

As Easy as ABC?

Multidimensional Screening in Public Finance*

Sander Renes** Floris T. Zoutman***

June 28, 2016

Abstract

We characterize the second-best allocation in a Mirrleesian optimal tax model where agents differ in multiple dimensions and the planner can tax multiple goods non-linearly. We develop a new method that allows us to solve the partial differential equations that describe the optimum regardless of the dimensionality of the problem. We derive four theoretical properties of the optimum. First, the optimal tax system is described by a multidimensional version of Diamond's (1998) and Saez' (2001) *ABC*-formula. Second, the Atkinson-Stiglitz theorem does not generalize to settings where the planner screens in multiple dimensions. Third, the optimal marginal tax rate on each good depends on the consumption level of multiple goods. Fourth, a no-distortion at the top/bottom result continues to hold. A calibrated simulation on taxation of couples shows a strong positive relationship between an individual's optimal marginal tax rate and the income earned by his spouse.

1 Introduction

How should a government combine taxes on labor income with health care subsidies? What is the optimal relation between capital and labor income taxes? When should

*The authors would like to thank Felix Bierbrauer, Robin Boadway, Eva Gavrilova, Aart Gerritsen, Nathan Hendren, Yasushi Iwamoto, Bas Jacobs, Laurence Jacquet, Etienne Lehmann, John Morgan, Nicola Pavoni, Emmanuel Saez, Dominik Sachs, Dirk Schindler, Bauke Visser, Casper de Vries, Hendrik Vrijburg and Nicolas Werquin for useful suggestions and comments on an earlier version of this paper. Furthermore, this paper benefited from comments and suggestions made by participants at the 2011 Nake Conference, Utrecht, the 2013 CESifo Area Conference on Public Economics, Munich, the 69th IIPF Conference, Taormina, the annual meeting of the European Economic Association 2015, Mannheim; and seminar participants at the Erasmus School of Economics, the Norwegian University of Science and Technology, the Norwegian School of Economics, the University of Konstanz, the University of Mannheim and the Center for European Economic Research (ZEW). All remaining errors are our own. Renes gratefully acknowledges financial support from the Netherlands Organisation for Scientific Research (NWO) under Open Competition grant 400-09-338. Zoutman gratefully acknowledges financial support from NWO under Open Competition grant 400-09-383.

**SFB 884 Political Economy of Reforms, University of Mannheim. Corresponding author. E-mail: s.renes@uni-mannheim.de. Sonderforschungsbereich 884, University of Mannheim, L13, 17 - Room 425, 68131 Mannheim, Germany.

***NHH Norwegian School of Economics, Department of Business and Management Science. E-mail: floris.zoutman@nhh.no

housing subsidies depend on wealth and income? The optimal interplay between these tax and subsidy instruments crucially depends on the dimensions in which agents differ. However, up to now, the literature on optimal taxation has almost exclusively focused on models where agents differ in only one dimension, namely ability (e.g. Mirrlees, 1971; Diamond, 1998; Saez, 2001; Bovenberg and Jacobs, 2005), and on models where agents do differ in multiple dimensions but the government can only tax labor income non-linearly (e.g. Saez, 2002; Choné and Laroque, 2010; Rothschild and Scheuer, 2013; Jacquet and Lehmann, 2014). One reason for this relatively narrow focus is provided in Mirrlees (1976, p.344) when he remarks:

"the attractive simplicity of the [unidimensional] marginal-tax-rate equation has to some extent been lost in its generalized [multi-dimensional] form A principal purpose of this section [on multi-dimensional taxation] has been to warn that the results of the previous section depend for their relative simplicity on the one-parameter assumption."

In other words, he considered solving the multidimensional optimal tax problem as considerably more complex than the unidimensional problem.

However successful the unidimensional approach has been, economists have since recognized that it is important to expand our models to allow for both multidimensional heterogeneity between agents, and non-linear taxation of multiple goods (see a.o. Sandmo, 1993; Saez, 2002; Judd and Su, 2006; Lockwood and Weinzierl, 2012). Recent policy discussions on health care subsidies in the US, leading up to and following the introduction of the affordable care act, have made the call to extend our models all the more urgent. This act introduces a universal health care insurance system that includes income-dependent subsidies. The act thus contains multiple instruments to redistribute from rich to poor, and from healthy to sick. Moreover, the best-selling book Piketty (2014), as well as recent developments in the literature, e.g. Farhi et al. (2012) and Piketty and Saez (2013), have fueled the discussion on taxation of capital income. In this case, the government combines a non-linear tax on labor and capital income in an effort to redistribute both from high to low ability, and from wealthy to poor.

In this paper, we return to the Mirrleesian model of taxation where the government taxes multiple goods non-linearly and screens in multiple characteristics of the agents. Our main contribution is the development of a new solution method that allows us to solve the partial differential equations (PDEs) that describe the optimal allocation analytically. Using our method we derive a formula for the optimal marginal tax rates that is similar to the *ABC*-formula for the unidimensional model (see Diamond, 1998; Saez, 2001). The formula allows us to derive several general, theoretical properties of the optimal allocations, and to perform a simple simulation calibrated to calculate the optimal non-linear tax system for couples.

Our model follows the basic set-up of Mirrlees (1976, sect. 4). Agents differ in multiple independent characteristics and their type is private information. Agents make a vector of continuous, taxable choices pertaining to labor income, consumption of health care products and savings, for instance. We will refer to these choices as goods, even though they can be both inputs and outputs to the production process. For technical convenience we restrict the preferences and choice set of the agents to ensure that the planner can (ex-post) deduce all the hidden characteristics of the agent by observing its bundle of goods in an incentive-compatible allocation. Moreover, we assume that the preference for at least one good is separable from the other goods as well as the type of the agent. This good acts as the numeraire (untaxed) commodity. The social planner maximizes a

(weakly) concave sum of the agents' utility levels.¹

We solve the optimization problem of the social planner by treating it as a multi-dimensional screening problem. In the screening problem, each agent sends a message concerning his type to the planner, and the planner assigns a bundle of goods to each agent on the basis of the message it receives. We use a first-order approach to characterize the planner's optimum. That is, the planner maximizes social welfare subject to the first-order incentive compatibility constraint, assuming that the identified optimal allocation satisfies second-order incentive constraints.²

The first-order conditions of the planner's problem form a set of PDEs. We show that the PDEs can be rewritten in the form of a Poisson equation, which in turn can be solved through a Green's function. For most relevant type spaces there exists an analytical Green's function that solves the Poisson equation. In this paper we provide the Green's function, and hence provide an analytical solution for the PDEs, for the cases where the type space equals \mathcal{R}_+ , \mathcal{R}_+^2 and \mathcal{R}^J . Analytical Green's functions for other type spaces can readily be obtained from the mathematical literature. To our knowledge, the application of the Green's function method is novel in the fields of mechanism design and optimal taxation. We are, hence, the first to provide a general approach to analytically solving multi-dimensional screening models.

We use the solution to the PDEs to describe the wedges between the marginal rate of substitution and the marginal rate of transformation in the optimal allocation. At each point in the type space, the optimal wedge for each good is the sum of characteristic-specific wedges. For example, if the wedge on labor income redistributes from high earnings ability to low earnings ability, as well as from healthy to sick types, the optimal wedge on income can be decomposed in a health-specific and an ability-specific wedge.

Each of these characteristic-specific wedges can be written as a generalized version of Diamond's (1998) and Saez' (2001) *ABC*-formula. In the formula, *A* measures the change in the elasticity of the marginal rate of substitution of the good w.r.t. the numeraire, when the characteristic changes. It serves as a measure of the quality of the signal obtained from the good. If the marginal rate of substitution is strongly affected by the characteristic, the good reveals a lot of information about the characteristic, and hence the wedge should be large. When utility is additively separable as in Armstrong (1996); Rochet and Choné (1998) the *A*-term can be rewritten in terms of empirically observable elasticities. *B* Measures the redistributive benefit of marginally distorting the price of the good. Its value is determined by the marginal social welfare weights (see Saez and Stantcheva, 2016) and a weighting function. The weighting function is the derivative of the Green's function. The weighting function describes how utility is redistributed along the type space when the wedge at a specific point in the type space is marginally increased without breaking incentive compatibility. For instance in the unidimensional

¹A recent and growing literature discusses optimal tax models where agents differ in multiple dimensions, but the planner can only observe labor income (e.g. Choné and Laroque, 2010; Jacquet and Lehmann, 2014). In that setting there is not enough information to deduce all hidden characteristics even if the allocation is incentive compatible.

²The reinterpretation of the optimal tax problem as a screening problem with only first-order constraints is not completely innocuous. In our derivation we do not explicitly consider the optimal-tax system that implements the allocation in a market economy, nor do we address second-order constraints. In our simulations we ensure that second-order incentive constraints are not binding, while we know from Renes and Zoutman (2014) that a direct relation between the taxes and wedges exists. For technical convenience we discuss both issues separately in subsections 4.5 and 4.6.

case, an increase in the wedge at n' transfers utility from types above this point to types below this point (see Diamond, 1998). Finally, C is the inverse of the size of the tax base for which the marginal choice is distorted. Like in the unidimensional model, the generalized ABC -formula thus facilitates an intuitive explanation of the wedges in terms of equity and efficiency, and can moreover serve as a convenient way to approach data.

A corollary to our main proposition shows that the optimal wedge on a good is zero if this good does not reveal information about any of the hidden characteristics. As a direct result, if all hidden characteristics only affect the preferences of a single good, the optimal tax rate on each of the other goods should equal zero. Under unidimensional heterogeneity this result is known as the Atkinson-Stiglitz (A-S) theorem (see Atkinson and Stiglitz, 1976). The corollary also implies that the A-S theorem does not generalize to a setting where the planner screens in several dimensions of heterogeneity. With at least two relevant sources of heterogeneity, the planner can never extract all the relevant private information from an agent through a single choice and hence will want to use several tax instruments.

Mirrlees (1976, sect. 3) shows that if agents are heterogeneous in only one dimension, the optimal wedge on each good can be written as a function that only depends on the consumption of that good. This is no longer possible in multidimensional screening models. In order to facilitate full revelation of J hidden characteristics, the optimal wedge is a function of J goods.³ This type of interdependency is common in real-world tax systems. For instance, health care subsidies often depend both on income as well as on the amount of health care consumed. These interdependencies are weakly sub-optimal in the unidimensional optimal tax literature, but optimal in our multi-dimensional tax model.

Intertemporal interdependencies between marginal tax rates have recently received a lot of attention in the New Dynamic Public Finance (NDPF) (see Golosov et al., 2007 for an overview). These models extend the Mirrleesian framework to a setting where agents' types evolve stochastically over time. One of the most striking outcomes in this literature is that the optimal marginal tax rate in any period depends on the income earned in past periods (see e.g. Kocherlakota, 2005). We show that this result generalizes to a deterministic dynamic setting. In particular, if agents know their entire (future) history of types at the beginning of the first period, the optimal tax rate in each period does not just depend on current and past earnings, but in addition it even depends on future earnings. Earnings in each period provide information on the type of the agent, and hence, allow the planner to screen more accurately. This also provides an intuitive explanation for the interdependencies identified in the NDPF literature. In NDPF models agents do not know the outcome of future shocks, but they do differ in both their current type and their history of types. To screen for all this heterogeneity the optimal tax rate in the current period must therefore depend on both current and previous periods' earnings.

We also derive a generalization of the no-distortion at the top result (see e.g. Sadka, 1976 and Seade, 1977). As in the unidimensional case, the optimal wedges at the extreme points of the type distribution are zero. If a type exists that has extreme values for all characteristics, his optimal marginal wedge on all goods equals zero. Intuitively, since there are no "more extreme" types, setting a wedge to separate out more extreme types yields no information to the planner. Hence, for any marginal distortion, the

³This result describes the general case, special cases may exist where the optimal tax rate can be written in a simpler form.

efficiency cost of the distortion is higher than the welfare gain at the extreme points of the distribution. Note that, unlike in the unidimensional case, in the multidimensional case such types do not necessarily exist. For instance, the healthiest person in the economy may not be the person with the highest ability in the economy. In that case, the health-specific wedges are zero for the healthiest type in the economy, while the ability-specific wedges equal zero for the type with the highest ability.

Finally, we use our *ABC*-formula to simulate a model of joint taxation of couples. Our simulations are based on a simple fixed-point algorithm in the spirit of the unidimensional algorithm in Mankiw et al. (2009). The important advantage of our algorithm over previous algorithms discussed in Mirrlees (1976) and Golosov et al. (2014), is that it does not rely on solving a partial differential equation numerically. Instead, we use the analytical solution obtained by means of the Green's function. We calibrate our model to high income couples in the US, in a setting where both spouses choose their labor supply to maximize a joint, quasi-linear utility function. The ability of the spouses is positively, but not perfectly correlated.

Even though we assume the couple's utility function is separable in each of the spouses labor supply choices, we nevertheless find that the optimal tax rate exhibits large interdependencies. Top income earners that are married to spouses with close to zero income face a marginal tax rate of about 25 percent. However, if both spouses are top income earners their marginal top tax rate should be around 65 percent. The simulation shows that interdependencies in tax rates, in this case between the income of the two partners, may be of quantitative importance.

The rest of this paper is organized as follows. The next section discusses related literature. Section 3 introduces the model and discusses incentive compatibility in our setting. The fourth section introduces the Green's function method and derives the optimal allocation and the *ABC*-formula. Section 5 discusses our generalization of the Atkinson-Stiglitz theorem. Section 6 derives boundary conditions. Section 7 compares our results to results obtained in the NDPF. Section 8 describes the simulations. The final section concludes. Long proofs are relegated to the appendix.

2 Related Literature

In a concurrent working paper, Golosov et al. (2014) also investigate an optimal tax model where agents choose multiple goods and have multiple hidden characteristics. In their model they specify the set of tax instruments available to the planner ex ante, and use a variational approach to derive the functional derivative towards a marginal perturbation of the tax-schedule. This derivative can serve as a theoretically grounded measure to evaluate the welfare effects of a tax reform. By setting this derivative equal to zero, one obtains a set of PDEs that implicitly determine the optimal-tax system. The variational approach adopted in Golosov et al. (2014) fundamentally differs from the mechanism-design approach we use in this paper. Our mechanism-design approach allows us to derive the allocation that maximizes welfare under a set of informational restrictions, without specifying the instruments required to attain this allocation. The variational approach allows one to find the optimal allocation for a given set of instruments, without specifying the informational restrictions. When the instrument set is restricted for reasons that are difficult to model, the variational approach is an attractive method to derive the

optimal restricted tax system.⁴ However, mechanism design has proven to be better suited in guiding policy makers in the choice of instruments to apply, since no a priori restrictions are necessary. One clear example of this approach is our investigation of the Atkinson-Stiglitz theorem in section 5. We show when one needs more than one non-linear tax-instrument and relate this to the fundamentals of our model economy. A similar analysis through a variational approach would be extremely cumbersome, since the set of instruments to use or exclude is not clearly defined ex ante. Hence, while the variational approach can have great practical value in deriving optimal tax systems when the instruments are given, the mechanism design approach is more useful in deriving fundamental properties of the optimal tax system. Finally, it should be emphasized that in many cases the mechanism-design approach and the variational approach yield identical outcomes (e.g. Saez, 2001). In those cases, the method we develop in this paper has the distinct advantage that we can solve the resulting PDEs, whereas Golosov et al. (2014) provide the optimal tax system in the form of a PDEs.

We rely on the first-order approach to derive the properties of the second-best allocation. Another approach to keep the model tractable is to discretize the type space.⁵ In a model with discretely distributed types it is possible to (numerically) verify which incentive constraints are binding, such that the optimal allocation can be derived without relying on the first-order conditions alone (see e.g. Cremer et al., 2001, 2015). The downside of a discrete distribution is that the optimal wedge can only be determined on a discrete number of points. Moreover, as the number of types increases, the problem becomes less tractable. In our solution it is possible to calculate the wedge for all goods in a tractable manner, because types and choices are continuously distributed. This provides the entire shape of the optimal tax system.

Kleven et al. (2009) study the taxation of couples in a setting where partners have different earnings ability. To maintain analytic tractability they assume the primary earner chooses labor supply on the intensive margin while the secondary earner chooses on the extensive margin. In our model agents only make intensive-margin choices. We argue that many economic decisions, such as savings and consumption choices, are more accurately portrayed as choices on the intensive margin. The best solution would be to combine both approaches by extending our model with extensive-margin decisions, as was done for the unidimensional model in Jacquet et al. (2013). However, we leave this for future research.

Several papers study multidimensional screening in a setting where the number of observable choice variables is smaller than the number of characteristics. In such a setting, it is not possible to reveal all hidden characteristics, even if the allocation is incentive compatible. Pass (2012) develop a method where agents of a different type are pooled together based on their marginal preferences, or some convex combination of their characteristics. For a given “pooling function” one can integrate out the excess dimensions in the type space, until it equals the dimensionality of the choice space. Choné and Laroque (2010) apply this method to calculate the optimal income tax rate in a setting where agents differ in both ability and disutility of labor. Jacquet and Lehmann (2014) extend the method further, by identifying the optimal pooling function in a setting where multiple dimensions of the type space affect the behavioral elasticities of the agents. This

⁴In the Ramsey-tax model, for example, the planner is exogenously restricted from using lump-sum taxation.

⁵See Armstrong and Rochet, 1999 for a user’s guide.

allows them to calculate the optimal tax rates in a setting where agents differ in both the elasticity of labor supply and their earnings ability. To limit the complexity of our model we restrict our attention to the case where all hidden characteristics can be revealed. However, in future research it may be possible to use a pooling function in the spirit of Pass (2012); Jacquet and Lehmann (2014) to extend our results to a setting where not all characteristics can be revealed.

Rothschild and Scheuer (2011, 2013, 2014); Gomes et al. (2014); Scheuer (2013, 2014) study a setting where agents face different productivity levels in different sectors of the economy, and hence differ in multiple dimensions. The planner can only tax labor income non-linearly. In these papers, general equilibrium effects exist because agents can shift effort from one sector to the other. This forces the planner to adjust his optimal tax rates, compared to the standard Mirrleesian tax rates, to reduce inefficient shifting. In our model we do not explicitly model different sectors, but it does nest a model where each agent decides on the intensive margin of effort in different sectors simultaneously.

3 The Model I, agents

We use a mechanism-design approach to find the optimal-tax system as has been the standard in the literature since Mirrlees (1971, 1976). We first find the second-best allocation using a direct mechanism. In the direct mechanism the agents send a message to the planner that reveals their type. On the basis of this message, the planner assigns the agents with a bundle of goods. The planner chooses the bundles in an effort to maximize a Bergson-Samuelson welfare function of agents' utility levels subject to resource and incentive-compatibility constraints. After that we use the characterization of the second-best allocation to identify the optimal tax system.

This section formally introduces the preferences of the agents and derives the incentive-compatibility constraints. The subsequent section introduces the preferences of the planner, solve its maximization problem, and discusses the optimal-tax equation.

3.1 Preferences

The economy is populated by a unit mass of agents whose preferences can be characterized by a twice-differentiable utility function:

$$u(\mathbf{x}, y, \mathbf{n}).$$

$\mathbf{x} \in \mathbf{X} \subseteq \mathcal{R}^I$ denotes a vector of I choice variables. These choice variables can, for instance, include effective labor supply, consumption of health care products or savings. Choices in \mathbf{x} are indexed by i . $y \in Y \subseteq \mathcal{R}$ is the numeraire commodity and is therefore untaxed. Throughout the paper we will refer to the choice variables in $\{\mathbf{x}, y\}$ as goods, even though they can be both inputs and outputs to the production process. The planner can observe each agent's consumption of goods \mathbf{x} and y perfectly.

$\mathbf{n} \in \mathbf{N} \subseteq \mathcal{R}^J$ Denotes an J -vector of characteristics that forms the type of the agent. Elements in the type vector are denoted as n_j and indexed by j . Characteristics can for instance include earnings ability, health status and preference parameters. We assume each characteristic denotes some independent aspect of the agents' personality, such that

no characteristic can be found as a deterministic function of the other characteristics.⁶

An agent's type is private information and can not be observed by the social planner. However, the distribution is assumed to be known by both the agents and the planner. For technical convenience we assume that the type space, \mathbf{N} , is convex and piecewise smooth. The distribution of \mathbf{n} is then given by the twice-differentiable cumulative density function $F(\mathbf{n})$ with $F : \mathbf{N} \rightarrow [0, 1]$ and associated probability density function $f(\mathbf{n})$. Both are defined over the closure of \mathbf{N} . We assume $f(\cdot) > 0$ in the interior of \mathbf{N} .⁷

For simplicity we assume y is a normal good, such that $u_y > 0$, $u_{yy} \leq 0$ for any value of $\{\mathbf{x}, y, \mathbf{n}\}$. This assumption implies non-satiation of the utility function everywhere. Moreover, we assume utility is separable in y such that cross-derivatives of $u(\cdot)$ with respect to y equal zero, $u_{yn_j} = u_{yx_i} = 0$ for all x_i and n_j . Since any good can be chosen as the numeraire, this assumption implies that there must exist at least one good for which the marginal utility is independent of all other goods and characteristics. These assumptions greatly simplify the multi-dimensional screening problem faced by the social planner (see also Basov, 2006 in this respect).

Note that we do not restrict ourselves to static models: different goods can be consumed in different periods. However, we do assume that agents know all characteristics of their type and the direct mechanism at the beginning of the first period.⁸

We need two additional assumptions to ensure full separation of types can occur in a direct mechanism. These assumptions are our screening conditions. If they hold, the government can screen for the unobserved characteristics via the choices made by the agents. This allows a direct link between the optimal screening solution and market outcomes. First, we assume that $I \geq J \geq 1$, such that there are at least as many goods in \mathbf{x} as characteristics in \mathbf{n} . This ensures the goods space is large enough to contain all information of the type space in an incentive compatible manner.

Second, let:

$$\mathbf{s}(\mathbf{x}, y, \mathbf{n}) \equiv -\frac{u_{\mathbf{x}}(\mathbf{x}, y, \mathbf{n})}{u_y(y)},$$

denote the vector of shadow prices, such that each element, $s_i(\cdot)$, denotes the marginal rate of substitution for decision variable x_i with respect to the numeraire good y . We assume the Jacobian matrix of partial derivatives of \mathbf{s} towards \mathbf{n} , denoted by \mathbf{s}_n , is of full rank for any combination $\{\mathbf{x}, y, \mathbf{n}\}$, i.e. $rank(\mathbf{s}_n) = J$. This assumption ensures there are no characteristics that do not influence marginal preferences. In addition, it excludes the possibility of having two characteristics that jointly affect the preference for only one good. An example of the latter is a model where agents differ in their degree of earnings ability and in their opportunity cost of work as in Choné and Laroque (2010). In their model the utility cost of providing a unit of effective labor supply is decreasing in ability and increasing in the opportunity cost of work. Since both characteristics act only on effective labor supply, it is fundamentally impossible to separate them both through

⁶Note that the conventional utility representation $\tilde{u}(y, L)$ with L denoting labor supply, is a special case of our utility representation (e.g. Mirrlees, 1971; Saez, 2001). If one takes the standard assumption that gross income equals $x_1 = n_1 L$ where n_1 is earnings ability, it can be seen that this utility function can be rewritten into our form: $\tilde{u}(y, L) = \tilde{u}\left(y, \frac{x_1}{n_1}\right) = u(x_1, y, n_1)$

⁷Ebert (1992) and Hellwig (2010b) show that in case of unidimensional heterogeneity of agents, bunching of types generally occurs when $f(\mathbf{n})$ is non differentiable, or zero on the interior.

⁸The model with unidimensional heterogeneity has often been used to describe a dynamic economy. See Golosov et al. (2013) for a recent example.

this one choice. By assuming \mathbf{s}_n is of full rank, we guarantee that there is always a second observable choice which can be used to disentangle the effect of ability and the opportunity cost of work. For example, if the planner could also observe the time spend on video games and the opportunity cost of working increases in the preference for video games, the screening problem can be solved. In that case, the planner can deduce both characteristics by jointly observing labor earnings and the time spend on video games.

For bookkeeping, the Jacobian of first-order derivatives $\phi'(\cdot)$ of any function $\phi(\cdot) : \mathcal{R}^a \rightarrow \mathcal{R}^b$, is of dimension $b \times a$. The second-order derivatives $\phi''(\cdot)$ are of dimension $ab \times a$. For any multi-vector functions $\psi(\mathbf{z}^1, \mathbf{z}^2, \dots) : \mathcal{R}^{a^1} \times \mathcal{R}^{a^2} \dots \rightarrow \mathcal{R}$ the vectors of first-order derivatives ψ_{z^k} are of dimension $a^k \times 1$ and the matrix of second-order derivatives $\psi_{z^k z^l}$ are of dimension $a^k \times a^l$, where the dimension of the matrix follows the order of the subscripts. Superscript T denotes the transpose operator. Vectors and multidimensional constructs are denoted in bold. Scalars and derivatives of scalar functions are in normal font throughout the paper.

3.2 Incentive Compatibility

Before we go to the problem faced by the social planner, we need to consider the problem of the agents in our economy. In particular, we derive conditions under which an allocation is incentive compatible. The incentive compatibility constraints will serve as constraints to the social planner when he chooses the optimal allocation.

In a direct mechanism the social planner offers bundles $\{\mathbf{x}(\mathbf{m}), y(\mathbf{m})\}$ for all $\mathbf{m} \in \mathbf{N}$. The set of all bundles, describing the complete pattern of consumption in the economy is the allocation. The allocation is described by functions $\mathbf{x}(\mathbf{m})$ and $y(\mathbf{m})$. Function $\mathbf{x}(\cdot)$ maps from the message space to the goods space, $\mathbf{x} : \mathbf{N} \rightarrow \mathbf{X} \subseteq \mathcal{R}^I$ and $y(\mathbf{m})$ maps from the message space to the numeraire-commodity space, $y : \mathbf{N} \rightarrow Y \subseteq \mathcal{R}^1$.

Each agent selects a bundle $\{\mathbf{x}(\mathbf{m}), y(\mathbf{m})\}$ from the allocation by sending a message $\mathbf{m} \in \mathbf{N}$ to the social planner. An allocation is incentive compatible if each agent truthfully reveals all his unobserved characteristics and receives the bundle designed for him:

$$\mathbf{n} = \arg \max_{\mathbf{m}} u(\mathbf{x}(\mathbf{m}), y(\mathbf{m}), \mathbf{n}) \quad \forall \quad \mathbf{n} \in \mathbf{N}. \quad (1)$$

Moreover, define indirect utility as follows:

$$V(\mathbf{n}) \equiv \max_{\mathbf{m}} u(\mathbf{x}(\mathbf{m}), y(\mathbf{m}), \mathbf{n}) \quad \forall \quad \mathbf{n} \in \mathbf{N}. \quad (2)$$

Proposition 1, which largely follows Mirrlees (1976) and McAfee and McMillan (1988), derives the first and second-order conditions to problem (1).

Proposition 1 *An interior twice-differentiable allocation is incentive compatible, and hence solves (1) if:*

$$\mathbf{y}'(\mathbf{n}) = \mathbf{s}(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})^T \mathbf{x}'(\mathbf{n}) \quad \forall \quad \mathbf{n} \in \mathbf{N}, \quad (3)$$

$$\mathbf{x}'(\mathbf{n})^T \mathbf{s}_n \leq \mathbf{0} \quad \forall \quad \mathbf{n} \in \mathbf{N}, \quad (4)$$

where $\leq \mathbf{0}$, signifies negative semi-definiteness of the matrix.

Through the envelope theorem a fully equivalent set of conditions can be derived:

$$V'(\mathbf{n}) = u_n(\mathbf{x}(\mathbf{n}), \mathbf{n})^T \quad \forall \quad \mathbf{n} \in \mathbf{N}, \quad (5)$$

$$u_{nn}(\mathbf{x}(\mathbf{n}), \mathbf{n}) - V''(\mathbf{n}) \leq 0 \quad \forall \quad \mathbf{n} \in \mathbf{N}. \quad (6)$$

Proof. *The proof can be found in the appendix.* ■

Equation (3) states that an agent should be indifferent between truth telling and mimicking at the margin for each of his characteristics. For each row j the left-hand side of the equation denotes the gain in y as a consequence of marginally changing the reported characteristic n_j . The right-hand side denotes the loss in \mathbf{x} , measured in y units of utility, for the same marginal change in the message. Therefore, equation (3) states that in equilibrium the marginal cost of mimicking should equal the marginal benefits for all characteristics and for all types in the economy. Equation (4) is the usual second-order condition as first derived in Mirrlees (1976). If the marginal rate of substitution for decision variable x_i is increasing (decreasing) in characteristic n_j , $\frac{\partial s_i}{\partial n_j} > 0$ ($\frac{\partial s_i}{\partial n_j} < 0$) and the allocated amount of the good is decreasing (increasing) in the characteristic, $\frac{\partial x_i}{\partial n_j} > 0$ ($\frac{\partial x_i}{\partial n_j} < 0$), the allocation induces self selection. The equation implies higher (lower) quantities of the good are assigned to people with a stronger (weaker) preference for the good.

Equations (5) and (6) are fully equivalent formulations of the incentive constraints derived through the envelope theorem. Together they imply that the utility of agents that truthfully reveal their type increases at least as fast as the utility of agents that mimic 'lower' types. Although their explanation is less intuitive, (5) and (6) are extremely convenient mathematical expressions in the derivations in subsequent sections.

If the inequality in (4) holds strictly, our initial assumption that \mathbf{s}_n is of rank J , implies that full separation occurs on the market. This is formalized in the next lemma:

Lemma 1 *If the allocation is incentive-compatible and satisfies (4) strictly, all characteristics are revealed through the bundles chosen by the agents in the direct revelation mechanism.*

Proof. Note that (4) can only hold strictly, i.e. be a definite inequality, if the product $\mathbf{x}'(\mathbf{n})^T \mathbf{s}_n$ is definite, and hence of full rank J . Since in a matrix product $\text{rank}(\mathbf{AB}) \leq \min(\text{rank}(\mathbf{A}), \text{rank}(\mathbf{B}))$, it follows that (4) can only hold strictly if the Jacobian of the allocation, $\mathbf{x}'(\mathbf{n})$, is of full rank J for all values $\mathbf{n} \in \mathbf{N}$. Since $I \geq J$ it follows that the allocation is locally invertible around type \mathbf{n} for all $\mathbf{n} \in \mathbf{N}$. Hence, at least one inverse function from the image of $\mathbf{x}(\mathbf{n})$ to the type space exists: $(\mathbf{x})^{\leftarrow} : \text{image}(\mathbf{x}(\cdot)) \rightarrow \mathbf{N}$. It follows that by observing the bundle chosen by an agent, one can deduce each of his hidden characteristics. ■

By lemma 1, if the second order incentive constraints (4) are strictly satisfied it follows that the type of the agent can be deduced by observing all his choice variables. This is convenient for our analysis, since it allows us to relate optimal policy to the unobserved underlying characteristics, as well as to observable goods.

4 Model II, The Second-Best Allocation

Now that we have established conditions for incentive compatibility, we can turn our attention to the social planner. In our derivation we apply the first-order approach, and assume that the second-order incentive compatibility conditions (4) are satisfied strictly such that lemma 1 holds. In subsection 4.6 we discuss potential violations of second-order constraints in more detail.

4.1 The social planner

We assume the social planner maximizes a (weakly) concave sum of agent utility:

$$SW = \int_{\mathbf{N}} W(u(\mathbf{x}, y, \mathbf{n})) dF(\mathbf{n}), \quad (7)$$

$$W'(\cdot) > 0, W''(\cdot) \leq 0, \quad (8)$$

where $\int_{\mathbf{N}} \cdots d\mathbf{n}$ is short-hand notation for the multiple integral with respect to each of the characteristics. The social planner commits to the allocation he offers and cannot alter the allocation after types are revealed.⁹ Redistribution is welfare increasing because of (at least one of) two reasons. First, concavity in the utility functions of the agents implies that agents with higher income, measured through the numeraire y , have a lower marginal utility of income. Second, $W''(\cdot) < 0$ implies the social planner gives a higher welfare weight to agents with lower utility. If, on the contrary, the utility function would be linear in income and the welfare function linear in utility, the first best would be attainable through lump-sum taxation.

The social planner is bound by the economy's resource constraint:

$$\int_{\mathbf{N}} y(\mathbf{n}) dF(\mathbf{n}) + R \leq \int_{\mathbf{N}} q(\mathbf{x}(\mathbf{n})) dF(\mathbf{n}), \quad (9)$$

where R denotes exogenous government expenditure and $q(\cdot)$ is the agent's production function which transforms goods in \mathbf{x} into units of y . A partial derivative q_{x_i} may be either positive or negative depending on whether good x_i is an input to, or an output of the production process. We assume the production technology exhibits diminishing marginal returns, i.e. $q_{x_i x_i} \leq 0$ for all goods x_i , to guarantee that an interior solution will be reached in laissez faire. This specification nests standard partial-equilibrium analysis where agents face a fixed set of prices for each of the goods. In that case the agent's production function can be written as $q(\mathbf{x}) = \mathbf{p}^T \mathbf{x}$, where \mathbf{p} denotes a price vector.

4.2 First-order conditions

In the first-order approach the social planner maximizes social welfare subject to the first-order incentive-compatibility constraint (5), the economy's resource constraint (9), and a constraint that allows us to introduce the indirect utility function as the state variable:

$$\max_{V(\mathbf{n}), \mathbf{x}(\mathbf{n}), y(\mathbf{n})} \int_{\mathbf{N}} W(V(\mathbf{n})) dF(\mathbf{n}),$$

s.t.

$$0 \geq R + \int_{\mathbf{N}} (y(\mathbf{n}) - q(\mathbf{x}(\mathbf{n}))) dF(\mathbf{n}),$$

$$V'(\mathbf{n}) = u_{\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n})^T, \quad \forall \mathbf{n} \in \mathbf{N} \quad (10)$$

$$V(\mathbf{n}) = u(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n}). \quad \forall \mathbf{n} \in \mathbf{N} \quad (11)$$

The Lagrangian to this problem is given by:

⁹See Roberts (1984) for a discussion on the issue of commitment.

$$\mathcal{L} = \int_{\mathbf{N}} [(W(V) - \lambda(R + y - q(\mathbf{x}))) f + \boldsymbol{\theta}^T (V'^T - u_{\mathbf{n}}(\mathbf{x}, \mathbf{n})) + \eta(u(\mathbf{x}, y, \mathbf{n}) - V)] d\mathbf{n},$$

where λ is the Lagrangian multiplier associated with the resource constraint, $\boldsymbol{\theta}(\mathbf{n})$ is a J -column vector of Lagrangian multipliers for the set of local incentive-compatibility constraints in (5), $\eta(\mathbf{n})$ is the Lagrangian multiplier that ensures maximized utility equals the utility function for each type. Note that in the Lagrangian \mathbf{x} , y , V , f , F , $\boldsymbol{\theta}$ and their derivatives depend on \mathbf{n} , but for clarity of exposition this notation is suppressed. Since we cannot work with the derivative of the state variable, V' , we apply the divergence theorem to rewrite the Lagrangian as:¹⁰

$$\begin{aligned} \mathcal{L} = & \int_{\mathbf{N}} \left[(W(V) - \lambda(R + y - q(\mathbf{x}))) f - V \sum_{j=1}^J \frac{\partial \theta_j}{\partial n_j} - \boldsymbol{\theta}^T u_{\mathbf{n}}(\mathbf{x}, \mathbf{n}) + \eta(u(\mathbf{x}, y, \mathbf{n}) - V) \right] d\mathbf{n} \\ & + \int_{\partial \mathbf{N}} [V \boldsymbol{\theta}^T \mathbf{e}] d\partial \mathbf{N}. \end{aligned} \quad (12)$$

Where $\partial \mathbf{N}$ denotes the boundary of \mathbf{N} and \mathbf{e} the outward unit surface normal vector to this boundary. We maximize this last equation pointwise for each type in the type space, yielding the following first-order conditions for the planner:

$$\frac{\partial \mathcal{L}}{\partial y} = 0 : -\lambda f + \eta u_y = 0 \quad \forall \quad \mathbf{n} \in N, \quad (13)$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{x}} = \mathbf{0}_I : \lambda q'^T f - u_{\mathbf{x}\mathbf{n}} \boldsymbol{\theta} + \eta u_{\mathbf{x}} = \mathbf{0}_I \quad \forall \quad \mathbf{n} \in N, \quad (14)$$

$$\frac{\partial \mathcal{L}}{\partial V} = 0 : W' f - \sum_{j=1}^J \frac{\partial \theta_j}{\partial n_j} - \eta = 0 \quad \forall \quad \mathbf{n} \in N, \quad (15)$$

$$\frac{\partial \mathcal{L}}{\partial V} = 0 : \boldsymbol{\theta}^T \mathbf{e} = 0 \quad \forall \quad \mathbf{n} \in \partial N, \quad (16)$$

These first-order conditions were previously derived in Mirrlees (1976). The quote in the introduction shows his response, a general warning that the attractive simplicity of the uni-dimensional model disappears in the general setting. The main complexities that arise, are caused by the first-order conditions for indirect utility $V(\mathbf{n})$ (15,16). In the unidimensional model these first-order conditions describe an ordinary differential equation, which can be solved analytically. However, with multi-dimensional heterogeneity these same conditions form a set of J PDEs for the Lagrange multipliers $\boldsymbol{\theta}(\cdot)$. In the next subsection we show that the solution to the PDEs depends on the shape and dimensionality of the type space \mathbf{N} . Moreover, we show that the PDEs have an analytical solution for many multidimensional type spaces that are relevant to economists.

¹⁰The divergence theorem is also known as the Gauss's theorem, or Ostrogradsky's theorem. Here it acts as a multi-dimensional version of integration by parts. Through it we can rewrite: $\int_{\mathbf{N}} \boldsymbol{\theta}^T \mathbf{V}'^T d\mathbf{n} = - \int_{\mathbf{N}} V \sum_{j=1}^J \frac{\partial \theta_j}{\partial n_j} d\mathbf{n} + \int_{\partial \mathbf{N}} [V \boldsymbol{\theta}^T \mathbf{e}] d\partial \mathbf{N}$.

4.3 Solving for multipliers on the incentive compatibility constraints

To solve for $\boldsymbol{\theta}(\cdot)$, first combine first-order conditions (13) and (15) to obtain:

$$\sum_{j=1}^J \frac{\partial \theta_j}{\partial n_j} = \frac{\lambda(g(\mathbf{n}) - 1)}{u_y(\mathbf{n})} f(\mathbf{n}) \quad (17)$$

where $g(\mathbf{n}) \equiv \frac{W'(V(\mathbf{n}))u_y(\mathbf{n})}{\lambda}$ denotes the social marginal welfare weight; the social value of redistributing an extra unit of y towards type \mathbf{n} (see Saez and Stantcheva, 2016), and $u_y(\mathbf{n})$ is used as short-hand notation for $u_y(y(\mathbf{n}))$. Equation (17) forms a set of J first-order PDEs for the vector $\boldsymbol{\theta}(\mathbf{n})$. Boundary conditions are provided by (16). In lemma 2 we show that the solution to this set of PDE exists provided:

$$\int_{\mathbf{N}} \frac{\lambda(g(\mathbf{n}) - 1)}{u_y(\mathbf{n})} f(\mathbf{n}) = 0. \quad (18)$$

This constraint can be satisfied by choosing λ as follows:

$$\begin{aligned} 0 &= \int_{\mathbf{N}} \left(W'(V) - \frac{\lambda}{u_y} \right) f d\mathbf{n}, \\ \lambda &= \frac{\int_{\mathbf{N}} W'(V) f d\mathbf{n}}{\int_{\mathbf{N}} \frac{f}{u_y} d\mathbf{n}} \end{aligned} \quad (19)$$

Hence, λ equals the average marginal increase in welfare when the economy's resource constraint is relaxed by 1 unit of y . This was already shown to hold true for the uni-dimensional model in Mirrlees (1971).

The lemma further shows that $\boldsymbol{\theta}(\mathbf{n})$ is a conservative vector field. That is, there exists a scalar function $\Theta(\mathbf{n})$ such that $\boldsymbol{\theta}(\mathbf{n}) = \Theta'(\mathbf{n})$ for all types in the type space.¹¹ The scalar function $\Theta(\cdot)$ allows us to write the set of J first-order partial differential equations as a single second order differential equation:

$$\nabla^2 \Theta(\mathbf{n}) = \frac{\lambda(g(\mathbf{n}) - 1)}{u_y(\mathbf{n})} f(\mathbf{n}) \quad \forall \mathbf{n} \in \mathbf{N}, \quad (20)$$

$$\Theta'(\mathbf{n}) \mathbf{e} = \mathbf{0} \quad \forall \mathbf{n} \in \partial \mathbf{N}, \quad (21)$$

where the operator ∇^2 denotes the Laplacian operator: $\nabla^2 = \frac{\partial^2}{\partial n_1^2} + \dots + \frac{\partial^2}{\partial n_J^2}$. Equation (20) is a type of elliptic PDE known as a Poisson equation. The set of equations (21) form boundary conditions to the gradient of $\Theta(\cdot)$, rather than to $\Theta(\cdot)$ itself and are known as Von Neumann boundary conditions in the mathematical literature. Hence, finding $\Theta(\cdot)$ amounts to solving a Poisson equation under Von Neumann boundary conditions.¹²

¹¹If we do not assume $u_{yn} \neq 0$, this is no longer the case. Without this assumption analytic solutions for the partial differential equation only exist in very special cases. See example 35 in Basov (2006) for a discussion.

¹²The properties of this class of PDEs are well-known in mathematics and physics. Good introductions to their properties and the solution technique used in this paper can be found in Strauss (1992); Evans (2010)

In the proof to lemma 2 we show how the partial differential equation for $\Theta(\cdot)$ can be solved using the method of Green's functions. Afterwards, to find $\boldsymbol{\theta}(\cdot)$ we simply take the gradient of the solution for $\Theta(\cdot)$. These results, as well as some properties of the solution for $\boldsymbol{\theta}(\cdot)$ are formalized in lemma 2.

Lemma 2 *The set of Lagrange multipliers $\boldsymbol{\theta}(\cdot)$ follow from the solution to (20) subject to boundary constraints (21). The solution has the following six properties:*

- (i) $\boldsymbol{\theta}(\cdot)$ is a conservative-field vector with potential function $\Theta(\cdot)$,
- (ii) the solution to the potential is given by:

$$\Theta(\mathbf{n}) = \lambda \int_{\mathbf{N}} \frac{(g(\mathbf{m}) - 1) f(\mathbf{m})}{u_y(\mathbf{m})} G(\mathbf{n}, \mathbf{m}) d\mathbf{m} + D$$

where $G(\cdot)$ is the Green's function corresponding to this problem, D denotes a constant, and $\mathbf{n}, \mathbf{m} \in \mathbf{N}$,

- (iii) in our model $G(\cdot)$ is fully determined by the specification of the type space \mathbf{N} ,
- (iv) $G(\cdot)$ is twice differentiable everywhere except at $\mathbf{m} = \mathbf{n}$,
- (v) the solution to the Lagrange multipliers is therefore given by:

$$\theta_j(\mathbf{n}) = \lambda \int_{\mathbf{N}} \frac{(g(\mathbf{m}) - 1) f(\mathbf{m})}{u_y(\mathbf{m})} \omega_j(\mathbf{m}, \mathbf{n}) d\mathbf{m}, \quad (22)$$

where $\omega_j(\mathbf{n}, \mathbf{m}) \equiv \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial n_j}$,

- (vi) $\omega_j(\cdot)$ is unique up to an additive constant.

Proof. The proof can be found in the appendix. ■

Lemma 2 shows that solving for the Lagrange multipliers $\boldsymbol{\theta}(\cdot)$ can be reduced to finding the appropriate Green's function. The advantage of rewriting the PDE in this form is that, for a given type space, the Green's function is independent of the specification of the model (e.g. the utility function, distribution of characteristics or the welfare function). A Green's function generally only depends on 3 things: i.) the differential operator, ii.) the type of boundary constraints and iii.) the space of integration. For our optimal-tax problem the differential operator is always the Laplacian and the boundary constraints are always of the Neumann type. Hence, the only primitive that affects the Green's function is the space of integration: the dimensionality and shape of the type space determine the appropriate $G(\cdot)$. This is an extremely useful property of our solution method since the Green's function for the Neumann problem has already been derived for many spaces that are relevant to economists. Lemma 2 allows us to solve for $\boldsymbol{\theta}(\cdot)$ in all of these known cases. For instance, if $\mathbf{N} = \mathcal{R}^J$ with $J \geq 2$ the Green's function is given by (see e.g. Seremet, 2002):

$$\begin{aligned} G(\mathbf{n}, \mathbf{m}) &= \frac{1}{2\pi} \log(\|\mathbf{n} - \mathbf{m}\|) \quad \text{if } J = 2, \\ G(\mathbf{n}, \mathbf{m}) &= \frac{\|\mathbf{n} - \mathbf{m}\|^{2-J}}{(2-J)\sigma_J} \quad \text{if } J > 2, \end{aligned}$$

where $\|x\|$ denotes the Euclidean distance of vector x , and σ_J is the surface area of the unit sphere in \mathcal{R}^J . The Green's function is somewhat more difficult to determine for bounded type spaces (we provide two examples below). However, Seremet (2002)

provides an extensive collection of Green's Functions for both bounded and unbounded spaces, which can readily be applied in any optimal-tax, or multi-dimensional screening problem. Hence, lemma 2 provides the solution to the partial differential equation that determines the optimal allocation for a large number of problems that involve multi-dimensional heterogeneity.

To interpret equation (22), recall that $g(\mathbf{m})$ denotes the monetized welfare gain of marginally increasing the transfer of y to type $\mathbf{m} \in \mathbf{N}$. The marginal cost of public funds equals 1 in the optimum (see Jacobs, 2010). Hence, $g(\mathbf{m}) - 1$ denotes the net welfare gain of transferring a marginal unit of public funds to type \mathbf{m} in monetary terms. Hence, $\frac{g(\mathbf{m})-1}{u_y(\mathbf{m})}$ measures the monetary welfare gain of transferring a unit of utils to type \mathbf{m} . Multiplication by λ translates the value into units of welfare. Through multiplication by $f(\cdot)$ population weights are given to each point in the integral.

Finally, $\omega_j(\mathbf{n}, \mathbf{m})$ is a weighting function. It gives a weight to each type within the integral, where the weight depends on type \mathbf{m} relative to the type \mathbf{n} . Intuitively, $\theta_j(\mathbf{n})$ is the Lagrange multiplier of the incentive constraint for characteristic j at point \mathbf{n} , and hence, measures the marginal welfare gain associated with a marginal relaxation of incentive constraint j at \mathbf{n} . When the incentive constraint at \mathbf{n} is marginally relaxed, this has incentive effects at all points in the vicinity of \mathbf{n} . The relaxation allows the planner to marginally decrease utils for some types and increase utils for others. $\omega_j(\mathbf{n}, \mathbf{m})$ provides the pattern of marginal gains and losses over the type space, i.e. for all $\mathbf{m} \in \mathbf{N}$, that occurs as a result of the relaxation in the constraint at one particular point \mathbf{n} for which the incentive effects on the rest of the space are off-set. The multiplication between $\omega_j(\cdot)$ and the other terms denotes the welfare effect of these extra utility transfers that occur when the incentive constraint is locally relaxed.

To provide more insight into the solution (22) we discuss two particular weighting functions $\omega_j(\mathbf{n}, \mathbf{m})$ in detail. We study the case where the type space is given by $\mathbf{N} = \mathcal{R}_+$, such that there is one hidden characteristic that can only take positive value, and the case where $\mathbf{N} = \mathcal{R}_+^2$. The first example is useful for instructional purposes. Since the results are identical to those obtained in the existing literature, it allows the reader to directly verify the results of our approach. The second case, with $\mathbf{N} = \mathcal{R}_+^2$, is the most obvious expansion of the unidimensional model, and is therefore useful for applied researchers.

4.3.1 Case 1: Unidimensional heterogeneity

With $\mathbf{N} = \mathcal{R}_+$ there is unidimensional heterogeneity in a single characteristic n and we are in the setting of Mirrlees (1971). The weighting function for this problem is given in example 1.

Example 1 *When $\mathbf{N} = \mathcal{R}_+$ the weighting function is:*

$$\omega(n, m) = H(n - m) \quad \forall \quad n, m \in \mathcal{R}_+, \quad (23)$$

where $H(x) = 0$ when $x \leq 0$ and $H(x) = 1$ when $x > 0$.

Proof. The proof can be found in the appendix. ■

Figure 1 graphically displays the weighting function for a fixed value of n . In the graph, m is on the horizontal axis. As can be seen the function takes value 1 for $m < n$

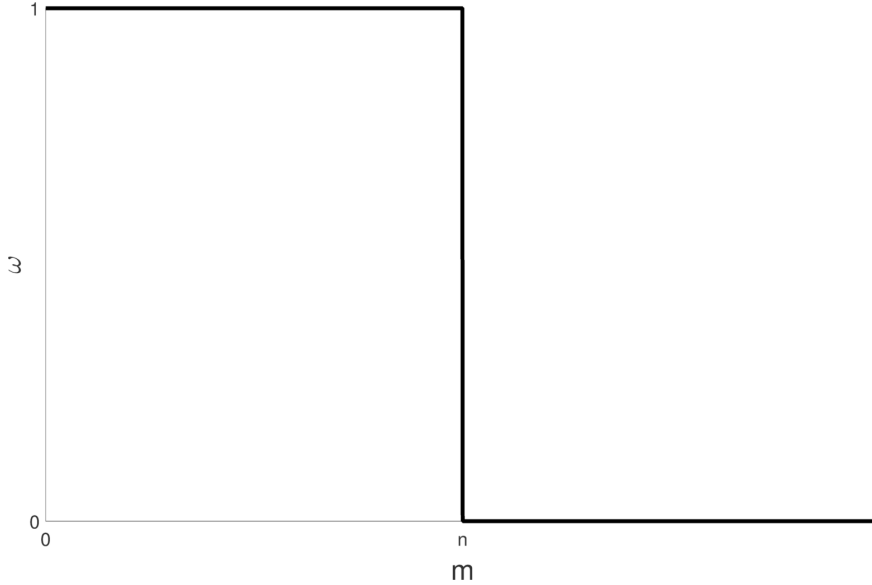


Figure 1: Weighting function for \mathcal{R}_+^1

and value zero thereafter. Combining weighting function (23) with equation (22) we arrive at the unidimensional solution for $\theta(\cdot)$:

$$\theta(n) = \lambda \int_0^\infty H(n-m) \frac{(g(m)-1)f(m)}{u_y(m)} dm. \quad (24)$$

Since, the weighting function takes value 0 for all $m \geq n$, we can simplify this expression to:

$$\theta(n) = \lambda \int_0^n \frac{(g(m)-1)f(m)}{u_y(m)} dm. \quad (25)$$

This solution for the multiplier is equivalent to the one found in Mirrlees (1971).¹³

As can be seen from equation (25), $\theta(n)$ measures the redistributive benefit when the transfer to all types $m < n$ is increased by a marginal unit. Intuitively, when the incentive-compatibility constraint is marginally relaxed at n , this allows the planner to increase the amount of resources transferred to all agents with $m < n$ uniformly. Hence, the weighting function takes value 1 for all $m < n$ and value zero for all $m > n$. With this interpretation in mind we now turn to the weighting function when $\mathbf{N} = \mathcal{R}_+^2$.

4.3.2 Case 2: Bidimensional heterogeneity

Assume $\mathbf{N} = \mathcal{R}_+^2$ such that each type \mathbf{n} has two characteristics, $n_1, n_2 \in [0, \infty]^2$. In this case the weighting function is given in the following example:

¹³Following Diamond (1998) the convention is to present the solution as:

$$\theta(n) = \lambda \int_n^\infty \frac{(1-g(m))f(m)}{u_y(m)} dm. \quad (26)$$

To arrive at that solution, we simply rewrite our weighting function $\omega(n, m)$ from eq. (23). Since the weighting function is unique up to an additive constant, we can rewrite as:

$$\hat{\omega}(n, m) = \omega(n, m) - 1,$$

and substitute this weighting function into (22) to obtain equation (26).

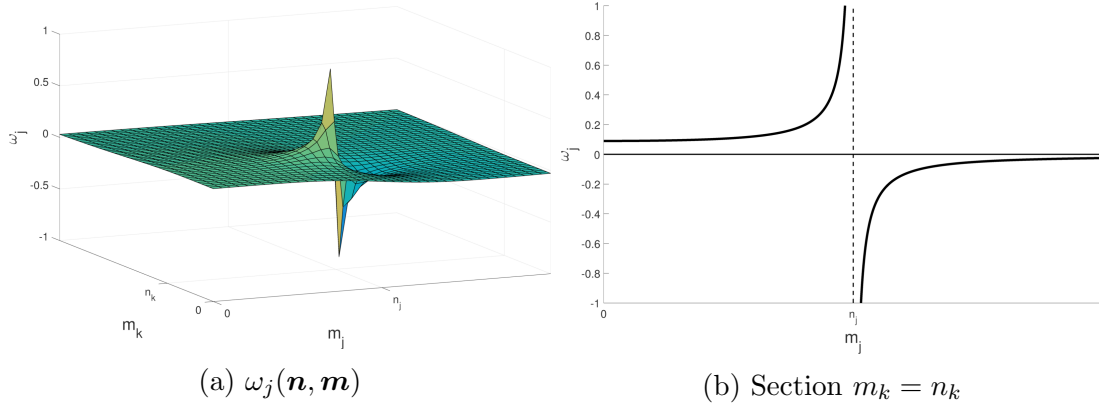


Figure 2: Weighting function for \mathcal{R}_+^2 for type $\mathbf{n} = (n_j, n_k)$

Example 2 When $\mathbf{N} = \mathcal{R}_+^2$ the weighting function is given by:

$$\omega_j(\mathbf{n}, \mathbf{m}) = \frac{n_j - m_j}{2\pi [(n_j - m_j)^2 + (n_k - m_k)^2]} + \frac{n_j - m_j}{2\pi [(n_j - m_j)^2 + (n_k + m_k)^2]} + \frac{n_j + m_j}{2\pi [(n_j + m_j)^2 + (n_k - m_k)^2]} + \frac{n_j + m_j}{2\pi [(n_j + m_j)^2 + (n_k + m_k)^2]}, \quad (27)$$

for all $\mathbf{n}, \mathbf{m} \in \mathcal{R}_+^2$ and $k, j \in \{1, 2\}$ $k \neq j$.

Proof. The proof can be found in the appendix. ■

In figure 2 we have drawn $\omega_j(\mathbf{n}, \mathbf{m})$ for a fixed vector $\mathbf{n} = (n_1, n_2)$. The vertical axis is the value of the weighing function, while the horizontal axes represent m_j and m_k . To interpret the figure first consider the right-hand panel of figure 2. It displays the cross section of the weighting function along the line where $m_k = n_k$. Recall that $\omega_j(\cdot)$ provides the pattern of marginal gains and losses over the type space as a result of the relaxation in incentive constraint j at point \mathbf{n} . Consider first the extreme values in dimension j . The weighting function is approximately zero for values of $m_j \gg n_j$, while for $m_j \ll n_j$ the weighting function takes on positive values. Hence, a relaxation in the incentive constraint allows the planner to transfer public funds to types $m_j \ll n_j$. Note the similarity with the unidimensional case where a relaxation also allowed the planner to redistribute public funds to types $m < n$.

Continuing our focus on the cross section $m_k = n_k$ in figure 2, we see that relaxing the incentive constraint allows the planner to additionally tax away large amounts of resources from types just above n_j , and to transfer those resources towards types just below n_j . Because the incentive constraint is relaxed, this causes no mimicking. However, because the incentive constraint is only relaxed locally, these large transfers have to converge to zero very quickly. If they did not, they would induce violations of incentive constraints at other points in the type space.

Turning our attention away from the line $m_k = n_k$, and to the left-hand picture, note that a relaxation in incentive constraint j has to be accompanied by transfers to types $m_k \neq n_k$ to ensure the incentive constraint in dimension k is not violated. If the planner would only redistribute funds on the line $m_k = n_k$, this would lead to a violation of the incentive constraint in dimension k . All types on the line $m_k = n_k$ that would loose funds, would pretend to be just adjacent to the line to avoid making losses. Simultaneously, all

types that are adjacent to the line in the vicinity of positive transfers, would pretend to have $n_k = m_k$ to receive the extra funds. Hence there is a propagation of the forces in the other direction and the transfer has to be smoothed out around the line $m_k = n_k$ to preserve incentive compatibility.

Generalizing from these two examples, relaxation of incentive constraint j at $\mathbf{n} = (n_1, \dots, n_j, \dots, n_p)$ allows the planner to transfer extra resources towards types $\mathbf{m} = (m_1, \dots, m_j, \dots, m_J)$ where $m_j < n_j$. Although the forces acting on incentives are the same in any dimension, the shape of this transfer scheme is strongly influenced by the propagation of incentive effects through the type space. This propagation is measured in the weighting function $\omega_j(\mathbf{n})$.

In the next subsection, we use lemma 2 to provide an implicit solution for the optimal wedges on each good in the second-best allocation.

4.4 The ABC-formula

The first-order conditions derived in the subsection 4.2 provide an implicit solution for the optimal allocation. With the solution we obtained for $\theta(\cdot)$ in lemma 2, we can express the optimal wedge on each good as a function of the allocation, in the spirit of the ABC-formula in Diamond (1998) and Saez (2001). With this formula we provide an implicit solution for the optimal wedge that is independent of the dimensions of the type space, except through the weighting function $\omega_j(\cdot)$. In the remainder of the paper we show how this formula can be applied to derive important properties of the optimal allocation, and to numerically simulate the optimal-tax system.

Proposition 2 *The optimal wedge on good i for type \mathbf{n} can be described by the following formula:*

$$\frac{q_{x_i}(\mathbf{x}(\mathbf{n})) - s_i(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})}{s_i(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})} = \sum_{j=1}^J A_{ij}(\mathbf{n}) B_{ij}(\mathbf{n}) C_{ij}(\mathbf{n}) \quad (\text{ABC})$$

$$\forall \quad i = 1, \dots, I; \mathbf{n} \in \mathbf{N},$$

where the derivatives of the production function, q_{x_i} , denote the marginal rate of transformation between x_i and y . The terms on the right-hand side of the equation are defined as:

$$A_{ij}(\mathbf{n}) \equiv \xi_{s_i n_j}(\mathbf{n}) = -\frac{\partial s_i}{\partial n_j} \frac{n_j}{s_i} \quad (28)$$

$$B_{ij}(\mathbf{n}) = u_y(\mathbf{n}) \int_{\mathbf{N}} \frac{(g(\mathbf{m}) - 1) f(\mathbf{m})}{u_y(\mathbf{m})} \omega_j(\mathbf{n}, \mathbf{m}) d\mathbf{m}, \quad (29)$$

$$C_{ij}(\mathbf{n}) = \frac{1}{n_j f(\mathbf{n})},$$

Proof. The proof can be found in the appendix. ■

As in the unidimensional case, the left-hand side of equation (ABC) represents the optimal wedge on good x_i for type \mathbf{n} . This wedge is measured as the normalized difference between marginal rate of substitution and transformation with respect to our numeraire. This optimal distortion is broken down into three factors of interest on the right-hand side.

The A_{ij} -term is equal to the elasticity of the marginal rate of substitution with respect to ability, $\xi_{s_i n_j}$. If $\xi_{x_i n_j}$ is large, it means that the marginal rate of substitution for good x_i , and hence the preferences of the agent, strongly increases in characteristic n_j . It follows that x_i is a very strong signal of characteristic n_j . Since the planner uses the wedges to screen the agents, the wedges should be positively related to the strength of the informational value of a good, and the magnitude of the optimal wedge therefore increase in A_{ij} .

The B_{ij} -term represents the redistributive benefit of increasing the wedge on x_i . Perhaps surprisingly, the B_{ij} -term in our multi-dimensional screening model is almost equivalent to its counterpart in the unidimensional model of Saez (2001). As in his model, the B_{ij} -term is determined by the welfare weights $g(\cdot)$, the marginal utility of the numeraire $u_y(\cdot)$ and the distribution $f(\cdot)$. The main difference arises in the weighting function $\omega_j(\cdot)$, which generalizes the formula to a setting with multi-dimensional heterogeneity.¹⁴ To see why $\omega_j(\cdot)$ appears in the formula of the optimal wedge recall from the previous subsection that $\omega_j(\cdot)$ provides the pattern of gains and losses when the incentive constraint is marginally relaxed. An increase in the wedge, leads a relaxation in the incentive-constraint, and hence allows the planner to make additional transfers according to $\omega_j(\cdot)$. The remaining terms in B_{ij} measure the welfare effects of such a transfer.

The C_{ij} -term measures the effect of the distribution of types across the type space. C_{ij} is inversely related to the size of the potential tax base for which a wedge on good x_i distorts marginal incentives. The larger this tax base is, the larger the distortion, and hence, the larger the efficiency cost associated with the distortion. Therefore, the optimal wedge decreases in the size of the tax base.

The most striking difference between the unidimensional and the multidimensional ABC -formula is the need to sum over all characteristics to get the optimal wedge for a good x_i . The optimal wedge on good x_i is the sum of the optimal wedges for each of the characteristics in the type. The fact that the wedge is additively separable in each characteristic, implies that it can easily be decomposed. These decomposition can be particularly useful for policy evaluation.

4.5 Optimal tax rates

So far we have assumed that the social planner offers bundles to the agents through a direct mechanism. In this subsection we consider the properties of a tax system that can implement the optimal allocation in a market economy. We first discuss the general tax system that implements the second-best allocation on the market. Afterwards, we impose additional additive separability constraints on the utility function and show that this allows us to derive an elasticity-based formulation for the optimal-tax rate.

In a market economy the agents maximize utility subject to a budget constraint, instead of sending a message to the planner. The planner can affect the budget constraint by choosing an appropriate tax function. The maximization problem of the agent in the

¹⁴Saez (2001) also normalizes the B_{ij} -term by dividing through $1 - F(\mathbf{n})$ and multiplying C_{ij} with the same factor. In the unidimensional model, this allows one to interpret B_{ij} as a conditional average. Although, we can clearly apply the same normalization in our model, we opt not to. The interpretation as a conditional average that makes this normalization appealing in unidimensional models is lost with multi-dimensional heterogeneity.

market can be written as:

$$\max_{\mathbf{x}, y} \{u(\mathbf{x}, y, \mathbf{n}) : y = q(\mathbf{x}) - \mathcal{T}(\mathbf{x}), \mathbf{x} \in \mathbf{X}, y \in Y\}, \quad (30)$$

where $\mathcal{T}(\cdot)$ represents the tax function. A tax function can implement the second-best allocation in a market economy when the bundle selected by each agent under his budget constraint equals the bundle that agent would have received in the direct mechanism. Mathematically, this condition can be written as follows:

$$\{\mathbf{x}(\mathbf{n}), y(\mathbf{n})\} = \arg \max_{\mathbf{x}, y} \{u(\mathbf{x}, y, \mathbf{n}) : y = q(\mathbf{x}) - \mathcal{T}(\mathbf{x}), \mathbf{x} \in \mathbf{X}, y \in Y\} \quad \forall \mathbf{n} \in \mathbf{N}. \quad (31)$$

The first-order condition to problem (31) can be written as:

$$\frac{\mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))}{1 - \mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))} = \frac{q_{x_i}(\mathbf{x}(\mathbf{n})) - s_i(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})}{s_i(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})} \quad \forall \mathbf{x}(\mathbf{n}) \in \mathbf{x}(\mathbf{n}), \mathbf{n} \in \mathbf{N}. \quad (32)$$

That is, the tax rate on each good should be equal to the wedge for all bundles that are assigned in the second-best allocation. The conditions in (32) form a necessary constraint on the marginal tax rates in the market economy. However, these conditions are not always sufficient. Intuitively, there may be bundles in $\mathbf{x} \in \mathbf{X}$ that are not assigned to any type in the direct mechanism. In that case, even if the second-best allocation is (first- and second-order) incentive compatible, problems can occur if agents prefer bundles outside of the allocation $\mathbf{x}(\mathbf{n})$ altogether. Then additional constraints may need to be imposed on the tax function to ensure it implements the allocation in the market. However, proposition 1 of Renes and Zoutman (2014) shows that conditions the conditions in (32) are locally sufficient, provided the allocation $\{\mathbf{x}(\mathbf{n}), y(\mathbf{n})\}$ is optimal to a welfarist planner. Since the social planner in our model is indeed welfarist, this proposition applies here. Therefore, condition (32) is both a necessary, and a locally sufficient condition for the optimal tax rate. Combining equation (32) with our proposition 2, the optimal marginal tax rates on each good are thus given by:

$$\frac{\mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))}{1 - \mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))} = \sum_{j=1}^J \xi_{s_i, n_j}(\mathbf{n}) \frac{u_y(\mathbf{n})}{n_j f(\mathbf{n})} \int_{\mathbf{N}} \frac{(g(\mathbf{m}) - 1) f(\mathbf{m})}{u_y(\mathbf{m})} \omega_j(\mathbf{n}, \mathbf{m}) d\mathbf{m}. \quad (33)$$

When we compare the expression in (33) to the optimal-tax expression in equation (19) of Golosov et al. (2014), the most important difference is that their expression forms a partial differential equation which is generally difficult to solve, even numerically.¹⁵ On the other hand, for type spaces where $\omega_j(\cdot)$ is known, equation (33) forms a fixed-point solution to the optimal-tax function, which is much easier to solve and simulate.

4.5.1 The optimal tax expression under additive separability

If we restrict the utility function further, the first-order condition of the household on the market allows us to express the optimal-tax expression in terms of the compensated elasticity of demand similar to Saez (2001) and Jacquet et al. (2013). In particular, if we are willing to assume utility is additively separable in each of the goods and characteristics:

$$u(\mathbf{x}, y, \mathbf{n}) = \psi(y) + \sum_{i=1}^I u^i(x_i, n_i), \quad (34)$$

¹⁵Equation (19) of Golosov et al. (2014) considers the optimal unrestricted tax system. This result is closest to our setting since we do not impose any restrictions on the instruments.

it follows that the optimal tax on each good is given by the following expression:

$$\frac{\mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))}{1 - \mathcal{T}_{x_i}(\mathbf{x}(\mathbf{n}))} = \frac{\varepsilon_{x_i, n_i}}{\varepsilon_{x_i}^c} \frac{u_y(\mathbf{n})}{n_i f(\mathbf{n})} \int_{\mathcal{N}} \frac{(g(\mathbf{m}) - 1) f(\mathbf{m})}{u_y(\mathbf{m})} \omega_i(\mathbf{n}, \mathbf{m}) d\mathbf{m}, \quad (35)$$

where $\varepsilon_{x_i, n_i} \equiv \frac{\partial x_i}{\partial n_i} \frac{n_i}{x_i}$ is the elasticity of the consumption of x_i with respect to characteristic n_i in the optimal allocation. While $\varepsilon_{x_i}^c \equiv \frac{\partial x_i}{\partial \tau_i} \Big|_{u=\bar{u}} \frac{q_{x_i} - T_{x_i}}{x_i}$ is the compensated own-price demand elasticity of the good.¹⁶ Additively separable utility functions are commonly applied in the multi-dimensional screening literature for technical convenience (e.g. Armstrong, 1996; Rochet and Choné, 1998). In many optimal-tax models the additive separability assumption has an intuitive interpretation, as it implies that the preferences for each good differs in only one dimension. For instance, in our simulation we calculate the optimal tax rate for couples where we assume that the labor supply of each spouse only directly depends on his/her own ability level.

Additive separability in the utility function allows us to simplify the optimal-tax expression significantly. If we compare equation (35) to equation (33) the most significant difference lies in the fact that there is no need to sum over the different characteristics of the agent. With additively separable utility the marginal rate of substitution of each good is only influenced by its “own” characteristic. Hence, all but one term in the sum cancels out. Additionally, separability allows us to decompose ξ_{s_i, x_i} into the elasticity of the demand of the good with respect to its characteristic, and the compensated price elasticity.¹⁷ This formulation also shows that the optimal-tax rate is inversely related to the compensated elasticity, which is a standard result in optimal taxation (see e.g. Ramsey, 1927).

4.6 Bunching

So far we have ignored possible violations of the second-order incentive constraints (4). However, in multi-dimensional screening models with both incentive constraints and participation constraints bunching at the participation constraint is a common phenomenon (e.g. Armstrong, 1996; Rochet and Choné, 1998; Zheng, 2000). Basov (2006, ch. 7) shows that whether or not bunching occurs depends on the joint convexity of the type space and the utility function. He argues that if a diffeomorphism of the type space exists that allows us to represent it such that the surplus function (utility of participating minus utility of non participation) is positive and concave in transformed types and the transformed type space is convex, bunching does not occur. This condition is rather difficult to check due to the interaction between type space and utility function, but it does indicate that a large class of problems could exist where bunching does not occur.

Although our model is different in the sense that it has no participation constraint, we cannot a priori exclude the possibility that the first-order approach will violate second-order conditions on part of the type space. However, even if bunching is optimal on a part of the type space, the optimality conditions in proposition 2 hold on the separating set, as long as the separating set is convex. Hence, even if the first-order approach does not yield the optimal wedge for every type, this does not necessarily invalidate our approach.

¹⁶A proof to equation (35) can be found in Appendix D.

¹⁷This formulation of the A -term is equivalent to the formulation for the unidimensional model in Jacquet et al. (2013).

For instance, consider the class of utility functions in (34) with linear sub-utility functions $u^i = n_i u^i(x_i)$. This is the same class discussed in Rochet and Choné (1998). In this case it can easily be shown that the separating set is convex. To see this, denote the part of the the type space where bunching occurs as \mathbf{N}_B . Denote the part of the type space that does separate out by $\mathbf{N}_S \equiv \mathbf{N} \setminus \mathbf{N}_B$. Define the 3 types $\mathbf{n}^1, \mathbf{n}^2, \mathbf{n}^3$, such that $\mathbf{n}^2 = k\mathbf{n}^1 + (1 - k)\mathbf{n}^3$ and assume that $\mathbf{n}^1, \mathbf{n}^3 \in \mathbf{N}_S$ and thus separate. If type \mathbf{n}^2 separates for all $k \in (0, 1)$, $\mathbf{n}^1, \mathbf{n}^3 \in \mathbf{N}_S$ the separating set is convex (or empty). Now let $\{\mathbf{x}, \mathbf{y}\}_B$ denote the bundle that maximizes \mathbf{n}^2 's utility among the bunching bundles. Define a loss function, $L(\mathbf{n}) = V(\mathbf{n}) - u(\{\mathbf{x}, \mathbf{y}\}_B, \mathbf{n})$. We know from individual rationality that $L(\mathbf{n}^1), L(\mathbf{n}^3) > 0$, otherwise they would not have separated out. From the fact that utility is linear in types and conditions (5) and (6), we know that $L(\cdot)$ is convex, smooth and has increasing differences over the separating set. It follows that $L(\mathbf{n}^2) > 0$. Hence, any convex combination of types in the separating set, also resides in the separating set. The set must thus be convex (or empty). Hence, for this utility function, which plays an important role in the multi-dimensional screening literature, proposition 2 continues to apply in the separating set.

5 The Atkinson-Stiglitz Theorem

The A-S theorem (Atkinson and Stiglitz, 1976) states that indirect taxation is superfluous in a setting where agents are heterogeneous in earnings ability, as long as preferences are homogeneous, and the utility function is weakly separable in labor. This result has subsequently been shown to hold with more general utility functions in Laroque (2005), Kaplow (2006), and Hellwig (2010a), all under the assumption of unidimensional heterogeneity. Gauthier and Laroque (2009) are the first to consider multi-dimensional heterogeneity. However, in their model the marginal rate of substitution between each of the goods, except labor income, is independent of the hidden characteristics. Therefore, it is impossible to screen for all characteristics in their model, whereas in our model the government can screen for all characteristics.

When the conditions for the A-S theorem apply, the planner can reach the second-best allocation by only taxing labor income. Equivalently the planner could tax labor income non-linearly and levy a uniform rate on all other commodities. The main application of the theorem is, perhaps, the fact that commodities should be taxed at the same rate over time. That is, the optimal capital income tax rate equals zero if the A-S theorem applies.

In this section we investigate whether the A-S theorem generalizes to a setting where the government screens in multiple dimensions. To that end, the next corollary to proposition 2 provides the conditions under which a good has a non-zero wedge in the optimum.

Corollary 1 *The optimal wedge on good x_i is zero for all types if $\xi_{s_i, n_j} = 0 \quad \forall \quad n_j \in \mathbf{N}$. That is, the optimal wedge is zero if the marginal rate of substitution for x_i does not depend on any characteristic n_j .*

Proof. If the marginal rate of substitution, s_i , is independent of all characteristics n_j , then $\xi_{s_i, n_j} = 0 \quad \forall \quad n_j \in \mathbf{N}$, such the optimal wedge on x_i is zero by equation (ABC). ■

Intuitively, corollary 1 shows that the marginal wedge on a good equals zero if the preference for this good is not directly influenced by any characteristic. In that case, the good does not provide any first-order information about characteristics, and distortions away from laissez faire yields an efficiency loss without an offsetting information gain.

From corollary 1 one can derive the A-S theorem in a setting with unidimensional heterogeneity. To see this note that weak separability implies that the marginal rates of substitution for all goods except income are independent of the type, such that all $\xi_{s_i n_j} = 0$ except for income.

It also follows that the A-S theorem does not hold in a setting where the planner screens in multiple dimensions. To see this remember that we have assumed \mathbf{s}_n has rank J , so that $\xi_{s_i n_j} \neq 0$ for at least J choices in our model. Hence, the A-S theorem cannot hold if $J \geq 2$. Intuitively, a planner that wants to redistribute in multiple dimensions cannot do so by distorting the price of only one good. In fact, under our full revelation assumption a Tinbergen rule applies. A planner that wants to redistribute over J characteristics will need to distort (at least) J goods.

In the literature many violations of the A-S theorem have been recorded. In Erosa and Gervais (2002) preferences are not weakly separable over time since consumption at old age is a stronger complement to leisure than consumption at a younger age. Therefore, the distortion of the labor income tax is reduced by taxing capital income. In Golosov et al. (2013) capital is optimally taxed if households with higher ability also have higher patience. Boadway and Song (2015) show that the theorem fails if households optimally choose corner solutions. Therefore, if some households are cash constrained and do not buy all commodities, the tax rate on these commodities should be different to identify the credit constrained households. Farhi and Werning (2010) and Kopczuk (2013) show that the bequest motive may generate a negative externality which can be remedied through the taxation of capital. The argument that is closest to our is derived in Cremer et al. (2001) and Saez (2002). These papers show that under two-dimensional heterogeneity commodity taxation is not superfluous. However, the former result is derived in a setting with discrete types and the latter is derived under the assumption that welfare weights are only correlated with earnings ability and commodity taxes are linear, which are restrictions we do not use.

Our result adds to this literature by showing generally that the A-S theorem cannot hold under multidimensional heterogeneity, provided that all (or at least two) underlying characteristics can be revealed. This has large implications for the evaluation of government policy. According to the A-S theorem we can obtain a second-best allocation if the only government intervention is the taxation of labor income. We show that if a government cares about redistribution in multiple dimensions, such as from healthy to sick and from rich to poor, it needs to distort multiple choices in order to attain the second-best allocation. Therefore, government interventions in many markets, like the health care and the housing market, are not just a empirical irregularities. These interventions may well be optimal. In addition, the optimal wedge on capital income may be non-zero if households differ in their investment skills, for no other reason than this difference in skill (see also Golosov et al., 2014).

6 Boundary Conditions

The boundary conditions (16) imply that $\theta_j(\cdot)$ equals zero on the boundary of the type space ∂N . Hence, we find:

$$\theta_j(\underline{n}_j) = \theta_j(\bar{n}_j) = 0, \quad (36)$$

where \underline{n}_j (\bar{n}_j) represents the type that has the lowest (highest) value for characteristic j . Define corner types \mathbf{n}_\lceil as agents that have either highest or lowest values for all their characteristics. In a two-dimensional type space, type $\mathbf{n} = (\underline{n}_1, \underline{n}_2)$ and $\mathbf{n} = (\bar{n}_1, \bar{n}_2)$ are obviously corner types, but so are the types that combine the lowest value of n_1 with the highest value of n_2 and vice versa: $\mathbf{n} = (\underline{n}_1, \bar{n}_2)$ and $\mathbf{n} = (\bar{n}_1, \underline{n}_2)$. Note that corner types do not necessarily exist. For instance, the characteristics may be distributed along a circle, or the type space may alternatively be unbounded. There are at most 2^J corner types, but there may be less, or none at all, depending on the shape of the type space. Corollary 2 establishes that the optimal wedge on all goods equals zero for each existing corner type.

Corollary 2 *The optimal wedge for any type \mathbf{n}_\lceil equals zero if the types exist.*

Proof. From the boundary conditions it follows that $\theta_j(\underline{n}_j) = \theta_j(\bar{n}_j) = 0$ for all $j = 1, \dots, J$. The optimal wedge at the corner types can be found by taking the limit of equation (ABC) if \mathbf{n} goes to an \mathbf{n}_\lceil :

$$\lim_{\mathbf{n} \rightarrow \mathbf{n}_\lceil} \frac{q_{x_i} - s_i}{s_i} = \lim_{\mathbf{n} \rightarrow \mathbf{n}_\lceil} \sum_{j=1}^J \varepsilon_{x_i, n_j} \frac{u_y \theta_j(\mathbf{n}) / \lambda}{n_j f(\mathbf{n})} = 0 \quad \forall \quad i \in 1, \dots, I,$$

provided $\mathbf{n}_\lceil \in \mathbf{N}$. That is, provided the corner type exists. ■

Corollary (2) shows that the no-distortion at the top and the bottom results, as derived in Sadka (1976) and Seade (1977), remains valid in a multidimensional framework as long as the type distribution is bounded. The redistributive benefit of taxation, B_{ij} , equals zero at the corners of the type space. Intuitively, the marginal tax rate on any good at type \mathbf{n} redistributes resources from types above \mathbf{n} to types below \mathbf{n} . For corner types at the top there is no redistributive benefit of redistributing resources from people above the type since they do not exist. On the bottom, there is no redistributive benefit to redistributing resources to types below that type since they do not exist. However, in both cases there is an efficiency loss of taxation. Hence, the optimal tax rate equals zero.

Our result is similar to Golosov et al. (2011) who derive the optimal tax rate at the boundary in a model where earnings ability follows a stochastic progress. They show that the optimal tax rate for types that persistently have the highest or lowest skill realization is zero, provided such types exist.

7 Interdependencies in the Tax System in Mirrleesian and New Dynamic Public Finance

The NDPF, pioneered in Golosov et al. (2003), generalizes the Mirrlees model with unidimensional heterogeneity to a dynamic stochastic setting. For each agent, earnings ability follows a stochastic process. Earnings ability in period t is revealed to the agent at the beginning of the period. Agents do not know what their earnings ability will be in future periods, but they form expectations about it based on the stochastic process and their history. The planner does not observe ability, but does observe all choices made by the agents in the economy, is aware of the stochastic process, and can keep records over time.

One of the most intriguing results in the NDPF is the complexity of the optimal tax system. Kocherlakota (2005) shows that the optimal tax on labor income in period t may depend on the entire history of labor income up to period t .¹⁸ However, so far there is no clear explanation why interdependencies are present in the NDPF in the first place, or alternatively, why they are typically absent in the classical (static) Mirrleesian public finance. In this section we show that interdependencies in the wedges generally do occur in static Mirrleesian public finance models, provided agents are heterogeneous in multiple dimensions. Subsequently, we give our model a dynamic interpretation, and provide a link between the result found in Kocherlakota (2005) and our model.

The next corollary uses proposition 2 and lemma 1 to show that the optimal tax rate on each good generally depends on J goods.

Corollary 3 *The optimal tax rate, \mathcal{T}_{x_i} , on any good x_i can be written as a function of J goods.*

Proof. By equation (33) the optimal tax on each good x_i depends on all J characteristics. To express the optimal tax system in terms of observables \mathbf{x} rather than unobservables \mathbf{n} we need a mapping from the goods space to the type space. By corollary 1 there exists at least one mapping from the optimal allocation to the type space: $(\mathbf{x})^\leftarrow : \mathbf{X} \rightarrow \mathbf{N}$ for all $\mathbf{n} \in \mathbf{N}$. Since the Jacobian $\mathbf{x}'(\mathbf{n})$ is of rank J everywhere, the inverse mapping depends on J goods everywhere on the type space. Therefore, we can write the optimal tax rates as a function of J goods. ■

By this corollary, a separable tax system where the tax rate of each good only depends on the consumption of that good will generally not exist if $J > 1$. Intuitively, the optimal wedge on each good depends on all characteristics in the type space. For instance, the optimal wedge decreases in the density $f(\mathbf{n})$ because a distortion becomes more costly as the size of the distorted group increases (see proposition 2). Most probability density functions, $f(\mathbf{n})$, are a function of all characteristics. This is for instance true even if those characteristics are independently distributed. In order to reveal all J characteristics that influence the probability density function, the tax needs to be conditioned on at least J goods. Similarly, the B_{ij} -term directly depends on the weighting function and welfare weights, all of which are likely to depend on each characteristic in the type. It may be possible to construct special cases where the optimal wedge on each good depends on fewer goods. However, in general we expect any optimal tax system that redistributes in multiple dimensions to have an amount of interdependencies equal to the dimensionality of the type space.

7.1 A Dynamic Interpretation

We now turn to the dynamic problem that is considered in the NDPF, and show that lemma 3 provides an explanation to the interdependencies found in that literature as well. Consider a T -period economy and assume that agents' abilities evolve over time. Denote ability in period t by n_t . Assume that the agents know their entire (future) history of

¹⁸Subsequent papers have made some progress on limiting the number of interdependencies in special cases. In particular, Albanesi and Sleet (2006); Brendon (2012) investigate conditions under which intertemporal interdependencies vanish. This is for instance the case if the stochastic process is iid. In that case, the optimal wedge in each period depends on current labor income and wealth, such that only an intratemporal interdependency remains between the optimal wedge on capital and labor income.

ability levels, the vector $\mathbf{n} = [n_1, \dots, n_T]$, at the beginning of the first period. Clearly, this is a strong assumption on the information available to the agents in this model. In contrast, in NDPF models agents only learn their ability in each period at the beginning of that period, such that their knowledge of future types is limited to knowledge about the stochastic process and their history. Note however, that this model can be seen as a special case of a NDPF model where ability evolves according to a fully deterministic process known to the agents, but not the planner. The planner cannot observe the agents' ability levels but is aware of the probability density function $f(\mathbf{n})$, and cumulative density function $F(\mathbf{n})$. In addition, assume for simplicity that each agent makes one independent choice each period, the amount of labor income he earns, x_t .

Our model can be used to calculate the welfare-optimizing wedge on each period's labor income x_t . By proposition 2 the optimal wedge is given by equation (ABC). From corollary 3 it follows that the optimal wedge in each period can be written as a function of the entire history of labor income (including future periods in this case). Intuitively, the optimal wedges depend on terms A , how much information a choice reveals, B , the redistributive value of the wedge, and C , the incidence of marginal distortions. Each of these terms may depend on the entire vector \mathbf{n} . In turn, in order to reveal the entire vector the planner needs to observe at least as many choices in \mathbf{x} as there are elements in \mathbf{n} . Given the assumptions in this subsection, the planner therefore has to condition the optimal tax rate in each period on the entire history of labor income.

It may seem surprising that an even stronger version of Kocherlakota's result can be constructed so easily in a deterministic setting. However, as was noted in Pavan et al. (2014), in a deterministic model agents can plan their entire choice vector \mathbf{x} with full knowledge about their type \mathbf{n} , regardless of the period the consumption is in. In a stochastic setting information concerning the type is revealed over time, such that agents always have weakly less information than in the deterministic setting. As such, a stochastic model may allow for less (profitable) deviations than a deterministic model. This implies the incentive constraints are more binding in our model than in the NDPF.

The practical implications of this result are less clear. A literal reading of the model described in this subsection would suggest intertemporal interdependencies be build into the tax system. One could let the labor income tax increase in previous income, for instance. However, it is unclear how much welfare is gained by introducing such interdependencies. Simulations on a NDPF model in Farhi and Werning (2013) show that much of the welfare gain of optimal taxation can be obtained by tax systems that do not exhibit any intertemporal interdependencies. If the cost of implementation and administration are high, implementing interdependencies might not be optimal in real tax systems.

8 Simulations

In this section we present our simulation results. We simulate the optimal income tax for couples that maximize a joint utility function. The calibration of our simulation is deliberately chosen to closely resemble the one-dimensional simulations on optimal income taxation for individuals presented in Saez (2001). To facilitate comparison, we will define the utility and welfare almost exactly as he does. This allows straightforward comparison between the results obtained in the unidimensional model of income taxation, and its closest bidimensional counterpart. Our main interest in these simulations lies in

the effect of bidimensional heterogeneity on the top tax rate. In particular, we try to quantify whether, and by how much, the optimal top-income tax rate depends on earnings of the other spouse in the couple. Below we first discuss the calibration of the model before turning to the results.

8.1 Calibration

We assume all agents in our model are couples. The spouses within the couple are assumed to maximize a joint utility function:

$$\begin{aligned} u(\mathbf{x}, y, \mathbf{n}) &= y - \sum_{i=1}^2 \frac{n_i}{\delta_i} (L_i)^{\delta_i} \\ &= y - \sum_{i=1}^2 \frac{n_i}{\delta_i} \left(\frac{x_i}{n_i} \right)^{\delta_i}, \end{aligned} \quad (37)$$

where n_i is the ability of spouse i , L_i is labor effort, and $x_i = n_i L_i$ is gross income, the product between ability and labor effort.¹⁹ With this joint utility function, there is no strategic interaction within the couple and the marginal utility of labor supply of one spouse is independent of the labor supply of the other spouse.

The compensated labor supply elasticity of spouse i is given by $\varepsilon_{x_i}^c = \frac{1}{\delta_i - 1}$. We choose $\delta_i = 3$ for both spouses such that $\varepsilon_{x_i}^c = 0.5$.

Note that we use a slightly altered version of the standard quasi-linear utility function, because we multiply the marginal utility of labor supply with ability. The advantage of this specification is that it ensures that labor supply L_i is bounded between 0 and 1 as long as each spouse faces a positive tax rate. In the Algorithm section in our Appendix we show that this allows us to simplify the couples' numerical optimization problem significantly. To our knowledge, this transformation of quasi-linear utility was first introduced in Brewer et al. (2010) in optimal-tax simulations for the UK.

The total vector of characteristics of a couple, the type of the agent, is given by $\mathbf{n} = (n_1, n_2)$. We assume matches between spouses are exogenous. The distribution of types follows a bivariate type I Pareto distribution:

$$f(n_1, n_2) = (\alpha + 1) \alpha (\underline{n}_1 \underline{n}_2)^{\alpha+1} (\underline{n}_2 n_1 + \underline{n}_1 n_2 - \underline{n}_1 \underline{n}_2)^{-(\alpha+2)} \quad \forall \quad n_i \in [\underline{n}_i, \infty). \quad (38)$$

This bivariate Pareto distribution has the property that the marginal distribution of each characteristic n_i follows a standard Pareto distribution with shape parameter α and scale parameter \underline{n}_i . It has been shown that this marginal distribution provides an excellent fit for top incomes in the US (see e.g. Saez, 2001; Atkinson et al., 2011), and it is hence a logical choice for inquiry into the optimal top tax rates. However, for lower- and middle income levels the Pareto distribution does not provide a good fit. Evaluating what the tax rates are at those income levels is therefore beyond the scope of this simulation. We leave the modeling of the overall shape of the optimal tax system to future research.

Consistent with US evidence we set $\alpha = 2$. The scaling parameters are chosen such that $\underline{n}_1 = \underline{n}_2 = 1$. This implies the marginal distribution of earnings ability for each of

¹⁹In this section we deviate from our convention to use i as a count variable for the goods, and j as a count variable for the characteristics. Instead, the count variable i here denotes both the characteristic and the choice variables of spouse i .

the two spouses is equal. In reality, women typically earn less than men. However, it is unclear whether this difference is driven by differences in ability, effort, occupational decisions or outright discrimination. We leave the modeling of this aspect of reality to future research.

The correlation between the two ability levels is given by $\rho = \frac{1}{\alpha} = 0.5$. This is consistent with positive but imperfect assortative matching. For comparison, we also show results with perfect positive assortative matching, i.e. $\rho = 1$. In that case there is effectively only a single dimension of heterogeneity, because there exists a one-to-one (linear) mapping between the ability levels of the spouses. By corollary 3 It follows that the income tax on the income of a spouse only depends on his or her own income. As a result, the optimal tax rates on the income of each of the two spouses follows the standard optimal-tax formula provided in Saez (2001).

We measure the production of each couple as their combined gross income, such that the production function within the couple is given by:

$$q(\mathbf{x}) = \sum_{i=1}^2 x_i. \quad (39)$$

y denotes the consumption of the couple and is given by the difference between gross income and the tax function:

$$y = \sum_{i=1}^2 x_i - \mathcal{T}(x_1, x_2).$$

The tax function $\mathcal{T}(\cdot)$ depends on both income levels, and is allowed to take on any non-linear form, consistent with our model.

Finally, we assume that the planner maximizes a logarithmic transformation of the couple's indirect utility function:

$$W(V(\mathbf{n})) = \int_N \log(V(\mathbf{n})) f(\mathbf{n}) d\mathbf{n}. \quad (40)$$

To find the optimal tax system we use a fixed-point algorithm which can be roughly subdivided into three stages. In the first stage, we define the primitives of the model and define an initial tax system. In the second stage, the algorithm uses the couple's maximization problem to find the allocation that is implemented by this tax system. In the third stages, the algorithm updates the tax system using the first-order condition for the planner provided in proposition 2 with the allocation that was calculated in the second stage. The algorithm then returns to the second stage to find the allocation under the updated tax system and so forth, until it converges. Since our description of the second best is derived through a first-order approach, we verify that the the second-order incentive constraints are met after the algorithm has converged. In our application, second-order incentive constraints are never violated. A detailed step-by-step description of the algorithm can be found in the Appendix.

8.2 Results

Our results are presented in figure 3. The top panel of the figure shows the entire shape of the marginal tax rate on the income of partner i , $\mathcal{T}_{x_i}(x_i, x_j)$. The bottom panel

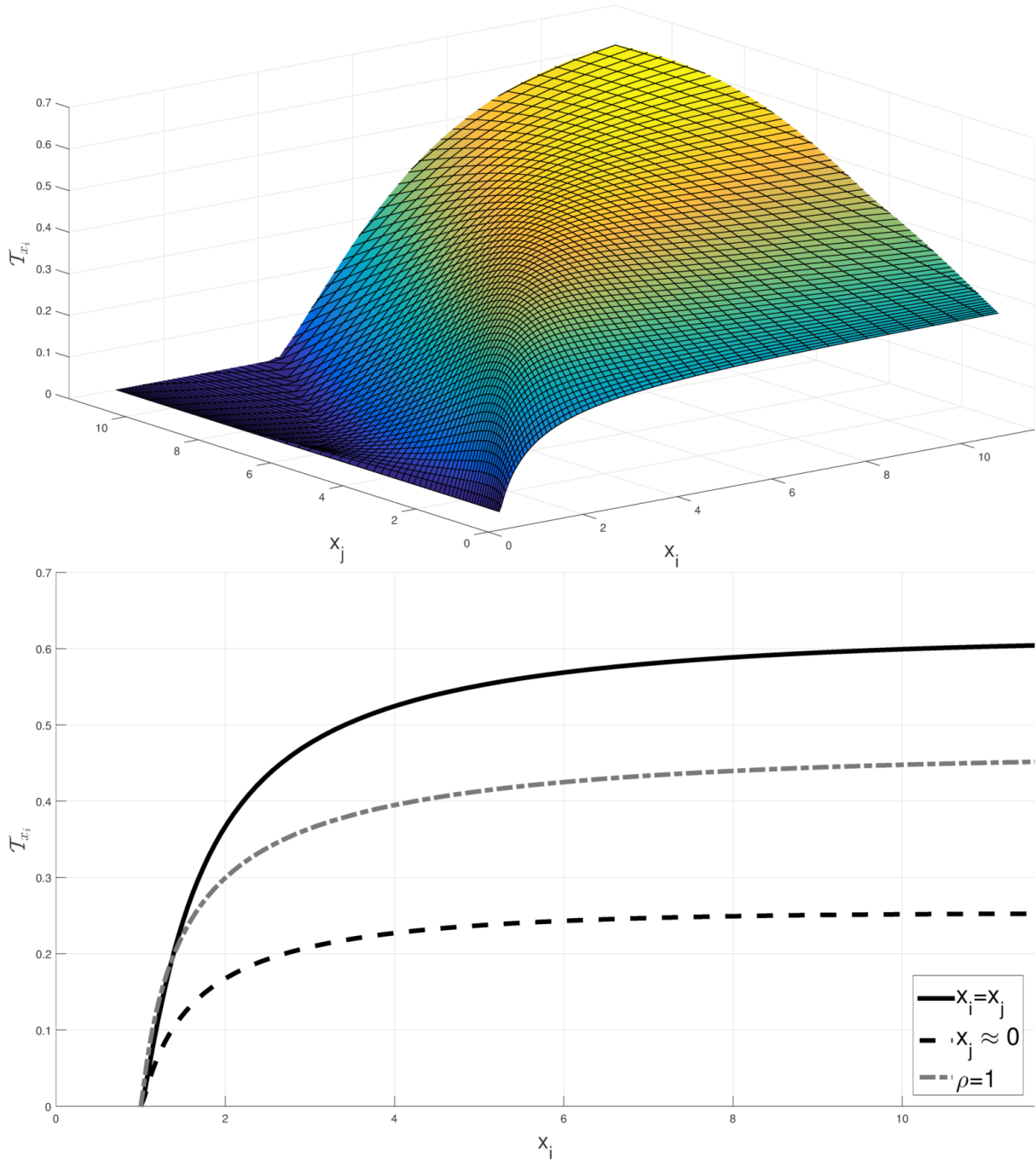


Figure 3: Optimal marginal tax rates

shows the cross section of the main figure where partner j earns the same gross income, $x_i = x_j$ (solid line) and the cross section where partner j earns close to zero income (dashed line). For comparison, the bottom panel also shows the tax rate in case ability is perfectly correlated between the two partners (grey dashed-dotted line).

As can be seen from the figure, marginal tax rates are strictly progressive in the sense that they increase in earned income.²⁰ In addition, they appear to converge to a constant at the top. When ability between the two partners is correlated perfectly, this constant coincides with the top tax rate in Saez (2001). However, with imperfect correlation the top tax rate strongly depends on the income earned by the partner. When both partners are top earners, the tax rate is around 65 percent. However, when only one partner is a top earner while the other earns very little, the top tax rate is only about 25 percent. Hence, the top tax rate strongly depends on the income earned by the partner.

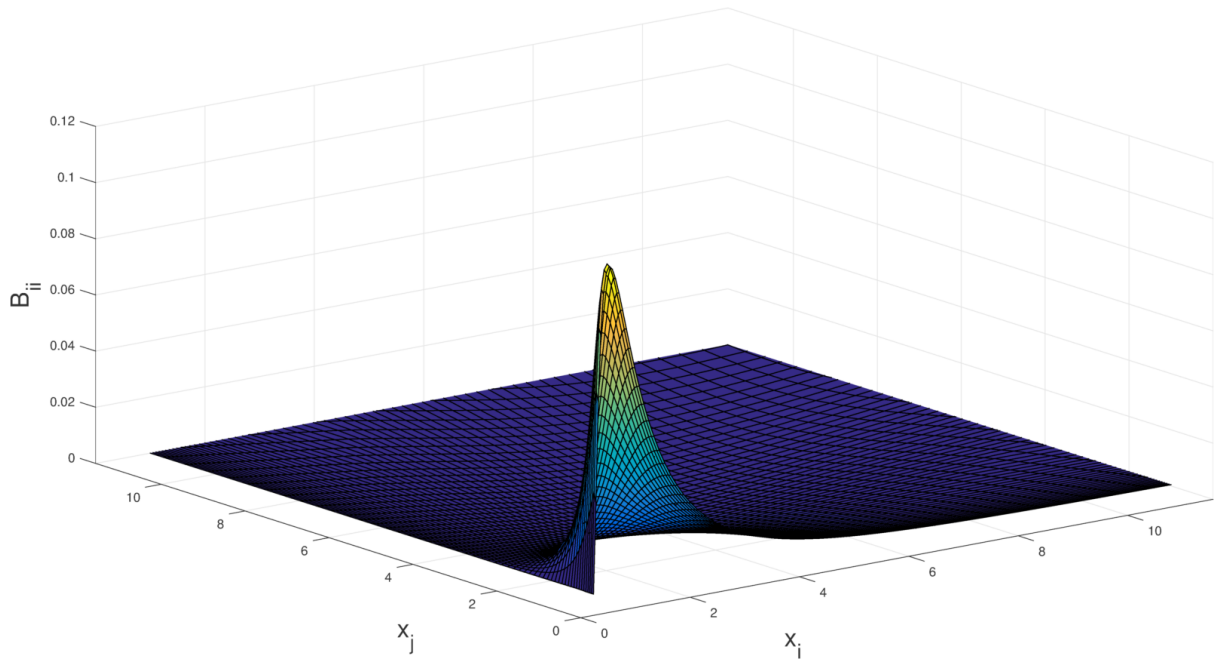
To assess why this interdependency plays such a prominent role in our simulations, figure 4 shows the B_{ii} and C_{ii} -term of proposition 2. Our utility function guarantees that the A_{ii} -term is a constant for all types. As can be seen, the B_{ii} -term is close to constant for top-income earners. Intuitively, the welfare weights $g(\mathbf{n})$ in equation (29) are close to zero for all couples when at least one of the two spouses is a top-income earner. However, the C_{ii} -term is much larger for couples in which both partners earn a high income, than in couples where only one of the two partners has a high income. The reason is that the density strongly decreases in the income of the partner. Intuitively, couples with two top income earners are extreme outliers. Even with positive assortative matching, it is still rather unlikely that two top earners match, unless matching is perfect. Hence, taxing couples with two top-earners causes less distortions than taxing couples with only one top-income earner, which provides an efficiency argument for increasing their marginal tax rate.

The point of this simulation is not to give a definitive answer on what the top income tax rate in the US, or any country for that matter, should be. Answering this question requires a far more comprehensive calibration of the model, which is beyond the scope of this paper. However, the simulation does show that the top income tax rate is not a fixed constant. It can (strongly) depend on other observable choice variables such as the income earned by the spouse. The reason is that these choice variables can provide additional information on preferences, welfare weights and densities, all of which plays an important role in calculating the optimal tax rate. Conditioning tax systems on these additional variables allows the government to redistribute more income at lower efficiency cost.

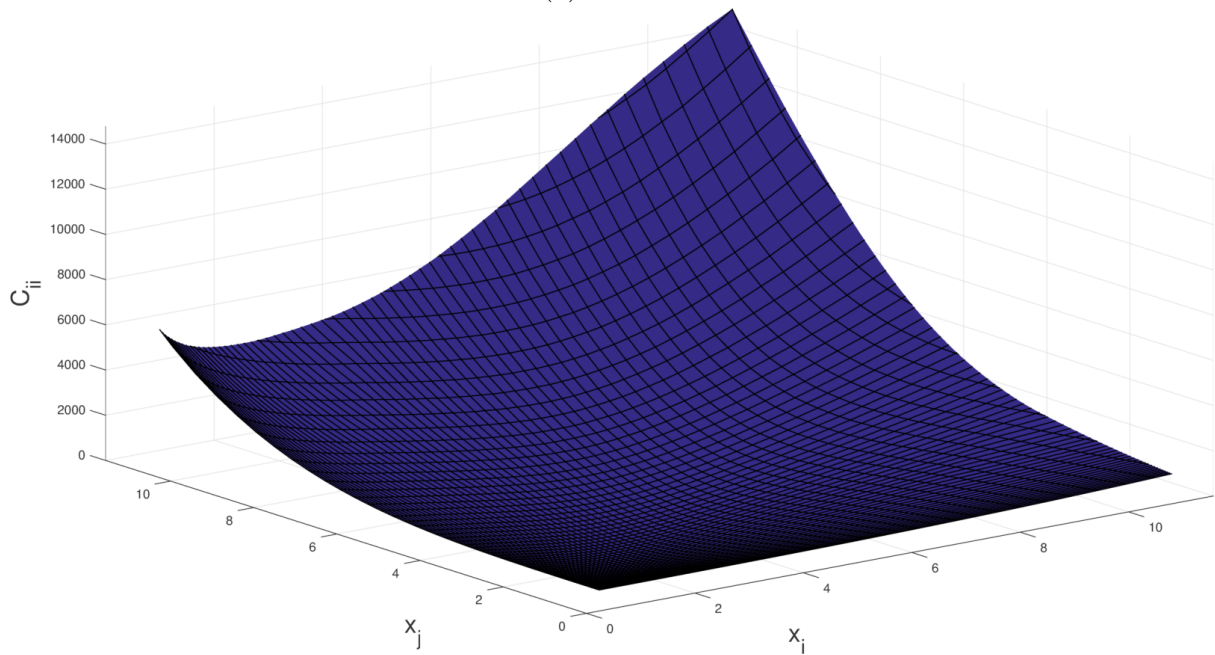
9 Concluding Remarks

Although significant progress has been made in multidimensional mechanism design, the equilibrium in a Mirrleesian optimal tax model with multiple goods and multiple characteristics has so far not been characterized. In this paper we characterize it and show some similarities and differences with the unidimensional Mirrleesian model. Our model can be used to study the relationship between several tax tools. Our characterization

²⁰Unlike in Saez (2001) the optimal tax schedule in our model does not follow a U-shape. The likely reason is that the ability of all spouses follows a Pareto distribution in our model, whereas this distribution only applies to the top agents in Saez (2001).



(a) B -term



(b) C -term

Figure 4: Decomposition of Optimal Tax Rates

of the second-best equilibrium shows that the government should condition taxes on consumption patterns that provide as much information on the underlying types as possible. Furthermore, we show how the equilibrium in this model relates to the equilibria in stochastic dynamic NDPF models and in the wider class of multidimensional screening models.

Since we use a first-order approach to characterize the optimum, we implicitly assume all types separate out in equilibrium in most of this paper. In reality, it may not be optimal to allow some bunching at the lower end of the type space. This prescription fits reasonably well with the tax systems observed in welfare states. The lowest earning individuals get a fixed amount of welfare assistance, or income subsidies, creating a bunch at the lower end of the income distribution. Moreover, most income assistance programs are conditioned on (the absence of) wealth, to prevent abuse. These are the kind of interdependencies between underlying characteristics (wealth and ability) our model shows to be optimal in case of multiple sources of heterogeneity. This observation is particularly relevant in the many welfare systems that (also) subsidize medical expenses and/or housing for a large group of individuals. These individual subsidies generally depend on both the expenses and on income. In models where agents only differ in earnings ability such interdependencies between subsidies are suboptimal. Since transfers to the agents need only be differentiated by income level, only distortions in income are necessary. Similarly, in models of health insurance where agents only differ in health status optimal transfers are independent of income. Separability of the income/subsidy scheme would follow from a direct application of the Atkinson-Stiglitz theorem. However, if the government wants to redistribute both from rich to poor and from healthy to sick, our results indicate that the government should use instruments that depend on both income and health care expenses (or other observables related to the same characteristics). This gives a clear theoretical explanation for the interdependencies often found in real world taxes and subsidies.

The equilibrium in our model depends on a set of partial differential equations that cannot be solved using traditional techniques applied in mechanism design. We rewrite these equations and use the method of Green's functions to solve these PDE, allowing us to write the equation for the optimal tax rate as a fixed point that can easily be solved numerically. In future work this method can be extended to other economic problems. Although this new approach does not give us closed form solutions for the optimum, the fixed-point equation obtained allow a much more general interpretation of the properties of the second best than the PDE do. Furthermore, we can identify some important properties that are independent of the dimensions of the problem.

Since this model contains the multidimensional type space that is also found in the NDPF, but without the dynamic stochasticity, it provides a convenient middle ground between the complex stochastic dynamics in these models, and the known intuitions of the classical Mirrlees model. The interdependencies in the optimal wedges that present difficulties in NDPF models are, for instance, also prevalent in our setting, but can be traced much more conveniently. We can gain intuition for these tax schedules from multidimensional screening models, in particular from our model. In fact, the discussion in section 7 already suggests that our findings might be generalized to dynamic settings. This would allow an elasticity approach and a new focus on screening and implementation separately in these models as well.

The next step in this line of research clearly is to find specific, realistic and relevant

settings and simulate the model for realistic parameter settings. The simulations reported in this paper show that such simulations are feasible, even though the multidimensional heterogeneity does set strong requirements on the optimization algorithm. In addition, the problem of implementation, which is discussed in Renes and Zoutman (2014), might add further difficulties. Implementation can prove especially difficult if there is bunching as well as separation. The instruments on the interior of the separating partition will likely require a differ from the instruments on the bunching partitions. However, the pay-off seems worth the effort. The model presented in this paper can be used to provide a more precise insight in the optimal relation between the income tax system and the myriad of social schemes like health care subsidies, housing subsidies, and welfare assistance that characterize modern welfare states.

References

- Albanesi, Stefania, and Christopher Sleet (2006) ‘Dynamic optimal taxation with private information.’ *Review of Economic Studies* 73(1), 1–30
- Armstrong, Mark (1996) ‘Multiproduct nonlinear pricing.’ *Econometrica* 64(1), 51–75
- Armstrong, Mark, and Jean-Charles Rochet (1999) ‘Multi-dimensional screening: A user’s guide.’ *European Economic Review* 43(4), 959–979
- Atkinson, Anthony B., and Joseph E. Stiglitz (1976) ‘The design of tax structure: Direct versus indirect taxation.’ *Journal of Public Economics* 6(1-2), 55–75
- Atkinson, Anthony B, Thomas Piketty, and Emmanuel Saez (2011) ‘Top incomes in the long run of history.’ *Journal of Economic Literature* 49(1), 3–71
- Basov, Suren (2006) *Multidimensional screening* (Heidelberg: Springer)
- Boadway, Robin, and Zhen Song (2015) ‘Indirect taxes for redistribution: Should necessity goods be favored?’ *Research in Economics* 70, 64–88
- Bovenberg, A. Lans, and Bas Jacobs (2005) ‘Redistribution and education subsidies are siamese twins.’ *Journal of Public Economics* 89(11-12), 2005–2035
- Brendon, Charles (2012) ‘General efficiency-equity trade-offs for dynamic Mirrleesian tax problems.’ mimeo Oxford Oxford
- Brewer, Mike, Emmanuel Saez, and Andrew Shephard (2010) ‘Means-testing and tax rates on earnings.’ In *The Mirrlees Review – Dimensions of Tax Design*, ed. James A. Mirrlees, Stuart Adam, Timothy J. Besley, Richard Blundell, Steven Bond, Robert Chote, Malcolm Gammie, Paul Johnson, Gareth D. Myles, and James M. Poterba (Oxford: Oxford University Press) chapter 3, pp. 202–274
- Choné, Philippe, and Guy Laroque (2010) ‘Negative marginal tax rates and heterogeneity.’ *American Economic Review* 100(5), 2532–2547
- Cremer, Helmuth, Jean-Marie Lozachmeur, Dario Maldonado, and Kerstin Roeder (2015) ‘Household bargaining and the design of couples’ income taxation.’ CESifo Working Paper Series No. 5205 Munich

- Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet (2001) ‘Direct versus indirect taxation: The design of the tax structure revisited.’ *International Economic Review* 42(3), 781–800
- Diamond, Peter A. (1998) ‘Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates.’ *American Economic Review* 88(1), 83–95
- Ebert, Udo (1992) ‘A reexamination of the optimal non-linear income tax.’ *Journal of Public Economics* 49(1), 47–73
- Erosa, Andres, and Martin Gervais (2002) ‘Optimal taxation in life-cycle economies.’ *Journal of Economic Theory* 105(2), 338–369
- Evans, Lawrence C. (2010) *Partial differential equations: Second Edition* (Providence, RI: American Mathematical Society)
- Farhi, Emmanuel, and Iván Werning (2010) ‘Progressive estate taxation.’ *Quarterly Journal of Economics* 125(2), 635–673
- (2013) ‘Insurance and taxation over the life cycle.’ *Review of Economic Studies* 80(2), 596–635
- Farhi, Emmanuel, Christopher Sleet, Ivan Werning, and Sevin Yeltekin (2012) ‘Non-linear capital taxation without commitment.’ *Review of Economic Studies* 79(4), 1469–1493
- Franklin, Jerrold (2012) ‘Green’s functions for neumann boundary conditions.’ arXiv:1201.6059
- Gauthier, Stephane, and Guy Laroque (2009) ‘Separability and public finance.’ *Journal of Public Economics* 93(11), 1168–1174
- Golosov, Mikhail, Aleh Tsyvinski, and Iván Werning (2007) ‘New dynamic public finance: A user’s guide.’ In *NBER Macroeconomics Annual 2006, Volume 21*, ed. Daron Acemoglu, Kenneth Rogoff, and Micheal Woodford (Cambridge-MA: MIT Press) pp. 317–388
- Golosov, Mikhail, Aleh Tsyvinski, and Matthew Weinzierl (2013) ‘Preference heterogeneity and optimal commodity taxation.’ *Journal of Public Economics* 97(1), 160–175
- Golosov, Mikhail, Aleh Tsyvinski, and Nicolas Werquin (2014) ‘A variational approach to the analysis of tax systems.’ NBER working paper No. 20780 Cambridge, MA
- Golosov, Mikhail, Maxim Troshkin, and Aleh Tsyvinski (2011) ‘Optimal dynamic taxes.’ NBER Working Paper No. 17642 Cambridge, MA
- Golosov, Mikhail, Narayana Kocherlakota, and Aleh Tsyvinski (2003) ‘Optimal Indirect and Capital Taxation.’ *Review of Economic Studies* 70(3), 569–587
- Gomes, Renato D, Jean-Marie Lozachmeur, and Alessandro Pavan (2014) ‘Differential taxation and occupational choice.’ CESifo Working Paper Series No. 5054 Munich
- Hellwig, Martin F. (2010a) ‘A generalization of the Atkinson–Stiglitz (1976) theorem on the undesirability of nonuniform excise taxation.’ *Economics Letters* 108(2), 156–158

- (2010b) ‘Incentive problems with unidimensional hidden characteristics: A unified approach.’ *Econometrica* 78(4), 1201–1237
- Jacobs, Bas (2010) ‘The marginal cost of public funds is one.’ CESifo Working Paper No. 3250 Munich
- Jacquet, Laurence, and Etienne Lehmann (2014) ‘Optimal nonlinear income taxation with multidimensional types: The case with heterogeneous behavioral responses.’ THEMA Université de Cergy-Pontoise No. 2014-01 Paris
- 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
- Judd, Kenneth, and Che-Lin Su (2006) ‘Optimal income taxation with multidimensional taxpayer types.’ In *Computing in Economics and Finance*, ed. Christopher F. Baum, vol. 471 Society for Computational Economics
- Kaplow, Louis (2006) ‘On the undesirability of commodity taxation even when income taxation is not optimal.’ *Journal of Public Economics* 90(6), 1235–1250
- Kleven, Henrik J., Claus T. Kreiner, and Emmanuel Saez (2009) ‘The optimal income taxation of couples.’ *Econometrica* 77(2), 537–560
- Kocherlakota, Narayana R. (2005) ‘Zero expected wealth taxes: A Mirrlees approach to dynamic optimal taxation.’ *Econometrica* 73(5), 1587–1621
- Kopczuk, Wojciech (2013) ‘Incentive effects of inheritances and optimal estate taxation.’ NBER Working Paper No. 18747
- Laroque, Guy R. (2005) ‘Indirect taxation is superfluous under separability and taste homogeneity: A simple proof.’ *Economics Letters* 87(1), 141–144
- Lockwood, Benjamin B., and Matthew C. Weinzierl (2012) ‘De gustibus non est taxandum: Theory and evidence on preference heterogeneity and redistribution.’ NBER Working Paper No. 17784 Cambridge, MA
- Mankiw, N.Gregory, Matthew Weinzierl, and Danny Yagan (2009) ‘Optimal taxation in theory and practice.’ *Journal of Economic Perspectives* 23(4), 147–174
- McAfee, R. Preston, and John McMillan (1988) ‘Multidimensional incentive compatibility and mechanism design.’ *Journal of Economic Theory* 46(2), 335–354
- Mirrlees, James A. (1971) ‘An exploration in the theory of optimum income taxation.’ *Review of Economic Studies* 38(2), 175–208
- (1976) ‘Optimal tax theory: A synthesis.’ *Journal of Public Economics* 6(4), 327–358
- Pass, Brendan (2012) ‘Convexity and multi-dimensional screening for spaces with different dimensions.’ *Journal of Economic Theory* 147(6), 2399–2418
- Pavan, Alessandro, Ilya Segal, and Juuso Toikka (2014) ‘Dynamic mechanism design: A myersonian approach.’ *Econometrica* 82(2), 601–653

- Piketty, Thomas (2014) *Capital in the 21st Century* (Cambridge, MA: Harvard University Press)
- Piketty, Thomas, and Emmanuel Saez (2013) ‘A theory of optimal inheritance taxation.’ *Econometrica* 81(5), 1851–1886
- Ramsey, Frank P (1927) ‘A contribution to the theory of taxation.’ *The Economic Journal* pp. 47–61
- Renes, Sander, and Floris T. Zoutman (2014) ‘When a price is enough: Implementation in optimal tax design.’ Rotterdam. mimeo
- Roberts, Kevin (1984) ‘The theoretical limits to redistribution.’ *Review of Economic Studies* 51(2), 177–195
- Rochet, Jean-Charles, and Philippe Choné (1998) ‘Ironing, sweeping, and multidimensional screening.’ *Econometrica* 66(4), 783–826
- Rothschild, Casey, and Florian Scheuer (2011) ‘Optimal taxation with rent-seeking.’ NBER Working Paper No. 17035 Cambridge, MA
- (2013) ‘Redistributive taxation in the Roy model.’ *Quarterly Journal of Economics* 128(2), 623–668
- (2014) ‘A theory of income taxation under multidimensional skill heterogeneity.’ NBER Working Paper No. 19822 Cambridge, MA
- 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), 205–229
- (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.’ *American Economic Review* 106(1), 24–45
- Sandmo, Agnar (1993) ‘Optimal redistribution when tastes differ.’ *FinanzArchiv/Public Finance Analysis* 50(2), 149–163
- Scheuer, Florian (2013) ‘Adverse selection in credit markets and regressive profit taxation.’ *Journal of Economic Theory* 148(4), 1333–1360
- (2014) ‘Entrepreneurial taxation with endogenous entry.’ *American Economic Journal: Economic Policy* 6(2), 126–163
- Seade, Jesus K. (1977) ‘On the shape of optimal tax schedules.’ *Journal of Public Economics* 7(2), 203–235
- Seremet, Victor D. (2002) *Handbook of Green’s Functions and Matrices* (Southampton: WIT press)

Showalter, RE (1977) *Hilbert space methods for partial differential equations* (San Marcos, TX: Electronic Monographs in Differential Equations)

Strauss, Walter A (1992) *Partial differential equations: An introduction* (New York, NY: John Wiley & Sons)

Zheng, Charles Z. (2000) ‘Optimal auction in a multidimensional world.’ Center for Mathematical Studies in Economics and Management Science, Northwestern University, Discussion Paper 1282 Evanston, IL

A Proof of proposition 1

Proof. The first order conditions for incentive compatibility are given by:

$$\begin{aligned}\mathbf{0}_J &= \left. \frac{\partial u(\mathbf{x}(\mathbf{m}), y(\mathbf{m}), \mathbf{n})}{\partial \mathbf{m}} \right|_{\mathbf{m}=\mathbf{n}}, \\ &= \mathbf{x}'(\mathbf{n})^T u_x(\mathbf{x}(\mathbf{n}), \mathbf{n}) + u_y(y(\mathbf{n})) y'(\mathbf{n})^T,\end{aligned}\quad (41)$$

where $\mathbf{0}_J$ denotes a J -column vector of zeros. This can be rewritten as follows:

$$y'(\mathbf{n}) = \mathbf{s}(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n})^T \mathbf{x}'(\mathbf{n}), \quad (42)$$

which is equation (3).

We can derive (5) from (2) using the envelope theorem:

$$\begin{aligned}V'(\mathbf{n})^T &= \mathbf{x}'(\mathbf{n})^T u_x(\mathbf{x}(\mathbf{n}), \mathbf{n}) + u_y(y(\mathbf{n})) y'(\mathbf{n})^T + u_n(\mathbf{x}(\mathbf{n}), \mathbf{n}), \\ V'(\mathbf{n}) &= u_n(\mathbf{x}(\mathbf{n}), \mathbf{n})^T,\end{aligned}\quad (43)$$

where the latter equality follows from the first-order condition (3).

The second-order conditions are :

$$\left. \frac{\partial^2 u(\mathbf{x}(\mathbf{m}), y(\mathbf{m}), \mathbf{n})}{\partial \mathbf{m}^2} \right|_{\mathbf{m}=\mathbf{n}} \preceq \mathbf{0}, \quad (44)$$

Where $\preceq \mathbf{0}$ denotes the negative semi-definiteness of the matrix. Taking the derivative of (43) with respect to \mathbf{m} gives:

$$\begin{aligned}\frac{\partial^2 u(\mathbf{x}(\mathbf{m}), y(\mathbf{m}), \mathbf{n})}{\partial \mathbf{m}^2} &= \left(u_x(\mathbf{x}(\mathbf{m}), \mathbf{n})^T \otimes I_J \right) \mathbf{x}''(\mathbf{m}) \\ &\quad + \mathbf{x}'(\mathbf{m})^T u_{xx}(\mathbf{x}(\mathbf{m}), \mathbf{n}) \mathbf{x}'(\mathbf{m}) \\ &\quad + u_{yy}(y(\mathbf{m})) y'(\mathbf{m})^T y'(\mathbf{m}) \\ &\quad + u_y(y(\mathbf{m})) y''(\mathbf{m}).\end{aligned}\quad (45)$$

where \otimes denotes the Kronecker product. To simplify this expression we take the total derivative of the first order condition (41):

$$D_n \mathbf{0}_J = D_n \left[\mathbf{x}'(\mathbf{n})^T u_x(\mathbf{x}(\mathbf{n}), \mathbf{n}) + u_y(y(\mathbf{n})) y'(\mathbf{n})^T \right] \quad (46)$$

$$\begin{aligned}\mathbf{0}_{J \times J} &= \left(u_x(\mathbf{x}(\mathbf{n}), \mathbf{n})^T \otimes I_J \right) \mathbf{x}''(\mathbf{n})^T + \mathbf{x}'(\mathbf{n})^T u_{xx}(\mathbf{x}(\mathbf{n}), \mathbf{n}) \mathbf{x}'(\mathbf{n}) + \\ &\quad u_{yy}(y(\mathbf{n})) y'(\mathbf{n})^T y'(\mathbf{n}) + u_y(y(\mathbf{n})) y''(\mathbf{n}) + \\ &\quad \mathbf{x}'(\mathbf{n})^T u_{xn}(\mathbf{x}(\mathbf{n}), \mathbf{n}),\end{aligned}\quad (47)$$

Now combine equations (44, 45, 47) to get the following expression:

$$0 \succ \mathbf{x}'(\mathbf{n})^T u_{x\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n}). \quad (48)$$

Partially differentiate the vector of shadow prices with respect to \mathbf{n} to express (48) in terms of this derivative:

$$\begin{aligned} \mathbf{s}_{\mathbf{n}} &= \frac{-u_{x\mathbf{n}}u_y}{(u_y)^2} \\ u_{x\mathbf{n}} &= -\mathbf{s}_{\mathbf{n}}u_y, \end{aligned}$$

and substitute this result and (42) into (48) to arrive at:

$$\begin{aligned} 0 &\succ \mathbf{x}'(\mathbf{n})^T [-\mathbf{s}_{\mathbf{n}}(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n}) u_y(y(\mathbf{n}))] \\ 0 &\leq \mathbf{x}'(\mathbf{n})^T \mathbf{s}_{\mathbf{n}}(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n}), \end{aligned} \quad (49)$$

where the final inequality, equation (49), follows from the fact that $u_y > 0$.

An equivalent expression can be derived by totally differentiating, equation (43) with respect to \mathbf{n} :

$$\begin{aligned} V''(\mathbf{n}) &= Du_{\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n}) \\ &= \mathbf{x}'(\mathbf{n})^T u_{x\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n}) \\ &\quad + u_{\mathbf{n}\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n}) \end{aligned}$$

Now combine this last expression with (48) to obtain the final equation of the proposition:

$$V''(\mathbf{n}) - u_{\mathbf{n}\mathbf{n}}(\mathbf{x}(\mathbf{n}), \mathbf{n}) \succ 0.$$

■

B The Green's function

B.1 Proof of lemma 2

Proof. To proof the lemma we solve the set of PDEs (17) subject to boundary constraints (16). During the proof we simply denote the right-hand side of equation (17) as $h(\mathbf{n})$ for simplicity. First, assume $\boldsymbol{\theta}(\cdot)$ is a conservative field vector. Under this assumption there exists a scalar potential function $\Theta(\cdot) : \mathbf{N} \rightarrow R$ such that:

$$\Theta'(\mathbf{n}) = \boldsymbol{\theta}(\mathbf{n}) \quad \forall \quad \mathbf{n} \in \mathbf{N}. \quad (50)$$

In that case we can write the PDEs (17) and the boundary constraints (16) as Poisson equation (20) under Neumann boundary constraints given by (21). During the remainder of the proof we refer to this as the Neumann Problem. If this Neumann problem has a solution, it follows that we were correct in assuming that $\boldsymbol{\theta}(\cdot)$ is a conservative field vector. Proving existence of a solution is beyond the scope of this paper. However, Showalter (1977, Th. I.4.5, III.1.1 and III.2.1) shows that the Neumann has a solution if the boundary of \mathbf{N} is piecewise smooth, as we have assumed, and the following condition holds:

$$\int_{\partial N} \Theta'(\mathbf{n}) \mathbf{e} d\mathbf{n} = \int_N h(\mathbf{n}) d\mathbf{n} = 0. \quad (51)$$

For our problem this implies we should choose λ according to equation (19) which is the standard first-order condition for λ . Hence, $\boldsymbol{\theta}(\cdot)$ is a conservative-field vector proving the first part of the lemma.

To derive the solution for $\boldsymbol{\theta}(\cdot)$ we now solve for the potential function $\Theta(\cdot)$. The Neumann problem, can be solved by means of a Green's function $G(\mathbf{n}, \mathbf{m}) : \mathbf{N} \times \mathbf{N} \rightarrow \mathcal{R}$. With a Green's function one can write the solution to the Poisson equation in integral form:

$$\Theta(\mathbf{n}) = \int_N h(\mathbf{m}) G(\mathbf{n}, \mathbf{m}) d\mathbf{m} + \text{boundary terms},$$

as we will show formally below. The Green's function $G(\mathbf{n}, \mathbf{m})$ for the Neumann problem has to satisfy the following properties (see e.g. Franklin, 2012):

$$\left(\frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial \mathbf{m}} \right)^T \mathbf{e} = 0 \quad \forall \mathbf{n} \in N, \mathbf{m} \in \partial N. \quad (52)$$

$$\nabla_{\mathbf{m}}^2 G(\mathbf{n}, \mathbf{m}) = -\frac{1}{V} + \delta(\mathbf{n} - \mathbf{m}) \quad \forall \mathbf{n}, \mathbf{m} \in N. \quad (53)$$

$$\nabla_{\mathbf{n}}^2 G(\mathbf{n}, \mathbf{m}) = \delta(\mathbf{n} - \mathbf{m}) \quad \forall \mathbf{n}, \mathbf{m} \in N. \quad (54)$$

Equation (52) states that the inner product between the gradient of $G(\cdot)$ and the outward normal to N should equal 0 on the boundary of N . Equation (53) states that the Laplacian of $G(\cdot)$ with respect to \mathbf{m} should be the difference between a Dirac-Delta function and $\frac{1}{V}$, where V is the Lebesgue measure of the entire space N : $V \equiv \int_N d\mathbf{m}$. Equation (54) states that the Laplacian of $G(\cdot)$ with respect to \mathbf{n} should be a Dirac-Delta function. The Dirac-Delta function is defined as follows:

$$\delta(\mathbf{n} - \mathbf{m}) : \mathcal{R}^J \rightarrow \mathcal{R}$$

$$\delta(\mathbf{x}) = 0 \quad \forall \mathbf{x} \neq 0,$$

$$\int_N \delta(\mathbf{n} - \mathbf{m}) d\mathbf{m} = 1 \quad \forall \mathbf{n} \in N.$$

That is, the Dirac-Delta function equals zero everywhere except at $\mathbf{0}$ and it should integrate to 1 over the entire space. Clearly, this is only possible if $\delta(\mathbf{0})$ is undefined. Intuitively, the Dirac-Delta function acts as a distribution function with a single mass point at $\mathbf{x} = 0$. Combining the properties of the Dirac-Delta function with conditions (53,54) we arrive at the following conditions:

$$\nabla_{\mathbf{m}}^2 G(\mathbf{n}, \mathbf{m}) = -\frac{1}{V} \quad \forall \mathbf{n} \neq \mathbf{m}, \quad (55)$$

$$\nabla_{\mathbf{n}}^2 G(\mathbf{n}, \mathbf{m}) = 0 \quad \forall \mathbf{n} \neq \mathbf{m}, \quad (56)$$

$$\int_N \nabla_{\mathbf{n}}^2 G(\mathbf{n}, \mathbf{m}) d\mathbf{m} = \int_N \delta(\mathbf{n} - \mathbf{m}) d\mathbf{m} = 1 \quad \forall \mathbf{n} \in N, \quad (57)$$

$$\int_N \nabla_{\mathbf{m}}^2 G(\mathbf{n}, \mathbf{m}) d\mathbf{m} = -\frac{1}{V} \int_N d\mathbf{m} + \int_N \delta(\mathbf{n} - \mathbf{m}) d\mathbf{m} = 0 \quad \forall \mathbf{n} \in N. \quad (58)$$

The final condition (58) is superfluous, because it is implied by condition (52). To see this apply the divergence theorem to $\nabla_{\mathbf{m}}^2 G(\cdot)$:

$$\int_N \nabla_{\mathbf{m}}^2 G(\mathbf{n}, \mathbf{m}) d\mathbf{m} = \int_{\partial N} \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial \mathbf{m}} \mathbf{e} d\mathbf{m} = 0.$$

where the second equality follows through condition (52). Hence, any function $G(\cdot)$ that satisfies conditions (52,55,56,57) is a Green's function for the Neumann problem.

By combining properties (55,58) one can show that:

$$\begin{aligned} \int_N \nabla_m^2 G(\mathbf{n}, \mathbf{m}) \Theta(\mathbf{m}) d\mathbf{m} &= \int_N \delta(\mathbf{n} - \mathbf{m}) \Theta(\mathbf{m}) d\mathbf{m} - \frac{1}{V} \int_N \Theta(\mathbf{m}) d\mathbf{m}, \\ &= \Theta(\mathbf{n}) - \bar{\Theta}, \end{aligned} \quad (59)$$

where $\bar{\Theta} = \frac{1}{V} \int_N \Theta(\mathbf{m}) d\mathbf{m}$ is the average value of $\bar{\Theta}$ in N . The simplification of the first term on the right-hand side follows because the Dirac-Delta function acts as a distribution function with a single mass point at $\mathbf{n} = \mathbf{m}$. The second term is the standard definition of an unweighted average. Note that the second term acts as a constant for all \mathbf{n} .

We can now use Green's second identity in conjunction with (59) to write the solution to the Neumann problem in terms of the Green's function:

$$\begin{aligned} \int_N \nabla_m^2 G(\mathbf{n}, \mathbf{m}) \Theta(\mathbf{m}) - \nabla^2 \Theta(\mathbf{m}) G(\mathbf{n}, \mathbf{m}) d\mathbf{m} &= \quad (60) \\ \int_{\partial N} \left(\Theta(\mathbf{m}) \left(\frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial \mathbf{m}} \right)^T - \Theta'(\mathbf{m}) G(\mathbf{n}, \mathbf{m}) \right) ed\partial N, \\ \Theta(\mathbf{n}) &= \int_N h(\mathbf{m}) G(\mathbf{n}, \mathbf{m}) d\mathbf{m} + \bar{\Theta}. \end{aligned} \quad (61)$$

The first equality in (60) is simply Green's second identity. The right hand side of this expression vanishes because of conditions (21) and (52). The second equality follows from equations (59) and (20). Equation (61) gives the general solution to the Poisson equation in terms of the Green's function. Hence, the problem of solving equations (20,21) can be reduced to finding a Green's function that satisfies equation (52,55,56,57) and then applying equation (61). This proves the second part of the lemma.

To solve for the third part of the lemma take the derivative of Equation (61) towards n_j :

$$\theta_j(\mathbf{n}) = \int_N h(\mathbf{m}) \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial n_j} d\mathbf{m}, \quad (62)$$

and define $\omega_j(\mathbf{n}, \mathbf{m}) \equiv \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial n_j}$ as the weighting function.

From condition (57) and (58) it follows that the weighting function is differentiable everywhere except at $\mathbf{n} = \mathbf{m}$ proving the fourth part of the lemma.

Moreover, conditions (52,55,56,57) only depend on the type space proving the fifth part of the lemma.

Finally, note that the weighting function is unique up to a constant C . To see this choose $\hat{\omega}_j(\mathbf{n}, \mathbf{m}) = \omega_j(\mathbf{n}, \mathbf{m}) + C$ where C is an arbitrary constant, and substitute the new weighting function into (62):

$$\begin{aligned} \int_N h(\mathbf{m}) \hat{\omega}_j(\mathbf{n}, \mathbf{m}) d\mathbf{m} &= \int_N h(\mathbf{m}) \omega_j(\mathbf{n}, \mathbf{m}) d\mathbf{m} + C \int_N h(\mathbf{m}) d\mathbf{m}, \\ &= \theta_j(\mathbf{n}), \end{aligned}$$

where the first equality follows through the sum rule, and the last equality follows through equation (18). This proves the sixth and final point of the lemma. ■

B.2 Case 1: The positive real line

Proof. It is useful to think of R_+ as the interval $[0, V]$ where we will let V approach ∞ at the end of the proof. In that case the boundary of \mathbf{N} is given by the two points $\partial\mathbf{N} = \{n = 0, n = V\}$.

The Green's function for this Neumann problem is given by Franklin (2012):²¹

$$G(n, m) = nH(n - m) + mH(m - n) - \frac{m^2}{2V} \quad \forall \quad n, m \in R_+, \quad (63)$$

Where $H(\cdot)$ is the heaviside step function. To proof this function is the Green's function for this problem, we show that it satisfies conditions (52,55,56,57). To see that this function satisfies (52) take the derivative with respect to m :

$$\frac{\partial G(n, m)}{\partial m} = H(m - n) - \frac{m}{V} \quad \forall \quad m \neq n.$$

The derivative evaluated at $m = 0$ and at $m = V$ both equal zero which satisfies condition (52). Moreover, the second derivative of the Green function with respect to m is given by:

$$\frac{\partial^2 G(n, m)}{\partial^2 m} = -\frac{H(n - m)}{V} - \frac{H(m - n)}{V} = -\frac{1}{V} \quad \forall \quad m \neq n.$$

Therefore, the second derivative takes value $-\frac{1}{V}$ everywhere except at $m = n$. This satisfies conditions (55). Finally, it is easy to verify that conditions (56,57) are satisfied, since the Green's function is symmetric in m and n except for its final term. Hence, (63) is a Green's function for this problem. The weighting function is the derivative of the Green's function with respect to n :

$$\omega(n, m) = \frac{\partial G(n, m)}{\partial n} = H(n - m),$$

which proofs the example. ■

B.3 Case 2: positive quadrant in 2D

Proof. It is again useful to think of the measure of the type-space as limiting to infinity. Therefore, let $n_1, n_2 \in [0, V^{\frac{1}{2}}]$ such that the total area of the type space is $\left(V^{\frac{1}{2}}\right)^2 = V$. We will again let the Lebesgue measure V go to infinity in the proof. The Green's function for the positive quadrant is then given by:

$$\begin{aligned} G(\mathbf{n}, \mathbf{m}) &= \frac{1}{2\pi} \ln(\|\mathbf{n} - \mathbf{m}\|) + \frac{1}{2\pi} \ln(\|\mathbf{n} - \mathbf{m}_*\|) \\ &\quad + \frac{1}{2\pi} \ln(\|\mathbf{n} - \mathbf{m}_{**}\|) + \frac{1}{2\pi} \ln(\|\mathbf{n} + \mathbf{m}\|) - \frac{m_1^2}{4V} - \frac{m_2^2}{4V}, \end{aligned} \quad (64)$$

where $\mathbf{m}_* = (m_1, -m_2)$ and $\mathbf{m}_{**} = (-m_1, m_2)$. To see that this Green's function satisfies conditions (52,55,56,57) note first that the border of R_+^2 can be partitioned into four parts:

$$\begin{aligned} \partial\mathbf{N} &= \{\mathbf{n} = (n_1, 0), n_1 \in R_+\} \cup \{\mathbf{n} = (0, n_2), n_2 \in R_+\} \cup \\ &\quad \left\{ \mathbf{n} = \left(n_1, V^{\frac{1}{2}}\right), n_1 \in R_+ \right\} \cup \left\{ \mathbf{n} = \left(V^{\frac{1}{2}}, n_2\right), n_2 \in R_+ \right\}. \end{aligned}$$

²¹Note that since $V \rightarrow \infty$ all terms that depend on V vanish. We include them here for completeness.

The outward unit normal vectors to each partition respectively are given by $(0, -1)$, $(-1, 0)$, $(0, 1)$ and $(1, 0)$. It follows that we can write (52) as a set of four equations:

$$\frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_1} \Big|_{m_1=0} = 0, \quad (65)$$

$$\frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_2} \Big|_{m_2=0} = 0, \quad (66)$$

$$\lim_{V \rightarrow \infty} \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_1} \Big|_{m_1=V^{\frac{1}{2}}} = 0 \quad (67)$$

$$\lim_{V \rightarrow \infty} \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_2} \Big|_{m_2=V^{\frac{1}{2}}} = 0, \quad (68)$$

To see that equation (64) indeed satisfies these restrictions we take its derivative with respect to m_1, m_2 . To facilitate derivation, note that the first 4 terms of the Green's function (64) all have the same form, and thus very similar derivatives. Denoting the first term of (64) by Γ , we can see the derivative with respect to m_1 is given by:

$$\begin{aligned} \Gamma &= \frac{1}{2\pi} \ln(\|\mathbf{n} - \mathbf{m}\|), \\ \frac{\partial \Gamma}{\partial m_1} &= \frac{1}{2\pi} \left[\frac{1}{\|\mathbf{n} - \mathbf{m}\|} \frac{\partial \|\mathbf{n} - \mathbf{m}\|}{\partial m_1} \right], \\ \frac{\partial \Gamma}{\partial m_1} &= \frac{1}{2\pi} \left[\frac{1}{\|\mathbf{n} - \mathbf{m}\|} \frac{n_1 - m_1}{((n_1 - m_1)^2 + (n_2 - m_2)^2)^{0.5}} \right], \\ \frac{\partial \Gamma}{\partial m_1} &= \frac{n_1 - m_1}{2\pi ((n_1 - m_1)^2 + (n_2 - m_2)^2)}. \end{aligned}$$

Hence, the derivatives of (64) with respect to m_1 and m_2 are given by:

$$\begin{aligned} \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_1} &= \frac{m_1 - n_1}{2\pi [(n_1 - m_1)^2 + (n_2 - m_2)^2]} + \frac{m_1 - n_1}{2\pi [(n_1 - m_1)^2 + (n_2 + m_2)^2]} \quad (69) \\ &+ \frac{n_1 + m_1}{2\pi [(n_1 + m_1)^2 + (n_2 - m_2)^2]} + \frac{n_1 + m_1}{2\pi [(n_1 + m_1)^2 + (n_2 + m_2)^2]} - \frac{m_1}{2V}, \\ \frac{\partial G(\mathbf{n}, \mathbf{m})}{\partial m_2} &= \frac{m_2 - n_2}{2\pi [(n_1 - m_1)^2 + (n_2 - m_2)^2]} + \frac{n_2 + m_2}{2\pi [(n_1 - m_1)^2 + (n_2 + m_2)^2]} \\ &+ \frac{m_2 - n_2}{2\pi [(n_1 + m_1)^2 + (n_2 - m_2)^2]} + \frac{n_2 + m_2}{2\pi [(n_1 + m_1)^2 + (n_2 + m_2)^2]} - \frac{m_2}{2V}. \end{aligned}$$

By substituting in $m_1 = 0$ and $m_2 = 0$ respectively, it can be verified that (65,66) are satisfied. Moreover, (67) and (68) are established by choosing $m_1, m_2 = V^{0.5}$ and taking the limit $V \rightarrow \infty$.

To show condition (55) is satisfied take the second derivative of (64) with respect to

m_1 and m_2 :

$$\begin{aligned}\frac{\partial^2 G(\mathbf{n}, \mathbf{m})}{\partial m_1^2} &= \frac{-(n_1 - m_1)^2 + (n_2 - m_2)^2}{2\pi [(n_1 - m_1)^2 + (n_2 - m_2)^2]^2} + \frac{-(n_1 - m_1)^2 + (n_2 + m_2)^2}{2\pi [(n_1 - m_1)^2 + (n_2 + m_2)^2]^2} \\ &\quad + \frac{-(n_1 + m_1)^2 + (n_2 - m_2)^2}{2\pi [(n_1 + m_1)^2 + (n_2 - m_2)^2]^2} + \frac{-(n_1 + m_1)^2 + (n_2 + m_2)^2}{2\pi [(n_1 + m_1)^2 + (n_2 + m_2)^2]^2} - \frac{1}{2V}, \\ \frac{\partial^2 G(\mathbf{n}, \mathbf{m})}{\partial m_2^2} &= \frac{(n_1 - m_1)^2 - (n_2 - m_2)^2}{2\pi [(n_1 - m_1)^2 + (n_2 - m_2)^2]^2} + \frac{(n_1 - m_1)^2 - (n_2 + m_2)^2}{2\pi [(n_1 - m_1)^2 + (n_2 + m_2)^2]^2} \\ &\quad + \frac{(n_1 + m_1)^2 - (n_2 - m_2)^2}{2\pi [(n_1 + m_1)^2 + (n_2 - m_2)^2]^2} + \frac{(n_1 + m_1)^2 - (n_2 + m_2)^2}{2\pi [(n_1 + m_1)^2 + (n_2 + m_2)^2]^2} - \frac{1}{2V}.\end{aligned}$$

Since $\frac{\partial^2 G(\mathbf{n}, \mathbf{m})}{\partial m_1^2} \Big|_{n \neq m} = -\frac{\partial^2 G(\mathbf{n}, \mathbf{m})}{\partial m_2^2} \Big|_{n \neq m}$, the Laplacian of G equals $-\frac{1}{V}$ for all $\mathbf{m} \neq \mathbf{n}$, satisfying (55). Verifying conditions (56,57) is straightforward, since the Green's function is symmetric between \mathbf{m} and \mathbf{n} except for its final two terms. Hence, (64) is a Green's function to the Neumann problem in the positive quadrant. Finally, the weighting function in the example is found by taking the first derivative of (64) and letting $V \rightarrow \infty$.

■

C Proof of proposition 2

Proof. Starting from the first order conditions (13), (14) and (15), first solve (13) for η :

$$\eta = \frac{\lambda f}{u_y}.$$

Now substitute this expression into (14):

$$\lambda q^{iT} f - u_{x\mathbf{n}} \boldsymbol{\theta} + \frac{\lambda f}{u_y} u_x = \mathbf{0}_k,$$

Consider equation i of this system of equations and simplify:

$$\begin{aligned}q_{x_i} - s_i &= \sum_{j=1}^J -\frac{\partial s_i}{\partial n_j} \theta_j \frac{u_y}{\lambda f}, \\ \frac{q_{x_i} - s_i}{s_i} &= \sum_{j=1}^J \xi_{s_i, n_j} \theta_j \frac{u_y}{\lambda} \frac{1}{n_j f}.\end{aligned}$$

Substitute in equation (22) to arrive at the ABC -formula. ■

D Proof to equation (35)

Proof. To proof equation (35) we start from equation (33) and show that $\xi_{s_i, n_i} = \frac{\varepsilon_{x_i, n_i}}{\varepsilon_{x_i}^c}$ and $\xi_{s_i, n_j} = 0$ for all $j \neq i$ when utility is additively separable. The second part of this

statement is trivial, since the cross-derivative $u_{x_i n_j} = 0$ for all $j \neq i$. To proof the first part we first derive ε_{x_i, n_i} and $\varepsilon_{x_i}^c$ in a setting where we do not restrict the utility function, beyond the restrictions discussed in section 3. Subsequently, we show that $\xi_{s_i, n_i} = \frac{\varepsilon_{x_i, n_i}}{\varepsilon_{x_i}^c}$ follows from additive separability of the utility function.

Take the agent's optimization problem on the market (30) and consider a small perturbation in the tax gradient around the optimum. Denote the perturbation by column vector $\boldsymbol{\tau}$. The agent's optimization problem including the perturbation is given by:

$$\max_{\mathbf{x}, y} u(\mathbf{x}, y, \mathbf{n}) \quad s.t. \quad y = q(\mathbf{x}) - \mathcal{T}(\mathbf{x}) + \boldsymbol{\tau}^T (\mathbf{x} - \mathbf{x}(\mathbf{n}))$$

We use the perturbation $\boldsymbol{\tau}$ to find the effect of a change in the marginal tax rate on the demand for good x_i . A change in $\boldsymbol{\tau}$ affects the marginal tax rate, but at the margin around $\mathbf{x} = \mathbf{x}(\mathbf{n})$ it has no influence on the average tax rate. As such, evaluating the effect of the perturbation at the optimum (i.e. at $\boldsymbol{\tau} = \mathbf{0}, \mathbf{x} = \mathbf{x}(\mathbf{n})$) allows us to find compensated demand effects. The agents' perturbed first-order condition is given by:

$$k(\mathbf{x}, \mathbf{n}, \boldsymbol{\tau}) \equiv u_{\mathbf{x}}(\mathbf{x}, \mathbf{n}) + u_y(y(\mathbf{n})) (q'^T(\mathbf{x}) - \mathcal{T}'^T(\mathbf{x}) + \boldsymbol{\tau}) = \mathbf{0}.$$

The function $k(\mathbf{x}, \mathbf{n}, \boldsymbol{\tau})$ implicitly defines \mathbf{x} as a function of \mathbf{n} and $\boldsymbol{\tau}$. We now totally differentiate $k(\mathbf{x}, \mathbf{n}, \boldsymbol{\tau})$ with respect to $\boldsymbol{\tau}$, \mathbf{n} and all the endogenous variables \mathbf{x} to arrive at the effect of a change in the exogenous variables on demand:

$$\begin{aligned} k_{\mathbf{x}} d\mathbf{x} + k_{\mathbf{n}} d\mathbf{n} + k_{\boldsymbol{\tau}} d\boldsymbol{\tau} &= \mathbf{0}, \\ d\mathbf{x} &= - (k_{\mathbf{x}})^{-1} (k_{\mathbf{n}} d\mathbf{n} + k_{\boldsymbol{\tau}} d\boldsymbol{\tau}) \end{aligned} \quad (70)$$

where $d\mathbf{x}$, $d\mathbf{n}$ and $d\boldsymbol{\tau}$ are column vectors containing the differentials for vectors \mathbf{x} , \mathbf{n} and $\boldsymbol{\tau}$. The second-order condition to the household maximization problem ensure $k_{\mathbf{x}}$ is negative definite. Hence, its inverse exists. Now consider the partial derivatives of $k(\cdot)$ with respect to the exogenous variables evaluated around the optimum, such that $\boldsymbol{\tau} = \mathbf{0}$ and $\mathbf{x} = \mathbf{x}(\mathbf{n})$:

$$\begin{aligned} k_{\mathbf{n}} &= u_{\mathbf{x}\mathbf{n}}, \\ k_{\boldsymbol{\tau}} &= u_y \mathbf{I}, \end{aligned}$$

The compensated own-elasticity depends on the compensated demand effect $\frac{\partial x_i}{\partial \tau_i}$. To arrive at this partial derivative, consider row i of the system of equations (70) and set all differentials of the exogenous variables $\mathbf{n}, \boldsymbol{\tau}$ except $d\tau_i$ equal to zero. We find:

$$\frac{\partial x_i}{\partial \tau_i} \Big|_{u=\bar{u}} = -u_y (k_{\mathbf{x}})_{ii}^{-1},$$

where the notation $(k_{\mathbf{x}})_{ii}^{-1}$ denotes the i -th element on the main diagonal of matrix $(k_{\mathbf{x}})^{-1}$. Normalize to arrive at the compensated elasticity:

$$\varepsilon_{x_i}^c \equiv \frac{\partial x_i}{\partial \tau_i} \Big|_{u=\bar{u}} \frac{q_{x_i}(\mathbf{x}) - \mathcal{T}_{x_i}(\mathbf{x})}{x_i} = \frac{-u_y (k_{\mathbf{x}})_{ii}^{-1} (q_{x_i}(\mathbf{x}) - \mathcal{T}_{x_i}(\mathbf{x}))}{x_i}.$$

In addition we need to calculate the elasticity of consumption of good i with respect to characteristic i . To that end again consider row i of the system of equations (70) and

now set all differentials of the exogenous variables $\mathbf{n}, \boldsymbol{\tau}$ except dn_i equal to zero. We then arrive at:

$$\frac{\partial x_i}{\partial n_i} = - (k_{\mathbf{x}})_i^{-1} u_{x n_i} = - (k_{\mathbf{x}})_{ii}^{-1} u_{x_i n_i},$$

where $(k_{\mathbf{x}})_i^{-1}$ denotes the i -th row of matrix $(k_{\mathbf{x}})^{-1}$, and the final equality follows from the fact that we assumed utility is additively separable, such that $u_{x_i, n_j} = 0$ for all $i \neq j$. Normalize to arrive at the elasticity of x_i with respect to characteristic n_i :

$$\varepsilon_{n_i, x_i} \equiv \frac{\partial x_i}{\partial n_i} \frac{n_i}{x_i} = \frac{-n_i (k_{\mathbf{x}})_{ii}^{-1} u_{x_i n_i}}{x_i}.$$

The final step is show that ξ_{s_i, n_i} can be written as $\frac{\varepsilon_{n_i, x_i}}{\varepsilon_{x_i}^c}$:

$$\begin{aligned} \frac{\varepsilon_{n_i, x_i}}{\varepsilon_{x_i}^c} &= \frac{n_i u_{x_i n_i}}{u_y (q_{x_i}(\mathbf{x}) - \mathcal{T}_{x_i}(\mathbf{x}))}, \\ &= - \frac{\partial s_i}{\partial n_i} \frac{n_i}{q_{x_i}(\mathbf{x}) - \mathcal{T}_{x_i}(\mathbf{x})}, \\ &= \frac{\partial s_i}{\partial n_i} \frac{n_i}{s^i} = \xi_{s_i, n_i}. \end{aligned}$$

where the second equality follows by noting that $\frac{\partial s_i}{\partial n_i} = -\frac{u_{x_i n_i}}{u_y}$, and the third equality follows by substituting in the agent's FOC.

■

E Simulation Algorithm

In this appendix we describe our simulation algorithm step-by-step.

1. We first specify the primitives of the model. These are i.) the utility function from equation (37):

$$u(\mathbf{x}, y, \mathbf{n}) = y - \sum_{i=1}^2 \frac{n_i}{\delta_i} \left(\frac{x_i}{n_i} \right)^{\delta_i},$$

- ii.) the welfare function from equation (40):

$$W(V(\mathbf{n})) = \log(V(\mathbf{n})),$$

- iii.) the couples' production function from equation (39):

$$q(\mathbf{x}) = x_1 + x_2,$$

- iv.) the distribution of unobserved characteristics from equation (38):

$$f(\mathbf{n}) = (\alpha + 1) \alpha (\underline{n}_1 \underline{n}_2)^{\alpha+1} (\underline{n}_2 n_1 + \underline{n}_1 n_2 - \underline{n}_1 \underline{n}_2)^{-(\alpha+2)} \quad \forall \quad n_i \in [\underline{n}_i, \infty), \quad (71)$$

where we set the lower bounds $\underline{n}_1 = \underline{n}_2 = 1$, and $\alpha = 2$, and

- v.) the weighting function $\omega(\mathbf{n}, \mathbf{m})$ that matches the type space $n_i \in [\underline{n}_i, \infty)$. Compared to weighting function (27) we make a slight adjustment, because the

lower bound of the type space is larger than zero. The weighting function for this type space is given by:

$$\begin{aligned} \omega_i(\mathbf{n}, m) = & \frac{n_i - m_i}{2\pi [(n_i - m_i)^2 + (n_j - m_j)^2]} + \frac{n_i - m_i}{2\pi [(n_i - m_i)^2 + (n_j + m_j - 2\underline{n}_j)^2]} \\ & + \frac{n_i + m_i - 2\underline{n}_i}{2\pi [(n_i + m_i - 2\underline{n}_i)^2 + (n_j - m_j)^2]} \\ & + \frac{n_i + m_i - 2\underline{n}_i}{2\pi [(n_i + m_i - 2\underline{n}_i)^2 + (n_j + m_j - 2\underline{n}_j)^2]}. \end{aligned} \quad (72)$$

2. With the primitives in place we divide the type space into a discrete grid for computational purposes. The grid, denoted by \mathcal{N} , consists of types $\mathbf{n}^1, \dots, \mathbf{n}^{K^2}$ where the bounds on the grid are chosen such that it covers most of the mass of types. We select the bottom type slightly larger than the bottom of the type space $\mathbf{n}^1 = (\underline{n}_1 + \epsilon, \underline{n}_2 + \epsilon)$, $\epsilon > 0$.²² The top type is denoted $\mathbf{n}^{K^2} = (n^K, n^K)$, where n^K is chosen such that the cumulative mass of types covered by our grid is $F(\mathbf{n}^{K^2}) = 0.9999$. This implies the grid contains about 99.99 percent of the mass of types. In between the top and bottom type we create a “full” grid of types. Having a full grid implies that for each characteristic n_1^k we have a full set of types ranging from (n_1^k, n_2^1) to (n_1^k, n_2^K) and vice versa. A full grid turns out to be extremely useful for numerical purposes. We space the grid log-linearly to account for the high density at the lower end. We choose $K = 100$ such that there are a total of $K^2 = 10,000$ types in our type space.
3. We define an additional grid of types \mathcal{M} consisting of types $\mathbf{m}^1, \dots, \mathbf{m}^{K_M^2}$, which we use as the integration variable in the government’s first-order condition. In equation (72) the weighting function is undefined at points where $\mathbf{n} = \mathbf{m}$. We therefore choose the grid \mathcal{M} such that between each type (n_1^k, n_2^j) and (n_1^{k+1}, n_2^{j+1}) there are four log-linearly spaced types in \mathcal{M} that satisfy $n_1^k < m_1 < n_1^{k+1}$ and $n_2^j < m_2 < n_2^{j+1}$. This guarantees that there exist no two grid points $\mathbf{n} \in \mathcal{N}$ and $\mathbf{m} \in \mathcal{M}$ for which $\mathbf{n} = \mathbf{m}$. Moreover, it makes sure that grid \mathcal{M} contains more types than grid \mathcal{N} to increase the precision of our numerical integration procedure.
4. The next step is to define an initial tax system. We choose a tax system with tax gradient $\mathcal{T}'(\mathbf{x}) = \mathbf{0}$ and the intercept $\mathcal{T}(\mathbf{0}) = -1$. We additionally define an iteration counter ι and set it to 0. We adopt the convention that for any set of variables or functions x, y, \dots curly brackets $\{x, y, \dots\}^\iota$ denotes the value of the variables in iteration ι . For instance the initial tax system can be written as $\{\mathcal{T}'(\mathbf{x}), \mathcal{T}(\mathbf{0})\}^0$.
5. The next step is to numerically approximate the allocation $\{x_1(\mathbf{n}), x_2(\mathbf{n}), y(\mathbf{n}), V(\mathbf{n})\}^\iota$ under the tax system $\{\mathcal{T}'(\mathbf{x}), \mathcal{T}(\mathbf{0})\}^\iota$. We do this by solving the household optimization problem for all types in \mathcal{M} and \mathcal{N} . Because the utility function is quasi-linear in y it is possible to solve for \mathbf{x} and y independently. We first solve for \mathbf{x} . The first-order conditions for a type $\mathbf{n}^k = (n_1^k, n_2^k) \in \mathcal{N}, \mathcal{M}$ and good x_i is given

²²The density function at the exact bottom of the type space is undefined.

by:

$$1 - \mathcal{T}_{x_i}(x_1, x_2) = \left(\frac{x_i}{n_i^k} \right)^{\delta_i - 1},$$

For positive tax rates, the left-hand side of the first-order condition of each good is bounded between 0 and 1. Hence, $L_i^k \equiv \frac{x_i}{n_i^k}$ is bounded between 0 and 1 as long as the tax rate is non-negative. This allows us to solve the first-order conditions for x_1 and x_2 using a grid-search approach. To be precise, we create a full grid L which consists of 750^2 linearly spaced values in the interval $[0, 1] \times [0, 1]$. For each type we then choose the combination $\{L_1, L_2\} \in L$ that minimizes, the Euclidian distance of the vector of FOCs. Formally: written as:

$$\min_{\{L_1, L_2\}} \left\{ \sum_{i=1}^2 \left((1 - \mathcal{T}_{x_i}(n_1^k L_1, n_2^k L_2)) - L_i^{\delta_i - 1} \right)^2 : \{L_1, L_2\} \in L \right\}. \quad (73)$$

The first-order condition depends on the tax gradient $\mathcal{T}'(\cdot)$ as a function of $n_i^k L_i$. Hence, in order to do a full grid search procedure we need to evaluate the tax gradient for each $\mathbf{n} \in \mathcal{N}, \mathcal{N}$ and each $L_1, L_2 \in L$. However, the planner's first-order condition below only provides information on the tax gradient for the optimal choices of types in \mathcal{N} . We use nearest-neighbor interpolation to approximate for the marginal tax rates on other points. This does not have a big impact on the results unless the tax function is highly non-linear or the types in \mathcal{N} are spaced very far apart, our grid L guarantees the later is not the case. The optimization problem (73) then allows us to calculate $\{x_1(\mathbf{n}), x_2(\mathbf{n})\}^t$ for all types in \mathcal{N}, \mathcal{M} .

We proceed by solving for indirect utility $V(\mathbf{n})$ and consumption $y(\mathbf{n})$. We first calculate the consumption of the bottom type, $y(\mathbf{n}^1)$. Since the marginal tax rate of the bottom type equals zero, its budget constraint is given by:

$$\begin{aligned} y(\mathbf{n}^1) &= q(\mathbf{x}(\mathbf{n}^1)) + \mathcal{T}_0 \\ &= \sum_{i=1}^2 x_i(\mathbf{n}^1) + \mathcal{T}_0, \end{aligned}$$

The indirect utility of the bottom type follows by substituting the optimal choice variables into the utility function:

$$\begin{aligned} V(\mathbf{n}^1) &= u(\mathbf{x}(\mathbf{n}^1), y(\mathbf{n}^1), \mathbf{n}^1) \\ &= y(\mathbf{n}^1) - \sum_{i=1}^2 \frac{\delta_i}{n_i^1} \left(\frac{x_i(\mathbf{n}^1)}{n_i^1} \right) \end{aligned}$$

This provides us with $\{x_1(\mathbf{n}^1), x_2(\mathbf{n}^1), y(\mathbf{n}^1), V(\mathbf{n}^1)\}^t$ for the bottom type. To find the allocation for the remaining types, note that the incentive-compatibility constraint (5) provides information on the gradient of the indirect utility function. The gradient of $V(\mathbf{n})$ is given by:

$$\begin{aligned} V_{n_i}(\mathbf{n}) &= u_{n_i}(\mathbf{x}(\mathbf{n}), y(\mathbf{n}), \mathbf{n}), \\ &= \left(1 - \frac{1}{\delta_i} \right) \left(\frac{x_i(\mathbf{n})}{n_i} \right)^{\delta_i}. \end{aligned}$$

Since the gradient only depends on variables that we have already determined, we can find the indirect utility of each of the agents by integrating over this gradient. We use two iterative methods to obtain the indirect utility of a type (n_1^k, n_2^j) . First, we consider the indirect utility of the previous type $V(n_1^{k-1}, n_2^j)$ and note that:

$$V(n_1^k, n_2^j) = V(n_1^{k-1}, n_2^j) + \int_{n_1^{k-1}}^{n_1^k} V_{n_1}(x, n_2^j) dx.$$

Alternatively, we can apply the same method now considering the indirect utility of type $V(n_1^k, n_2^{j-1})$:

$$V(n_1^k, n_2^j) = V(n_1^k, n_2^{j-1}) + \int_{n_2^{j-1}}^{n_2^j} V_{n_2}(n_1^k, x) dx.$$

We use the trapezoidal method to approximate the integral using both approaches. In the absence of an approximation error, both approaches yield the same result. However, the gradient of $V(\mathbf{n})$ is typically non-linear, and hence, the trapezoidal method contains an approximation error. In our application the trapezoidal approximation tends to underestimate the true value of the integral. We therefore use the maximum of both approximations as the value $V(n_1^k, n_2^j)$.²³

Finally, to find $y(\mathbf{n})$ we use the relationship between the direct and the indirect utility function again:

$$\begin{aligned} V(\mathbf{n}^k) &= u(x_1(\mathbf{n}^k), x_2(\mathbf{n}^k), y(\mathbf{n}^k)), \\ &= y(\mathbf{n}^k) - \sum_{i=1}^2 \frac{\delta_i}{n_i^k} \left(\frac{x_i(\mathbf{n}^k)}{n_i^k} \right), \\ y(\mathbf{n}^k) &= V(\mathbf{n}^k) + \sum_{i=1}^2 \frac{\delta_i}{n_i^k} \left(\frac{x_i(\mathbf{n}^k)}{n_i^k} \right). \end{aligned}$$

This provides us with the entire allocation $\{x_1(\mathbf{n}), x_2(\mathbf{n}), y(\mathbf{n}), V(\mathbf{n})\}^t$.

6. The next step is to calculate the optimal tax rate for allocation $\{x_1(\mathbf{n}), x_2(\mathbf{n}), y(\mathbf{n}), V(\mathbf{n})\}^t$ using the *ABC*-formula in proposition 2. For our utility function the *A*-term simplifies to:

$$\begin{aligned} A_{ii}(\mathbf{n}) &= 1 - \frac{1}{\delta_i} \quad \forall \quad i = 1, 2, \\ A_{ij}(\mathbf{n}) &= 0 \quad \forall \quad j \neq i. \end{aligned}$$

Hence, the *ABC*-formula for each good k can be written as:

$$\frac{\mathcal{T}_{x_i}(x_1(\mathbf{n}), x_2(\mathbf{n}))}{1 - \mathcal{T}_{x_i}(x_1(\mathbf{n}), x_2(\mathbf{n}))} = \left(1 - \frac{1}{\delta_i}\right) B_{ii} C_{ii}. \quad (74)$$

Using our welfare function (40) and utility function (37) the B_{ii} simplifies to:

$$B_{ii}(\mathbf{n}) = \int_{m \in \mathbf{N}} \left(\frac{1}{V(\mathbf{m}) \lambda} - 1 \right) f(\mathbf{m}) \omega_i(\mathbf{n}, \mathbf{m}) dm. \quad (75)$$

²³If $k = 1$ the first approximation approach does not work. In that case we choose the second approximation by default, and vice versa if $j = 1$. We do not need to consider the case where $k = j = 1$ since we have already calculated indirect utility for the bottom type.

To calculate λ we use equation (19). With our welfare and utility function this equation can be written as:

$$\lambda = \int_{\mathbf{m} \in \mathcal{N}} \frac{f(\mathbf{m})}{V(\mathbf{m})} d\mathbf{m}. \quad (76)$$

We approximate the integrals in (75) and (76) over grid \mathcal{M} using Simpson's rule. Finally, the C -term is simply given by:

$$C_{ii}(\mathbf{n}) = \frac{1}{n_i f(\mathbf{n})}. \quad (77)$$

Combining equations (74-77) we arrive at the following optimal marginal tax rates under the given allocation:

$$\mathcal{T}_{x_i}(x_1(\mathbf{n}), x_2(\mathbf{n})) = \frac{\frac{(1-\frac{1}{\delta_i})}{n_i f(\mathbf{n})} \int_{\mathbf{m} \in \mathcal{N}} \left(\frac{1}{V(\mathbf{m})\lambda} - 1 \right) f(\mathbf{m}) \omega_i(\mathbf{n}, \mathbf{m}) d\mathbf{m}}{1 + \frac{(1-\frac{1}{\delta_i})}{n_i f(\mathbf{n})} \int_{\mathbf{m} \in \mathcal{N}} \left(\frac{1}{V(\mathbf{m})\lambda} - 1 \right) f(\mathbf{m}) \omega_i(\mathbf{n}, \mathbf{m}) d\mathbf{m}},$$

which we can evaluate for all agents in the grid \mathcal{N} . To arrive at a new tax gradient we take a weighted average between the old tax gradient and the optimal tax gradient:

$$\{\mathcal{T}'(x_1(\mathbf{n}), x_2(\mathbf{n}))\}^{\iota+1} = \kappa \mathcal{T}'(x_1(\mathbf{n}), x_2(\mathbf{n})) + (1 - \kappa) \{\mathcal{T}'(x_1(\mathbf{n}), x_2(\mathbf{n}))\}^{\iota},$$

where $\kappa \in (0, 1)$ denotes the speed with which the algorithm updates. Higher values of κ can increase the rate of convergence, but may in some cases lead to a less stable algorithm. In our algorithm we set $\kappa = 0.1$.

We use the resource constraint measured as a population average to update the value of the intercept $\mathcal{T}(0)$. The resource deficit in the economy is given by:

$$DEF = \int_{\mathbf{n} \in \mathcal{N}} (y(\mathbf{n}) - x_1(\mathbf{n}) - x_2(\mathbf{n})) f(\mathbf{n}) d\mathbf{n}.$$

We update $\mathcal{T}(0)$ by setting:

$$\{\mathcal{T}(0)\}^{\iota+1} = \{\mathcal{T}(0)\}^{\iota} + DEF.$$

This provides us with a new tax system $\{\mathcal{T}'(\mathbf{x}), \mathcal{T}(0)\}^{\iota+1}$.

7. Finally, we calculate convergence measures. In the *ABC*-formula for the optimal tax rate the only part that updates between subsequent iterations are the welfare weights $g(\cdot)$.

We calculate the average distance between the welfare weights in the previous iteration and the current iteration for all types in \mathcal{M} . Denote this distance by $|g|$. In addition, we calculate the absolute value of the resource deficit relative to overall income:

$$RDEF = \frac{|DEF|}{\int_{\mathbf{n} \in \mathcal{N}} (x_1(\mathbf{n}) + x_2(\mathbf{n})) f(\mathbf{n}) d\mathbf{n}}.$$

If $\max(|g|, RDEF) < 0.001$ the algorithm has converged. If not, we set $\iota = \iota + 1$ and the algorithm returns to step 5.