - \mathbf{x})}\right)}{\mathcal{A} X^l(\mathbf{x})}, \quad (40)$$

where I use Slutsky symmetry $X_{T_l}^j(\mathbf{x}) = X_{T_j}^l(\mathbf{x})$. This is a different way of expressing the government's balancing exercise between efficiency and equity. The left-hand side of this equation is the proportional reduction of the l -th commodity along the compensated demand schedule. The normalized covariance on the right-hand side is an extension of the *distributional characteristic* that was introduced by Feldstein (1972a,b).

This optimal-tax equation resembles closely a well-known result from the literature. Suppose that the government could not use the nonlinear, non-separable tax function T , but instead it had to resort to separate linear tax rates t_l . Adapting the results of Atkinson and Stiglitz (1980, p.386-390) to my notations, denoting as $\bar{X}^l$ the population aggregate demand for the l -th component of the tax base, the optimal tax rates would be determined by the following equation [TODO: do we really need the bar on $\mathbf{X}$?]:²⁰

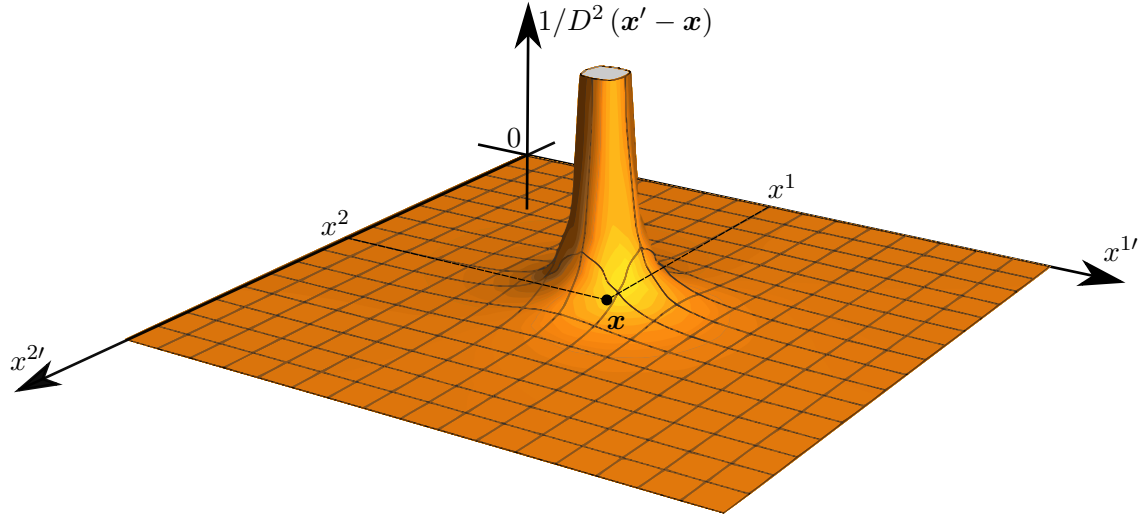
$$\forall l : \frac{\sum_j t_j \bar{X}_{t_j}^l}{\bar{X}^l} = \frac{\text{cov}\left(\bar{\alpha}(\mathbf{x}'), x'^l\right)}{\mathcal{A} \bar{X}^l}. \quad (41)$$

The covariance on the right-hand side is again an extension of the distributional characteristic introduced by Feldstein (1972a,b). Since the different marginal tax rates are

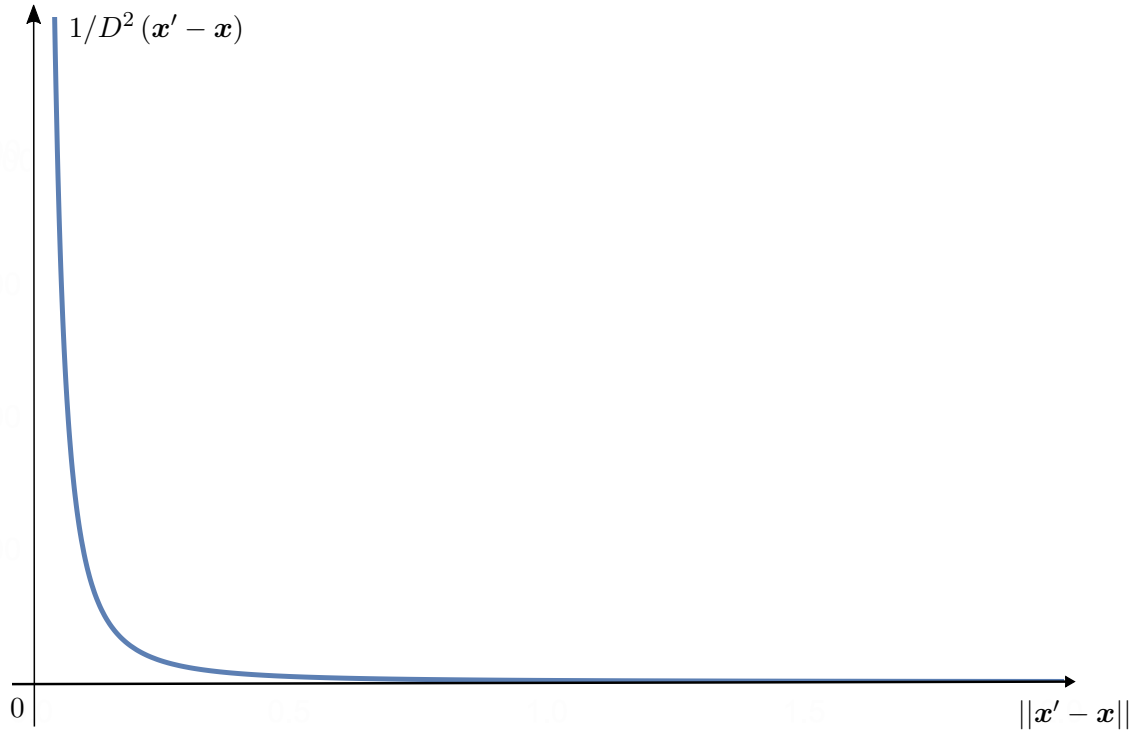
²⁰This result is a bit simpler than the one found by Atkinson and Stiglitz (1980, p.388), since I allow for a non-zero tax intercept. This implies that net social welfare weights average to one, similar to transversality condition (28).

Figure 8: The distance weight $1/D^2(\mathbf{x}' - \mathbf{x}) \equiv 1/[2\pi((x^{1'} - x^1)^2 + (x^{2'} - x^2)^2)]$ is a measure of the inverted distance of a tax base value $\mathbf{x}'$ to the value $\mathbf{x}$. My extended notion of the distributional characteristic is a normalized covariance of the average net marginal social welfare weights with the tax base components, distance-weighted by $1/D^2$.

- (a) The weights are highest if the value $\mathbf{x}'$ is close to $\mathbf{x}$. There is a singularity at $\mathbf{x}' = \mathbf{x}$. At values removed from $\mathbf{x}$ this weight decreases, steeply so at closer distances, and more gradually further away. They converge to zero at infinite distance, without ever reaching it.



- (b) The distance weights $1/D^2$ are symmetric, in the sense that they depend only on the length $\|\mathbf{x}' - \mathbf{x}\|$ of the vector $\mathbf{x}' - \mathbf{x}$.



now separable and constant, this equation is no longer specific to one value of the tax base. Moreover, there is no distance weighting in the distributional characteristic. My result uses a localized distributional characteristic for each component of the tax base, whereas the result for linear tax rates uses global distributional characteristics. Result (40) is a direct extension of (41), from linear, separable tax rates to a non-separable, nonlinear tax function.

4.3. Higher-Dimensional Case

The derivations for the higher-dimensional case are entirely analogous to those in the two-dimensional case. Again assuming that the tax base space coincides with the real vector space, $\mathcal{X} = \mathbb{R}^L$, introduce the L -dimensional distance function $D^L(\mathbf{v}) \equiv LV^L(\|\mathbf{v}\|)$, with $V^L(r)$ the volume of an L -dimensional sphere with radius r . The optimum can then be formulated as in the following theorem.

Theorem 2. *The tax optimum with an L -dimensional tax base with multidimensional heterogeneity of the agents, in absence of bunching and when the tax base space coincides with the real vector space, $\mathcal{X} = \mathbb{R}^L$, complies to the following necessary condition:*

$$\forall \mathbf{x}, l : \left[\sum_j \left(T_j \overline{x_{T_l}^j} \right) f^{\mathcal{X}} \right] (\mathbf{x}) = \text{cov} \left(\overline{\alpha}(\mathbf{x}'), \frac{x'^l - x^l}{D^L(\mathbf{x}' - \mathbf{x})} \right), \quad (42)$$

with transversality condition:

$$\int_{\mathbb{R}^L} \overline{\alpha}(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) d\mathbf{x} = 1. \quad (43)$$

Rewriting condition (42), it again follows that the proportionate reduction of aggregate demand along compensated demand curves should equal a distance-weighted distributional characteristic, as stated in the following corollary.

Corollary 2. *If we assume that we can use Slutsky symmetry $x_{T_l}^j = x_{T_j}^l$ and that the tax base space coincides with the L -dimensional real vector space, $\mathcal{X} = \mathbb{R}^L$, then in absence of bunching the tax optimum complies to the following necessary condition:*

$$\forall \mathbf{x}, l : \frac{\sum_j T_j(\mathbf{x}) X_{T_j}^l(\mathbf{x})}{X^l(\mathbf{x})} = \frac{\text{cov} \left(\overline{\alpha}(\mathbf{x}'), \frac{x'^l - x^l}{D^L(\mathbf{x}' - \mathbf{x})} \right)}{\mathcal{A} X^l(\mathbf{x})}. \quad (44)$$

The interpretation of these results is entirely analogous to the two-dimensional case. It is a sufficient statistics formulation that can be applied to a wide range of problems, e.g. the joint progressive taxation of couples, the mixed nonlinear taxation of non-transferable commodities, the optimal mixed progressive taxation of labour income and capital income, and so on. The necessary sufficient statistics can in principle be estimated empirically.

A disadvantage of formulation (44) is that it obfuscates the underlying economics. It is not clear how the Atkinson-Stiglitz result, which states that under a number of conditions

commodities should be taxed at uniform rates, could be derived from this equation, nor how it would follow that heterogeneous discount rates could lead to non-zero optimal taxes on capital income. Moreover, if individuals would not report their true incomes, because they evade taxes or because they shift income between different bases, the form of the optimal-tax equation in terms of declared incomes would not change at all. To study such questions, we need to delve deeper into the micro-foundations of the model. This will be the topic of the next section. But first I will apply the results of this section to the optimal taxation of couples.

4.4. Application to the Taxation of Couples

To illustrate my result, let us consider the example of the optimal taxation of couples. Contrary to the example in subsection 3.4, I no longer impose boundaries on the tax base space. Taxable incomes can become negative, and there is no conceptual difference between the primary and the secondary earner in the tax code. In order to guarantee the anonymity of the members of the couple, they are randomly assigned to the first or the second component of the tax base. The individuals earn gross incomes x^1 and x^2 , summarized in the vector $\mathbf{x}$. This procedure imposes symmetry on the welfare weights and sufficient statistics, and thus on the tax function.²¹

Introduce the following aggregate tax elasticities:

$$\forall l, j = 1, 2 : \bar{\mathbf{e}}_j^l(\mathbf{x}) \equiv - (1 - T_j(\mathbf{x})) \frac{\overline{x_{T_j}^l}(\mathbf{x})}{x^l}. \quad (45)$$

Applying condition (44) then tells us that the following conditions are necessary for the optimum:

$$\forall \mathbf{x} : \begin{cases} \frac{T_1(\mathbf{x})}{1-T_1(\mathbf{x})} \bar{\mathbf{e}}_1^1(\mathbf{x}) + \frac{T_2(\mathbf{x})}{1-T_2(\mathbf{x})} \bar{\mathbf{e}}_2^1(\mathbf{x}) &= \frac{1}{f^{\mathbf{x}}(\mathbf{x})} \frac{\text{cov}\left(\bar{\alpha}(\mathbf{x}'), \frac{x'^1 - x^1}{D^L(\mathbf{x}' - \mathbf{x})}\right)}{\mathcal{A}x^1}, \\ \frac{T_1(\mathbf{x})}{1-T_1(\mathbf{x})} \bar{\mathbf{e}}_1^2(\mathbf{x}) + \frac{T_2(\mathbf{x})}{1-T_2(\mathbf{x})} \bar{\mathbf{e}}_2^2(\mathbf{x}) &= \frac{1}{f^{\mathbf{x}}(\mathbf{x})} \frac{\text{cov}\left(\bar{\alpha}(\mathbf{x}'), \frac{x'^2 - x^2}{D^L(\mathbf{x}' - \mathbf{x})}\right)}{\mathcal{A}x^2}. \end{cases} \quad (46)$$

The left-hand sides contain the relative reductions of the tax bases. They are a measure of the distortions that are caused by the tax system. These quantities can in principle be measured empirically. The same is valid for the density function $f^{\mathbf{x}}$.

The right-hand side contains the localized distributional characteristics of the respective tax bases. Their values depend on the normative judgments of the government. The government could for example favour specialization within the couple, where one of both partners spends more time at home (e.g. caring for family members) while the other supplies labour in the market. Alternatively, the government could favour equal sharing of both market and other activities, or it could not care at all about the distribution

²¹Note that for the range $x^1 < x^2$, this setup must produce the exact same optimal tax function as the setup where distinction is made between the primary and secondary earners, with $x^P \geq x^S$. This observation will play a central role in subsection 6.1, where I apply the *method of images* to deal with restrictions to the tax base space.

within the family. Possibly the government would want to equalize incomes, taking into account scale effects from the size of the family. These important questions are beyond the scope of this paper.²²

5. Connection to Economic Fundamentals

In order to find more insight into optimal-tax equation (42), we need to restate our results in terms of the economic fundamentals of the problem. In the present section I do this for the case where the dimension of the type space is equal to the dimension of tax base space, $K = L$. I assume that the allocation $\mathbf{x}(\boldsymbol{\theta}, T)$ is a continuous, one-to-one function, strictly monotonous in all characteristics θ^k . I also assume that the tax function is smooth. I will extend these assumptions in section 6 about robustness.

I start in subsection 5.1 by restating the Euler-Lagrange equation in the type space, and I formulate the corresponding optimality conditions. Next I account for preference optimization in subsection 5.2 to find a characterization of the optimum in terms of economic fundamentals. I apply these findings to the joint taxation of couples in subsection 5.3. In subsection 5.4 I extend this characterization, allowing for a type-dependent individual budget constraint, accounting e.g. for differences in bequests received by different individuals. I demonstrate these results for the optimal mixed taxation of capital income and labour income in subsection 5.5.

5.1. Euler-Lagrange Formalism in the Type Space

A first step to reformulate the tax optimum in terms of economic fundamentals is to reformulate Euler-Lagrange equation (12) as a partial differential equation in the type space. Introduce the Jacobian matrix containing the partial derivatives of the allocation:

$$\mathcal{J} \equiv \begin{pmatrix} x_{\theta^1}^1(\boldsymbol{\theta}, T) & \cdots & x_{\theta^K}^1(\boldsymbol{\theta}, T) \\ \vdots & \ddots & \vdots \\ x_{\theta^1}^L(\boldsymbol{\theta}, T) & \cdots & x_{\theta^K}^L(\boldsymbol{\theta}, T) \end{pmatrix}. \quad (47)$$

This matrix extends the notion of a gradient to vector-valued functions. The following theorem then shows how the Euler-Lagrange equation for the tax optimum can be reformulated.

²²An important question, which I have been unable to solve, is what happens to the right-hand sides of these equations as the value of a tax base component converges to infinity. Denote for example the limit of $\bar{\alpha}(x^1, x^2)$ for $x^1 \rightarrow \infty$ as $\bar{\alpha}^{1+}(x^2)$. The right-hand side of the optimal-tax equation for the l -th component then becomes:

$$-\bar{\alpha}^{1+}(x^2) \frac{\int_{\mathbb{R}^2} \frac{x'^l - x^l}{D^L(\mathbf{x}' - \mathbf{x})} d(1 - F^{\mathcal{X}}(\mathbf{x}'))}{x^l f^{\mathcal{X}}(\mathbf{x})}.$$

The fraction in this equation appears to be a localized, multi-dimensional extension of the Pareto-term $[1 - F(x)]/xf(x)$ which occurs in the optimality condition for the one-dimensional problem (note that the divergence of $\int_{\mathbb{R}^2} [(\mathbf{x}' - \mathbf{x})/D^L(\mathbf{x}' - \mathbf{x})] d(1 - F^{\mathcal{X}}(\mathbf{x}'))$ equals $-f^{\mathcal{X}}(\mathbf{x})$, just like the derivative of $1 - F(x)$ equals $-f(x)$). The precise interpretation of this term, and under which circumstances it converges to a constant value, as of yet remains unclear.

Theorem 3. Suppose the type space and the tax base space have equal dimensions, respectively $K = L$, and suppose that the mapping between both spaces is one-to-one. The tax optimum, in absence of bunching, then complies to the following partial differential equation, referred to as the Euler-Lagrange condition in the type space:

$$\begin{aligned} \forall \boldsymbol{\theta} \in \Theta : (1 - \alpha(\mathbf{x}(\boldsymbol{\theta}, T))) f^\Theta(\boldsymbol{\theta}) \\ = \sum_{k=1}^K \frac{\partial}{\partial \theta^k} \left\{ \left[\nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla_{\mathbf{x}} T} \cdot (\mathcal{J}^{-1})^\top \right]^k (\mathbf{x}(\boldsymbol{\theta}, T)) f^\Theta(\boldsymbol{\theta}) \right\}, \end{aligned} \quad (48)$$

subject to the boundary conditions:

$$\forall \boldsymbol{\theta} \in \Gamma(\Theta) : \left\{ \left[\nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla_{\mathbf{x}} T} \cdot (\mathcal{J}^{-1})^\top \right] (\mathbf{x}(\boldsymbol{\theta}, T)) \cdot \hat{\boldsymbol{\theta}} \right\} f^\Theta(\boldsymbol{\theta}) = 0, \quad (49)$$

and the government budget constraint:

$$\int_{\mathbb{R}^K} T(\mathbf{x}(\boldsymbol{\theta}, T)) f^\Theta(\boldsymbol{\theta}) d\boldsymbol{\theta} = 0. \quad (50)$$

Proof. See appendix C. □

The derivation of these equations is not so straightforward. The intuition though is clear. Since the instrument of our optimization is defined in the tax base space rather than the type space, we can not simply state a traditional Euler-Lagrange equation in the type space. An additional term, the transpose of the inverse of the Jacobian matrix, needs to be added to account for the fact that we are working in a different vector space. The intuition of equation (48) remains the same as before: the social welfare effect of a tax increase for all individuals of type $\boldsymbol{\theta}$, captured by the left-hand side, should be exactly compensated by the compensated effects experienced by the individuals in the immediate environment of type $\boldsymbol{\theta}$, captured by the right-hand side.

The Euler-Lagrange equation (48) in the type space can be seen as a partial differential equation of the same form as Euler-Lagrange equation (12) in the tax base space:

$$\forall \boldsymbol{\theta} : A'(\boldsymbol{\theta}) = \sum_{k=1}^K \frac{\partial B'^k(\boldsymbol{\theta})}{\partial \theta^k}.$$

Assume now that the population density $f^{\mathcal{X}}(\mathbf{x})$ is strictly positive for any $\mathbf{x} \in \mathbb{R}^L$, so the tax base space coincides with the real vector space: $\mathcal{X} = \mathbb{R}^L$. This means for example that I do not exclude negative values of the tax bases. Using the same methods as in section 4, we then find the tax optimum in the K -dimensional type space:

$$\forall \boldsymbol{\theta}, k : \left[\nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla_{\mathbf{x}} T} \cdot (\mathcal{J}^{-1})^\top \right]^k (\mathbf{x}(\boldsymbol{\theta}, T)) = \frac{\text{cov} \left(\alpha(\mathbf{x}(\boldsymbol{\theta}', T)), \frac{\theta'^k - \theta^k}{D^K(\boldsymbol{\theta}' - \boldsymbol{\theta})} \right)}{\mathcal{A} \theta^k f^\Theta(\boldsymbol{\theta})}. \quad (51)$$

This equation again reflects the balancing exercise between efficiency and equity. The right-hand side captures the equity aspects of the government's balancing exercise.

Rather than being a distributional characteristic of a component of the tax base though, it is now a distributional characteristic for the k -th component of the type of the individuals. It is an indication of how much the government cares about that particular individual trait. Perhaps it values more highly individuals with a lower ability, individuals who are more willing to work, or those who receive a lower inheritance. Again this distributional characteristic is localized around the type θ under consideration, weighted by distance function $1/D^K$, with a similar interpretation to that formulated in section 4.

The term $\mathbf{x}_{\nabla_{\mathbf{x}}}(\mathbf{x}(\theta, T))$ on the left-hand side is an indication of how strong behavioural responses are when there is a change to the gradient of the tax function. The inverse $\mathcal{J}^{-1}$ of the Jacobian matrix indicates how the types θ are related to the tax base values $\mathbf{x}$. If we slightly change the tax base value $\mathbf{x}$ that we are studying, then there is a change in the underlying type θ that is choosing the tax base value under consideration, determined by the elements $(x_{\theta^k}^l)^{-1}$ of the inverted Jacobian matrix.²³ The product $\mathbf{x}_{\nabla_{\mathbf{x}}}(\mathbf{x}(\theta, T)) \cdot (\mathcal{J}^{-1})^T$ combines these two terms to give information about how strong is the change in the type choosing an allocation $\mathbf{x}$, if the tax gradient at that point changes:

$$\mathbf{x}_{\nabla_T}(\theta, T) \cdot (\mathcal{J}^{-1})^T = - \left. \frac{d\theta}{d\nabla_{\mathbf{x}}T(\mathbf{x}(\theta, T))} \right|_{\mathbf{x}}^T.$$

Formally this property follows from the *analytic implicit function theorem*, which is a multidimensional extension of the implicit function theorem in one dimension.

Substituting this relation into Euler-Lagrange equation (51) in the type space leads to an interesting reformulation of our optimal tax condition:

$$\forall \theta, k : - \sum_l \frac{T_l(\mathbf{x}(\theta, T))}{\theta^k f^\Theta(\theta)} \cdot \left. \frac{d\theta^k}{dT_l(\mathbf{x}(\theta, T))} \right|_{\mathbf{x}} f^\Theta(\theta) = \frac{\text{cov}\left(\alpha(\mathbf{x}(\theta', T)), \frac{\theta'^k - \theta^k}{D^K(\theta' - \theta)}\right)}{\mathcal{A}\theta^k f^\Theta(\theta)}. \quad (52)$$

The left-hand side contains the proportional change in the k -th component of the type residing at the value of the tax base under consideration. The term $d\theta^k/dT_l|_{\mathbf{x}}$ tells us how the k -th component of the type of the individuals choosing tax base $\mathbf{x}$ changes when there is a reform of the l -th component of the tax gradient. The larger is this term, the less able will be the government to target this trait using this component of the tax base. The left-hand side in its entirety is an indication of the cost that is associated to revealing information about that characteristic of the individual.

Optimal-tax condition (52) then shows how the government faces a balancing exercise between the social advantage of redistribution between individuals with different values of the trait θ^k , against the cost of targeting that trait.

Even though equation (52) has an intuitive interpretation, and even though it involves information about the types that cannot be observed by the government, it is still formulated in terms of summary statistics, which depend on the underlying behavioural

²³The *inverse function theorem* states that the elements of the inverse of the Jacobian matrix are the inverses of the elements of the original Jacobian matrix.

model. Indeed, the results that I have formulated until now are stated independently of what drives individual behaviour.²⁴ We gain more insight in the following subsection by finding an explicit formulation for the term $\left. \mathrm{d}\theta^k / \mathrm{d}T_l \right|_x$ in terms of more fundamental economic quantities.

5.2. Accounting for Preference Optimization

I will now become more explicit about the model driving individual behaviour. Assuming that individuals maximize a utility function, it becomes possible to find explicit formulations for the behavioural effects $\mathbf{x}_{\nabla T}(\boldsymbol{\theta}, T)$ and for the Jacobian $\mathcal{J}$.

Introduce a numéraire commodity c , and denote the individual budget constraint as follows:

$$c = B(\mathbf{x}|T), \quad (53)$$

where the function $B(\mathbf{x}|T)$ depends on the tax system. A common example would be the case where individuals earn gross income z , and they consume out of their net-of-tax earnings. In this case the budget constraint would be $c = z - T(z)$.

I assume that individuals maximize a utility function $u(c, \mathbf{x}|\boldsymbol{\theta})$ subject to budget constraint (53). Substituting the latter constraint for the numéraire commodity, introduce the constrained utility function:

$$U(\mathbf{x}|\boldsymbol{\theta}, T) \equiv u(B(\mathbf{x}|T), \mathbf{x}|\boldsymbol{\theta}). \quad (54)$$

The individual first-order conditions can then be formulated as follows:

$$\forall l : U_l = u_c B_l + u_l = 0, \quad (55)$$

where the subscripts denote partial derivatives with respect to the respective arguments, e.g. $U_l \equiv \partial U / \partial x^l$. The individuals thus choose the values of the tax base such that small variations of their choice do not affect their utility.

Suppose now that there is a small change in the marginal tax rate T_j at base $\mathbf{x}$. Individuals will update their behaviour such that condition (55) remains fulfilled. Similarly, if an individual experiences an exogenous change in his trait θ^k , he will update his behaviour. This observation leads to the following condition:

$$\forall l, \forall \nu = T_j, \theta^k : \frac{\mathrm{d}U_l}{\mathrm{d}\nu} = 0. \quad (56)$$

The total derivative in this equation can be expanded as follows:

$$\forall l, \forall \nu = T_j, \theta^k : \sum_{j=1}^L U_{lj} \frac{\mathrm{d}x^j}{\mathrm{d}\nu} + U_{l\nu} = 0. \quad (57)$$

²⁴One exception is the reformulation of result (39) in the form (40), and more generally from theorem (2) to corollary (2). I use Slutsky symmetry to arrive at the latter formulation.

Define the Hessian H of U as the matrix containing its second-order partial derivatives:

$$H \equiv \begin{bmatrix} U_{11} & \cdots & U_{1L} \\ \vdots & \ddots & \vdots \\ U_{L1} & \cdots & U_{LL} \end{bmatrix}.$$

This allows rewriting result (57) in matrix notation:

$$\forall \nu = T_j, \theta^k : \frac{d\mathbf{x}}{d\nu} = -H^{-1} \cdot U_{\mathbf{x}\nu},$$

where I multiply both sides from the left by the inverse of the Hessian.

It now follows that the Jacobian matrix is given by:

$$\mathcal{J} = -H^{-1} \cdot U_{\mathbf{x}\theta}, \quad (58)$$

and the matrix with compensated effects looks as follows:

$$\forall l : \mathbf{x}_{T_l} = -H^{-1} \cdot U_{\mathbf{x}T_l}. \quad (59)$$

Before applying these results, I will give more explicit formulations for the terms $U_{\mathbf{x}T_l}$ and $U_{\mathbf{x}\theta}$. First assume that there is a reform to the marginal tax rate of size dT_j at tax base value $\mathbf{x}^R$, keeping constant the liability at $\mathbf{x}^R$ itself. Denoting by $\mathbf{x}$ the value of the tax base chosen after the reform by someone who originally chooses $\mathbf{x}^R$, we obtain a new budget constraint:

$$c = B(\mathbf{x}|T) - (x^j - x^{R,j}) dT_j.$$

It is clear that when the individual does not change his behaviour, $x^j = x^{R,j}$, his tax liability remains unaltered. The change in the marginal tax rate applies only to his responses to the reform. Note that taking the derivative of this budget constraint, for given value of the tax base, yields the following relation:

$$\left. \frac{dc}{dT_j} \right|_{\mathbf{x}} = - (x^j - x^{R,j}).$$

Using definition (54) we now find:

$$U_{lT_j} = \frac{\partial}{\partial x^l} \left(u_c \left. \frac{dc}{dT_j} \right|_{\mathbf{x}} \right) = - \frac{\partial}{\partial x^l} \left(u_c (x^j - x^{R,j}) \right). \quad (60)$$

Evaluating this expression in the situation with no tax reform, when $x^j = x^{R,j}$, it follows:

$$U_{lT_j} = \begin{cases} -u_c & \text{if } l = j, \\ 0 & \text{if } l \neq j. \end{cases} \quad (61)$$

Similarly, again using definition (54):

$$U_{l\theta^k} = u_{l\theta^k} + u_{c\theta^k} B_l = u_{l\theta^k} - u_{c\theta^k} \frac{u_l}{u_c} = u_c \frac{\partial (u_l/u_c)}{\partial \theta^k} \Big|_x, \quad (62)$$

where I substitute individual first-order condition (55) in the second step.

Substituting results (58), (59), (61), and (62) into optimal-tax condition (51) then yields the following necessary condition for the optimum:

$$\begin{aligned} \forall \theta, l : T_l(\mathbf{x}(\theta, T)) \\ = - \frac{1}{f^\Theta(\theta)} \sum_k \frac{\partial (u_l/u_c)}{\partial \theta^k} \Big|_x \text{cov} \left(\alpha(\mathbf{x}(\theta', T)), \frac{\theta'^k - \theta^k}{D^K(\theta' - \theta)} \right). \end{aligned} \quad (63)$$

This equation is a direct extension of a result obtained by Mirrlees (1976) for the case with a one-dimensional type space.²⁵ The l -th component of the optimal tax gradient, T_l , is higher for a particular type θ if there are fewer individuals at that point whose behaviour is impacted, or if the government particularly values compensation for traits that can be better targeted using that component of the tax base.

To gain more intuition about the term $\partial (u_l/u_c) / \partial \theta^k \Big|_x$, consider for example the following Atkinson-Stiglitz (1976) type of model with one-dimensional type n . Suppose individuals with ability n supply labour ℓ , which yields earn gross labour income $z = n\ell$. They consume a numéraire good c and a taxed good y . They maximize utility $u(c, y, \frac{z}{n})$, subject to budget constraint $c = z - p \cdot y - T(z, y)$, with p the producer price of good y . We are now interested in the following term:

$$\frac{\partial (u_y/u_c)}{\partial n} \Big|_{z,y} = - \frac{\ell}{n} \frac{1}{u_c} [u_{y\ell} - u_{c\ell} (p + T_y)],$$

where I substitute the first-order condition $u_y = (p + T_y) u_c$ on the right-hand side. It immediately shows that when preferences are separable between leisure and consumption, such that $u_{y\ell} = u_{c\ell} = 0$, this expression becomes zero. Conditional on the tax base (z, y) , the marginal rate of substitution u_y/u_c does not reveal any information about the individual type, and there is no point in taxing good y . Doing so would introduce a distortion, without introducing the counterbalancing advantage of being able to differentiate between individuals with different abilities. If, on the other hand, utility is not separable, then equation (63) tells us that the marginal tax rate on y will be higher when the term $\partial (u_y/u_c) / \partial n \Big|_{z,y}$ is larger, as suggested by Corlett and Hague (1953). Optimal-tax condition (63) shows how this reasoning extends to multi-dimensional characteristics.

5.3. Application to Taxation of Couples

I will now apply the findings of the previous subsection to the optimal taxation of couples, following a model introduced by Kleven, Kreiner, and Saez (2007). Let each

²⁵The difference with the formulation by Renes and Zoutman (2016a) stems from the fact that the net social welfare weight α already incorporates the income effects. This makes that formulation (63) has a more intuitive interpretation.

couple consist of two individuals with respective labour abilities n^1 and n^2 , with domain $[\underline{n}^1, \bar{n}^1] \times [\underline{n}^2, \bar{n}^2]$. As before, labour supply of the individuals is denoted by ℓ^1 and ℓ^2 , and the corresponding gross labour incomes are z^1 and z^2 . The budget constraint for the household is given as follows:

$$c = z^1 + z^2 - T(z^1, z^2).$$

Utility is separable, and linear in consumption:

$$u\left(c, \frac{z^1}{n^1}, \frac{z^2}{n^2}\right) \equiv c - n^1 h^1\left(\frac{z^1}{n^1}\right) - n^2 h^2\left(\frac{z^2}{n^2}\right).$$

Define the following elasticities:

$$\varepsilon^k \equiv \frac{1 - T_k}{z^k} \frac{\partial z^k}{\partial (1 - T_k)} = \frac{h_{\ell^k}^k}{\ell^k h_{\ell^k \ell^k}^k}.$$

The effect of the respective abilities on the marginal rates of substitution then look as follows:

$$\begin{aligned} \frac{\partial (u_{z^1}/u_c)}{\partial n^2} \Big|_z &= \frac{\partial (u_{z^2}/u_c)}{\partial n^1} \Big|_z = 0, \\ \frac{\partial (u_{z^1}/u_c)}{\partial n^1} \Big|_z &= \frac{1}{\varepsilon^1}, \quad \frac{\partial (u_{z^2}/u_c)}{\partial n^2} \Big|_z = \frac{1}{\varepsilon^2}. \end{aligned}$$

Substituting these results into equations (58), (59), (61), and (62), and applying theorem 3, yields the following divergence equation for the tax optimum:

$$\begin{aligned} \forall n^1, n^2 : (\alpha - 1) f^\Theta(n^1, n^2) &= \frac{\partial}{\partial n^1} \left(\frac{T_1}{1 - T_1} \frac{h_{\ell^1}^1}{\ell^1 h_{\ell^1 \ell^1}^1} n^1 f^\Theta(n^1, n^2) \right) \\ &+ \frac{\partial}{\partial n^2} \left(\frac{T_2}{1 - T_2} \frac{h_{\ell^2}^2}{\ell^2 h_{\ell^2 \ell^2}^2} n^2 f^\Theta(n^1, n^2) \right). \end{aligned} \quad (64)$$

This is exactly the optimality condition found by Kleven, Kreiner, and Saez (2007) in their proposition 4. Using the methods in the present paper, we can now state the optimum for the case where the domain of the skills coincides with the real vector space $\mathbb{R}^2$:

$$\forall k : \frac{T_k}{1 - T_k} = \frac{1}{\varepsilon^k} \text{cov} \left(\alpha', \frac{n'^k - n^k}{D^2(n'^1 - n^1, n'^2 - n^2)} \right) \frac{1}{n^k f^\Theta(n^1, n^2)}. \quad (65)$$

This solution, to the best of my knowledge, had not been stated before in the literature.

Furthermore, apply theorem 3 to find the following boundary conditions:

$$\forall k : \forall n^k = \underline{n}^k, \bar{n}^k : \left\{ \frac{T_k}{1 - T_k} \varepsilon^k n^k f^\Theta(n^1, n^2) \right\} = 0.$$

This again confirms the findings stated by Kleven, Kreiner, and Saez (2007) in their proposition 4.

5.4. Type-dependent Budget Constraint

Much of the optimal-tax literature, including the contributions of Mirrlees (1971), Mirrlees (1976) and Atkinson and Stiglitz (1976), assumes that the budget constraint does not depend directly on the individual traits. Still it is interesting to allow for this possibility. It is possible for example that individuals receive an unobserved inheritance, as studied by Cremer, Pestieau, and Rochet (2001), or that they receive a skill-dependent return on their investment, as described by Gerritsen et al. (2015).

Assume thus that the individual budget constraint can be expressed as follows:

$$c = B(\mathbf{x}|\boldsymbol{\theta}, T).$$

For example, suppose that individuals have gross labour income z , and they receive an unobserved bequest b . In this case the budget constraint would be $c = z - T(z) + b$.

The derivations from subsection 5.2 remain valid. The value (62) of second-order partial derivative $U_{l\theta^k}$ changes though:

$$U_{l\theta^k} = u_c \frac{\partial}{\partial \theta^k} \left[\frac{u_l}{u_c} (B(\mathbf{x}|\boldsymbol{\theta}, T), \mathbf{x}|\boldsymbol{\theta}) + B_l \right] \Big|_{\mathbf{x}}.$$

Note that in the argument of the marginal rate of substitution u_l/u_c , I substitute the budget constraint for the numéraire. This formulation takes into account the effects of the changes in the type on the budget constraint. The term within square brackets indicates how strongly the individual marginal rate of substitution would deviate from its optimum if there were an exogenous change in his trait θ^k , if the individual would not alter his behaviour.

Optimal-tax condition (51) then becomes:

$$\begin{aligned} \forall \boldsymbol{\theta}, l : T_l(\mathbf{x}(\boldsymbol{\theta}, T)) = & -\frac{1}{f^\Theta(\boldsymbol{\theta})} \sum_k \frac{\partial}{\partial \theta^k} \left[\frac{u_l}{u_c} (B(\mathbf{x}|\boldsymbol{\theta}, T), \mathbf{x}|\boldsymbol{\theta}) + B_l \right] \Big|_{\mathbf{x}} \\ & \cdot \text{cov} \left(\alpha(\mathbf{x}(\boldsymbol{\theta}', T)), \frac{\theta'^k - \theta^k}{D^K(\boldsymbol{\theta}' - \boldsymbol{\theta})} \right). \end{aligned} \quad (66)$$

The intuition is similar to that in subsection 5.2. The l -th component of the tax gradient is higher for a particular type $\boldsymbol{\theta}$ if there are less individuals whose behaviour is affected. It is also higher if the government cares more about traits which have a stronger influence on the marginal rate of substitution in the direction of that tax base component – now taking into account the effect of a change in the type on the individual budget constraint. If a small change in an individual trait causes a strong deviation in a particular direction of the tax base, then the government can use that tax base component more easily to target that particular characteristic.

5.5. Application to Mixed Taxation of Labour and Capital Incomes

I will apply the findings from the previous subsection to the optimal mixed taxation of capital income and labour income. I will describe a model where individuals differ in two

dimensions: they have different labour productivities n , and they have different discount rates ρ . This distinction was also made by Diamond and Spinnewijn (2011), who find desirability conditions in a model with just two or three values of each characteristic. I extend their model here to continuous characteristics. Moreover, to be able to demonstrate my findings for a type-dependent budget constraint, I assume that more patient individuals obtain higher returns on their savings, and that more able individuals obtain higher returns.

There are two periods in the model. In the first period, individuals supply labour ℓ , which yields a gross income $z \equiv n\ell$, they consume an amount c^1 , and they save an amount a . Their savings yield a gross capital income $y \equiv aR(n, \rho)$, which depends on the amount saved and on the discount rate and the ability of the individual.

In order to keep the model tractable, I assume that individuals pay a tax $T(z, y)$ at the end of the first period. It depends on their labour income and their capital income. Their first-period budget constraint thus looks as follows:

$$c^1 + a + T(z, y) = z.$$

Second-period consumption then equals the amount saved, plus the capital income obtained from it:

$$c^2 = (1 + R(n, \rho)) a.$$

To be able to apply the findings of the previous subsection, I write the intertemporal budget constraint in terms of the tax bases:

$$c^1 = B(z, y|n, \rho, T) \equiv z - T(z, y) - \frac{y}{R(n, \rho)}.$$

Individuals obtain utility from their first-period and their second-period consumption, and they experience a disutility from their labour supply. I immediately write the utility function in terms of the numéraire commodity c^1 and the values of the tax base:

$$u(c^1, z, y|\rho, n) \equiv v(c^1) + \rho v\left(\left[\frac{1}{R(n, \rho)} + 1\right]y\right) - w\left(\frac{z}{n}\right).$$

Preferences are thus intertemporally separable, and they are separable between leisure and consumption. This allows deriving the first-order conditions:

$$0 = \frac{u_z}{u_{c^1}}(B(z, y|n, \rho, T), z, y|n, \rho) = (1 - T') - \frac{1}{n} \frac{w_\ell}{v_{c^1}}, \quad (67)$$

$$0 = \frac{u_y}{u_{c^1}}(B(z, y|n, \rho, T), z, y|n, \rho) = -\left(T_y + \frac{1}{R(n, \rho)}\right) + \rho \frac{v_{c^2}}{v_{c^1}} \left(1 + \frac{1}{R(n, \rho)}\right). \quad (68)$$

The relevant derivatives of the budget constraint are as follows:

$$B_{zn} = B_{z\rho} = 0, \quad B_{yn} = \frac{R_n}{R^2}, \quad B_{y\rho} = \frac{R_\rho}{R^2}. \quad (69)$$

and the relevant derivatives of the marginal rates of substitution:

$$\left. \frac{\partial}{\partial \rho} \left(\frac{u_z}{u_{c^1}} \right) \right|_{z,y} = 0, \quad (70)$$

$$\left. \frac{\partial}{\partial n} \left(\frac{u_z}{u_{c^1}} \right) \right|_{z,y} = \frac{1}{n} \frac{w_\ell}{n v_{c^1}} \left(1 + \frac{w_{\ell\ell} \ell}{w_\ell} \right), \quad (71)$$

$$\left. \frac{\partial}{\partial \rho} \left(\frac{u_y}{u_{c^1}} \right) \right|_{z,y} = \frac{R_\rho}{R^2} + \frac{v_{c^2}}{v_{c^1}} \left(1 + \frac{1}{R(\rho)} \right) - \rho \frac{v_{c^2}}{v_{c^1}} \frac{R_\rho}{R^2} \left[1 + \left(1 + \frac{1}{R(n, \rho)} \right) \frac{v_{c^2} c^2}{v_{c^2}} \right], \quad (72)$$

$$\left. \frac{\partial}{\partial n} \left(\frac{u_y}{u_{c^1}} \right) \right|_{z,y} = \frac{R_n}{R^2} - \rho \frac{v_{c^2}}{v_{c^1}} \frac{R_n}{R^2} \left[1 + \left(1 + \frac{1}{R(n, \rho)} \right) \frac{v_{c^2} c^2}{v_{c^2}} \right]. \quad (73)$$

We now have the necessary elements to derive the optimal-tax equations. First apply optimal-tax condition (66) to find the optimal marginal tax on labour income:

$$\forall \theta : \frac{T_z}{1 - T_z} = \frac{1}{n f^\Theta(n, \rho)} \left(1 + \frac{w_{\ell\ell} \ell}{w_\ell} \right) \text{cov} \left(\alpha(n', \rho'), \frac{n' - n}{D^2(n' - n, \rho' - \rho)} \right),$$

where I used first-order condition (67) to rearrange some terms. The first two terms on the right-hand side of this equation look a bit similar to the result found by Mirrlees (1971). The first term incorporates properties of the skill distribution, and the second term is the inverse of the Frisch elasticity of labour supply (controlling for capital income). The third term though looks completely different in this case. Rather than looking at the net social welfare weights of all individuals with higher skills, and weighing them equally, as in the B -term in equation (8), it now incorporates a distance-weighted covariance of the net social welfare weights with the tax base. The distance-weighting is symmetric, and also tax base levels *below* the tax base under consideration contribute to the optimal tax on labour income.

For the optimal tax on capital income we obtain the following equation:

$$\begin{aligned} \forall \theta : f^\Theta T_y = & - \left[\left. \frac{\partial}{\partial \rho} \left(\frac{u_y}{u_{c^1}} \right) \right|_{z,y} + B_{y\rho} \right] \text{cov} \left(\alpha(n', \rho'), \frac{\rho' - \rho}{D^2(n' - n, \rho' - \rho)} \right) \\ & - \left[\left. \frac{\partial}{\partial n} \left(\frac{u_y}{u_{c^1}} \right) \right|_{z,y} + B_{yn} \right] \text{cov} \left(\alpha(n', \rho'), \frac{n' - n}{D^2(n' - n, \rho' - \rho)} \right). \end{aligned}$$

We can learn a number of things. Note first that if the discount rates are not heterogeneous, then the covariance on the first line will be zero. Assume that the covariance on the second line is not zero, e.g. because the government would like to redistribute away from individuals with high innate abilities. The only reason then why the optimal marginal tax rate on capital income might differ from zero, in this model, is that the term within square brackets on the second line differs from zero. This is the case when an exogenous change in the labour ability n of an individual, somehow causes his savings no longer to be in optimal in his original bundle. My model is such that preferences between leisure and consumption do not depend directly on labour ability. One reason is that I assumed separability between leisure and consumption, thus excluding the motives

found by Corlett and Hague (1953). Another reason is that I exclude direct dependence of the discount rate on labour ability ($\rho \neq \rho(n)$), thus excluding the motives found by Saez (2002). Still, if the interest rate obtained by the individuals is skill-dependent ($R_n \neq 0$), this term will differ from zero because of results (69) and (73), and the optimal marginal tax on capital income will differ from zero. This complements the findings of Gerritsen et al. (2015). It is thus important not only to take into account the direct impact of the traits on individual preferences, but also to take into account the effects on the individual budget constraint.

If neither of above reasons is true, then the second line will be zero. Still there might be reasons to tax capital income. If not all individuals have the same discount rate, then by definition the marginal rate of substitution between the two periods will depend on it. The term between square brackets on the first line thus differs from zero. If the covariance between the welfare weights and the discount rates differs from zero, then the optimal marginal tax rate on capital income will differ from zero, confirming the findings of Diamond and Spinnewijn (2011). Furthermore, if the rates of return $R(n, \rho)$ depend directly on the discount rate, such that more patient individuals obtain higher returns, this further affects the value of the term between square brackets on the first line, further contributing to a non-zero optimal tax on capital income. Again it turns out to be important to incorporate the effects of the individuals traits on their budget constraints.

Note that even when the second line in this optimal-tax equation is zero, and even when the welfare weights are forced to be independent of the discount rate because individuals are held responsible for their own preferences (so $\alpha_\rho = 0$), even then the first line on the right-hand side might differ from zero, for the mere reason that the discount rates ρ are locally correlated with the labour skills n . So, even if we deem the preference for savings to be irrelevant for moral reasons, still we might end up using it to differentiate between individuals because it contains imperfect information about the innate abilities.

6. Robustness

There are a number of circumstances where the assumptions of the above derivations are violated. For example, I always assumed that the relevant functions and their derivatives are sufficiently smooth. When characterizing the optimum in section 4, I assumed that the tax base space spans the entire real vector space. And when connecting to economic fundamentals in section 5, I assumed that each allocation in the tax base space was chosen by exactly one type.

There are a number of ways in which these conditions may be violated. A first possibility is that the tax base space is restricted, for example because there are some legal or physical restrictions on the values that the tax base may take. It may be impossible to consume negative quantities of certain goods, or it may not be allowed to report negative values. When characterizing the optimum in section 4, I used the assumption that the tax base space coincides with the entire real vector space ($\mathcal{X} = \mathbb{R}^2$), so I could use Green

functions with boundary conditions for $\|\mathbf{x}\| \rightarrow \infty$. I will argue in subsection 6.1 that a restricted tax base space impels us to add some correction terms to the optimality conditions.

Another possible violation of my assumptions occurs when the type space has a higher dimension than the tax base space. In this case individuals of different types will inevitably pool at the same allocation, even when a strict monotonicity assumption is maintained for the allocation, i.e. $\partial x^l / \partial \theta^k > 0$ for all components. For example, for a given gross labour income, one might simultaneously observe talented individuals who prefer to work fewer hours, and less productive individuals who prefer to work more hours. This possibility is allowed for by the Euler-Lagrange equation in theorem 1 and the characterization of the optimum in terms of sufficient statistics in equation (44), where behavioural responses are averaged over the individuals pooling at the tax base value under consideration, but section 5 explicitly excludes this possibility when it assumes a one-to-one allocation. Subsection 6.2 explains how to deal with this.

Another situation where individuals of different types choose the same allocation, is when bunching occurs. Consider an infinitesimal volume $d\mathbf{x}$ around a tax base value $\mathbf{x}$. The number of individuals choosing a value in that volume equals $f(\mathbf{x}) d\mathbf{x}$. If we shrink this volume to a single point, $d\mathbf{x} \rightarrow 0$, then in absence of bunching the number of individuals in it will converge to zero. Since any real volume contains infinitely many points, removing a single point will not make a difference when we integrate over that volume. We say that a volume of measure zero contains a zero mass of individuals. The situation changes when there is bunching. Suppose for example that 10% of the population chooses not to work: they choose the corner solution $z = 0$. Then the number of individuals bunching on an infinitesimal interval dz around labour income $z = 0$, $f(0) dz$, will converge to 10% as $dz \rightarrow 0$. For this to be possible, there needs to be an infinite spike in the density function at $z = 0$. We say that there is a mass of individuals bunching at this point. The density function makes a discontinuous jump at the edge, and the objective function is no longer sufficiently smooth for my derivations to hold without modifications. Seade (1977) shows how in this situation the zero-marginal-tax result at the bottom disappears. I study this possibility in subsection 6.3.

Bunching may also occur at an interior point of the tax base space. Ebert (1992) gives an example for the labour income tax with a utilitarian government objective, where individual preferences and the skill distribution coalesce such that the individual second-order conditions become binding at a some range of skills. Mirrlees (1976) shows that with a one-dimensional type θ , when a single-crossing condition is fulfilled for the indifference curves, the individual second-order condition is equivalent to the weak monotonicity condition $dz/d\theta \geq 0$. If this condition is binding, so $dz/d\theta = 0$, it follows that a range of individuals of consecutive skill levels will bunch at the same labour income.

Bunching differs from pooling. In the case of bunching, the strict monotonicity condition is violated: individuals of consecutive types choose the same allocation, forming a mass point. In the case of pooling, the strict monotonicity condition remains satisfied; it usually occurs when the type space has a higher dimensionality than the tax base space. There is no infinite spike in the density function, and thus there is no mass point in the

income distribution.

Kleven, Kreiner, and Saez (2007) show for the nonlinear joint taxation of couples, that the no-bunching assumption is reasonable for a wide range of social objectives. Still there are cases where bunching is important. One example is the Rawlsian objective, where all weight falls on a single individual. Another example where bunching may occur is the case where most weight is given to a type in the interior of the type space (e.g. attaching a higher social welfare weight to the working poor than to those out of work).²⁶

An optimal allocation that includes an interior bunching range, is implemented using a kinked tax function. That is, the tax function itself remains continuous, but there is a discontinuity in its gradient. Suppose for example with a one-dimensional labour income tax base, that at a given point the marginal tax rate jumps from 40% to 45%. Individuals with a marginal rate of substitution between labour income and consumption lower than 40% ($-(1 - T') < -60\%$) will choose a labour income below the kink, while individuals with a marginal rate of substitution above 45% ($-(1 - T') > -55\%$) will choose a labour income above the kink. Those with a marginal rate of substitution between 40% and 45% will bunch at the kink. There will be a mass of individuals at that point. Again the objective function is no longer sufficiently smooth and I need to adapt my derivations. I do so in subsection 6.3.

Another difficulty occurs when the dimension of the type space is lower than the dimension of tax base space. This is typically the case in the literature that studies problems involving multidimensional tax bases with one-dimensional types. Examples include the seminal papers of Mirrlees (1976) and Atkinson and Stiglitz (1976). In this case the allocation forms a one-dimensional path in a multidimensional tax base space. A typical complication in this case, when solving the optimal-tax problem using the mechanism design approach, is that of *double deviations*. The problem is that the incentive compatibility constraint that is taken into account, guarantees only that each individual chooses the bundle that is meant for him among the bundles that are situated on that one-dimensional path. If the optimum is decentralized though using a tax system, there is no a priori reason why an individual would not prefer a bundle that is situated outside the one-dimensional allocation. Most authors, to circumvent the problem, simply set the tax liability infinitely high at these alternative values (see e.g. Rochet, 1985). Renes and Zoutman (2016b) show that it suffices that the individual second-order conditions for the decentralized optimization problem are fulfilled in the bundles that are assigned to the different types: it suffices that the tax liability increases strong enough when individuals deviate from the one-dimensional path of the allocation, to discourage them from doing so.

In my approach, where we are given a tax function with corresponding density functions and marginal behavioural responses, the problem of double deviations becomes

²⁶Armstrong (1996) and Rochet and Choné (1998) study the problem of a monopolist choosing nonlinear prices for a set of potential qualities of its product. They show that when individuals choose which qualities to buy, and they have an option not to buy the product at all, then the presence of bunching cannot be ruled out without making unreasonable assumptions. These results are less relevant for the present paper, as I rule out the existence of an outside option.

irrelevant. If we observe an income distribution corresponding to a given tax function, that means that individuals have already chosen their tax base values, conforming to their incentives. If we observe that individuals pool on a one-dimensional allocation, that automatically implies that the second-order conditions given by Renes and Zoutman (2016a) are fulfilled on that line. Still, I cannot simply ignore this case: I need to demonstrate how the multidimensional optimal-tax equation (42) collapses to the more traditional equations found e.g. by Mirrlees (1976) and Atkinson and Stiglitz (1976). This will be the topic of subsection 6.4.

6.1. Tax Base Restrictions

The optimal-tax condition in theorem 2 is based on the assumption that the tax base space coincides with the real vector space: $\mathcal{X} = \mathbb{R}^L$. As mentioned, the tax base space might be limited to a subset, e.g. for legal or physical reasons. One typical example is that one or more components of the tax base are restricted to be non-negative. Another relevant example is the joint taxation of couples, where both partners should be treated anonymously. As explained in subsection 3.4, the latter limits the tax base space to a bisection of the real space.

Theorem 2 is no longer correct when the tax base space is limited to a subset of $\mathbb{R}^L$. The reason is that it is based on a Green function of the form in equation (34). The problem is that the boundary condition (37) is no longer fulfilled when the boundary is situated at finite values of the tax base.

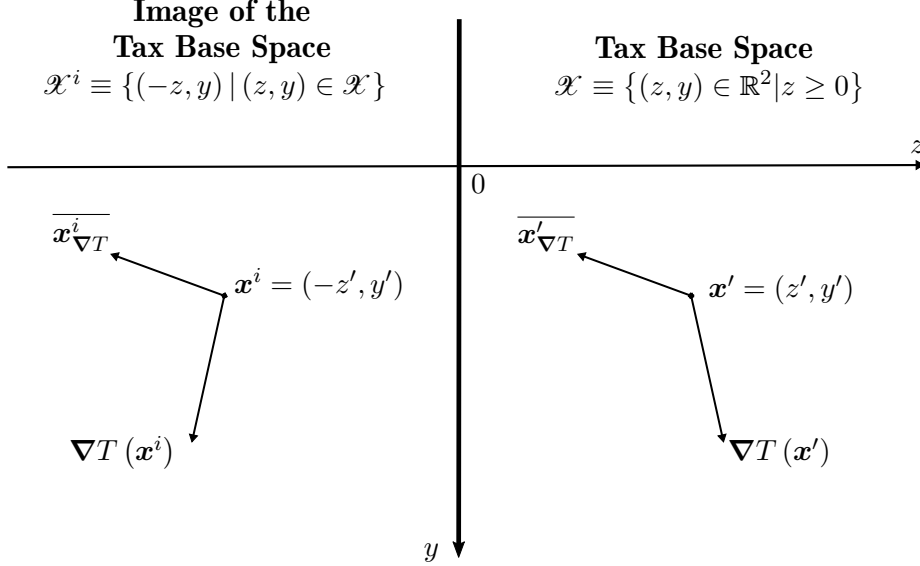
Remember how in subsection 4.4 for the joint taxation of couples, I got rid of the limitation that the income of the primary earner should be higher than the income of the secondary earner, by randomly assigning the partners in each couple to either the first or the second position in the definition of the tax base. This led to symmetric sufficient statistics, and if the welfare weights respected anonymity, it led to a tax function which was symmetric in the identity of the individuals. Using this procedure, the domain of the problem was extended to include the entire real vector space, rather than the half where the primary earner had a higher income than the secondary earner. We were thus able to solve the problem using theorem 2, based on a Green function with boundary conditions at infinite levels of the tax base.

The procedure set forth in subsection 4.4 is an application of a more general method, the *method of images*. I created an *image* of the economy under consideration where the values of z^P and z^S were switched for all individuals. This method is typically used in combination with Green functions, e.g. to solve problems in the field of electromagnetism. I will attempt to explain the solution procedure in economics terms. A more traditional treatment of the subject can be found e.g. in Lax et al. (1998, p.235-238).

I will focus on a two-dimensional tax base with values $\mathbf{x} \equiv (z, y)$, where y can take any value, but z is restricted to non-negative values ($z \geq 0$). The tax base space thus coincides with a half plane in the real vector space: $\mathcal{X} \equiv \{\mathbf{x} \equiv (z, y) \in \mathbb{R}^2 | z \geq 0\}$. Theorem 1, stating the Euler-Lagrange equation, remains valid, but we can no longer use theorem 2 to characterize its solution.

The trick to solve this problem is again to create an image of this economy. Assign

Figure 9: The tax base space $\mathcal{X}$ is limited to those points $(z, y) \in \mathbb{R}^2$ which have non-negative values of z . The fact that $\mathcal{X}$ does not coincide with $\mathbb{R}^2$ makes that we can no longer apply theorem 2. A solution is to create an “image world”, a mirror image of the true economy around the line $z = 0$. The tax base space of this image world has elements $(z^i, y^i) \in \mathcal{X}^i$, with non-positive values for z^i . Since the union $\mathcal{X} \cup \mathcal{X}^i$ coincides with $\mathbb{R}^2$, we can use theorem 2 to optimize the tax function on this extended economy. The values of this optimal-tax function on the set $\mathcal{X}$ coincide with the solution for the original problem.



all units in the population randomly to one of two groups. The first group reports their true gross income as usual, but the second group reports a negative value of z . For a member of the second group, if his tax base originally took the value $\mathbf{x}' = (z', y')$, he will now report the *image value* $\mathbf{x}^i = (-z', y')$. Now let the government optimize a new tax function on this extended domain. The new tax base space coincides with the entire real vector space $\mathbb{R}^2$, as illustrated in figure 9.

The derivations in section 3 are valid in this case. Implementing a localized reform similar to those used in subsections 3.1 and 3.2, should still leave social welfare unchanged. Theorem 1 remains valid, and the Euler-Lagrange equation looks exactly the same.²⁷ What is more, since the image economy is the exact mirror of the original economy, the optimal tax function on the extended domain should be symmetric around the line defined by the equation $z = 0$. It follows that the entity $\nabla_{\mathbf{x}} T(\mathbf{x}') \cdot \overline{\mathbf{x} \nabla_{\mathbf{x}} T}(\mathbf{x}')$ should be the exact mirror of the entity $\nabla_{\mathbf{x}} T(\mathbf{x}^i) \cdot \overline{\mathbf{x} \nabla_{\mathbf{x}} T}(\mathbf{x}^i)$.

Now consider a point $\underline{\mathbf{x}} \equiv (0, \underline{y})$ which lies on the line $z = 0$. Note that this point is necessarily equal to its image: $\underline{\mathbf{x}} = \underline{\mathbf{x}}^i$. It follows that the vector $\nabla_{\mathbf{x}} T(\underline{\mathbf{x}}) \cdot \overline{\mathbf{x} \nabla_{\mathbf{x}} T}(\underline{\mathbf{x}})$

²⁷I ignore the fact that the density function is halved, since half of the population is moved to the image world. The factor 1/2 drops out in all results anyway.

should be equal to its own image. The only way that this is possible, is that it is parallel to the line $z = 0$. It follows that in the optimum, necessarily $\nabla_{\underline{x}} T(\underline{x}) \cdot \overline{\nabla_{\underline{x}} T}(\underline{x}) \cdot \hat{\underline{x}} = 0$. This is exactly the requirement of boundary condition (13) in theorem 1.

If we can thus solve this new problem on domain $\mathbb{R}^2$, then we have a tax function which conforms to all conditions of theorem 1. We thus immediately obtain a solution for the original problem, by using the values of the optimal tax function on the positive half of $\mathbb{R}^2$.

The solution to the original optimal tax problem thus looks as follows:

$$\begin{aligned} \forall \underline{x} \in \mathcal{X}, l : \sum_j \left(T_j(\underline{x}) \overline{x_{T_l}^j} \right) f^{\mathcal{X}}(\underline{x}) \\ = \text{cov}_{\mathcal{X}} \left(\bar{\alpha}(\underline{x}'), \frac{x^l - x^l}{D^2(\underline{x}' - \underline{x})} \right) + \text{cov}_{\mathcal{X}} \left(\bar{\alpha}(\underline{x}'), \frac{x^{il} - x^l}{D^2(\underline{x}^i - \underline{x})} \right). \end{aligned} \quad (74)$$

To see more formally that this solution is correct, note that this is the solution that we would have found if I had used the following Green function in subsection 4.2:

$$G(\underline{x}, \underline{x}') \equiv \frac{\underline{x} - \underline{x}'}{D^2(\underline{x} - \underline{x}')} + \frac{\underline{x} - \underline{x}^i}{D^2(\underline{x} - \underline{x}^i)}.$$

If I can show that this Green function fulfills conditions (31) and (32), then I have formally shown the validity of solution (74).

The derivations in appendix B lead to the following property:

$$\sum_l \frac{\partial G^l(\underline{x}', \underline{x})}{\partial x^l} = \delta^2(\underline{x}' - \underline{x}) + \delta^2(\underline{x}^i - \underline{x}).$$

Since for all elements $\underline{x}$ in the tax base space $\mathcal{X}$ we necessarily have $\underline{x}^i \neq \underline{x}$, the second term on the right-hand side is zero, $\delta^2(\underline{x}^i - \underline{x}) = 0$, and we find that condition (31) is fulfilled:

$$\sum_l \frac{\partial G^l(\underline{x}', \underline{x})}{\partial x^l} = \delta^2(\underline{x}' - \underline{x}).$$

Furthermore, let $\underline{x} = (0, y)$ be any point on the boundary $z = 0$. The unit normal vector at this point equals $\hat{\underline{x}} = (-1, 0)$. We thus find:

$$G(\underline{x}, \underline{x}') \cdot \hat{\underline{x}} = -G^z(\underline{x}, \underline{x}') = - \left(\frac{0 - z'}{D^2(\underline{x} - \underline{x}')} + \frac{0 + z'}{D^2(\underline{x} - \underline{x}^i)} \right).$$

From the fact that $\underline{x}^i$ is the mirror image of $\underline{x}'$ around the line $z = 0$, and since the point $\underline{x}$ lies on that line, we have $D^L(\underline{x} - \underline{x}') = D^L(\underline{x} - \underline{x}^i)$. We find:

$$G(\underline{x}, \underline{x}') \cdot \hat{\underline{x}} = 0.$$

This shows that condition (32) is fulfilled and I have shown the formal validity of optimal-tax condition (74).

The question now is how the correction term on the second line of equation (74) should be interpreted. Given that the distance function D^2 declines strongly at larger distances, and given that the image point $\mathbf{x}^i$ lies outside of the domain of the tax base, this term will be most relevant very close to the boundary. If this term becomes infinitesimal at larger distances, then its mere function is to enforce the boundary conditions. Whether this term has a significant effect on the optimal-tax conditions elsewhere is an important empirical question.

With more complex restrictions to the tax base, e.g. multiple components that should be non-negative, more images should be added to find the tax optimum. For example, with the boundaries shown in figure 5 for the taxation of couples, the tax-base space is restricted to an eight of the real vector space. Seven images will be needed to solve the problem.

6.2. Pooling

I argued that pooling usually occurs when the dimension K of the type space is higher than the dimension L of the tax base space. In this situation, multiple individuals will inevitably end up choosing the same allocations. Note that theorem 2 already accounts for the possibility of pooling. This subsection concentrates on characterizations of the tax optimum in the type space. I assume in this subsection that there is no bunching, so the tax base components are strictly monotonous in all characteristics: $|\partial z^l / \partial \theta^k| > 0$. I will first explain the procedure for a one-dimensional tax base, then I will extend it to a multidimensional base.

6.2.1. One-Dimensional Tax Base

Figure 10 demonstrates the problem for two-dimensional types (θ^1, θ^2) and a one-dimensional tax base z . One could think of the characteristic θ^1 as intrinsic labour productivity, for example, while θ^2 could be an intrinsic disutility from work, and z would be gross labour income. In the left-hand side of the figure I have connected types with the same gross labour incomes using lines. Each line represents a single point in the tax base space. For a given productivity, individuals with a lower disutility of work will obtain a higher gross income. Conversely, for a given disutility of work, individuals with a higher productivity will also obtain a higher gross income.

Now pick one type on each of these connected lines that will represent the corresponding tax base value, in such a way that these representative types are connected by a continuous line, as demonstrated in figure 11. I denote this set of types as Φ . It forms a one-dimensional representation of the tax base space in the type space.

I will now reformulate Euler-Lagrange equation (6) as a partial differential equation in the set Φ , following the derivations set forth in subsection 5.1. Define $\varphi : \Theta \rightarrow \Phi$ as the function that maps any multidimensional type θ on the one-dimensional element $\varphi(\theta, T) \in \Phi$ which chooses the same value of the tax base. With a slight abuse of notation, I parametrize this element by a real number: $\varphi \in \mathbb{R}$. I will come back to the choice of this parametrization below. The function φ depends on tax function T . The

Figure 10: An example of pooling, with a two-dimensional type space and a one-dimensional tax base space. Each value z of the tax base corresponds to a one-dimensional set of types. For example, types a and b choose the same value $z(a) = z(b)$ of the tax base, and types c and d choose the same value $z(c) = z(d)$. The one-dimensional subsets of the type space who choose the same values of the tax base, are connected using lines.

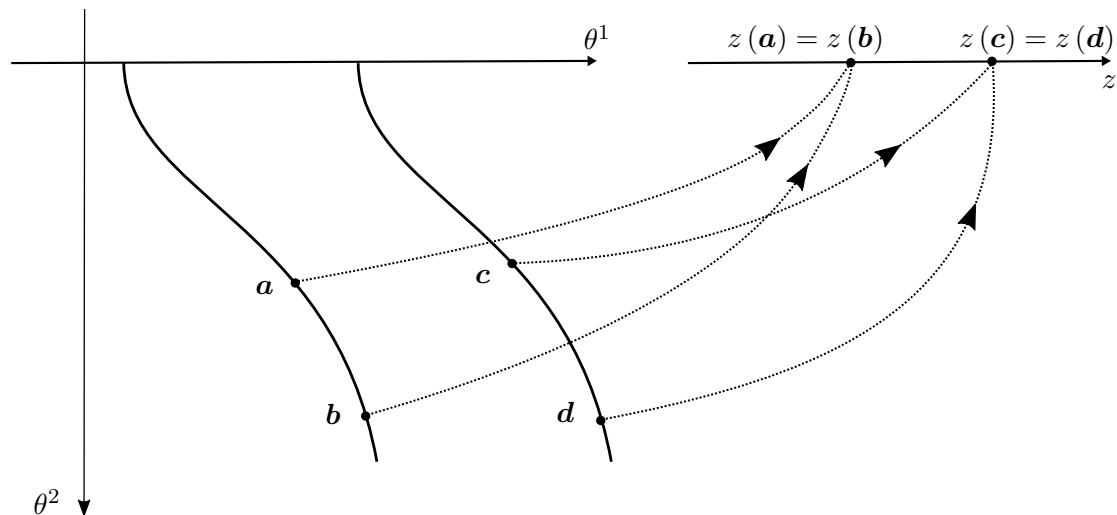
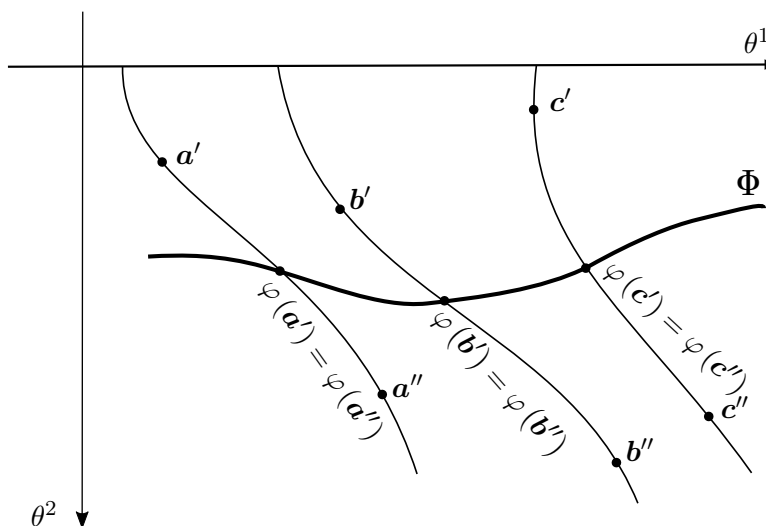


Figure 11: Since multiple types end up pooling at the same tax base values, we pick one type to represent each tax base value, such that the one-dimensional set Φ of these representative types forms a line in the type space. I parametrize the coordinates on this line as a scalar φ . For each type $\theta \in \Theta$, the scalar function $\varphi(\theta, T)$ then indicates which point in the set of representative types Φ chooses the same value for the tax base.



reason is that different types who pool together at one level of the tax base in presence of a given tax function, might choose different values of the tax base when another tax function is in place.

Denote the cumulative distribution function on the set Φ as $F^\Phi(\varphi)$, with corresponding density function $f^\Phi(\varphi)$. Define the function z^Φ that maps any element in Φ on the tax base value in $\mathcal{X}$ that it represents:

$$\forall \theta : z^\Phi(\varphi(\theta, T), T) \equiv z(\theta, T).$$

The procedures set forth in subsection 5.1 then imply that Euler-Lagrange equation (6) can then be rewritten as follows:

$$\forall \theta : \frac{\partial}{\partial \varphi} \left\{ \frac{[T_z \bar{z}_{T_z}](z^\Phi(\varphi(\theta, T), T))}{z^\Phi(\varphi(\theta, T), T)} f^\Phi(\varphi(\theta, T)) \right\} = [1 - \bar{\alpha}(z(\theta, T))] f^\Phi(\varphi(\theta, T)).$$

The main difference with theorem 3, besides the dimensionality, is that the welfare weights $\bar{\alpha}$ and the behavioural responses $\bar{z}_{T_z}$ are now averaged over the multitude of individuals pooling at the given value of the tax base.

Integrate the latter result and write the average behavioural response at z as an integral over the type space:

$$\begin{aligned} \forall \theta : T_z(z(\theta, T)) f^\Phi(\varphi(\theta, T)) & \left(\int_{\Theta} \frac{z_{T_z}(\theta')}{z^\Phi(\varphi(\theta, T), T)} dF^{\Theta|\mathcal{X}}(\theta' | z(\theta, T)) \right) \\ & = - \int_{\varphi(\theta, T)}^{\infty} [1 - \bar{\alpha}(z^\Phi(\varphi', T))] f^\Phi(\varphi') d\varphi'. \end{aligned} \quad (75)$$

The interpretation of this result is similar to the interpretation of equation (51) in section 5. The right-hand side represents the distributional advantage of taxation, the left-hand side represents the cost of redistribution.

One question now is how the function $\varphi(\theta, T)$ should be parametrized. The pooling functions introduced by Jacquet and Lehmann (2016) are one possibility.²⁸ If preferences are separable between consumption and leisure, equation (75) corresponds to the result found by Jacquet and Lehmann (2016) in their proposition 1.

A different possibility is to simply parametrize $\varphi(\theta, T)$ by the gross tax income that it represents. Substituting z for φ in equation (75) recovers optimal-tax equation (7) in the tax base space. It is not clear at this point why one parametrization should be preferred over the other. Different parametrizations might lead to different insights, depending on the problem being studied. The question of the optimal parametrization remains a topic for further research.

²⁸Jacquet and Lehmann (2016) divide the type vector into a one-dimensional wage, and a vector of other characteristics. They choose a reference value of these other characteristics. Keeping constant these other characteristics at the reference value, and varying the wage, they trace out a straight, one-dimensional path in the type space. Assuming that different wages yield different gross labour incomes, this one-dimensional path is a representation of the gross incomes. In terms of my approach, the value of the wage is then used as a parametrization of the different points on this line.

We can further reinterpret equation (75), using similar derivations as those in subsection 5.1. Use the implicit function theorem to find:

$$\forall \theta : T_z(z(\theta, T)) \left. \frac{d\varphi}{dT_z} \right|_{z(\theta, T)} = \frac{1}{f^\Phi(\varphi)} \int_{\varphi(\theta, T)}^{\infty} [1 - \bar{\alpha}(z^\Phi(\varphi', T))] f^\Phi(\varphi') d\varphi', \quad (76)$$

where I rewrite the integral on the left-hand side as an average at $z(\theta, T)$. This expression can again be interpreted as an information cost, similar to what we found in subsection 5.1. Individuals who originally pool at $z(\theta, T)$ are represented by the same type $\varphi(\theta, T)$. If there is a tax reform, these individuals will respond, and they will become represented by a different type in Φ . The derivative on the left-hand side of equation (76) denotes the average of these responses. The larger are these responses, the more difficult it will be to target specific types, and the lower will be the optimal marginal tax rate.

Similarly, using methods and notations from subsection 5.2, we can reformulate equation (75) as follows:

$$\forall \theta : T_z(z(\theta, T)) = \frac{1}{f^\Phi(\varphi)} \left. \frac{\partial(u_z/u_c)}{\partial \varphi} \right|_{z(\theta, T)} \left\{ \int_{\varphi(\theta, T)}^{\infty} [1 - \bar{\alpha}(z^\Phi(\varphi', T))] f^\Phi(\varphi') d\varphi' \right\}.$$

Again the term $\left. \frac{\partial(u_z/u_c)}{\partial \varphi} \right|_z$ is an indication of how able the government is to target specific types. If a characteristic has a large effect on the marginal rate of substitution, then the government can use small changes in the marginal tax rate to precisely target the individuals it would like to target.

6.2.2. Multidimensional Tax Base

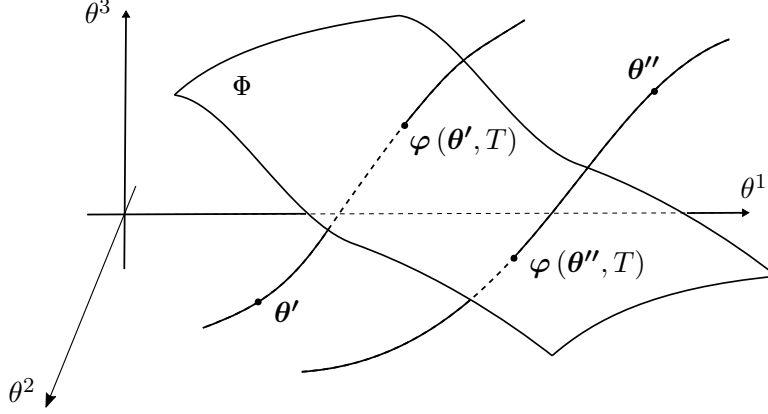
We now have the necessary tools to study the phenomenon of pooling with a multidimensional tax base. Still assuming that the type space Θ has a higher dimension than the tax base space $\mathcal{X}$, $K > L$, the subsets of Θ that connect types who pool at the same tax base value will have dimension $K - L$. For example, with four-dimensional types and a two-dimensional tax base, each point of the two-dimensional tax base would correspond to a two-dimensional plane in the four-dimensional type space.

Figure 12 demonstrates the situation for a two-dimensional tax base and a three-dimensional type space, where again types sharing the same value of the tax base form one-dimensional sets. The set Φ representing the tax base space in the type space is a two-dimensional plane. The derivations below are valid for any combination $K > L$. Denote the cumulative distribution function on the set Φ as $F^\Phi(\varphi)$, with corresponding density function $f^\Phi(\varphi)$.

Define $\varphi : \Theta \rightarrow \Phi$ as the vector-valued function that maps any K -dimensional type θ on the L -dimensional element $\varphi(\theta) \in \Phi$, which is parametrized as a real vector ($\varphi \in \mathbb{R}^L$). Introduce the function $\mathbf{x}^\Phi : \Phi \rightarrow \mathcal{X}$ which maps any element $\varphi \in \Phi$ on the tax base value $\mathbf{x} \in \mathcal{X}$ that it represents:

$$\forall \theta : \mathbf{x}^\Phi(\varphi(\theta, T), T) \equiv \mathbf{x}(\theta, T).$$

Figure 12: With three-dimensional types and two-dimensional tax base values, there are one-dimensional sets in the type space who end up choosing the same value of the tax base. On each of these lines we pick one type, which represents the corresponding value of the tax base, such that the two-dimensional set Φ of these representative types forms a plane in the type space. The coordinates on this plane are parametrized by a vector φ . For each type $\theta \in \Theta$, the function $\varphi(\theta, T)$ then indicates which point in the set of representative types Φ chooses the same value for the tax base.



Introduce the Jacobian matrix for this function:

$$\mathcal{J}^\Phi \equiv \begin{pmatrix} x_{\varphi^1}^{\Phi,1} & \cdots & x_{\varphi^L}^{\Phi,1} \\ \vdots & \ddots & \vdots \\ x_{\varphi^1}^{\Phi,L} & \cdots & x_{\varphi^L}^{\Phi,L} \end{pmatrix}.$$

Using again the derivations from subsection 5.1, we can now reformulate Euler-Lagrange equation (12) in the representative space Φ :

$$\begin{aligned} \forall \theta \in \Theta : & \left(1 - \bar{\alpha} \left(x^\Phi(\varphi(\theta, T), T) \right)\right) f^\Phi(\varphi(\theta, T)) \\ &= \sum_{l=1}^L \frac{\partial}{\partial \varphi^l} \left\{ \left[\nabla_x T \cdot \overline{x_{\nabla_x T}} \cdot \left[(\mathcal{J}^\Phi)^{-1} \right]^\top \right]^l \left(x^\Phi(\varphi(\theta, T), T) \right) f^\Phi(\varphi(\theta, T)) \right\}. \end{aligned} \quad (77)$$

The solution for this differential equation is entirely analogue to what we found in the previous subsection:

$$\begin{aligned} \forall \theta : & \left[\nabla_x T \cdot \overline{x_{\nabla_x T}} \cdot \left[(\mathcal{J}^\Phi)^{-1} \right]^\top \right] (x(\theta, T)) \\ &= \frac{1}{f^\Phi(\varphi(\theta, T))} \text{cov} \left(\bar{\alpha} \left(x^\Phi(\varphi', T) \right), \frac{\varphi' - \varphi}{D^K(\varphi' - \varphi)} \right). \end{aligned}$$

We can rewrite the left-hand side using derivations analogue to those in the previous

subsection:

$$\forall \theta : \nabla_x T(x(\theta, T)) \cdot \overline{\frac{d\varphi}{d\nabla_x T}} \Big|_x^\top = -\frac{1}{f^\Phi(\varphi(\theta, T))} \text{cov} \left(\bar{\alpha}(x^\Phi(\varphi', T)), \frac{\varphi' - \varphi}{D^K(\varphi' - \varphi)} \right),$$

or accounting for preference optimization:

$$\begin{aligned} \forall \theta : \nabla_x T(x(\theta, T)) \cdot \left(\overline{\frac{\partial (u_x/u_c)}{\partial \varphi}} \Big|_x^{-1} \right)^\top \\ = -\frac{1}{f^\Phi(\varphi(\theta, T))} \text{cov} \left(\bar{\alpha}(x^\Phi(\varphi', T)), \frac{\varphi' - \varphi}{D^K(\varphi' - \varphi)} \right). \end{aligned}$$

The interpretation of these equations is fully analogue to those in the previous subsections: the government balances the advantage of redistribution of increasing the marginal tax rate in a given direction, against the cost of targeting specific characteristics of the individuals.

6.3. Bunching

When individuals of consecutive types all choose the same level of the tax base, we say that they are bunching on that point. A range of points in the type space will be allocated to a single mass point in the tax base space. Ebert (1992) shows that it is possible to find combinations of social preferences, individual preferences and population densities such that this situation occurs in the optimum.

Bunching can also occur on a higher-dimensional subset of the tax base space. Say for example that the tax base is two-dimensional, taxing capital income and labour income. It is possible that different individuals with different capital incomes, all choose a corner solution where labour income is zero. The bunching range in this case is a one-dimensional line in the two-dimensional tax base space.

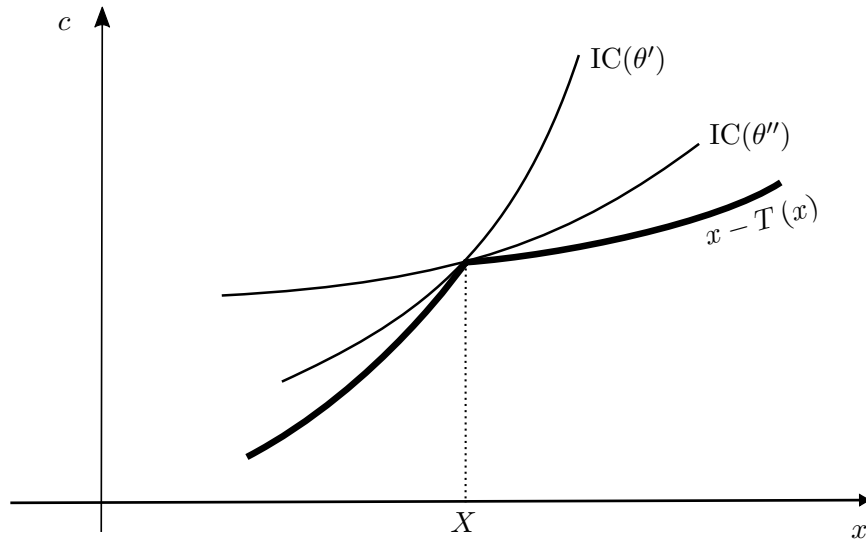
I show in subsection 6.3.1 how a bunching range in the interior of the tax base space can only be implemented through a kink in the tax function, and how this should be taken into account in the Euler-Lagrange formalism. In subsection 6.3.2 I show how transversality condition (43) continues to apply in this situation. I show how the tax optimum can then be characterized in presence of bunching in subsection 6.3.3. Finally in subsection 6.3.4, I treat the possibility of bunching on an edge of the tax base space.

6.3.1. Validity of Euler-Lagrange Equation

Suppose that there is a one-dimensional type θ , a one-dimensional tax base x , and a numéraire commodity c . Suppose that individuals optimize a utility function $u(c, x, \theta)$, subject to the budget constraint $c = x - T(x)$. If the tax function is differentiable, then the first-order condition for the individual is as follows:

$$\frac{u_x}{u_c} = -(1 - T'(x)). \quad (78)$$

Figure 13: Bunching at tax base value X . The tax function $T(x)$ has a kink at X , causing individuals of different types θ' and θ'' , with indifference curves of different slopes, to choose the same value of the tax base.



Assume that the following single-crossing condition is fulfilled:

$$\frac{\partial (u_x/u_c)}{\partial \theta} > 0. \quad (79)$$

Mirrlees (1976) then shows that the second-order condition for the individual optimization problem is equivalent to the following weak monotonicity condition:

$$\frac{dx}{d\theta} \geq 0.$$

Ebert (1992) shows that it is possible to find combinations of social preferences, individual preferences and population densities such that the second-order condition is binding on some range of the type space: $dx/d\theta = 0$. It immediately follows from conditions (78) and (79) that this allocation cannot be implemented using a smooth tax function. In order to implement it, there needs to be a kink in the tax function. Figure 13 shows how this works: both types θ' and θ'' , with indifference curves of different slopes, choose the same value X of the tax base. The problem with a kink in the tax function is that the Euler-Lagrange equation (12) in theorem 1 is no longer well-defined: the marginal tax rate T_x is not defined in the bunching point.

It is straightforward to see though that outside of the bunching point, the tax reforms constructed in subsection 3.1 should still leave social welfare unchanged. It follows that Euler-Lagrange equation (6) remains valid everywhere, except at the kink in the tax function. This reasoning is straightforward to extend to higher dimensions: Euler-Lagrange equation (12) remains valid outside of the bunching ranges.

Furthermore, as long as bunching does not occur on the edge of the tax base space, also boundary condition (13) remains valid. In one dimension this is easy to see. There still is no reason to distort behaviour of the top individual, if there are no individuals with higher incomes to levy additional taxes from. The condition $T_x(\bar{x}) = 0$ remains valid. Similarly, the condition $T_x(\underline{x}) = 0$ remains valid: there is no point in distorting behaviour of the individuals at the bottom, since there are no individuals with lower incomes to redistribute to. These findings are summarized in the following proposition.

Proposition 1. *The tax optimum with an L -dimensional tax base, if bunching occurs only on a set $\mathcal{K}$ on the interior of the tax base space $\mathcal{X}$, complies to the following partial differential equation, referred to as the Euler-Lagrange condition:*

$$\forall \mathbf{x} \in \mathcal{X} \setminus \mathcal{K} : (1 - \bar{\alpha}(\mathbf{x})) f^{\mathcal{X}}(\mathbf{x}) = \sum_{l=1}^L \frac{\partial}{\partial x^l} \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}_{T_l}}) f^{\mathcal{X}} \right](\mathbf{x}), \quad (80)$$

subject to the boundary conditions:

$$\forall \mathbf{x} \in \Gamma(\mathcal{X}) : \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}_{\nabla_{\mathbf{x}} T}} \cdot \hat{\mathbf{x}}) f^{\mathcal{X}} \right](\mathbf{x}) = 0, \quad (81)$$

and the government budget constraint:

$$\int_{\mathbb{R}^L} T(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) d\mathbf{x} = 0.$$

Proof. See appendix D. □

6.3.2. Validity of the Transversality Condition

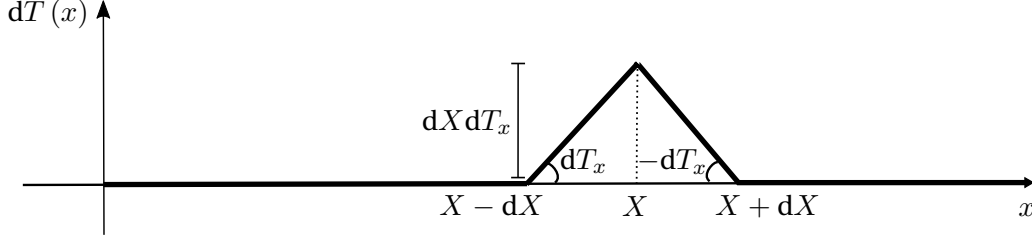
If we wish to use the methods of section 4 to solve the Euler-Lagrange equation for the part of $\mathcal{X}$ where it is still defined, we need to be able to take integrals over the entire tax base space, including the bunching ranges. I will show here that we can just *naively integrate* over the entire tax base space.

Let us first study the situation for a one-dimensional tax base x , with a single bunching point X . The dimension K of the type space can still be larger than one. Denote the domain of the tax base as $[\underline{x}, \bar{x}]$.

Introduce a tax reform around the bunching point X . Let the marginal tax rate increase by dT_x at an interval $[X - dX, X]$ of width dX below the kink, and let the tax rate decrease by the same dT_x over an interval $(X, X + dX]$ of the same width above the kink. This reform is shown in figure 14.

This reform has a number of effects. Below the kink, in the interval $[X - dX, X]$, individuals experience an increase dT_x in their marginal tax rate, which causes behavioural effects of size $\overline{x_{T_x}} dT_x$. Their tax liability thus changes by $dT_x [T_x \overline{x_{T_x}}](X - dX)$. The number of individuals experiencing this change is $f^{\mathcal{X}}(X - dX) dX$, such that the total impact on government revenue equals $[T_x \overline{x_{T_x}} f^{\mathcal{X}}](X - dX) dT_x dX$. Similarly, on the interval $(X, X + dX]$, the change in government revenue is $-[T_x \overline{x_{T_x}} f^{\mathcal{X}}](X + dX) dT_x dX$. Finally, on the bunching point X itself, there is a number of individuals $\int_{X-dX}^{X+dX} f^{\mathcal{X}}(x) dx$

Figure 14: The size of the tax reform around the bunching point X . The marginal tax rate is increased by dT_x over the interval $[X - dX, X)$, and it is decreased again by dT_x over the interval $(X, X + dX]$. Below $X - dX$ and above $X + dX$, the tax liability remains unaltered. The change in the tax liability at X equals $dXdT_x$.



who experience an increase $dXdT_x$ in their tax liability.²⁹ The effect on social welfare equals $dT_x dX \int_{X-dX}^{X+dX} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx$. Sum all these effects, divide by $dT_x dX$ and require that the result is zero:

$$\left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (X + dX) - \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (X - dX) = \int_{X-dX}^{X+dX} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx. \quad (82)$$

Since the derivative of the tax function at X is not defined, we can no longer obtain the traditional Euler-Lagrange equation from this condition.

What we are interested in now is how to calculate the integral of the entity $(1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x)$ from a value of the tax base x' below the kink, to a value x'' above the kink. Note first what happens if we integrate on the part below the kink:

$$\begin{aligned} \int_{x'}^{X-dX} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx &= \int_{x'}^{X-dX} \frac{d}{dx} \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x) dx \\ &= \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (X - dX) - \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x'), \end{aligned}$$

where I use the validity of the Euler-Lagrange equation on this range (by proposition 1). Similarly, above the kink:

$$\begin{aligned} \int_{X+dX}^{x''} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx &= \int_{X+dX}^{x''} \frac{d}{dx} \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x) dx \\ &= \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x'') - \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (X + dX). \end{aligned}$$

Adding the last two equations and using result (82), taking the limit $dX \rightarrow 0$:

$$\begin{aligned} \lim_{dX \rightarrow 0} & \left[\int_{x'}^{X-dX} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx + \int_{X+dX}^{x''} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx \right] \\ &= - \lim_{dX \rightarrow 0} \left[\int_{X-dX}^{X+dX} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx \right] + \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x'') - \left[T_x \bar{x}_{T_x} f^{\mathcal{X}} \right] (x'). \end{aligned}$$

²⁹Note that the imprecisions in this derivation disappear as $dX \rightarrow 0$. These results follow more formally from equation (104) in appendix D.

This shows that we do nothing wrong by *naively* integrating over the bunching range, writing e.g.:

$$\int_{x'}^{x''} (1 - \bar{\alpha}(x)) f^{\mathcal{X}}(x) dx = \left[T_x \overline{x_{T_x}} f^{\mathcal{X}} \right] (x'') - \left[T_x \overline{x_{T_x}} f^{\mathcal{X}} \right] (x').$$

This reasoning immediately extends to the multidimensional case. This follows more formally from condition (104) in appendix D. We can extend corollary 1 as follows:

Corollary 3. *The tax optimum with an L -dimensional tax base, with a bunching range $\mathcal{K}$, complies to the following condition, for any compact volume $V \subseteq \mathcal{X}$ with piecewise smooth boundary $\Gamma(V)$ that does not intersect with $\mathcal{K}$, and with $\hat{\mathbf{x}}$ the unit vector normal to that boundary at point $\mathbf{x} \in \Gamma(V)$:*

$$\int_V (1 - \bar{\alpha}(\mathbf{x})) dF^{\mathcal{X}}(\mathbf{x}) = \int_{\Gamma(V)} \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}_{\nabla_x T}} \cdot \hat{\mathbf{x}}) f^{\mathcal{X}} \right] (\mathbf{x}) d\Gamma.$$

It follows immediately from boundary condition (81) that transversality condition (43) remains valid:

$$\int_{\mathcal{X}} \bar{\alpha}(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) d\mathbf{x} = 1.$$

6.3.3. Solving the Euler-Lagrange Equation

In one dimension, we can now simply integrate Euler-Lagrange equation (80) to find the following optimal-tax condition:

$$\begin{aligned} \forall x : \overline{x_{T_x}} T_x(x) f^{\mathcal{X}}(x) &= \int_{\underline{x}}^x (1 - \bar{\alpha}(x')) dF^{\mathcal{X}}(x') \\ &= - \int_x^{\bar{x}} (1 - \bar{\alpha}(x')) dF^{\mathcal{X}}(x'). \end{aligned}$$

It follows that optimality condition (7) and the ensuing results remain valid everywhere beside the kinks. Similarly, outside of the kinks, the derivations that lead to the multi-dimensional optimum (42) remain valid with naive integration at the bunching points.³⁰ We thus find the following proposition.

Proposition 2. *The tax optimum with an L -dimensional tax base, with bunching on a set $\mathcal{K}$, and when the tax base space coincides with the real vector space, $\mathcal{X} = \mathbb{R}^L$, complies to the following necessary condition:*

$$\forall \mathbf{x} \in \mathcal{X} \setminus \mathcal{K}, \forall l : \left[\sum_j \left(T_j \overline{x_{T_l}^j} \right) f^{\mathcal{X}} \right] (\mathbf{x}) = \text{cov} \left(\bar{\alpha}(\mathbf{x}'), \frac{x^l - x^l}{D^L(\mathbf{x}' - \mathbf{x})} \right), \quad (83)$$

³⁰In a mechanism-design approach this can be understood as follows. In order to correctly account for bunching, not only the first-order incentive compatibility constraint needs to be taken along in the constrained objective function, but also the second-order constraint $dx/d\theta \geq 0$. The multiplier of this additional term equals zero as long as the second-order constraint is not binding. Thus with strict inequality $dx/d\theta > 0$, outside of the bunching range, the first-order conditions for the government optimization problem will be exactly the same as in the situation where the second-order condition for the individual optimization problem is not taken into account.

with transversality condition:

$$\int_{\mathbb{R}^L} \bar{\alpha}(\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) d\mathbf{x} = 1. \quad (84)$$

Remember that without bunching points, the optimal tax function is pinned down when the Euler-Lagrange equation is given, together with the transversality condition and the government budget constraint. If the Euler-Lagrange equation is no longer valid in a bunching point, this creates a degree of indeterminacy in the problem. Even if the marginal tax rates comply to the Euler-Lagrange equation in all other points, it is not clear what happens in the kink. Since in this paper I have always assumed that the permissible tax reforms are continuous functions, I assume here that the tax function is continuous in the kink. This assumption suffices to solve the indeterminacy.

6.3.4. Bunching on the Edge

A special case occurs when individuals bunch on an edge of the tax-base space. Take for example the situation of figure 9, where the tax base consists of labour income z and capital income y , so $\mathbf{x} \equiv (z, y)$, with the value of z restricted to be positive. If there is bunching on the edge, so a mass of individuals chooses a corner solution at $z = 0$, this does not change the form of the optimal-tax equation at the interior of the tax base space. The method of images can be applied, and equation (74) remains valid.

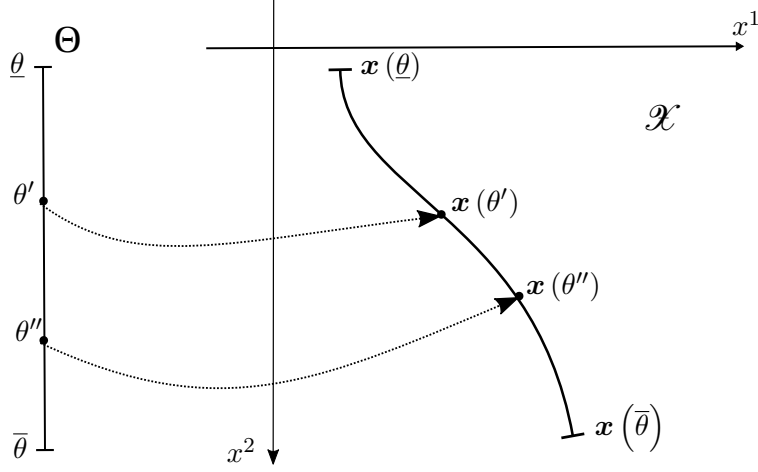
Without bunching at the bottom, we know that the boundary condition at $z = 0$ would be that total tax wedge caused by the tax on labour income, $T_z z T_z + T_y y T_z$, equals zero. This changes when there is bunching at $z = 0$. To see this, imagine that there is an increase dT in the tax liability for all individuals who have labour income $z = 0$, which phases out over the infinitesimal interval $z = [0, dz]$. The effect on social welfare of the increased tax liability equals $dT \int_0^{dZ} \int_{-\infty}^{+\infty} (1 - \bar{\alpha}(z', y')) f^{\mathcal{X}}(z', y') dy' dz'$, while the ensuing change in the marginal tax rate has an effect $dT \int_{-\infty}^{+\infty} (T_z \bar{z}_{T_z} + T_y \bar{y}_{T_z}) f^{\mathcal{X}}(dz, y') dy'$ on government revenue. Demanding that the total effect of this reform sums to zero and taking the limit $dz \rightarrow 0$ yields:³¹

$$\begin{aligned} & \lim_{dz \rightarrow 0} \int_0^{dZ} \int_{-\infty}^{+\infty} (1 - \bar{\alpha}(z', y')) f^{\mathcal{X}}(z', y') dy' dz' \\ &= - \lim_{dz \rightarrow 0} \int_{-\infty}^{+\infty} (T_z \bar{z}_{T_z} + T_y \bar{y}_{T_z}) f^{\mathcal{X}}(dz, y') dy'. \end{aligned}$$

The left-hand side of this equation would go to zero in the case without bunching at $z = 0$. If there is a mass of individuals bunching in this corner though, this is no longer the case. It follows that the total tax wedge for the workers with the lowest non-zero income no longer converges to zero. Assuming that welfare weights are larger than one at the bottom, we find $T_z \bar{z}_{T_z} + T_y \bar{y}_{T_z} > 0$ at the bottom. This extends the findings of Seade (1977) to multiple dimensions.

³¹A more formal derivation would be analogue to that of condition (104) in appendix D.

Figure 15: The type space Θ has only one dimension, while the tax base space $\mathcal{X}$ has two dimensions. The allocation, mapping types θ on tax base values $(x^1(\theta), x^2(\theta))$, forms a line, i.e. a one-dimensional subset of $\mathcal{X}$.



6.4. Lower-Dimensional Allocations

When the type space has a lower dimensionality than the tax base space, $K < L$, then the optimal allocation in the tax base space will also be K -dimensional, and as such it will not be surjective on the real vector space $\mathbb{R}^L$. Even if individuals are free to choose any bundle in $\mathbb{R}^L$, the government's objective (1) is an integral over all types in the population, and thus, after changing variables, over the actually chosen bundles in the tax base space. Since the domain of the optimization is no longer the entire real vector space $\mathbb{R}^L$, the optimal-tax condition in theorem 2 is no longer valid. The trick to finding a necessary condition for the tax optimum is again to find a suitable Green function.

Assume thus that we observe a tax function T , defined on an L -dimensional tax base space $\mathcal{X}$. We observe an income distribution which is restricted to a one-dimensional set, a line within $\mathcal{X}$ which is the image of a one-dimensional type space Θ . We also observe marginal behavioural responses $\mathbf{x}_{\nabla T}$ everywhere on this line. Denote the one-dimensional types as θ , with corresponding tax base values $\mathbf{x}(\theta, T)$. Denote the boundary values of the type space as $[\underline{\theta}, \bar{\theta}]$. This situation is illustrated for a two-dimensional tax base in figure 15.

The Green function for this problem is the vector-valued mapping defined by the following equation:

$$\begin{aligned} \forall \mathbf{x}, \mathbf{x}', l : G^l(\mathbf{x}, \mathbf{x}') \equiv & \delta(x^1 - x'^1) \dots \delta(x^{l-1} - x'^{l-1}) \\ & \cdot H\left(\frac{x^l - x'^l}{L}\right) \delta(x^{l+1} - x'^{l+1}) \dots \delta(x^L - x'^L), \end{aligned}$$

with H the unit step function defined in equation (22), and δ the Dirac delta function. To see that this is indeed the Green function for our problem, note first that condition

(31) is fulfilled:

$$\forall \mathbf{x}, \mathbf{x}' : \sum_l \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} = \delta^L(\mathbf{x} - \mathbf{x}'),$$

where I use definition (29) for the multidimensional Dirac delta function and property (23) of the unit step function. Second, note that $\mathbf{G}$ becomes zero at the edges, because of the unit step function H in its definition:

$$\forall \mathbf{x}', l : G^l(\mathbf{x}(\underline{\theta}), \mathbf{x}') = G^l(\mathbf{x}(\bar{\theta}), \mathbf{x}') = 0.$$

Now that we found the Green function for this optimal-tax problem, we can substitute it into equation (33). Use property (18) of the Dirac delta function to reduce the multidimensional integral to a one-dimensional integral over x^l , and find the following necessary optimal-tax condition:

$$\forall \mathbf{x}, l : \sum_j T_j X_{T_l}^j = \int_{x^l}^{\bar{x}^l} (1 - \alpha) dF^l(x^l),$$

with F^l the cumulative density function for tax base component x^l , and X^j aggregate demand for component x^j . Note that if x^l is monotonous, $\partial x^l / \partial \theta > 0$, then the integral on the right-hand side is the same for all components of the tax base. What follows is a traditional Ramsey-style proportional-reduction equation:

$$\forall \theta : \sum_j T_j X_{T_1}^j(\theta) = \dots = \sum_j T_j X_{T_L}^j(\theta) = \int_{\theta}^{\bar{\theta}} (1 - \alpha) dF^{\Theta}(\theta'),$$

which corresponds to the solution traditionally found in the literature (see e.g. Jacobs and Boadway, 2014). Note that the optimal tax function in this case is separable, confirming the findings of Renes and Zoutman (2016b).

7. Conclusion

I introduced the Euler-Lagrange method to characterize the optimum for problems with multidimensional tax bases and multiple dimensions of heterogeneity of the agents. I showed the general applicability of the Euler-Lagrange equation as a first-order condition. I introduced the localized distributional characteristic, and I used it to provide a number of characterizations for the optimum, both in terms of sufficient statistics and of economic fundamentals. I applied my findings to the Diamond and Spinnewijn (2011) optimal mixed labour and capital income tax problem and to the problem of the optimal joint taxation of households as described by Kleven, Kreiner, and Saez (2007). Finally, I showed that the method is robust to a number of potential complications.

A drawback of this paper, that is not easily solved using the formalism at hand, is that all responses to tax reform are assumed to be local. However, this is not uncommon in the literature, it is even generally assumed in most papers referred to in the present

publication. Still it would be interesting to allow for a participation margin – as in Jacquet, Lehmann, and Van der Linden (2013) – or for more general discrete jumps – e.g. allowing for a discrete-choice behavioural model. There is also potential for extensions towards different objective functions, taking into account different normative convictions, political-economy considerations, or alternative behavioural models.

8. Acknowledgements

This paper is part of the dissertation of the author to obtain the degree of Doctor in Economics. I am most grateful to my supervisor André Decoster, and to the members of my commission: Robin Boadway, Bas Jacobs, Laurence Jacquet, Erik Schokkaert, Frans Spinnewijn and Jo Van Biesebroeck. This paper has benefited from useful discussions with Geert Dhaene, Aart Gerritsen, Etienne Lehmann, Sašo Polanec and Floris Zoutman.

I acknowledge the financial support of the Belgian Federal Science Policy Office (BEL-SPO) via the BRAIN.be project BR/121/A5/CRESUS, and of the Flemish Research Foundation (FWO) via travel grant V464515N.

References

- Arfken, George B and Hans J Weber. 2005. *Mathematical methods for physicists international student edition*. Academic press.
- Armstrong, Mark. 1996. “Multiproduct nonlinear pricing.” *Econometrica: Journal of the Econometric Society* :51–75.
- Atkinson, A.B. and J.E. Stiglitz. 1980. *Lectures on public economics*. McGraw-Hill College.
- Atkinson, Anthony B and Joseph E Stiglitz. 2015. *Lectures on public economics*. Princeton University Press.
- Atkinson, Anthony Barnes and Joseph E Stiglitz. 1976. “The design of tax structure: direct versus indirect taxation.” *Journal of Public Economics* 6 (1-2):55–75.
- Boadway, Robin. 2012. *From optimal tax theory to tax policy*. Munich Lectures in Economics. The MIT Press.
- Boháček, Radim and Michal Kejak. 2016. “Optimal Government Policies in Models with Heterogeneous Agents.” .
- Christiansen, Vidar. 1984. “Which commodity taxes should supplement the income tax?” *Journal of public economics* 24 (2):195–220.
- Christiansen, Vidar and Matti Tuomala. 2008. “On taxing capital income with income shifting.” *International Tax and Public Finance* 15 (4):527–545.

- Corlett, W.J. and D.C. Hague. 1953. "Complementarity and the excess burden of taxation." *The Review of Economic Studies* 21 (1):21–30.
- Courant, R and D Hilbert. 1953. "Methods of Mathematical Physics Vol. I. Translated and revised from the German original. First English ed."
- Cremer, H., P. Pestieau, and J.C. Rochet. 2001. "Direct versus indirect taxation: the design of the tax structure revisited." *International Economic Review* 42 (3):781–800.
- Diamond, Peter and Johannes Spinnewijn. 2011. "Capital income taxes with heterogeneous discount rates." *American Economic Journal: Economic Policy* 3 (4):52–76.
- Diamond, Peter A. 1975. "A many-person Ramsey tax rule." *Journal of Public Economics* 4 (4):335–342.
- . 1998. "Optimal income taxation: an example with a U-shaped pattern of optimal marginal tax rates." *American Economic Review* :83–95.
- Ebert, Udo. 1992. "A Reexamination of the Optimal Nonlinear Income Tax." *Journal of Public Economics* 49 (1):47–73.
- Feldstein, Martin S. 1972a. "Distributional equity and the optimal structure of public prices." *The American Economic Review* 62 (1/2):32–36.
- . 1972b. "Equity and efficiency in public sector pricing: the optimal two-part tariff." *The Quarterly Journal of Economics* :176–187.
- Gerritsen, Aart. 2016. "Optimal nonlinear taxation: the dual approach."
- Gerritsen, Aart, Bas Jacobs, Alexandra Rusu, and Kevin Spiritus. 2015. "Optimal Taxation of Capital Income when Individuals have Different Returns." Mimeo, Max Planck Institute Munich/Erasmus University Rotterdam/KU Leuven.
- Jacobs, Bas. 2013. "The marginal cost of public funds is one at the optimal tax system." *Erasmus University Rotterdam* .
- Jacobs, Bas and Robin Boadway. 2014. "Optimal linear commodity taxation under optimal non-linear income taxation." *Journal of Public Economics* 117:201–210.
- Jacquet, Laurence and Etienne Lehmann. 2016. "Optimal Taxation with Heterogeneous Skills and Elasticities: Structural and Sufficient Statistics Approaches."
- Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden. 2013. "Optimal redistributive taxation with both extensive and intensive responses." *Journal of Economic Theory* 148 (5):1770–1805.
- Kleven, Henrik Jacobsen, Claus Thustrup Kreiner, and Emmanuel Saez. 2007. "The optimal income taxation of couples as a multi-dimensional screening problem." (2092).

- Lax, Peter D, Michael E Taylor, Michael E Taylor, and Michael E Taylor. 1998. “Partial Differential Equations I: Basic Theory.”
- Mirrlees, J.A. 1971. “An exploration in the theory of optimum income taxation.” *The review of economic studies* 38 (2):175–208.
- . 1976. “Optimal tax theory: A synthesis.” *Journal of Public Economics* 6 (4):327–358.
- Renes, Sander and Floris T Zoutman. 2016a. “As Easy as ABC? Multi-Dimensional Screening in Public Finance.” Mimeo, Erasmus Universiteit Rotterdam/Norwegian School of Economics.
- . 2016b. “When a Price is Enough.” .
- Rochet, Jean-Charles. 1985. “The taxation principle and multi-time Hamilton-Jacobi equations.” *Journal of Mathematical Economics* 14 (2):113–128.
- Rochet, Jean-Charles and Philippe Choné. 1998. “Ironing, sweeping, and multidimensional screening.” *Econometrica* :783–826.
- Sadka, Efraim. 1976. “On Income Distribution, Incentive Effects and Optimal Income Taxation.” *Review of Economic Studies* 43 (2):261–267.
- Saez, Emmanuel. 2001. “Using Elasticities to Derive Optimal Income Tax Rates.” *Review of Economic Studies* 68 (1):205–29.
- . 2002. “The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes.” *Journal of Public Economics* 83 (2):217–230.
- Saez, Emmanuel and Stefanie Stantcheva. 2016. “Generalized social marginal welfare weights for optimal tax theory.” *The American Economic Review* 106 (1):24–45.
- Scheuer, Florian and Iván Werning. 2016. “Mirrlees meets Diamond-Mirrlees.” Tech. rep., National Bureau of Economic Research.
- Seade, Jesus K. 1977. “On the Shape of Optimal Tax Schedules.” *Journal of Public Economics* 7 (2):203–235.
- Werquin, Nicolas, Aleh Tsyvinski, and Mikhail Golosov. 2015. “A Variational Approach to the Analysis of Tax Systems.” Tech. rep., Society for Economic Dynamics.

A. Proof for the Euler-Lagrange Equation in Theorem 1

The derivations in this section extend the proof for the traditional Euler-Lagrange equation, formulated in standard texts on the calculus of variations (see e.g. Arfken and

Weber, 2005, chapter 17). I extend it to our context with private information about the individual types.

We are maximizing social welfare (1) subject to government budget constraint (2). If the tax function T is to be optimal, any small perturbation to it should leave social welfare unchanged. We are thus interested in the effects of perturbations of T .

For each value of the tax base $\mathbf{x}$, introduce a reform of size $\varepsilon\eta(\mathbf{x})$. Here ε is a small parameter that allows varying the size of the reform, and $\eta(\mathbf{x})$ is an arbitrary, nonlinear but smooth tax reform function. The tax gradient at each point thus changes by $\varepsilon\nabla_{\mathbf{x}}\eta(\mathbf{x})$.

For a given tax function T , I will seek the optimal value of the reform parameter ε that optimizes social welfare. If the tax function T is optimal, then the optimal value of ε will be zero for any function η .

I have assumed that when there is a marginal tax reform, the wellbeing of an individual is affected only by the change in the tax liability at the value of the tax base that he chose before the reform. Similarly, I have assumed that behaviour is affected only by the changes in the tax liability and the tax gradient at the original value of the tax base. Taking into account the reform, the constrained objective function for the government is as follows:

$$\begin{aligned}\mathcal{L}(\varepsilon) \equiv & \int_{\Theta} [v(\boldsymbol{\theta}, T + \varepsilon\eta(\mathbf{x}))] dF^{\Theta}(\boldsymbol{\theta}) \\ & + \lambda \int_{\Theta} [T(\mathbf{x} + \mathbf{x}_T\varepsilon\eta(\mathbf{x}) + \mathbf{x}_{\nabla T} \cdot (\varepsilon\nabla_{\mathbf{x}}\eta(\mathbf{x}))^{\mathbf{T}}) + \varepsilon\eta(\mathbf{x})] dF^{\Theta}(\boldsymbol{\theta}).\end{aligned}\quad (85)$$

For the tax function $T(\cdot)$ to be optimal, it is necessary that the effect of a marginal change to the parameter ε , evaluated in $\varepsilon = 0$, is zero for any reform function $\eta(\mathbf{x})$:

$$\begin{aligned}0 = \frac{d\mathcal{L}^{\varepsilon}}{d\varepsilon}\Big|_{\varepsilon=0} = & \int_{\Theta} \left\{ \left(1 + \frac{v_T}{\lambda} + \nabla_{\mathbf{x}}T \cdot \mathbf{x}_T \right) \eta(\mathbf{x}) \right\} dF^{\Theta}(\boldsymbol{\theta}) \\ & + \int_{\Theta} \{ \nabla_{\mathbf{x}}T \cdot \mathbf{x}_{\nabla T} \cdot (\nabla_{\mathbf{x}}\eta(\mathbf{x}))^{\mathbf{T}} \} dF^{\Theta}(\boldsymbol{\theta}).\end{aligned}\quad (86)$$

I will now change variables in the integral. The law of iterated expectations tells us that for any function $\mathbf{g} : \Theta \rightarrow \mathbb{R}^D$, with D some dimensionality, a total expectation can be rewritten as a total expectation over conditional expectations:

$$\int_{\Theta} \mathbf{g}(\boldsymbol{\theta}) dF^{\Theta}(\boldsymbol{\theta}) = \int_{\Theta} \int_{\mathcal{X}} \mathbf{g}(\boldsymbol{\theta}) dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}(\boldsymbol{\theta}', T)) dF^{\Theta}(\boldsymbol{\theta}'), \quad (87)$$

where I introduce the conditional cumulative density function $F^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x})$.

The law of the unconscious statistician guarantees that a change of variables from the type space to the tax-base space is possible in the outer integral without worrying about the form of $\mathbf{x}(\boldsymbol{\theta}, T)$, as long as the distribution function $F^{\Theta}(\boldsymbol{\theta})$ is replaced by the distribution function $F^{\mathcal{X}}(\mathbf{x})$. This allows rewriting condition (86):

$$\begin{aligned}0 = & \int_{\mathcal{X}} \int_{\Theta} \left\{ \left(1 + \frac{v_T}{\lambda} + \nabla_{\mathbf{x}}T \cdot \mathbf{x}_T \right) \eta(\mathbf{x}) \right\} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) dF^{\mathcal{X}}(\mathbf{x}) \\ & + \int_{\mathcal{X}} \int_{\Theta} \{ \nabla_{\mathbf{x}}T \cdot \mathbf{x}_{\nabla T} \cdot (\nabla_{\mathbf{x}}\eta(\mathbf{x}))^{\mathbf{T}} \} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) dF^{\mathcal{X}}(\mathbf{x}).\end{aligned}\quad (88)$$

Use integration by parts to rewrite the second term on the right:

$$\begin{aligned}
& \int_{\mathcal{X}} \int_{\Theta} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla T} \cdot (\nabla_{\mathbf{x}} \eta(\mathbf{x}))^{\top} \} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) dF^{\mathcal{X}}(\mathbf{x}) \\
&= \int_{\Gamma(\mathcal{X})} \int_{\Theta} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla T} \cdot \hat{\mathbf{x}} f^{\mathcal{X}}(\mathbf{x}) \} \eta(\mathbf{x}) dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) d\Gamma \\
&\quad - \int_{\mathcal{X}} \int_{\Theta} \sum_l \frac{\partial}{\partial x^l} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{T_l} f^{\mathcal{X}}(\mathbf{x}) \} \eta(\mathbf{x}) dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) d\mathbf{x}.
\end{aligned} \tag{89}$$

Substitute this into (88):

$$\begin{aligned}
0 &= \int_{\mathcal{X}} \int_{\Theta} \left\{ \left(1 + \frac{v_T}{\lambda} + \nabla_{\mathbf{x}} T \cdot \mathbf{x}_T \right) f^{\mathcal{X}}(\mathbf{x}) \right. \\
&\quad \left. - \sum_l \frac{\partial}{\partial x^l} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{T_l} f^{\mathcal{X}}(\mathbf{x}) \} \right\} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) \eta(\mathbf{x}) d\mathbf{x} \\
&\quad + \int_{\Gamma(\mathcal{X})} \int_{\Theta} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla T} \cdot \hat{\mathbf{x}} \} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) \eta(\mathbf{x}) d\Gamma.
\end{aligned} \tag{90}$$

The *fundamental lemma of the calculus of variations* states that this expression can only be zero for all functions η if the integrands between curly brackets are zero for all tax-base levels $\mathbf{x}$.

B. Two-Dimensional Green Functions

Suppose $\mathbf{x} \neq \mathbf{x}'$. We find partial derivatives (using the fact that the surface area of a circle with radius r equals $V^2(r) = \pi r^2$):

$$\begin{aligned}
\forall \mathbf{x} \neq \mathbf{x}', \forall l : \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} &= \frac{1}{2\pi \|\mathbf{x} - \mathbf{x}'\|^2} + \frac{x^l - x'^l}{2\pi} \frac{\partial}{\partial \|\mathbf{x} - \mathbf{x}'\|} \left(\frac{1}{\|\mathbf{x} - \mathbf{x}'\|^2} \right) \frac{\partial \|\mathbf{x} - \mathbf{x}'\|}{\partial x^l} \\
&= \frac{1}{2\pi \|\mathbf{x} - \mathbf{x}'\|^2} - \frac{x^l - x'^l}{\pi} \frac{1}{\|\mathbf{x} - \mathbf{x}'\|^3} \frac{\partial \sqrt{(x^1 - x'^1)^2 + (x^2 - x'^2)^2}}{\partial x^l} \\
&= \frac{1}{2\pi \|\mathbf{x} - \mathbf{x}'\|^2} - \frac{(x^l - x'^l)^2}{\pi \|\mathbf{x} - \mathbf{x}'\|^4}.
\end{aligned} \tag{91}$$

Sum these components over l to find property (35).

Define $S^2(\mathbf{x}', R)$ as the area enclosed by the circle around the point $\mathbf{x}'$ with finite radius R . Since the partial derivative $\partial G^l(\mathbf{x}, \mathbf{x}') / \partial x^l$ is zero everywhere outside the point $\mathbf{x}'$, the following integral over the entire real space equals the integral over $S^2(\mathbf{x}', R)$:

$$\int_{\mathbb{R}^2} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x} = \int_{S^2(\mathbf{x}', R)} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x}. \tag{92}$$

Using the divergence theorem, we can rewrite the right-hand side of this equation as follows:

$$\int_{\mathbb{R}^2} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x} = \int_{\Gamma(S^2(\mathbf{x}', R))} \frac{\mathbf{x} - \mathbf{x}'}{2V^2(\|\mathbf{x} - \mathbf{x}'\|)} \cdot \hat{\mathbf{x}} d\Gamma, \quad (93)$$

where I substitute formulation (34) for the Green function $\mathbf{G}$. Note that everywhere on the boundary $\Gamma(S^2(\mathbf{x}', R))$, we have $\|\mathbf{x} - \mathbf{x}'\| = R$, so we can move the term $V^2(\|\mathbf{x} - \mathbf{x}'\|)$ out of the integral:

$$\int_{\mathbb{R}^2} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x} = \frac{1}{2V^2(R)} \int_{\Gamma(S^2(\mathbf{x}', R))} (\mathbf{x} - \mathbf{x}') \cdot \hat{\mathbf{x}} d\Gamma. \quad (94)$$

Again using the divergence theorem on the right-hand side:

$$\begin{aligned} \int_{\mathbb{R}^2} \left(\sum_{l=1}^2 \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l} \right) d\mathbf{x} &= \frac{1}{2V^2(R)} \int_{S^2(\mathbf{x}', R)} \sum_{l=1}^2 \frac{\partial (x^l - x'^l)}{\partial x^l} d\mathbf{x} \\ &= \frac{2 \int_{S^2(\mathbf{x}', R)} d\mathbf{x}}{2V^2(R)}. \end{aligned} \quad (95)$$

Since $\int_{S^2(\mathbf{x}', R)} d\mathbf{x}$ equals the area of any circle with radius R , this proves property (36).

This reasoning extends to any dimensionality $L \geq 2$, leading to the results of subsection 4.3. The Green function generally takes the following form:

$$\mathbf{G}(\mathbf{x}, \mathbf{x}') \equiv \frac{\mathbf{x} - \mathbf{x}'}{LV^L(\|\mathbf{x} - \mathbf{x}'\|)},$$

which solves the following partial differential equation:

$$\forall \mathbf{x} \in \mathbb{R}^L : \delta^L(\mathbf{x} - \mathbf{x}') = \sum_{l=1}^L \frac{\partial G^l(\mathbf{x}, \mathbf{x}')}{\partial x^l}, \quad (96)$$

complying to the following boundary condition:

$$\forall \underline{\mathbf{x}} \in \Gamma(\mathcal{X}), \forall \mathbf{x}' \in \mathcal{X} : \mathbf{G}(\underline{\mathbf{x}}, \mathbf{x}') \cdot \hat{\underline{\mathbf{x}}} = 0. \quad (97)$$

Adding any divergence-free vector field to $\mathbf{G}$ also solves the problem, since this leaves the divergence of $\mathbf{G}$ unaltered. Partial differential equation (96) thus has many solutions. For our specific problem, the above formulation is the one that we are interested in. To see this, note that the vector field $\mathbf{B}$ in equation (26) is conservative, being a gradient of the government budget constraint. We are thus interested in irrotational Green functions. Suppose now that two different irrotational functions solve divergence equation (96) subject to boundary condition (97). The difference between these two functions has divergence zero. Since it is also irrotational, it is a harmonic function. Since it disappears at the boundary, because of condition (97), it is zero everywhere. It follows that if two irrotational functions solve partial differential equation (96) subject to condition (97), they are necessarily equal.

C. Euler-Lagrange Equation in Type Space

Start from equation (86). Use the chain rule to find the following relation:

$$\nabla_x \eta(x) = \nabla_\theta \eta(x) \cdot \mathcal{J}^{-1}. \quad (98)$$

Substitute this into equation (86):

$$\begin{aligned} 0 = \frac{d\mathcal{L}^\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} &= \int_{\Theta} (1 - \alpha) \eta(x) dF^\Theta(\theta) \\ &\quad + \int_{\Theta} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top \cdot (\nabla_\theta \eta(x))^\top \right\} dF^\Theta(\theta). \end{aligned} \quad (99)$$

Use integration by parts to rewrite the right-hand side:

$$\begin{aligned} &\int_{\Theta} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top \cdot (\nabla_\theta \eta(x))^\top \right\} dF^\Theta(\theta) \\ &= \int_{\Gamma(\Theta)} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top \cdot \hat{\theta} \right\} \eta(x) f^\Theta(\theta) d\theta \\ &\quad - \int_{\Theta} \sum_k \frac{\partial}{\partial \theta^k} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top f^\Theta(\theta) \right\} \eta(x) d(\theta). \end{aligned} \quad (100)$$

Substitute this into (99):

$$\begin{aligned} 0 &= \int_{\Theta} \left\{ (1 - \alpha) f^\Theta(\theta) - \sum_k \frac{\partial}{\partial \theta^k} \left(\nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top f^\Theta(\theta) \right) \right\} \eta(x) d\theta \\ &\quad + \int_{\Gamma(\Theta)} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\mathcal{J}^{-1})^\top \cdot \hat{\theta} \right\} \eta(x) f^\Theta(\theta) d\theta. \end{aligned} \quad (101)$$

The *fundamental lemma of the calculus of variations* states that this expression can only be zero for all functions η if the integrands between curly brackets are zero for all types θ .

D. Proof of Euler-Lagrange Equation with Bunching

Denote as $\mathcal{B} \subset \Theta$ the set of types who belong to a bunching range. I assume that there is only one such range, so the set $\mathcal{B}$ is connected. This proof readily extends to a situation with multiple bunching ranges. The tax gradient $\nabla_x T$ is not defined on the set $\mathcal{B}$.

Optimality condition (86) can be adapted to this situation:

$$\begin{aligned} 0 = \frac{d\mathcal{L}^\varepsilon}{d\varepsilon} \Big|_{\varepsilon=0} &= \int_{\Theta \setminus \mathcal{B}} \left\{ \left(1 + \frac{v_T}{\lambda} + \nabla_x T \cdot x_T \right) \eta(x) \right\} dF^\Theta(\theta) \\ &\quad + \int_{\Theta \setminus \mathcal{B}} \left\{ \nabla_x T \cdot x_{\nabla T} \cdot (\nabla_x \eta(x))^\top \right\} dF^\Theta(\theta) \\ &\quad + \int_{\mathcal{B}} \left(1 + \frac{v_T}{\lambda} \right) \eta(x) dF^\Theta(\theta). \end{aligned} \quad (102)$$

Remember that we denote the corresponding bunching range in the tax base space as $\mathcal{K}$. Note that the set $\mathcal{K}$ will generally have Lebesgue measure zero in $\mathbb{R}^L$. In order to use results from multivariable calculus, we should first expand it to a set of non-zero volume with infinitesimal measure, apply our derivations, and then let that volume shrink to the true bunching range. For example, if the set $\mathcal{K}$ is a point in the tax base space, we should encircle it by a small sphere, and let the radius of this sphere go to zero. To avoid needless clutter, I will continue with the set $\mathcal{K}$, implicitly assuming that the above steps have been followed.

Following the same steps as appendix A, we find optimality condition:

$$\begin{aligned}
0 = & \int_{\mathcal{X} \setminus \mathcal{K}} \int_{\Theta} \left\{ \left(1 + \frac{v_T}{\lambda} + \nabla_{\mathbf{x}} T \cdot \mathbf{x}_T \right) f^{\mathcal{X}}(\mathbf{x}) \right. \\
& \left. - \sum_l \frac{\partial}{\partial x^l} \left\{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{T_l} f^{\mathcal{X}}(\mathbf{x}) \right\} \right\} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) \eta(\mathbf{x}) d\mathbf{x} \\
& + \int_{\Gamma(\mathcal{K})} \int_{\Theta} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla T} \cdot \hat{\mathbf{x}} \} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) \eta(\mathbf{x}) d\Gamma \\
& - \int_{\Gamma(\mathcal{K})} \int_{\Theta} \{ \nabla_{\mathbf{x}} T \cdot \mathbf{x}_{\nabla T} \cdot \hat{\mathbf{x}} \} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) f^{\mathcal{X}}(\mathbf{x}) \eta(\mathbf{x}) d\Gamma \\
& + \int_{\mathcal{K}} \int_{\Theta} \left\{ \left(1 + \frac{v_T}{\lambda} \right) f^{\mathcal{X}}(\mathbf{x}) \right\} dF^{\Theta|\mathcal{X}}(\boldsymbol{\theta}|\mathbf{x}) \eta(\mathbf{x}) d\mathbf{x}.
\end{aligned} \tag{103}$$

We can derive a number of properties from this equation. Applying the fundamental lemma of the calculus of variations to the first three lines, proves proposition 1. Next, consider the last two lines. A crucial point is that the tax liability must be equal for all types bunching on the set $\mathcal{K}$. The reform $\eta(\mathbf{x})$ should thus take a constant value on the set $\mathcal{K} \cup \Gamma(\mathcal{K})$. Since equation (103) must apply for any such value of $\eta(\mathbf{x})$, we find the following condition for the bunching range:

$$\int_{\Gamma(\mathcal{K})} \left[(\nabla_{\mathbf{x}} T \cdot \overline{\mathbf{x}_{\nabla T}}) f^{\mathcal{X}} \right](\mathbf{x}) \cdot \hat{\mathbf{x}} d\Gamma = \int_{\mathcal{K}} (1 - \bar{\alpha}(\mathbf{x})) dF^{\mathcal{X}}(\mathbf{x}). \tag{104}$$