# Redistribution, Distortion and Implementation: Unpacking the Optimal Two-Dimensional Tax Schedule

Kevin Spiritus and Floris Zoutman

February 12, 2021

**Abstract**

We analyze the optimal marginal excess burdens for the nonlinear taxation of two-dimensional incomes when households have multidimensional characteristics. We show that the optimal marginal excess burdens can be decomposed into two terms. The first component is a distributional characteristic for the income under consideration: the larger is the (negative) local covariance between the income and the welfare weights, the higher will be the optimal marginal excess burden. This component guarantees that the combination of marginal excess burdens is among the many to comply to the divergence equation found by Golosov et al. (2015). The second component guarantees that the combination of marginal excess burdens is the unique solution that can be implemented using a tax function. This term also captures the further distortions caused by the reform used to construct the distributional characteristic. We use simulations to show the relative importance of each term.

Rough draft, do not quote

A more polished version will be available soon

## 1 Introduction

Mirrlees (1976) was the first to characterize the optimal multidimensional tax schedule by a partial differential equation (PDE), using a mechanism design approach. The unknown function in his PDE was the allocation $\mathscr{U}(\boldsymbol{w})$ of the utilities of the different types. Following Mirrlees' (1976) assumption that a one-to-one correspondence exists between $\mathscr{U}(\boldsymbol{w})$ and its gradient on one hand, and the allocation of the incomes $\boldsymbol{x}(\boldsymbol{w})$ on the other hand, the optimal tax wedges follow directly from the solution to his PDE. Golosov et al. (2015) use a perturbation approach to characterize the optimal tax schedule more directly. They find that the optimum is characterized by a PDE that has the tax function $T(\boldsymbol{x})$ itself as unknown. Lehmann et al. (2021) show that, under the assumptions of our model, both approaches are equivalent, and the solution is unique. Our aim in this paper is to integrate the PDE found by Golosov et al. (2015), to gain more insight into the characteristics of the optimal tax schedule. We show that the optimal marginal excess burdens consists of a distributional characteristic of the income under consideration, and a term that guarantees the implementability of the marginal excess burdens using a tax function. We use simulations to show the relative importance of both terms.

## 2 PDE characterizations of the optimal tax schedule

The partial differential equation for the optimum as found by Golosov et al. (2015) is:

$$\forall \boldsymbol{x} \in \mathcal{X} : \big(1 - g(\boldsymbol{x})\big)\, h(\boldsymbol{x}) = \sum_i \frac{\partial}{\partial x_i} \left( \sum_j T_{x_j}(\boldsymbol{x}) \frac{\mathrm{d} x_j}{\mathrm{d} T_{x_i}} h(\boldsymbol{x}) \right), \tag{1}$$

with corresponding boundary condition:

$$\forall \boldsymbol{x} \in \partial \mathcal{X} : \sum_{i,j} T_{x_j}(\boldsymbol{x}) \frac{\mathrm{d} x_j}{\mathrm{d} T_{x_i}} e_i(\boldsymbol{x}) h(\boldsymbol{x}) = 0, \tag{2}$$

where $e_i(\boldsymbol{x})$ is the $i$-th component of the vector normal to the boundary at $\boldsymbol{x}$. Given the endogeneity of the behavioral effects and the welfare weights, Lehmann et al. (2021) show that (1) is a third-order PDE in the function $T(\boldsymbol{x})$.

Denote the compensated elasticity of income $x_j$ with respect to the net-of-tax rate $1 - T_{x_i}(\boldsymbol{x})$:

$$\forall \boldsymbol{x} : \varepsilon_i^j(\boldsymbol{x}) \equiv \frac{1 - T_{x_i}(\boldsymbol{x})}{x_j} \frac{\mathrm{d} x_j}{\mathrm{d}(1 - T_{x_i}(\boldsymbol{x}))}.$$

Denote the ratio of the net-of-tax incomes $(1 - T_{x_j})x_j$ and $(1 - T_{x_i})x_i$:

$$\forall \boldsymbol{x} : \sigma_i^j(\boldsymbol{x}) \equiv \frac{(1 - T_{x_j}(\boldsymbol{x}))x_j}{(1 - T_{x_i}(\boldsymbol{x}))x_i}.$$

Denote the marginal excess burdens of the tax on the $i$-th income at $\boldsymbol{x}$:

$$\forall i = 1,2 : MEB_i(\boldsymbol{x}) \equiv -\sum_j T_{x_j}(\boldsymbol{x}) \frac{\mathrm{d} x_j}{\mathrm{d} T_{x_i}} h(\boldsymbol{x}) = \sum_j \frac{T_{x_j}(\boldsymbol{x})}{1 - T_{x_j}(\boldsymbol{x})} \sigma_i^j(\boldsymbol{x}) \varepsilon_i^j(\boldsymbol{x}) x_i h(\boldsymbol{x}). \tag{3}$$

Grouping the excess burdens in a vector and denoting the divergence of $\boldsymbol{MEB}(\boldsymbol{x})$ as $\nabla \cdot \boldsymbol{MEB}(\boldsymbol{x}) \equiv \sum_i \partial MEB_i(\boldsymbol{x}) / \partial x_i$, rewrite equation (1):[1]

$$\forall \boldsymbol{x} : \big(1 - g(\boldsymbol{x})\big)\, h(\boldsymbol{x}) = -\nabla \cdot \boldsymbol{MEB}(\boldsymbol{x}), \tag{4}$$

with boundary condition:

$$\forall \boldsymbol{x} \in \partial \mathcal{X} : \sum_i MEB_i(\boldsymbol{x}) e_i(\boldsymbol{x}) = 0. \tag{5}$$

Before we attempt to integrate equation (4), we first explain why doing so is difficult when the tax base is multi-dimensional.

[1] We introduce some relevant mathematical concepts, including the divergence, in our mathematical appendix.

## 3 The problem of finding the optimal marginal excess burdens

Following the tax perturbation approach described in Section 2, the problem of optimal multidimensional taxation consists of finding a tax function $T(\boldsymbol{x})$ that implies a vector of marginal excess burdens $\boldsymbol{MEB}(\boldsymbol{x})$, which complies to partial differential equation (4) with boundary condition (5). In one dimension, it is easy to integrate equation (4) taking into account (5):

$$MEB(x) = \int_{x}^{\overline{x}} \big(1 - g(x)\big) h(x) \mathrm{d}x. \tag{6}$$

Equivalently, we can use definition (3) to find the more familiar expression:

$$\frac{T_x(x)}{1 - T_x(x)} = \frac{1}{\varepsilon(x)} \frac{\int_{x}^{\overline{x}} \big(1 - g(\hat{x})\big) h(\hat{x}) \mathrm{d}\hat{x}}{1 - H(x)} \frac{1}{\alpha(x)}, \tag{7}$$

with Pareto parameter $\alpha(x) \equiv x h(x)/1 - H(x)$. This solution reflects the work of Mirrlees (1971), Diamond (1998) and Saez (2001). Equation (6) expresses the optimal marginal excess burden, and thus the marginal tax rates $T_x(\boldsymbol{x})$, as a function of sufficient statistics that have a clear economic interpretation. Note that by treating the marginal excess burden as the unknown function in the optimal tax problem, Equation (4) is treated as if it were a first-order differential equation.

Suppose that we attempt to follow a similar approach for the multidimensional problem, treating Equation (4) as if it were a first-order PDE in the marginal excess burdens. We will see that infinitely many vector-valued functions $\boldsymbol{MEB}(\boldsymbol{x})$ then exist that comply to partial differential equation (4) with boundary condition (5). Starting from one such solution $\boldsymbol{MEB}(\boldsymbol{x})$, we can use the system of equations (3) to solve for the corresponding marginal tax rates $T_{x_j}(\boldsymbol{x})$. Here, a difficulty arises: there is no guarantee that a tax function $T(\boldsymbol{x})$ exists that implements the combination of marginal tax rates implied by the tentative solution $\boldsymbol{MEB}(\boldsymbol{x})$. A necessary condition for the obtained marginal tax rates to correspond to a tax function $T(x)$, is that the second-order partial cross derivatives are symmetric:

$$\frac{\partial T_{x_1}(\boldsymbol{x})}{\partial x_2} = \frac{\partial T_{x_2}(\boldsymbol{x})}{\partial x_1}. \tag{8}$$

If Condition (8) is fulfilled, the vector-valued function $(T_{x_1}(\boldsymbol{x}), T_{x_2}(\boldsymbol{x}))$ is *conservative*, implying that a corresponding tax schedule $T(\boldsymbol{x})$ does exist. A vector-valued function $\boldsymbol{MEB}(\boldsymbol{x})$ can thus be implemented using a tax function $T(\boldsymbol{x})$ if marginal tax rates $T_{x_j}(\boldsymbol{x})$ exist that solve the system of equations (3) and that satisfy Condition (8). Since the solution to our optimal tax problem is unique (Lehmann et al., 2021), only one of the multitude of functions $\boldsymbol{MEB}(\boldsymbol{x})$ that solve (4) with boundary condition (5) can effectively be implemented using a tax function.

In the remainder this paper, we shed more light on the quest to identify the $\boldsymbol{MEB}(\boldsymbol{x})$ function that solves equations (2)–(3) and that can be also implemented using a tax function. We do so both from an analytical perspective and using simulations.

## 4 Characterizing the marginal excess burdens

A central result in multivariable calculus is that every vector-valued function $\boldsymbol{MEB}(\boldsymbol{x})$ is uniquely determined by its divergence, its curl and by conditions at the boundary of the domain. Moreover, every vector-valued function $\boldsymbol{MEB}(\boldsymbol{x})$ can be decomposed into an irrotational component and a solenoidal component:

$$\boldsymbol{MEB}(\boldsymbol{x}) = \underbrace{\boldsymbol{I}(\boldsymbol{x})}_{\text{irrotational component}} + \underbrace{\boldsymbol{R}(\boldsymbol{x})}_{\text{solenoidal component}}. \tag{9}$$

The irrotational component $\boldsymbol{I}(\boldsymbol{x})$ has the same divergence as $\boldsymbol{MEB}(\boldsymbol{x})$: $\nabla \cdot \boldsymbol{I}(\boldsymbol{x}) = \nabla \cdot \boldsymbol{MEB}(\boldsymbol{x})$. The curl of the irrotational component equals zero, hence the name: $\nabla \times \boldsymbol{I}(\boldsymbol{x}) \equiv \partial I_1/\partial x_2 - \partial I_2/\partial x_1 = 0$. The *solenoidal* component $\boldsymbol{R}(\boldsymbol{x})$ has divergence zero: $\nabla \cdot \boldsymbol{R}(\boldsymbol{x}) = 0$. The curl of $\boldsymbol{R}(\boldsymbol{x})$ equals that of $\boldsymbol{MEB}(\boldsymbol{x})$: $\nabla \times \boldsymbol{R}(\boldsymbol{x}) = \nabla \times \boldsymbol{MEB}(\boldsymbol{x})$, which generally differs from zero unless $\boldsymbol{MEB}(\boldsymbol{x})$ happens to be conservative.

When the optimal tax schedule is in place, we know that the vector of marginal excess burdens complies to partial differential equation (4) with boundary condition (5). Given that $\nabla \cdot \boldsymbol{I}(\boldsymbol{x}) = \nabla \cdot \boldsymbol{MEB}(\boldsymbol{x})$, the irrotational component of the vector of marginal excess burdens then complies to the following condition:

$$\nabla \cdot \boldsymbol{I}(\boldsymbol{x}) = -\big(1 - g(\boldsymbol{x})\big)\, h(\boldsymbol{x}). \tag{10}$$

If we interpret equation (4) as a first-order partial differential equation, then it tells us that the irrotational component of $\boldsymbol{MEB}(\boldsymbol{x})$ complies to (10), but it gives us no information about the solenoidal component. Suppose that some function $\widehat{\boldsymbol{MEB}}(\boldsymbol{x})$ solves first-order PDE (4) with boundary condition (5). Then for any divergence-free function $\widehat{\boldsymbol{R}}(\boldsymbol{x})$ that to complies to boundary condition $\forall \boldsymbol{x} \in \partial\mathscr{X} : \sum_i \widehat{R}_i(\boldsymbol{x}) e_i(\boldsymbol{x}) = 0$, the function $\widehat{\boldsymbol{MEB}}(\boldsymbol{x}) + \widehat{\boldsymbol{R}}(\boldsymbol{x})$ also complies to (4) with boundary condition (5). Since infinitely many such divergence-free functions $\widehat{\boldsymbol{R}}(\boldsymbol{x})$ exist which comply to boundary condition $\forall \boldsymbol{x} \in \partial\mathscr{X} : \sum_i \widehat{R}_i(\boldsymbol{x}) e_i(\boldsymbol{x}) = 0$, there exist infinitely many solutions to first-order PDE (4) with boundary condition (5). Given the uniqueness of the optimal tax schedule, for any irrotational function $\boldsymbol{I}(\boldsymbol{x})$ that solves (10), there exists exactly one solenoidal function $\boldsymbol{R}(\boldsymbol{x})$ that yields a vector of marginal excess burdens $\boldsymbol{MEB}(\boldsymbol{x}) = \boldsymbol{I}(\boldsymbol{x}) + \boldsymbol{R}(\boldsymbol{x})$ that can be implemented using a tax function $T(\boldsymbol{x})$. The solenoidal component of the vector of marginal excess burdens is thus fixed by the implementation condition (8).

Decomposition (9) is known as a *Helmholtz decomposition.* We first study the Helmholtz decomposition of the marginal excess burdens on an unbounded domain, in Subsection 4.1, because in that case the decomposition is unique and its components have a clear economic interpretation. Next, in Subsection **??**, we generalize this *natural* Helmholtz decomposition to the case with a bounded domain, showing that the interpretation of both components remains unaltered.

## 4.1 Unbounded domain

We provide in Lemma 1 the unique Helmholtz decomposition of the vector of marginal excess burdens on a boundless income space $\mathscr{X} = \mathbb{R}^2$. Boundary condition (5) is replaced by the condition that $\boldsymbol{MEB}(\boldsymbol{x})$ vanishes faster than $1/x_i$ as any income $x_i$ becomes infinitely large. Given that the density function $h(\boldsymbol{x})$ integrates to one, the assumption that $\boldsymbol{MEB}(\boldsymbol{x})$ vanishes sufficiently rapidly is innocuous.[2]

**Lemma 1.** *Suppose that the income space is boundless: $\mathscr{X} = \mathbb{R}^2$, and the vector of marginal excess burdens $\boldsymbol{MEB}(\boldsymbol{x})$ vanishes at least as quickly as $1/x_i$ as one of the incomes $x_i$ becomes infinitely large. Then the Helmholtz decomposition of $\boldsymbol{MEB}$ into an irrotational and a solenoidal component:*

$$\boldsymbol{MEB}(\boldsymbol{x}) = \boldsymbol{I}(\boldsymbol{x}) + \boldsymbol{R}(\boldsymbol{x}), \tag{12}$$

*is unique, with components:*

$$\forall i = 1,2: I_i(\boldsymbol{x}) \equiv \iint_{\boldsymbol{x}' \in \mathscr{X}} \frac{x_i - x_i'}{2\pi|\boldsymbol{x} - \boldsymbol{x}'|^2} \left(\nabla' \cdot \boldsymbol{MEB}(\boldsymbol{x}')\right) \mathrm{d}\boldsymbol{x}', \tag{13}$$

$$\forall i, j = 1,2; j \neq i: R_i(\boldsymbol{x}) \equiv (-1)^j \iint_{\boldsymbol{x}' \in \mathscr{X}} \frac{x_j - x_j'}{2\pi|\boldsymbol{x} - \boldsymbol{x}'|^2} \left(\nabla' \times \boldsymbol{MEB}(\boldsymbol{x}')\right) \mathrm{d}\boldsymbol{x}'. \tag{14}$$

*where:*

$$\nabla \cdot \boldsymbol{I}(\boldsymbol{x}) = \nabla \cdot \boldsymbol{MEB}(\boldsymbol{x}'),\ \nabla \times \boldsymbol{I}(\boldsymbol{x}) = 0,$$

$$\nabla \cdot \boldsymbol{R}(\boldsymbol{x}) = 0,\ \nabla \times \boldsymbol{R}(\boldsymbol{x}) = \nabla \times \boldsymbol{MEB}(\boldsymbol{x}').$$

Lemma 1 is a standard result in vector analysis. It shows us how for given boundary conditions, knowledge of the divergence and the curl allows to construct the vector of marginal excess burdens. For completeness, we prove Lemma 1 in Appendix [TODO]. In what follows, we study the irrotational and the solenoidal components in turn.

Lemma 1 shows that the irrotational component of the vector of marginal excess burdens is uniquely determined by its divergence. When the optimal policy is in place, we can substitute the divergence of $\boldsymbol{MEB}(\boldsymbol{x})$ from (4):

$$\forall i = 1,2: I_i(\boldsymbol{x}) \equiv -\iint_{\boldsymbol{x}' \in \mathscr{X}} \frac{x_i - x_i'}{2\pi|\boldsymbol{x} - \boldsymbol{x}'|^2} \left(1 - g(\boldsymbol{x}')\right) h(\boldsymbol{x}') \mathrm{d}\boldsymbol{x}'. \tag{15}$$

---

[2] A back-of-an-envelope calculation illustrates this point. Assume that the marginal tax rates $T_{x_j}$, the elasticities $\varepsilon_j^i$ and the conditional Pareto parameters $\alpha_1(x_1|x_2) \equiv x_1 h_{x_1|x_2}(x_1|x_2)/(1 - H_{x_1|x_2}(x_1|x_2))$ become constant as income $x_1$ becomes infinitely large. Definition (3) then yields the following components for the marginal excess burden:

$$\forall i = 1,2: MEB_i(\boldsymbol{x}) = \frac{T_{x_1}}{1 - T_{x_i}} \varepsilon_i^1 \alpha_1(x_1|x_2)(1 - H_{x_1|x_2}(x_1|x_2)) h_{x_2}(x_2) + \frac{T_{x_2}}{1 - T_{x_i}} \varepsilon_i^2 x_2 h_{x_1|x_2}(x_1|x_2) h_{x_2}(x_2). \tag{11}$$

As $x_1$ increases, the first term of (11) converges to zero as fast as $1 - H_{x_1|x_2}(x_1|x_2)$, which for the Pareto distribution converges as fast as $1/x_1^{\alpha_1}$. As long as $\alpha_1 > 1$, which is empirically the case, the first term of (11) indeed vanishes at least as quickly as $1/x_1$. The second term of (11) vanishes as quickly as $h_{x_1|x_2}(x_1|x_2)$, which for the Pareto distribution converges as quickly as $1/x_1^{\alpha_1+1}$. The second term of (11) thus also vanishes at least as quickly as $1/x_1$.

We can use the divergence theorem to show that the average welfare weight on the income domain equals one:

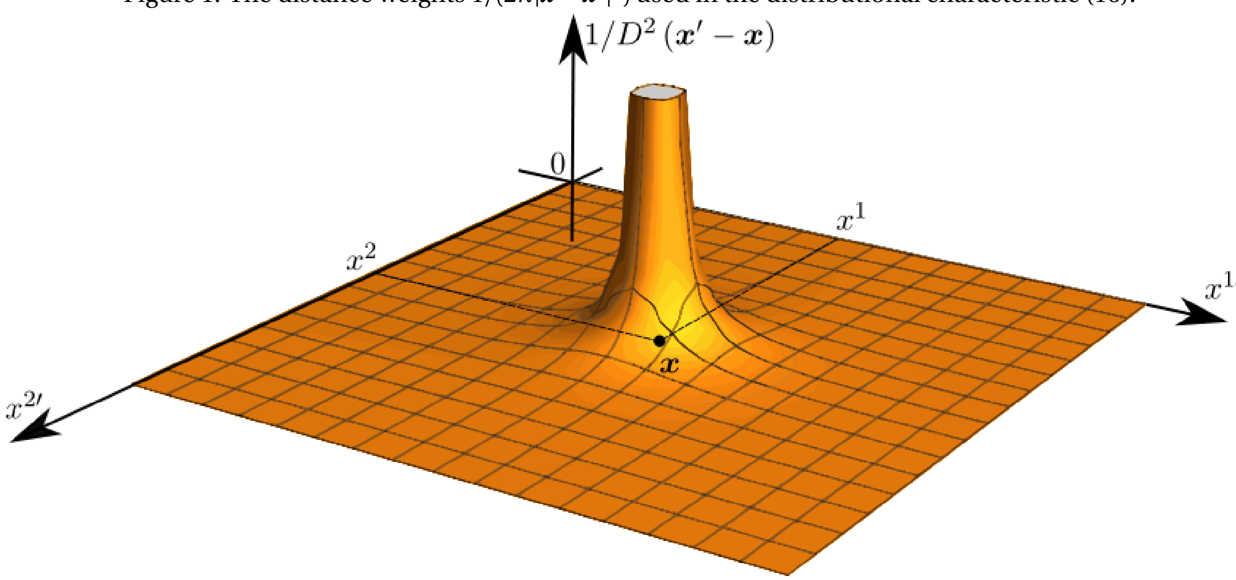
$$\begin{aligned}\iint_{\boldsymbol{x}'\in\mathscr{X}} \left(1-g(\boldsymbol{x}')\right) h(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}' &= -\iint_{\boldsymbol{x}'\in\mathscr{X}} \nabla'\cdot \boldsymbol{MEB}(\boldsymbol{x})\mathrm{d}\boldsymbol{x}' \\ &= -\iint_{\boldsymbol{x}'\in\partial\mathscr{X}} \sum_i MEB_i(\boldsymbol{x})e_i\mathrm{d}\boldsymbol{x}' \\ &= 0,\end{aligned}$$

where we use optimality condition (4) in the first step, the divergence theorem in the second step and boundary condition (5) in the third. Substituting the latter result into (15), we find:

$$\forall i=1,2: I_i(\boldsymbol{x}) \equiv -\mathrm{cov}\left(\frac{x_i'-x_i}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}, g(\boldsymbol{x}')\right). \tag{16}$$

One way to interpret (16) is that $I_i(\boldsymbol{x})$ is a distributional characteristic for income $x_i$ at the combination of incomes $\boldsymbol{x}$. At a given $\boldsymbol{x}$, the optimal marginal excess burden on the marginal tax on either income is larger if the welfare weights decrease faster with that particular income. The covariance in (16) is taken over the entire income domain $\mathscr{X}$. However, the weights $1/(2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2)$ ensure that the focus is on the covariance of the incomes $x_i$ and the welfare weights $g$ in the immediate environment of the income combination $\boldsymbol{x}$ that we are interested in. Figure 1 shows how the weights used are very large near the income combination $\boldsymbol{x}$, and quickly become very small as one moves away from $\boldsymbol{x}$.

Figure 1: The distance weights $1/(2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2)$ used in the distributional characteristic (16).




Turning now to the solenoidal term (14), things are more complicated. Contrary to the divergence of the vector of marginal excess burdens $\boldsymbol{MEB}(\boldsymbol{x})$ in (13), we do not have an explicit formula for the curl of $\boldsymbol{MEB}(\boldsymbol{x})$ in (14). What we do know is that with the additional information that the second-order partial derivatives of the tax function $T(\boldsymbol{x})$

should be equal, as stated in (8), the curl of $\boldsymbol{MEB}(\boldsymbol{x})$ is fixed everywhere, because of the uniqueness of the solution to our optimal tax problem. Sadly, we are unable to write the curl $\nabla \times \boldsymbol{MEB}(\boldsymbol{x})$ in terms of sufficient statistics alone. We are however able to rewrite the solenoidal term (13) in a way that has a clear economic interpretation.

Before we start rewriting the solenoidal term (13), it helps to turn to a more familiar problem, that of one-dimensional income taxation. In one dimension, the divergence of a function is simply its derivative, and divergence equation (4) turns into an ordinary differential equation (ODE):

$$\big(1-g(x)\big)\,h(x) = -\frac{\mathrm{d}MEB(x)}{\mathrm{d}x}. \tag{17}$$

The traditional way to solve (17) is to simply integrated it and to use the boundary condition. We now explore a different way to integrate (17) that also works in multiple dimensions.

We are looking for a scalar-valued function that has $(1-g(x))h(x)$ as its divergence. One can see the term $(1-g(x))h(x)$ as the distributional effect of a tax reform $\mathrm{d}T(x')$ which adds a liability $\delta(x-x')$ to the existing tax function at any income $x'$, where $\delta(\cdot)$ is the Dirac delta function introduced in the Mathematical Appendix. The Dirac delta function $\delta(x-x')$ is such that the change in the tax liability equals zero for any $x' \neq x$, and the change in tax liability at $x$ is such that $\int_{\mathscr{X}} \delta(x-x')\mathrm{d}x' = 1$. The tax reform $\mathrm{d}T(x')$ thus adds a spike to the tax function at income $x$, such that the mechanical revenue effect for a household equals one, and the welfare effect equals $-g(x)$. To find a function that has the term $(1-g(x))h(x)$ as its divergence, we first seek a tax reform $G(x-x')$ which has $\delta(x-x')$ as its divergence. One such function is the Heaviside function $H(x-x')$, which equals zero for $x' > x$, and one for $x' < x$. The Heaviside function can be seen as the following integral of the Dirac delta function:

$$H(x-x') \equiv \int_{-\infty}^{x-x'} \delta(x'')\mathrm{d}x''.$$

Verify that the divergence of the Heaviside function is the Dirac delta function: $\mathrm{d}H(x-x')/\mathrm{d}x = \delta(x-x')$. The distributional benefit of a tax reform of size $H(x-x')$ equals $\int_{\mathscr{X}} H(x-x')(1-g(x'))h(x')\mathrm{d}x'$. Taking the divergence of the latter, we obtain the distributional benefit of a tax reform of size $\delta(x-x')$:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{\mathscr{X}} H(x-x')\big(1-g(x')\big)\,h(x')\mathrm{d}x'\right) &= \int_{\mathscr{X}} \frac{\mathrm{d}H(x-x')}{\mathrm{d}x}\big(1-g(x')\big)\,h(x')\mathrm{d}x' \\
&= \int_{\mathscr{X}} \delta(x-x')\big(1-g(x')\big)\,h(x')\mathrm{d}x' \\
&= \big(1-g(x)\big)\,h(x)\mathrm{d}x.
\end{aligned}$$

We thus found a function with divergence $(1-g(x))h(x)$, and consequently a different way to integrate (17):

$$\begin{aligned} MEB(x) &= -\int_{\mathscr{X}} H(x-x')\big(1-g(x')\big)h(x')\mathrm{d}x' \\ &= -\int_{-\infty}^{x} \big(1-g(x')\big)h(x')\mathrm{d}x' \\ &= \int_{x}^{\infty} \big(1-g(x')\big)h(x')\mathrm{d}x'. \end{aligned}$$

Summarizing, to find a function which has divergence $(1-g(x))h(x)$ at income $x$, we first sought a tax reform $\mathrm{d}T(x')$ which has the Dirac delta function $\delta(x-x')$ as its divergence. The Heaviside function, which has this characteristic, is referred to as a *Green's function* or a *fundamental solution* of the problem.[3] The function $\int_{-\infty}^{x}\big(1-g(x')\big)h(x')\mathrm{d}x'$, which indicates the distributional benefits of the tax reform of size $H(x-x')$, is the function that has $(1-g(x))h(x)$ as its divergence, and thus solves our problem.

The above procedure, using fundamental solutions to find a function that has $(1-g(x))h(x)$ as its divergence, of course seems convoluted for solving the one-dimensional tax problem. For the multidimensional tax problem, however, the described procedure is useful. We show in Appendix [TODO] that in two dimensions, the vector of tax reforms of sizes $G_i(\boldsymbol{x}-\boldsymbol{x}') \equiv (x_i - x_i')/(2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2)$ has the Dirac delta function as its divergence:

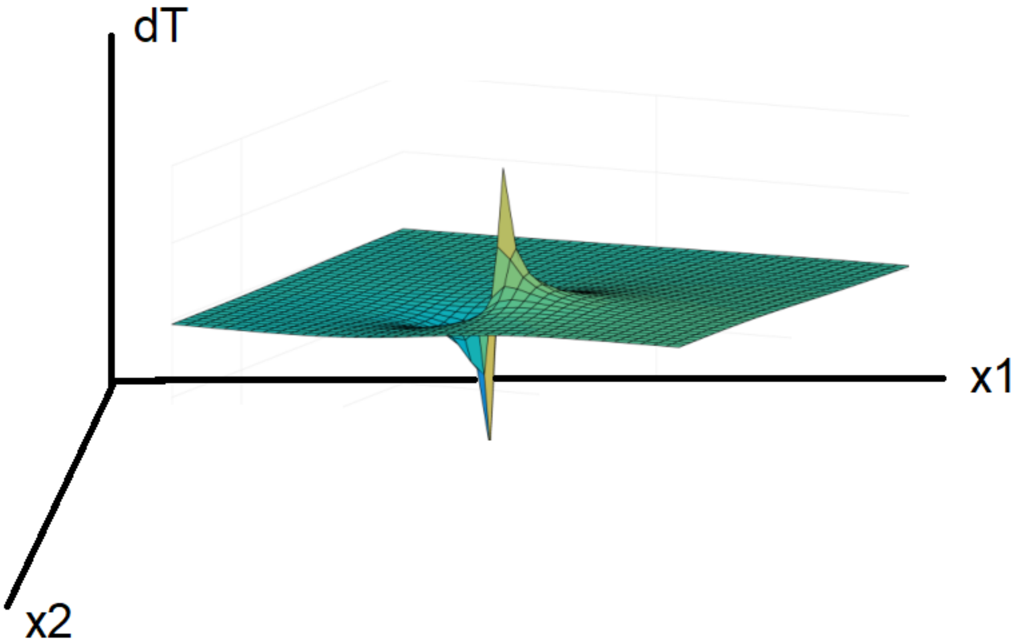
$$\nabla \cdot \left( \frac{x_1 - x_1'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}, \frac{x_2 - x_2'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2} \right) = \delta(\boldsymbol{x}-\boldsymbol{x}'). \tag{18}$$

Going back to Expression (15), we see that the term $G_i(\boldsymbol{x}-\boldsymbol{x}') = (x_i - x_i')/(2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2)$ appears in the integral in the term $I_i(\boldsymbol{x})$. In constructing the irrotational component of the vector of marginal excess burdens that has the prescribed divergence, we have implicitly used the fundamental solutions approach, constructing tax reforms which at any income $\boldsymbol{x}'$ have size $G_i(\boldsymbol{x}-\boldsymbol{x}')$. The term $I_i(\boldsymbol{x})$ is the distributional effect of such a tax reform.

The problem now is that the term $\boldsymbol{MEB}(\boldsymbol{x})$ on the left-hand side of (12) does not fully capture the compensated effects of a tax reform of size $G_i(\boldsymbol{x}-\boldsymbol{x}')$. To see this, consider the reform $G_1(\boldsymbol{x}-\boldsymbol{x}')$ as plotted in Figure (2). At the combination of incomes $\boldsymbol{x}$ there is a sharp increase in marginal tax rate $T_{x_1}(\boldsymbol{x})$. However, contrary to the Heaviside function in the one-dimensional problem, it is impossible in the two-dimensional problem to construct such a sharp increase in the marginal tax rate at one point without affecting the marginal tax rates elsewhere and in other directions. We instead use reforms that satisfy Property (18). This means that we should not only take into account the compensated effects of the increase in the marginal tax rate at the combination of incomes $\boldsymbol{x}$ itself, but also the compensated effects of the ensuing changes in the marginal tax rates at the at the surrounding incomes. We now show that this is exactly what the term $\boldsymbol{R}(\boldsymbol{x})$ accomplishes.

[3] See e.g. [TODO] for more information about this approach.

Figure 2: An extremely local increase in the marginal tax rate $T_{x_1}$




First note the following vector calculus identity, for any scalar function $f(\boldsymbol{x})$ and vector-valued function $\boldsymbol{v}(\boldsymbol{x})$:

$$\nabla\times(f\cdot\boldsymbol{v})=(\nabla f)\times\boldsymbol{v}+f\cdot(\nabla\times\boldsymbol{v}).$$

Applying the latter identity to (14) we obtain:

$$\begin{aligned}\forall i,j=1,2;j\neq i:R_i(\boldsymbol{x})\equiv&(-1)^j\iint_{\boldsymbol{x}'\in\mathscr{X}}\nabla'\times\left(\frac{x_j-x_j'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\cdot\boldsymbol{MEB}(\boldsymbol{x}')\right)\mathrm{d}\boldsymbol{x}'\\&-(-1)^j\iint_{\boldsymbol{x}'\in\mathscr{X}}\nabla'\left(\frac{x_j-x_j'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)\times\boldsymbol{MEB}(\boldsymbol{x}')\mathrm{d}\boldsymbol{x}'.\end{aligned}\tag{19}$$

Applying the divergence theorem and using the fact that $\boldsymbol{MEB}(\boldsymbol{x})$ vanishes for large incomes, the first term on the right-hand side equals zero. We obtain:

$$\forall i,j=1,2;j\neq i:R_i(\boldsymbol{x})\equiv-(-1)^j\iint_{\boldsymbol{x}'\in\mathscr{X}}\left\{\frac{\partial}{\partial x_2'}\left(\frac{x_j-x_j'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)MEB_1(\boldsymbol{x}')-\frac{\partial}{\partial x_1'}\left(\frac{x_j-x_j'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)MEB_2(\boldsymbol{x}')\right\}\mathrm{d}\boldsymbol{x}'.\tag{20}$$

To continue, first note the following identities (see Appendix [TODO]):

$$\begin{aligned}\frac{\partial}{\partial x_1'}\left(\frac{x_1-x_1'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)+\frac{\partial}{\partial x_2'}\left(\frac{x_2-x_2'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)&=-\delta(\boldsymbol{x}-\boldsymbol{x}'),\\\frac{\partial}{\partial x_2'}\left(\frac{x_1-x_1'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)-\frac{\partial}{\partial x_1'}\left(\frac{x_2-x_2'}{2\pi|\boldsymbol{x}-\boldsymbol{x}'|^2}\right)&=0.\end{aligned}$$

Substitute these identities into (20):

$$\forall i = 1,2 : R_i(\boldsymbol{x}) \equiv -\iint_{\boldsymbol{x}' \in \mathscr{X}} \left\{ \sum_j \frac{\partial}{\partial x'_j} \left( \frac{x'_i - x_i}{2\pi |\boldsymbol{x} - \boldsymbol{x}'|^2} \right) MEB_j(\boldsymbol{x}') \right\} \mathrm{d}\boldsymbol{x}' + MEB_i(\boldsymbol{x}). \tag{21}$$

Equation (21) now shows clearly how the solenoidal term $\boldsymbol{R}(\boldsymbol{x})$ captures the compensated effects of the tax reform of size $G_i(\boldsymbol{x} - \boldsymbol{x}')$, beyond the compensated effects of the reform at $\boldsymbol{x}$ itself.

Another way of looking at the problem is as follows. Suppose that we are given the income densities, the welfare weights and the elasticities in the optimum. The question then is what the corresponding optimal marginal excess burdens look like. Partial Differential Equation (4) gives us the divergence of the vector of marginal excess burdens. It allows us to uniquely determine the irrotational component $\boldsymbol{I}(\boldsymbol{x})$, which can be interpreted as a localized distributional characteristic. This distributional characteristic is a local version of the classical distributional characteristic found in optimal multidimensional linear tax models (e.g. Atkinson and Stiglitz, 1980). The construction of the solenoidal component $\boldsymbol{R}(\boldsymbol{x})$ is more complicated, since we have no direct information about the curl of the optimal vector of marginal excess burdens. Instead, the function $\boldsymbol{R}(\boldsymbol{x})$ must be such that the sum $\boldsymbol{I}(\boldsymbol{x})+\boldsymbol{R}(\boldsymbol{x})$ forms a vector of marginal excess burdens that can be implemented using a tax function $T(\boldsymbol{x})$, i.e. that the implied marginal tax rates satisfy Condition (8). Reformulation (21) of the function $\boldsymbol{R}(\boldsymbol{x})$ does not by itself identify the solenoidal component of the unique vector of marginal excess burdens that corresponds to the optimum. Formulation (21) identifies the solenoidal component of *any* vector-valued function $\boldsymbol{MEB}(\boldsymbol{x})$ which vanishes sufficiently quickly for large incomes, even if it cannot be implemented using a tax function. Condition (8) must additionally be satisfied to identify the unique optimum. We summarize these findings in the following proposition.

**Proposition 1.** *Suppose that the income space has no boundaries, so $\mathscr{X} = \mathbb{R}^2$. Then the optimal vector of marginal excess burdens is uniquely identified by the following symmetry condition on the implied marginal tax rates:*

$$\frac{\partial T_{x_1}(\boldsymbol{x})}{\partial x_2} = \frac{\partial T_{x_2}(\boldsymbol{x})}{\partial x_1}, \tag{22}$$

*and the following equation:*

$$\boldsymbol{MEB}(\boldsymbol{x}) = \boldsymbol{I}(\boldsymbol{x}) + \boldsymbol{R}(\boldsymbol{x}), \tag{23}$$

*with distributional characteristic:*

$$\forall i = 1,2 : I_i(\boldsymbol{x}) = -\mathrm{cov}\left( \frac{x'_i - x_i}{2\pi |\boldsymbol{x} - \boldsymbol{x}'|^2}, g(\boldsymbol{x}') \right), \tag{24}$$

*and additional compensated effects:*

$$\forall i = 1,2 : R_i(\boldsymbol{x}) = -\iint_{\boldsymbol{x}' \in \mathscr{X}} \left\{ \sum_j \frac{\partial}{\partial x'_j} \left( \frac{x'_i - x_i}{2\pi |\boldsymbol{x} - \boldsymbol{x}'|^2} \right) MEB_j(\boldsymbol{x}') \right\} \mathrm{d}\boldsymbol{x}' + MEB_i(\boldsymbol{x}). \tag{25}$$

Even if we are able to provide an economic interpretation for the solenoidal term $\boldsymbol{R}(\boldsymbol{x})$, it remains a source of dissatisfaction that we are unable to rewrite it in terms of sufficient statistics, such that the marginal tax rates no longer appear on the right-hand side of equation (12). It is not exceptional that the right-hand side of an optimal tax equation depends on the marginal tax rates. Even in the one-dimensional problem, characterization (7) of the optimum explicitly depends on the marginal tax rates through the income effects in the term $g(x')$. Furthermore, the different sufficient statistics such as the gross welfare weights, the elasticities and the distributional parameters are endogenous to the tax system in place. Still, we provide some simulations in Section 5 to determine the quantitative importance of the solenoidal component $\boldsymbol{R}(\boldsymbol{x})$.

# 5 Simulations

## 5.1 Helmholtz decomposition

We saw that from an analytical perspective, it is informative to look at the Helmholtz decomposition of the excess burden. From a computational perspective, it is more convenient to follow a finite differences approach to solve partial differential equation (1) subject to boundary condition(2) and the government's budget constraint. Lehmann et al. (2021) simulate an optimum for the multidimensional taxation of couples' incomes. We will perform a Helmholtz decomposition on the output of such simulations, to get an idea of the importance of the rotational and irrotational components of the vector of optimal marginal excess burdens.

Figure 3: Marginal tax rates with and without the rotational component (proof of concept)