Value Preserving Welfare Weights for Social Optimization Problems*

Alexis Anagnostopoulos, Eva Carceles-Poveda, Yair Tauman

March 13, 2018

Abstract. In this paper, we propose a method that determines endogenous individual weights for social optimization problems, which maximize the weighted sum of individual utilities subject to certain constraints. We first provide three axioms which uniquely determine, for any welfare function, the contribution of a bundle of goods to that welfare function. We then define weights to be *value preserving* (VP) if the contribution of an individual's initial endowments to the social welfare function is proportional to the contribution of the final consumption allocation to that individual's welfare function. We show that VP weights coincide with Negishi weights and the corresponding VP allocations coincide with Walrasian allocations in Arrow-Debreu economies. In contrast to Negishi weights, VP weights can also be used in economies with frictions. In the context of standard optimal taxation problems, we compare the optimal tax scheme under VP weights with the one under exogenously assumed equal weights. By muting the redistribution motive inherent in the equal weights assumption, VP weights can highlight the aspects that derive from taxation motives other than redistribution and can lead to very different implications regarding the optimal taxes. We also show how to extend our general methodology for computing social welfare weights under other normative principles of justice.

Keywords: Social Welfare Function, Heterogeneous Households, Redistribution, Optimal Taxation

JEL Classification: D63, E62, H21

*We would like to thank Marina Azzimonti, Dirk Krueger, Fransesc Ortega, Victor Rios-Rull, Warren Sanderson and Matthew Weinzierl for useful comments and suggestions.

1. Introduction

A typical social optimization problem maximizes the weighted sum of individual utilities over feasible allocations that satisfy certain constraints. Such problems are often used to understand the properties of constrained efficient allocations and to characterize optimal government policy. The purpose of this paper is to provide a method that can be used to determine endogenously the set of individual weights for any social optimization problem.

When the fundamental welfare theorems hold, Negishi (1960) provides a method which endogenously determines welfare weights by requiring that the resulting allocations coincide with the Walrasian allocations. Although the Negishi method was not developed with a view to achieving justice, some authors have argued that market allocations are indeed just.[1] Mankiw's (2010) Just Deserts Theory is a recent example. An alternative view on achieving fair allocations that has been commonly used in modern macroeconomics calls for choosing equal weights (EW) for every agent, which is often justified on the grounds of justice behind the veil of ignorance. However, the literature has pointed out several flaws with utilitarism, as well as empirical evidence against EW in the data[2]. In this paper, we suggest an alternative to the EW approach that is closer to the Negishi approach but that can be applied to environments where the fundamental welfare theorems do not hold. Specifically, the paper defines value preserving (VP) welfare weights to be weights which ensure that the value of an individual's contribution to social welfare is in proportion to the value the individual receives in the final allocation of the social optimization problem.

To define the VP principle, we first characterize axiomatically a mechanism that allows us to compute the per unit contribution of a good to a welfare function. The mechanism we propose is uniquely characterized by a set of three axioms. Given a bundle of commodities, the first axiom (rescaling) requires that the per unit contribution of a commodity should be independent of the units of measurement. The second axiom (separability) asserts that if the welfare function can be decomposed into the sum of different welfare functions, each of which is generated by different and disjoint sets of commodities, then the per unit contribution of a commodity depends only on the corresponding welfare function. The third axiom (continuity) requires that the mechanism is continuous in a neighborhood of both the welfare function and the initial bundle. It is shown that the unique mechanism that satisfies these three axioms equates the per unit contribution of a commodity to the marginal welfare with respect to that commodity.

We use this mechanism in defining the contribution of a bundle of goods to a welfare function as the inner product of the per unit contributions and the bundle. In turn, this provides a way to measure, for any set of individual welfare weights, the contribution of an individual's initial endowment of goods to social welfare (social value) and the contribution

[1] There is a wide range of views on what constitutes a fair approach. For earlier references see Rawls (1971), Noszick (1974) and Aumann (1975). Saez and Stantcheva (2014), Fleurbaey and Maniquet (2014) and Weinzierl (2015) provide authoritative reviews of both classic and recent contributions to that literature.

[2] For example, Weinzierl (2014) provides survey evidence indicating that only a minority of individuals consider the EW approach and its implications for taxation as fair. Moreover, Chang et al (2017) uncover the Pareto weights that justify current income tax systems as optimal in different OECD countries and show that these deviate considerably from EW.

of the individual's final allocation of goods to his individual welfare (individual value). The VP weights are the ones that equate the ratio of social to private (STP) contributions across agents.

Our second contribution lies in the characterization of the VP weights and allocations. We show that VP weights coincide with Negishi weights, and the corresponding VP allocations coincide with Walrasian allocations, for economies in which the welfare theorems hold. Intuitively, the reason is that, in such environments, the gradient of the welfare function at the initial bundle is a competitive equilibrium price vector. Hence, the social value of an individual's endowments coincides with the market value of his initial endowments and, by the budget constraint, also equals the value of his final allocation. This result provides some support for Mankiw's argument regarding the justice of market allocations. The result, however, is only true in a frictionless environment and will not, in general, hold in the presence of frictions. Nevertheless, VP weights and allocations can still be computed for economies in which the equilibrium is not Pareto optimal and they can therefore be seen as a natural extension of the Negishi method to these types of economies. A nice property of the VP approach is that the resulting allocations are invariant under positive affine transformations of one (or more) individual utilities, in contrast with the EW approach, which has the undesirable property that a rescaling of utilities changes the prescribed allocation.

The VP approach and the EW approach make implicit assumptions regarding the social preference for equality and redistribution that stand, in some sense, at opposite extremes[3]. We first illustrate this by characterizing theoretically optimal tax policy in a benchmark first best scenario where the government uses lump sum taxes. In this case, a VP government raises revenues from each individual in proportion to the present (market) value of their initial endowments, implying *no redistribution* of income, whereas a government with EW prescribes taxes that implement *perfect equalization of income* and consumption across individuals. When we depart from a first best environment, however, the different distributional motives have to be weighted againts other concerns, such as efficiency, insurance provision etc and the VP approach can provide a method to abstract from the redistribution effect. We illustrate this through two different examples that have been commonly used in the literature on optimal taxation.

First, we first study a two period example with wealth inequality, a flat distortionay tax on capital income and a flat tax on labor income that is not distortionary. With no initial inequality, both the VP and the EW weights prescribe no taxes on capital and positive taxes on labor to finane government spending. When inequality increases, however, EW taxes become higher for capital even with moderate levels of inequality purely for redistribution reasons and despite the distortions introduced, while VP taxes try to approach the ones that do not redistribute income at all, prescribing capital taxes that never become higher than labor taxes even with extreme levels of inequality. Introducing uninsured idiosyncratic labor income risk does not change the VP tax prescription, since the insurance motive for taxation also calls for high labor taxes to reduce the risky part of income. In sum, the different

[3] Throughout the paper, we use the term redistribution to refer to ex-ante redistribution and distinguish this from insurance which can be thought of as redistribution after the resolution of uncertainty.

distributional concerns of EW and VP weights generate an optimal tax reversal, illustrating that VP weights could potentially lead to very different implications regarding optimal taxes.

Whereas higher labor taxes under VP weights imply redistribution towards the rich in the previous example, we show that this is not a general property of VP weights. To do this, we then study a second example in which individuals also differ in their endowments of skill, while governments can only use a linear tax rate on income to fund a transfer and an exogenous level of government spending. In the absence of government spending, we can establish some interesting implications analytically. First, with no taxes or transfers, the ratios of social to private (STP) contributions, which a VP government will equalize across agents through tax policy, depend positively on the level of effort and ability, and negatively on initial wealth (and on the capital to ability ratio of the individual). Everything else equal, a VP government will not redistribute income as long as agents exert the same level of effort, but otherwise will redistribute towards the agent exerting a higher effort. We consider the fact that VP weights reward effort a desirable property. Second, whereas a utilitarian government always redistributes towards the poor (regardless of whether this is because of low ability or low initial wealth), the type of redistribution of a VP government is not always towards the wealth rich and in fact depends on the reason for why an agent is poor. In particular, if inequality in initial wealth is bigger than inequality in income, as in the US data, and preferences are such that all households exert the same level of effort, the STP contribution is higher for low wealth agents. A VP government will therefore redistribute towards low wealth individuals, who typically have a lower VP welfare weight, also consistent with US data. Whereas other effects play a role with positive government spending, such as how much agents contribute from a tax revenue perspective, an illustrative numerical example that is calibrated to match the Gini of income and wealth in the US data delivers similar implications.

Our paper is related to, and establishes a link between, several distinct strands of literature. It is motivated by the macroeconomic literature that uses EW in environments where the welfare theorems do not hold.[4] Instead, we obtain endogenous welfare weights using the VP principle.[5] We also use standard optimal taxation examples to show how VP weights can be computed and to illustrate how the choice of weights could affect the conclusions in that literature.

In more recent contributions on optimal taxation, Heathcote, Storesletten and Violante (2016), Weinzierl (2014), Saez and Stantcheva (2015) and Fleurbaey and Maniquet (2014) have considered alternatives to the EW approach.[6] Heathcote et al (2014) present an analyt-

[4]Aiyagari (1995), Domeij and Heathcote (2004), Abraham and Carceles-Poveda (2010) and Anagnostopoulos et al (2012) use this social objective for taxation problems. Davila et al (2012) use it to compare the constrained efficient allocations to the competitive equilibrium allocations in an economy with incomplete markets.

[5]Endogenous and time varying Pareto weights are also computed in the literature on endogenous incomplete markets arising from the presence of limited commitment (see for example Kehoe and Perri (2002). In that literature, however, the set of *initial* Pareto weights are exogenously given and it is exactly those initial weights that our approach aims to determine endogenously.

[6]In an earlier contribution, Benabou (2002) uses a social welfare function that focuses on efficiency and abstracts from equity concerns to study taxation and education policy.

ically tractable model with several frictions and they compare the EW implications to those when using the Negishi weights from the corresponding first best economy to determine the optimal tax progressivity in their environment with frictions. In contrast, our VP weights take into account how the value of endowments change depending on the friction and in that sense extend the Negishi approach to economies with frictions. Weinzierl (2014) shows that a social welfare function which includes a utilitarian component but penalizes deviations from equal sacrifice can explain a number of features of US tax policy. In contrast we follow a normative approach and we consider a different principle. Saez and Stantcheva (2015) propose a general non welfarist approach to optimal tax theory that can accommodate different principles or redistribution preferences by applying marginal social welfare weights, inspired by different fairness principles, directly to earning levels. Instead, Fleurbaey and Maniquet (2014) show that one can incorporate different fairness principles into the standard social welfare framework to do optimal taxation by treating utilities as normative indices that embed these ethical principles. Our work is complementary to these two papers. We use a welfarist approach but focus on a specific principle that takes a stand on social preferences about redistribution. Moreover, we develop an axiomatic approach that provides a way to measure the contribution and we also show that one could potentially use it to determine social welfare weights that satisfy other social justice principles and show how to apply our method to potentially dynamic economies with endogenous prices.

The VP principle is related to the literature on values of cooperative games, such as Shapley (1969) and Aumann (1975), but our approach is different, since we measure the value of an agent via the contribution of his initial endowments to welfare without the use of any game theoretic notion. Finally, our axiomatic approach is related to the literature on cost allocation as in Mirman and Tauman (1982) and Samet and Tauman (1982), but we use a different set of axioms and a different functional space to accomodate social welfare functions.[7]

The paper is organized as follows. Section 2 defines in general terms the social optimization problem we are interested in and Section 3 defines and discusses the VP principle. Section 4 presents the axiomatic approach to define the contribution of a bundle of goods to a welfare function and proves the main theorem. Section 5 provides a characterization of the VP weights for an Arrow-Debreu economy. Section 6 discusses how to use our approach with alternative principles. Section 7 provides an application of the theory to optimal taxation examples and Section 8 summarizes and concludes.

2. The Economy

Consider an economy with L goods that are indexed by $l \in \{1, ..., L\}$, I agents indexed by $i \in \{1, 2, ..., I\}$ and J firms that are indexed by $j \in \{1, 2, ..., J\}$.[8] Let $\mathbb{R}^L$ be the L-dimensional Euclidean space, let $\mathbb{R}^L_+$ be the nonnegative orthant of $\mathbb{R}^L$ and let $\mathbb{R}^L_{++}$ be the

[7] Importantly, we replace the axiom of additivity, which can be justified as an accounting convention in the context of cost functions, but is harder to justify in the context of welfare functions. From a technical perspective, we also provide a much simpler proof that does not rely on the Riesz representation theorem.

[8] In order to reduce notation we also use L to denote the *set* of goods $\{1, ..., L\}$. We follow a similar convention for I and J throughout the paper.

positive orthant of $\mathbb{R}^L$. We let $w^i = \left(w_1^i, ...w_L^i\right) \in \mathbb{R}_+^L$ be the vector of initial good allocations (or endowments) of agent $i \in I$ and $w = \left(w^1, ..., w^I\right) \in \mathbb{R}_+^{LI}$ be the vector of initial good allocations. Similarly, $x^i = \left(x_1^i, ..., x_L^i\right) \in \mathbb{R}_+^L$ represents the vector of final allocations of agent i and $x = \left(x^1, ..., x^I\right) \in \mathbb{R}_+^{LI}$ is the vector of final allocations.

A firm is identified with a production plan $z^j \in \mathbb{R}^L$. Let $z = \left(z^1, ..., z^J\right) \in \mathbb{R}^{LJ}$ and let $\theta_j^i \in \mathbb{R}_+$ be the initial share of agent i in firm j so that $\sum_{i=1}^I \theta_j^i = 1$ for all j. Finally, let $u_i\left(x^i\right)$ be the utility function of agent i. It is assumed that each u_i is continuous on $\mathbb{R}_+^L$.

Definition 1. Social Optimization Problem. A social optimization problem (SOP) is one that maximizes the weighted sum of utilities under constraints, where the individual weights are $\lambda = (\lambda_i)_{i \in I}$, $\lambda_i \geq 0$ for all i and $\sum_{i=1}^I \lambda_i = 1$. The resulting value function $F_\lambda(w) : \mathbb{R}_+^{LI} \to \mathbb{R}$ is a social welfare function (SWF) provided that the maximum exists, which is the case if utilities are continuous and the constraints define a compact subset of $\mathbb{R}_+^{LI} \times \mathbb{R}_+^{LJ}$ for every $w \in \mathbb{R}_+^{LI}$. Formally,

$$
\begin{aligned}
F_\lambda(w) &\equiv \max_{x,z} \sum_{i=1}^I \lambda_i u_i\left(x^i\right) \text{ s.t.} \\
x^i &\in \mathbb{R}_+^L,\ i = 1, ..., I \\
g_s(x, z, w) &\geq 0,\ s = 1, ..., S
\end{aligned}
\tag{1}
$$

where $g_s : \mathbb{R}_+^{LI} \times \mathbb{R}_+^{LJ} \times \mathbb{R}_+^{LI} \to \mathbb{R}$. The inequalities $g_s(x, z, w) \geq 0$ are constraints on the economy and the sets $\{(x, z, w) | g_s(x, z, w) \geq 0 \text{ for all } s\}$ are compact. Examples of social optimization problems will be provided in section 6.

3. Value Preserving (VP) Welfare Weights

Our objective is to determine welfare weights λ or, equivalently, to choose one out of the constrained Pareto optimal allocations. The approach we propose is inspired by Aumann (1975) and Shapley (1969) who select weights so that each individual's allocation is related to their contribution to society. It differs substantially from the aforementioned papers in terms of how this contribution is measured and in terms of how the weights are chosen.

We measure the social value or social contribution of agents through their initial endowments, i.e. through what they bring with them. Similarly, the value of a final allocation of an individual is the contribution of the allocated bundle to their individual welfare. The next section provides a mechanism that can be used to compute the contribution of a bundle of goods to a welfare function, regardless of whether it is an individual utility function or a social welfare function. In this section, we take such a mechanism as given and describe the value preserving principle.

Denote $\tilde{w}^i = \left(0, ..., 0, w^i, 0, .., 0\right) \in \mathbb{R}_+^{LI}$ and notice that $\tilde{w}^i \leq w$ and $\sum_i \tilde{w}^i = w$. Let $\lambda = (\lambda_i)_{i \in I}$ and let $C\left(F_\lambda, \tilde{w}^i, w\right)$ be the contribution of the bundle $\tilde{w}^i$ to the social welfare function F_λ generated by the weights λ and the vector of initial bundles w. Similarly, the contribution of the bundle x^i to individual i's welfare function $\lambda_i u_i$ generated by the bundle x^i is $C\left(\lambda_i u_i, x^i, x^i\right)$. In this case, we can simplify notation and write $C\left(\lambda_i u_i, x^i\right)$.[9] Given the

[9] We have assumed that the individual welfare function depends only on x^i and not any x^h, $h \neq i$. It

private and social contributions, welfare is allocated so that the value of the final allocation to an individual is proportional to the value of the individual's initial endowment bundle to society.

Definition 2. Value Preserving (VP) Weights and Allocations. Consider a SOP with weights $\lambda = (\lambda_i)_{i \in I}$, let $x_\lambda \in \mathbb{R}_+^{LI}$ be a maximizer and $F_\lambda(w)$ be the corresponding maximized value. The weights $\lambda = (\lambda_i)_{i \in I}$ and the final bundle allocation x_λ are value preserving iff

$$\frac{C\left(F_\lambda, \tilde{w}^i, w\right)}{C\left(\lambda_i u_i, x_\lambda^i\right)} = \frac{C\left(F_\lambda, \tilde{w}^h, w\right)}{C\left(\lambda_h u_h, x_\lambda^h\right)} \text{ for all } i, h \in I \tag{2}$$

The contribution of the initial allocation $\tilde{w}^i$ of agent $i \in I$ to social welfare is $C\left(F_\lambda, \tilde{w}^i, w\right)$. Similarly, the contribution of the final bundle x_λ^i to the private welfare function $\lambda_i u_i$ is $C\left(\lambda_i u_i, x_\lambda^i\right)$. Thus, under VP weights, the ratio of the social to the private contribution is equalized across all individuals. To put it differently, the private contribution of an individual is proportional to his social contribution, where the proportion is the same for all individuals.

In the following section, we characterize the contribution mechanism by a set of three axioms. Subsequently, we show that, for Arrow-Debreu complete market economies, VP weights coincide with the Negishi weights and VP allocations coincide with competitive equilibrium allocations. Thus, our principle can be thought of as weighing individuals according to what they could obtain through voluntary trade. There is a range of views on whether such a principle can be thought of as just.[10] Recently, Mankiw (2010) has argued, in the libertarian tradition, that this is indeed the case. Referring to a similar principle, Aumann (1975) suggests instead the term "reasonable compromise" as opposed to "equitable solution". A very different perspective underlies the commonly used approach of assigning equal weights to all individuals, which is often justified using an argument based on the original position 'behind the veil of ignorance'.[11] Our paper does not attempt to offer new arguments in favor of libertarian principles. Heterogeneity in endowments and utilities is assumed exogenously and any concept of justice would necessarily have to address the causes of this initial heterogeneity. The VP principle could be deemed just if one believed that all of the initial heterogeneity is deserved. Equal weights would be easier to justify if all heterogeneity were due to luck. We view these two assumptions regarding initial heterogeneity as the two extremes of a spectrum. The equal weights approach has been extensively used in macroeconomics. VP weights formalize an alternative extreme.

Mechanically, condition (2) provides equations that can be used to endogenously solve for the weights λ. Alternative principles, such as the equal sacrifice principle or the Rawlsian principle, which could be used instead of the VP principle to determine weights are discussed in Section 6. We note here that the VP principle differs from the principle of equal sacrifice used by Weinzierl (2014) and, earlier by Young (1988, 1990) in a fundamental way: private and social values are defined for *the same* economy. In contrast, the equal sacrifice principle

would be straightforward to define u_i on $\mathbb{R}_+^{LI}$, i.e. to allow for individual utilities that depend on the whole distribution x. In that case, the contribution would be denoted $C\left(\lambda_i u_i, \tilde{x}^i, x\right)$.

[10] See the introduction and, specifically, footnote 1.

[11] See Harsanyi (1953, 1955).

defines sacrifice by comparing allocations to a benchmark economy that is different than the economy in question. This introduces a certain level of arbitrariness in the choice of how to define the benchmark economy. For example, in tax applications this could be the economy with frictions but without a government or it could be the frictionless economy. Our approach avoids this difficult choice because it does not rely on such a benchmark economy.

4. Axiomatic Approach

In this section, we introduce an axiomatic approach to characterize the per unit contribution of a good to a welfare function. Let m be the number of goods and let $w = (w_1, ..., w_m) \in \mathbb{R}_+^m$ denote a bundle of these goods.[12] Let $\mathcal{F}^m$ be the set of all functions $F : R_+^m \to R \cup \{-\infty, +\infty\}$ which are continuously differentiable (*cd*) on $\mathbb{R}_+^m \setminus \{0\}$.

Definition 3. Per unit Contribution Mechanism. A per unit contribution mechanism for commodity j, $1 \leq j \leq m$, is a function $\widehat{C}_j(\cdot, \cdot)$ which associates with every integer $m \geq 1$, and every (F, w), $F \in \mathcal{F}^m$, $w \in \mathbb{R}_+^m \setminus \{0\}$ an element $\widehat{C}_j(F, w) \in \mathbb{R}$.

The function $\widehat{C}_j(F, w)$ measures the per unit contribution of the j^{th} commodity to the welfare function F when the overall bundle of goods is w. Next, we present three axioms that uniquely determine the per unit contribution of a good.

Axiom 1: Rescaling. Let $F \in \mathcal{F}^m$ and $G \in \mathcal{F}^m$. Suppose that

$$F(x) = qG(r_1 x_1, .., r_m x_m) + c$$

where $q \neq 0$ and c are real numbers and $r \in \mathbb{R}_{++}^m$. Then for every j, $1 \leq j \leq m$ and all $w \in \mathbb{R}_+^m \setminus \{0\}$

$$\widehat{C}_j(F, w) = q r_j \widehat{C}_j(G, r_1 w_1, .., r_m w_m).$$

This axiom requires that the per unit contribution is independent of the units of measurement of the goods. Consider an economy where the only good is apples ($m = 1$), let x denote apple allocations in kilograms and let $F(x)$ be the welfare function. Let $G(x)$ represent the same welfare as F but with the argument x measured in grams. In this example $r = 1000$, $q = 1$, $c = 0$ and $F(x) = G(1000x)$. Then, $\widehat{C}(F, w)$ is the per kilogram contribution of w kilograms of apples to the welfare F and $\widehat{C}(G, 1000w)$ is the per gram contribution of $1000w$ grams (=w kilograms) of apples to the welfare G. The axiom requires that the per kilogram contribution of w kilograms is the same as 1000 times the per gram contribution of $1000w$ grams of apples

$$\widehat{C}(F, w) = 1000\widehat{C}(G, 1000w)$$

We also allow for the rescaling of the units of utils (which measure the level of welfare). If we change every original util into q new utils, the contribution in terms of the new utils should be q times that of the original utils. For example, consider an economy in which the value of a welfare function G is measured in dollars. Let F represent the same welfare as

[12] This means the welfare function has m arguments. To relate to the notation of the previous section, in the case of an individual welfare function, or in the case where the social welfare function depends only on aggregate endowments, $m = L$. When the social welfare function depends on the whole distribution of endowments, then $m = LI$.

G except that F is measured in cents, namely $F(x) = 100G(x)$. In this case, the axiom requires that $\widehat{C}_j(F,w) = 100\widehat{C}_j(G,w)$ for $1 \leq j \leq m$.

Axiom 2 (Separability). Let $w \in \mathbb{R}_+^m \setminus \{0\}$ and let $(A_r)_{r=1}^k$ be a partition of $\{1,...,m\}$ with cardinality $|A_r| = m_r$ and $\sum_{r=1}^k m_r = m$. Let $F \in \mathcal{F}^m$ and $F^r \in \mathcal{F}^{m_r}$, $1 \leq r \leq k$. Suppose that for all $x \in \mathbb{R}_+^m$

$$F(x) = F^1\left(x^1\right) + ... + F^k\left(x^k\right)$$

where x^r denotes the projection of x on the coordinates of A^r. Then for every $1 \leq r \leq k$ and $j \in A^r$.

$$\widehat{C}_j(F,w) = \widehat{C}_j(F^r, w^r)$$

This axiom refers to the special case where the set of goods can be separated into groups that are independent from each other, in the sense that the effect of one group on the welfare function is independent from the effect of other groups. In a typical economic optimization problem, this happens when all utilities and constraints are separable with respect to different groups of goods. For example, consider the individual welfare function in a multi-period endowment economy with perishable goods, no trade and utility that is additively separable across time. Then each period's bundle of endowments adds to individual welfare (the sum over time of period utilities) only through that period's utility function. The axiom requires that, in this case, the per unit contribution of a good in a given period to the individual's welfare equals the per unit contribution of that good to that period's utility.

Let $X \subseteq R^m$ be a compact set and let $\|F\|_X^1$ represent the $C^1(X)$-norm of F, defined by

$$\|F\|_X^1 = \max_{x \in X}\left[|F(x)| + \sum_{j=1}^m \left|\frac{\partial F}{\partial x_j}(x)\right|\right]$$

Definition 4. Let $F \in \mathcal{F}^m$ and $w \in \mathbb{R}_+^m \setminus \{0\}$. We say that $\widehat{C}_j(F,w)$ is continuous at (F,w) if for every $\epsilon > 0$ there exists $\delta > 0$ and $\eta > 0$ such that for all $\widetilde{w}$, $w_j \leq \widetilde{w}_j \leq w_j + \eta$, $1 \leq j \leq m$, and for all $G \in \mathcal{F}^m$

$$\|F - G\|_{X(w,\eta)}^1 < \delta \rightarrow \left|\widehat{C}_j(F,w) - \widehat{C}_j(G,\widetilde{w})\right| < \epsilon$$

where $X(w,\eta)$ is the box defined by $X(w,\eta) = \Pi_{j=1}^m [w_j, w_j + \eta]$ [13].

Axiom 3: (Continuity). $\widehat{C}_j(\cdot,\cdot)$ is continuous at (F,w) for all $F \in \mathcal{F}^m$ and $w \in \mathbb{R}_+^m \setminus \{0\}$.

Theorem 1 asserts that there is a unique per unit contribution mechanism that satisfies the above three axioms up to a scalar multiplication.

Theorem 1. A per unit contribution mechanism $\widehat{C}_j(\cdot,\cdot)$ satisfies Axioms 1-3 for $1 \leq j \leq m$ iff there exists $\alpha \in \mathbb{R}_{++}$ such that for all $m \geq 1$, $1 \leq j \leq m$, $F \in \mathcal{F}^m$ and $w \in \mathbb{R}_+^m \setminus \{0\}$

$$\widehat{C}_j(F,w) = \alpha \frac{\partial F}{\partial w_j}(w) \tag{3}$$

[13] Since some of the coordinates of w may be zero, we only consider the right neighbourhood of w.

The (positive) constant α is universal, namely, it is the same constant for all pairs (F, w). If $I : \mathbb{R} \rightarrow \mathbb{R}$ is the identity function ($I(x) \equiv x$) then $\alpha = \widehat{C}_1(I, 1)$. By Theorem 1, $\widehat{C}(F, w) = \alpha \nabla F(w)$, where

$$\widehat{C}(F, w) = \left(\widehat{C}_1(F, w), ..., \widehat{C}_m(F, w)\right)$$

The theorem asserts that the per contribution of a good defined by (3) satisfies the three axioms and vice versa, any contribution mechanism that satisfies the three axioms must be given by (3). The proof of Theorem 1 is provided in Appendix A.

Even though the setup of the theorem makes no reference to markets, an interpretation that we find useful in what follows is to think of the gradient $\nabla F(w)$ of the welfare function as a vector of (shadow) prices.

Finally, we define the overall contribution of a bundle of goods as follows.

Definition 5. Contribution Mechanism. Define the set D as follows:

$$D = \left\{ \left(F, w', w\right) | F \in \mathcal{F}^m \text{ for some } m,\ w \in \mathbb{R}^m_+ \setminus \{0\},\ 0 \leq w' \leq w \right\}$$

A contribution mechanism is a function $C(\cdot, \cdot, \cdot)$ which associates with every $(F, \widetilde{w}, w) \in D$ an element $C(F, w', w) \in \mathbb{R}$, where

$$C\left(F, w', w\right) = \sum_{j=1j}^{m} w'_j \widehat{C}_j(F, w) \tag{4}$$

Note that the contribution mechanism $C(F, w', w)$ represents the overall contribution of the bundle $\widetilde{w} \leq w$ to any welfare function $F \in \mathcal{F}^m$ when the initial bundle is $w \in \mathbb{R}^m_+ \setminus \{0\}$.

5. Value Preserving Weights for Arrow-Debreu Economies

When the fundamental welfare theorems hold, the Negishi approach yields welfare weights such that a Pareto optimal allocation chosen by maximizing a weighted sum of individual welfare functions (utilities) coincides with the Walrasian allocation. We show in this section that, in such environments, our approach also yields the Walrasian allocation. Thus, VP weights coincide with Negishi weights and VP allocations coincide with Walrasian allocations in Arrow-Debreu complete market economies. We prove this result for economies with homothetic technologies. The result can be extended to the general case of non-homothetic technologies with a straightforward modification of the VP concept which we discuss at the end of the section.

5.1. Competitive Equilibrium. Each firm j has a convex production set $Z_j \subseteq \mathbb{R}^L$. A production plan z^j belongs to the production set Z_j iff $f_j\left(z^j\right) \leq 0$, where $f_j\left(z^j\right)$ summarizes the technological constraints. We assume that $f_j : \mathbb{R}^L \rightarrow \mathbb{R}$ is twice continuously differentiable, $f_j(0) \leq 0$ and $\bigtriangledown f_j\left(z^j\right) >> 0$ for all $j \in J$. Each firm j maximizes profits and solves:

$$\max_{z^j} pz^j \text{ s.t. } f_j\left(z^j\right) \leq 0$$

where $p \in \mathbb{R}^L_+$ is the vector of commodity prices. Since f_j is continuous, $f_j\left(z^j\right) = 0$ for any maximizer z^j.

Each household is initially endowed with $w^i \in \mathbb{R}_+^L$ units of the goods and $\theta_j^i \geq 0$ shares of each firm j, with $\sum_{i=1}^I \theta_j^i = 1$ for all $j \in J$. Each household's utility is represented by a twice continuously differentiable, strictly increasing and concave function $u_i : \mathbb{R}_+^L \to \mathbb{R}$. Household i takes p, w^i as well as θ_j^i, z^j for all $j \in J$ as given and solves:

$$\max_{x^i} \; u_i\left(x^i\right) \text{ s.t. } px^i = pw^i + \sum_{j=1}^J \theta_j^i pz^j$$

Assuming an interior solution, the first order conditions for $i = 1, ..., I$ with respect to x_l^i, $l = 1, ...L$ imply:

$$\frac{\partial u_i\left(x^i\right)}{\partial x_l^i} = p_l \mu_i \tag{5}$$

where μ_i is the Lagrange multiplier on the individual budget constraint. Let E be the set of all economies described above. For each $e \in E$, we describe next the family of corresponding social optimization problems.

5.2. Social Optimization Problem. The corresponding social optimization problem with welfare weights $\lambda = (\lambda_i)_{i=1}^I$ is given by:

$$\begin{aligned} F_\lambda(w) &= \max_{x,z} \sum_{i=1}^I \lambda_i u_i\left(x^i\right) \text{ s.t.} \\ \sum_{i=1}^I x^i &= \sum_{i=1}^I w^i + \sum_{j=1}^J z^j \\ f_j\left(z^j\right) &\leq 0 \text{ for } j = 1, ..., J \end{aligned}$$

Denote the maximizing production plan by $z_\lambda(w) = \left(z_\lambda^1(w), ..., z_\lambda^J(w)\right)$ and the maximizing consumption allocation by $x_\lambda(w) = \left(x_\lambda^1(w), ..., x_\lambda^I(w)\right)$.[14] Assuming again an interior solution, if η_l denotes the Lagrange multiplier on the resource constraint for good l, it must be that

$$\frac{\partial u_i\left(x_\lambda^i(w)\right)}{\partial x_l^i} = \frac{\eta_l}{\lambda_i} \tag{6}$$

5.3. Equivalence of VP and Negishi weights. Consider first constant returns to scale technologies so that firms make zero profits. That is, assume that $z \in Z_j$ implies $qz \in Z_j$ for all $q \in \mathbb{R}_+$. Proposition 1 states the result for such homothetic technologies.

Proposition 1. Consider an economy $e \in E$ and suppose that technologies are homothetic. Let $\Delta^I = \left\{\lambda \in \mathbb{R}_+^I | \sum_{i=1}^I \lambda_i = 1\right\}$. Then, $\lambda \in \Delta^I$ is a value preserving weight iff it is a Negishi weight for the initial endowments $w = \left(w^1, ..., w^I\right) \in \mathbb{R}_+^{LI} \setminus \{0\}$ of the economy.

The proof is provided in Appendix A. It is straightforward to extend this result to general technologies by extending the definition of the initial endowments. To ensure the equivalence of VP and Negishi weights even for environments where firms can have positive profits,

[14] Note that both the value F_λ and the allocations x_λ, z_λ are written as functions of the whole distribution of endowments w in order to conform to the notation used in previous sections. Clearly, in this case, these can be written as functions of the aggregate endowment only.

one needs to include in the contribution of an individual not only their initial bundle of commodities w^i but also their ownership of firms. The way to do this is to consider the equilibrium production plans of the firms $\left(z^j(w)\right)_{j=1}^J$, multiplied by the shares of individual i in each firm $\left(\theta_j^i\right)_{j\in J}$ as part of that individual's initial endowments. To be precise, define the modified initial endowment of individual i as

$$y^i \equiv w^i + \sum_{j=1}^J \theta_j^i z^j(w), \ i \in I$$

and let $\tilde{y}^i = \left(0, ..., 0, y^i, 0, .., 0\right) \in \mathbb{R}_+^{LI}$. Then measure the contribution of i to social welfare $F_\lambda(w)$ through these modified initial endowments $\tilde{y}^i$ so that this contribution is now $C\left(F_\lambda, \tilde{y}^i, w\right)$. The value preserving weights $\lambda = (\lambda_1, ..., \lambda_I)$ are now the solution to:

$$\frac{C\left(F_\lambda, \tilde{y}^i, w\right)}{C\left(\lambda_i u_i, x_\lambda^i(w)\right)} = \frac{C\left(F_\lambda, \tilde{y}^h, w\right)}{C\left(\lambda_h u_h, x_\lambda^h(w)\right)} \text{ for all } i, h \in I$$

With this modification, replacing $\tilde{w}^i = \left(0, ..., 0, w^i, 0, .., 0\right)$ by $\tilde{y}^i$ it is easy to verify that Proposition 1 holds true and the VP weights coincide with the Negishi weights for production technologies that are convex, even if the technologies are not homothetic.

Proposition 1 implies that, in a first best environment a planner allocating goods to individuals by maximizing a weighted sum of utilities with VP (or Negishi) weights chooses a Walrasian allocation, i.e. individuals obtain an allocation that they can achieve through voluntary trade. One can compare such an allocation to another Pareto optimal allocation that arises from equal weights (EW) and can be justified by appealing to a "behind the veil of ignorance" argument. In the latter, the planner would equalize marginal utility of wealth across consumers under standard assumptions (see Mas-Colell et al (1995)). Clearly, the two methods are very different with regard to the desirability of redistribution and can lead to very different allocations, as we will see in the next section.

Moving away from first best environments, it has been standard practice amongst macroeconomists to take the equal weights approach, partly because the Negishi approach is not applicable. An important exception to this is Heathcote, Storesletten and Violante (2014) who compute the Negishi weights for a first-best version of their model and apply the same weights to their model which incorporates externalities, incomplete markets and distortionary taxes.[15] This would be equivalent to the VP approach if the value of a good in the economy with frictions were the same as in the corresponding first best economy, i.e. equal to the competitive price in the frictionless economy. The VP approach differs because it incorporates the value of each good in the actual (distorted) economy, where the value is measured using the marginal increase in social welfare arising from an increase in the available good in question.

6. Alternative Principles

In this paper, we use a specific principle, the value-preserving principle, to determine endogenously the welfare weights for social optimization problems. Our approach is closely related

[15] Other exceptions include Benabou (2002), Weinzierl (2014), Saez and Stantcheva (2015) and Fleurbaey and Maniquet (2014).

to Mankiw's (2010) "just deserts" theory, which calls for individuals receiving compensation that is congruent with their contribution and for the contribution to be measured according to marginal productivity theory. Mankiw (2010) argues that individuals would exercise the right to leave society and live on their own if they felt their contributions were insufficiently rewarded. This ensures that, under standard assumptions, allocations will be in the core and, for sufficiently large economies, they will be sufficiently close to a Walrasian competitive equilibrium, in which the factors of production are paid their marginal products[16]. In what follows, we take our contribution mechanism as given and discuss how it can be used in the application of principles other than the VP principle. In particular, we look at two popular alternatives to utilitarianism that have been discussed often in the literature, namely the equal sacrifice principle and the Rawlsian principle.

Consider first the equal sacrifice principle. We have used it previously to determine social welfare weights by defining the sacrifice as the difference in utility with respect to the no tax allocation. To see how this principle could be used to determine social welfare weights using something that is more consistent with our approach, we again assume that the starting point from which the sacrifice is calculated is the first best (no tax) allocation.[17] The equal sacrifice condition can then be stated as follows:

$$\frac{C\left(\lambda_i u_i, x_\lambda^i\right)}{C\left(\lambda_i^N u_i, x_N^i\right)} = \frac{C\left(\lambda_h u_h, x_\lambda^h\right)}{C\left(\lambda_h^N u_h, x_N^h\right)} \text{ for all } i, h \in I \tag{7}$$

where $\lambda^N = \left(\lambda_i^N\right)_{i \in I}$ are the Negishi weights corresponding to the first best allocation, which we denote by $x_N \in \mathbb{R}_+^{LI}$, and $x_\lambda \in \mathbb{R}_+^{LI}$ is the final bundle in the allocation with taxes for a given a set of weights $\lambda = (\lambda_i)_{i \in I}$. Condition (7) states that the ratio of the private value of the tax allocation relative to the private value of the first best allocation has to equalize across agents under the equal sacrifice weights. Here, it is important to note that this formulation uses our contribution mechanism to compute the private value to the agent of different allocations and the resulting allocations are invariant under non-uniform affine transformations of utilities. In contrast, this desirable property will not be satisfied if one uses instead the ratio of weighted utility levels $\lambda_i u_i$ in the tax and first best allocations to compute the welfare weights, as is common in the literature.

Our contribution mechanism can also be used to determine social welfare weights that would correspond to a Rawlsian principle**.** These weights would be determined by:

$$\max_\lambda \min_i \left\{C\left(\lambda_i u_i, x_\lambda^i\right)\right\} \tag{8}$$

where x_λ^i is the allocation with taxes for a given set of weights $\lambda = (\lambda_i)_{i \in I}$ and $C\left(\lambda_i u_i, x_\lambda^i\right)$ is the private value to the agent of this allocation. Essentially the Rawlsian weights would maximize the private value to the individual with the lowest payoff. As with the previous

[16] In Section 1 of the computational appendix, we also provide a discussion of altenative contribution mechanisms that have been used in the cost allocation literature and do not relate directly to marginal product accounting.

[17] Note that, although common in the literature, this choice is arbitrary and is made for illustrative purposes. One of the benefits of our VP principle relative to the equal sacrificie principle is that it naturally avoids having to make this choice of a reference economy altogether.

mechanism, the allocations resulting from this problem are invariant under non-uniform affine transformations of utilities. However, if we use the weighted utility levels $\lambda_i u_i$ instead of the contribution to welfare $C\left(\lambda_i u_i, x^i_\lambda\right)$, this property will fail, since the allocations will not be invariant under transformations that add different constants to different utilities.

7. Applications to Optimal Taxation

In this section, we discuss different examples in which the government chooses optimal taxes to maximize a social welfare function based on individual utilities. Although the models are stylized, they are rich enough to allow the analysis of optimal tax policy as an instrument for revenue-raising, distributional concerns and potential insurance provision while under the restrictions imposed by the loss of efficiency from distortionary taxation. First, characterize the optimal VP lump sum taxes in a first best environment. Second, we discuss several settings that depart from the first best. In the first set of examples, individuals differ in their initial level of wealth and are potentially subject to idiosyncratic labor income shocks[18]. In the second set of examples, individuals potentially differ in their preferences, wealth and skill endowments, none of which are tradable[19].

7.1. A Theoretical Result with Lump Sum Taxation. In this section, we consider an Arrow Debreu economy with lump sum taxation. We let $G \in \mathbb{R}^L_+$ be the total lump sum revenue needed to be collected by the government. Let $\widehat{\tau} \in \mathbb{R}^{LI}$ denote a tax policy, where $\widehat{\tau}^i_l$ is the lump sum tax (or transfer) of good l for agent i. We first define the competitive equilibrium given a feasible tax policy $\widehat{\tau} \in \mathbb{R}^{LI}$ such that $\sum_{i\in I} \widehat{\tau}^i = G$.

Definition 3. Given G, a competitive equilibrium with respect to a lump sum tax policy $\widehat{\tau}$ is a tuple $(\widehat{x}, \widehat{z}, \widehat{\tau}, \widehat{p})$, where $\widehat{x}^i \in \mathbb{R}^L_+$, $\widehat{z}^j \in \mathbb{R}^L$, $\widehat{\tau}^i \in \mathbb{R}^L$, $\widehat{p} \in \mathbb{R}^L_+$, $i \in I$, $j \in J$ such that:

$$\sum_{i\in I} \widehat{x}^i = \sum_{i\in I} w^i - G + \sum_{j=1}^{J} \widehat{z}^j \tag{9}$$

$$f_j\left(\widehat{z}^j\right) \leq 0, \text{ for all } j \in J \tag{10}$$

$$\sum_{i\in I} \widehat{\tau}^i = G \tag{11}$$

$$\widehat{z}^j \text{ maximizes } \widehat{p} z^j \text{ over } \left\{z^j | f_j\left(z^j\right) \leq 0\right\}, j \in J \tag{12}$$

$$\widehat{x}^i \text{ maximizes } u_i\left(x^i\right) \text{ over } \left\{x^i \in \mathbb{R}^L_+ | \widehat{p} x^i \leq \widehat{p}\left(w^i - \widehat{\tau}^i + \sum_{j\in J} \theta^j_i \widehat{z}^j\right)\right\}, i \in I \tag{13}$$

We refer to $\widehat{p}$ as a competitive price with respect to $\widehat{\tau}$.

For the study of the optimal taxation problem, we consider all feasible lump sum tax policies and their corresponding competitive equilibria. Namely, the planner faces the following set of constraints $\widehat{L}(G)$ on allocations:

$$\widehat{L}(G) = \left\{(\widehat{x}, \widehat{z}) | \text{there exist } \widehat{\tau} \in \mathbb{R}^{IL} \text{ and } \widehat{p} \in \mathbb{R}^L_+ \text{ s.t. } (\widehat{x}, \widehat{z}, \widehat{\tau}, \widehat{p}) \text{ satisfies (9)-(13)}\right\}$$

[18] This environment have been used extensively in the macroeconomic literature following Bewley (1977, 1983) and Aiyagari (1994).

[19] This type of setting has been studied in the optimal taxation literature following Mirrless (1971).

The social welfare function of the Ramsey taxation problem can then be defined as[20]

$$\widehat{F}_\lambda (w) = \max_{x,z} \sum_{i \in I} \lambda_i u_i \left(x^i\right) \text{ s.t. } (x, z) \in \widehat{L}(G) \tag{14}$$

A weight λ is a VP weight with respect to $\widehat{F}_\lambda (w)$ iff

$$\frac{\bigtriangledown \widehat{F}_\lambda (w) \tilde{w}^i}{\bigtriangledown \widehat{F}_\lambda (w) \tilde{w}^h} = \frac{\lambda_i \bigtriangledown u_i \left(\widehat{x}^i_\lambda\right) \widehat{x}^i_\lambda}{\lambda_h \bigtriangledown u_h \left(\widehat{x}^h_\lambda\right) \widehat{x}^h_\lambda}, i, h \in I$$

where $\widehat{x}_\lambda$ is a maximizer of (14).

As in Proposition 1, we maintain the assumptions of smooth and concave utilities, smooth and convex technologies as well as technologies that are homothetic.[21]

Proposition 2. Let $W - G >> 0$ where $W = \sum_{i \in I} w^i$. A lump sum tax policy $\widehat{\tau}$ is a VP tax and $\lambda \in \Delta^I$ is the corresponding VP weight iff $\widehat{p}\widehat{\tau}^i = \frac{\widehat{p}w^i}{\widehat{p}W}\widehat{p}G$, where $\widehat{p}$ is the competitive price with respect to $\widehat{\tau}$.

The proof of Proposition 2 appears in Appendix A. The proposition shows that VP taxes require each agent's overall tax liability to be proportional to that agent's share of initial wealth $\widehat{p}w^i$. It is easy to verify that

$$\frac{\hat{p}w^i}{\hat{p}w^h} = \frac{\hat{p}\left(w^i - \hat{\tau}^i\right)}{\hat{p}\left(w^h - \hat{\tau}^h\right)} = \frac{\hat{p}x^i}{\hat{p}x^h}, i, h \in I$$

which means that the before-tax and after-tax distributions of wealth (and, hence, expenditure) are the same. Hence, the VP government will not do any redistribution of income. In the special case where the government does not need to raise revenue ($\widehat{p}G = 0$), $\widehat{p}\widehat{\tau}^i = 0$ for all i is a VP tax.[22] In what follows, we apply pur methodology to examples that move away from a first best environment[23].

7.2. Capital versus Labor Income Taxation. We consider an economy with two periods $t = 1, 2$, a continuum of identical firms and three traded goods in each period: a final (consumption) good, capital services and labor. In each period t, a representative firm rents capital K_t and labor N_t at competitive prices R_t and ω_t and uses them as inputs to produce a consumption good Y_t using a constant returns to scale technology

$$Y_t = K_t^\theta N_t^{1-\theta},\ t = 1, 2 \tag{15}$$

where $0 < \theta < 1$.

The economy is populated by a continuum of households represented by the interval $[0, 1]$. Households are of two types which are indexed by $i = 1, 2$. A proportion p_i of households,

[20] This formulation is known as the primal approach to the Ramsey taxation problem, where the government chooses allocations subject to implementability constraints.

[21] We can dispense with homotheticity by modifying the initial bundle as is done right after Proposition 1.

[22] Note that the VP tax $\widehat{\tau}$ is not uniquely determined in this case unless the tax is only on one good. Every τ such that $\widehat{p}\widehat{\tau}^i = \widehat{p}\tau^i$ for all $i \in I$ is a VP tax as well.

[23] In Section 2 of the computational appendix accompanying the paper, we provide a fully analytical example of the model in the next section with lump sum taxation.

with $p_1 + p_2 = 1$, is initially endowed with k_{i1} units of capital. This is the only source of heterogeneity across the two types. Households are also endowed with $T_{it} = T$ units of time in each period which are allocated to labor, which is assumed to be inelastic[24]. In each period, households decide on the supply of labor n_{it} and capital services k_{it+1} to that period's firm and on their demand for final goods. The final goods in period 1 can be consumed c_{i1} or transformed one-to-one to capital k_{i2} for period two production. In period 2, the final goods bought are consumed c_{i2}.

For simplicity, we assume that the government only needs to raise revenue in the second period to finance an exogenously given level of spending G_2.[25] The government uses flat rate taxes τ_k and τ_n on capital and labor income respectively, implying that the government budget constraint is given by $G_2 = \tau_k R_2 K_2 + \tau_n \omega_2 N_2$. Household i solves:

$$\max_{\{c_{i1}, c_{i2}, k_{i1}, n_{i1}, n_{i2}\}} U(c_{i1}, c_{i2}) \text{ s.t.}$$
$$\begin{aligned} c_{i1} + k_{i2} &= R_1 k_{i1} + \omega_1 n_{i1} \\ c_{i2} &= R_2 k_{i2} (1 - \tau_k) + \omega_2 n_{i2} (1 - \tau_n) \qquad (16) \\ 0 &\leq n_{it} \leq T_{it} \; t = 1, 2 \text{ and } c_{it} \geq 0, \; t = 1, 2 \end{aligned}$$

Note that, in this example, labor taxes are not distortionary. Therefore, the efficient tax scheme, in the sense of maximizing aggregate production Y_2, is to tax only labor income. However, the choice of welfare weights will determine societal preference for equality and this can conflict with a pure efficiency objective. Let $\tau = (\tau_k, \tau_n)$ and denote the vector of endowments in the economy by $w = (T_{11}, T_{12}, T_{21}, T_{22}, k_{11}, k_{21})$. Given a set of welfare weights $\lambda = \{\lambda_i\}_{i=1,2}$, with $\lambda_i \geq 0$ with $\sum_i p_i \lambda_i = 1$ and the CE allocations $c^*_{it}(\tau, w)$, the government solves:[26]

$$\begin{aligned} V_\lambda(w) &= \max_{\tau_k} \sum_{i=1}^{2} \lambda_i p_i U\left(c^*_{i1}(\tau, w), c^*_{i2}(\tau, w)\right) \qquad (17) \\ \text{s.t. } G_2 &= \tau_k R_2 K_2 + \tau_n \omega_2 N_2 \end{aligned}$$

Let the consumption allocations that solve this problem be denoted by c_{i1}, c_{i2}. The VP optimal taxes and social welfare weights are then determined by the optimality condition of the government problem, the government budget constraint and the VP condition, which in this example is given by:

$$\frac{\frac{\partial V_\lambda(w)}{\partial k_{11}} k_{11} + \frac{\partial V_\lambda(w)}{\partial T_{11}} T_{11} + \frac{\partial V_\lambda(w)}{\partial T_{12}} T_{12}}{\lambda_1 p_1 \left[U_{c_{11}} c_{11} + \beta U_{c_{12}} c_{12}\right]} = \frac{\frac{\partial V_\lambda(w)}{\partial k_{21}} k_{21} + \frac{\partial V_\lambda(w)}{\partial T_{21}} T_{21} + \frac{\partial V_\lambda(w)}{\partial T_{22}} T_{22}}{\lambda_2 p_2 \left[U_{c_{21}} c_{21} + \beta U_{c_{22}} c_{22}\right]} \qquad (18)$$

In what follows, we compare the prescribed optimal tax schemes under VP and EW weights. Our first result is stated below.

[24] We do not include the analysis of the endogenous labor supply case because it does not add new insights, since the capital tax is still more distortionary than the labor tax. In the next section, we consider an example with endogenous labor supply that also has ability and preference heterogeneity.

[25] The results in this section go through if we impose taxes in the two periods. The analysis of this case can be provided by the authors upon request.

[26] Since G_2 is given, the government just needs to determine τ_k and τ_n will be obtained as a residual.

Proposition 3. If $G_2 = 0$, then the VP taxes are $\tau_k = 0$ and $\tau_n = 0$.

Proposition 3, which is proven in Appendix A, asserts that zero taxes are VP in the absence of government spending. To see why this is the case note that the VP condition in (18) can be simplified to:

$$\frac{\lambda_1 p_1 U_{c_{11}}\left[R_1 k_{11} + w_1 T_{11} + \frac{(1-\tau^n) w_2}{(1-\tau^k) R_2^{RA}} T_{12}\right] + \Phi_1}{\lambda_1 p_1 u U_{c_{11}}\left[R_1 k_{11} + w_1 T_{21} + \frac{(1-\tau^n) w_2}{(1-\tau^k) R_2} T_{22}\right]} = \frac{\lambda_2 p_2 U_{c_{21}}\left[R_1 k_{21} + w_1 T_{11} + \frac{(1-\tau^n) w_2}{(1-\tau^k) R_2} T_{12}\right] + \Phi_2}{\lambda_2 p_2 U_{c_{21}}\left[R_1 k_{21} + w_1 T_{21} + \frac{(1-\tau^n) w_2}{(1-\tau^k) R_2} T_{22}\right]} \tag{19}$$

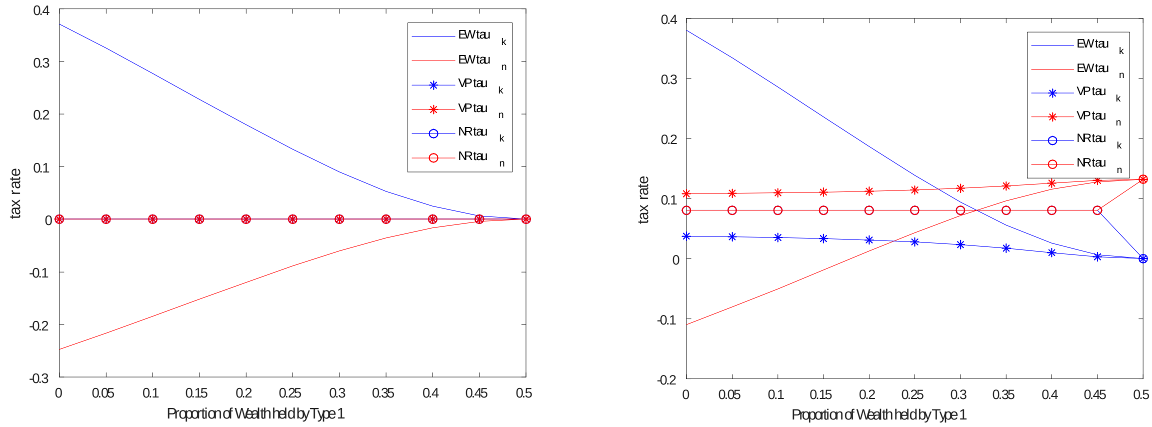
where the term Φ_i represents the the indirect effect on social welfare of a change in the agent's endowments through their effect on the endogenous prices and tax revenues, and it can be writen as:

$$\begin{aligned} \Phi_i &= \Omega \theta Y_2 \left[\theta\left(\tau_k - \tau_n\right) + \tau_n\right] \frac{p_i \left[R_1 k_{i1} + w_1 T_{i1}\right]}{R_1 K_1 + w_1 N_1} + \\ & (\theta - 1)\, \theta \left[\frac{p_i k_{i1}}{K_1} - \frac{p_i T_{i1}}{N_1}\right] \sum_j \lambda_j u_{cj1} \left[\left[\frac{\pi_j k_{j1}}{K_1} - \frac{\pi_j T_{j1}}{N_1}\right] Y_1 + \left[\frac{\pi_j k_{j2}}{K_2} - \frac{\pi_j T_{j2}}{N_2} \frac{(1-\tau^n)}{(1-\tau^k)}\right] K_2\right] \end{aligned} \tag{20}$$

The left and right hand sides of (19) represent the social to private contribution (STP) of type 1 and type 2 agents respectively. As we see, the social contribution of an individual reflects the direct effect of a change in the endowments on social welfare, which happens to be equal to their private contribution, plus the indirect effect captured by Φ_i. As shown in the appendix, these price and tax effects are equal to zero if taxes are zero, implying that the left and right hand side of the VP condition equalizes at zero taxes and these are therefore VP. Intuitively, the model with no taxes is essentially an Arrow Debreu economy, and the VP government can therefore achieve the no redistribution outcome as long as it does not need to raise revenue to finance government spending.

With positive government spending, however, zero taxes are not a feasible solution to the government problem and the tax and price effects captured by Φ_i are not equal to zero any more. In particular, since the time endowments are the same across agents, we can see from (20) that these effects are bigger for the agent that has a bigger share of the initial capital stock. At the same time, the private contribution is also higher for this agent, and it is therefore not clear how the VP government will redistribute income and how close the taxes will be from the ones that do not redistribute income at all. Since the solution for this case is not analytical, we compute it numerically. To do this, we assume a standard time separable constant relative risk aversion utility function $U(c_{i1}, c_{i2}) = \frac{c_{i1}^{1-\sigma}}{1-\sigma} + \beta \frac{c_{i2}^{1-\sigma}}{1-\sigma}$, with a risk aversion of $\sigma = 2$. For the other parameters, we set $\beta = 0.9$, $\theta = 0.4$, $p_i = 0.5$ and $T_{it} = 1$ for $i = 1, 2$. The initial capital is K_1 is normalized to one, and we assume that type 1 agents own a fraction $\phi (= \frac{p_1 k_{11}}{K_1})$, while type 2 agents own a fraction $1 - \phi$ initially. The results for different levels of ϕ are displayed on the right panel of Figure 1 below. For comparison, the left panel displays the results with $G_2 = 0$.

Figure 1: Taxes with $G_2 = 0$ (left) and $G_2 > 0$ (right)



Note that, when $\phi = 0.5$ there is no heterogeneity. As ϕ is decreased towards zero, type 1 agents become progressively more wealth-poor relative to type 2 agents[27]. In the two panels, we also display the EW taxes and the taxes that would achieve no redistribution (NR), in the sense that they equalize the ratio of the present value of tax liabilities to the ratio of the present value of endowments:

$$\frac{\frac{1}{R_2}\left(\tau_k R_2 k_{12} + \tau_n w_2 n_{12}\right)}{R_1 k_{11} + w_1 n_{11} + \frac{w_2}{R_2} n_{12}} = \frac{\frac{1}{R_2}\left(\tau_k R_2 k_{22} + \tau_n w_2 n_{22}\right)}{R_1 k_{21} + w_1 n_{21} + \frac{w_2}{R_2} n_{22}} \tag{21}$$

As discussed earlier, when $G_2 = 0$, VP taxes coincide with the no redistribution zero taxes for all levels of inequality. Equal weights, on the other hand, prescribe increasing tax rates on capital income and subsidy rates on labor income as wealth inequality increases. While this scheme redistributes from the higher wealth types to the lower wealth types, it does not achieve full redistribution due to the presence of distortionary capital income taxes that are not type dependent, which implies that the government has to trade-off the redistribution motive with efficiency considerations.

Consider now the case with $G_2 > 0$ shown on the right panel of the figure. As discussed earlier, the efficient way to raise revenue is to taxing labor income and not capital income, which is exactly the optimal tax scheme when agents are identical ($\phi = 0.5$) and for small levels of inequality, regardless of the government regime (EW or VP). As inequality increases, however, the stronger the incentive of a utilitarian government to tax capital and lower taxes on labor income, purely for redistribution purposes. In contrast, VP tax rates move towards the ones that do not redistribute income (NR), with an increase in capital taxes and a decrease in labor taxes, as inequality increases. Moreover, they are always higher for labor income regardless of the level of inequality, even for the most extreme case in which one of the households is initially endowed with the entire capital stock[28].

[27] We only plot the figures for levels of inequality between 0 and 0.5, since the picture is symmetric for inequality levels between 0.5 and 1.

[28] With taxes in the second period only, it is easy to show that the taxes that do not redistribute income are constant across inequality levels and equal to each other. Moreover, if taxes are imposed in the two periods, the NR taxes are also constant across inequality levels but they will be higher for labor. In either case, the VP taxes still move towards their NR level as inequality increases.

In sum, in this example, both the EW and VP governments have to tradeoff distributional versus efficiency concerns. Whereas a government with EW tries to get as close as possible to full redistribution as inequality increases, the VP government tries get as close as possible to the no redistribution outcome, although it cannot achieve it fully due to the presence of distortionary taxes. This opposite distributional concern leads to a tax reversal with respect to the standard EW approach in this example, even for moderate levels of inequality. This illustrates that it could potentially have very different implications regarding the level of optimal income taxes.

Two additional remarks are worth noting. First, note that another motive for the use of taxation can be provided by the lack of insurance markets in the presence of idiosyncratic risk, since tax policy could play a role in providing such insurance. When we introduce idiosyncratic risk in the second period as in Aiyagari (1994, 1995), however, the overall predictions regarding taxes do not change. The reason is that, given the government's available instruments τ_n and τ_k, the way to provide insurance in this model is to shift the burden of taxation to labor income (opposite to what redistribution would require), since this reduces the risky part of income and increases the safe part. Given this, when markets are incomplete, labor taxes are still higher than capital taxes with VP weights, while the strong redistribution motive under EW dominates the insurance motive and still leads to higher capital taxes when inequality increases in this case.[29]

Second, whereas the government redistributes towards the wealthy for all levels of inequality in this particular example, this is not a general property of VP weights. As we will see in the next section, this can be reversed if the model exhibits different types of heterogeneity.

7.3. Skill and Preference Heterogeneity. In the previous section, we have assumed that agents only differ in their initial wealth ex ante and potentially in their risky income shocks ex-post. In this section, we consider a different set of examples in the spirit of the Mirleess (1971) literature with other potential sources of heterogeneity.

The economy is still populated by two type of agents whose proportion is denoted by p_i. As before, we assume that $T_1 = T_2 = T$ but agents potentially differ in their preferences for leisure, their wealth k_i and their ability or skills ω_i, which the government also takes into account when calculating the social contribution of an agent. An individual's labor income is given by $\omega_i l_i$, where l_i is labor or effort and an individual's total income is equal to $\omega_i l_i + k_i$. The government tax policy consists of a tax rate τ that is proportional to income and a lump sum transfer R. Given the government's policy (τ, R), individuals choose consumption c_i and leisure $T_i - l_i$ to maximize:

$$\max_{c_i, l_i} U^i\left(c_i, T - l_i\right) \text{ s.t.} \tag{22}$$
$$c_i = \omega_i l_i (1-\tau) + k_i (1-\tau) + R$$

Denote the vector of endowments by $w = (\omega_1, \omega_2, T_1, T_2, k_1, k_2)$ and denote the solution

[29] In the interest of space, we do not provide the results with incomplete markets here, but they can be provided by the authors upon request and are included in Section 3 of the computational appendix.

to this problem as a function of the government's tax policy and the individual endowments by c_i^* and l_i^*. Given a set of welfare weights λ, the Ramsey planner chooses the pair (τ, R) to maximize:

$$\begin{aligned} V_\lambda(w) &= \max_\tau \sum_i \lambda_i p_i \left[U^i\left(c_i^*, T - l_i^*\right)\right] \\ G + R &\leq \tau\left[p_1\left(\omega_1 l_1^* + k_1\right) + p_2\left(\omega_2 l_2^* + k_2\right)\right] \end{aligned} \tag{23}$$

where G is an exogenously given level of government spending and $V_\lambda(w)$ is the social welfare function. Denote the solution to this Ramsey problem for a given set of welfare weights by τ and R and the corresponding allocations by c_i and l_i. The government policy (τ, R) and the social welfare weights are determined by the governmnet budget constraint, the first order condition of the Ramsey problem and the VP condition, which in this model is given by

$$\frac{\frac{\partial V_\lambda(w)}{\partial \omega_1}\omega_1 + \frac{\partial V_\lambda(w)}{\partial T_1}T + \frac{\partial V_\lambda(w)}{\partial k_1}k_1}{\lambda_1 p_1\left[U_c^1 c_1 + U_{T-l}^1\left(T_1 - l_1\right)\right]} = \frac{\frac{\partial V_\lambda(w)}{\partial \omega_2}\omega_2 + \frac{\partial V_\lambda(w)}{\partial T_2}T + \frac{\partial V_\lambda(w)}{\partial k_2}k_2}{\lambda_2 p_2\left[U_c^2 c_2 + U_{T-l}^2\left(T - l_2\right)\right]} \tag{24}$$

As shown in the computational appendix, the VP condition above can be simplified to:

$$\frac{g_1(1-\tau)\left[l_1 + \frac{k_1}{\omega_1} + T\right] + \Phi_1}{g_1\left[R + \left(\frac{k_1}{\omega_1} + T\right)(1-\tau)\right]} = \frac{g_2(1-\tau)\left[l_2 + \frac{k_2}{\omega_2} + T\right] + \Phi_2}{g_2\left[R + \left(\frac{k_2}{\omega_2} + T\right)(1-\tau)\right]} \tag{25}$$

where the two sides represent the STP contributions of each agent respectively. Moreover, $g_i = \frac{\lambda_i U_{c_i}^i}{\Omega}$ represents the marginal social welfare weight (MSWW) of agent i defined in Saez et al (2016), measuring the value that society puts on providing an additional dollar of consumption to any given individual and Ω is the social marginal cost of government revenues (or the multiplier of the government budget constraint in the Ramsey problem). Finally„ the terms Φ_i, representing the indirect effects of a change on the endowments of agent i on his social contribution through changes in tax policy, are given by[30]:

$$\Phi_i = \left(\frac{\partial R}{\partial \omega_i}\omega_i + \frac{\partial R}{\partial T_i}T_i + \frac{\partial R}{\partial k_i}k_i\right)\sum_j \lambda_j p_j U_{c_j}^j - \left(\frac{\partial \tau}{\partial \omega_i}\omega_i + \frac{\partial \tau}{\partial T_i}T_i + \frac{\partial \tau}{\partial k_i}k_i\right)\sum_j \lambda_j p_j U_{c_j}^j\left(\omega_j l_j + k_j\right)$$

Several observations can be established using the VP condition when government spending is equal to zero, $G = 0$. First, an important difference with the example of the previous section is that, even though the tax effects Φ_i are equal to zero if $\tau = R = 0$, this policy does not guarantee the equalization of the social to private contributions across agents due to the fact that the government takes into account ability when calculating the social contribution of individuals[31]. As stated in the following proposition, this implies that the VP government will redistribute income even in the absence of government spending and distortions unless households are sufficiently similar, in the sense that they have the same capital to ability ratios and work the same amount of hours.

[30] Note that there are no endogenous prices in this example and therefore no price effects due to changes in endowments.

[31] In constrast, and consistent with the example in the previous section, if the government did not take into account ability when calculating the social contribution, then the STP would equalize across agents with $\tau = R = 0$ and there would not be any redistibution with VP weights in the absence of government spending.

Proposition 3: Assume that $G = 0$ and that there exists a set of social welfare weights λ such that the solution to the Ramsey problem under those weights is $\tau = R = 0$. Then this solution is value preserving (VP) iff:

$$\frac{T + \frac{k_1}{\omega_1} + l_1}{T + \frac{k_1}{\omega_1}} = \frac{T + \frac{k_2}{\omega_2} + l_2}{T + \frac{k_2}{\omega_2}} \text{ or } \frac{l_1}{T + \frac{k_1}{\omega_1}} = \frac{l_2}{T + \frac{k_2}{\omega_2}} \tag{26}$$

The proposition, which follows directly from the VP condition in (25) evaluated at $\tau = R = 0$, illustrates when such a policy is VP. As reflected in (26), the STP contributions depend positively on effort and ability and negatively on initial wealth (and the capital to ability ratio). This has several important implications.

First, $\tau = R = 0$ is VP only if households have the same capital to ability ratios (and in particular if they have the same capital and ability) and preferences are such that both agents exert the same amount of effort (or if labor supply is inelastically supplied)[32]. Intuitively, if there is no heterogeneity or if heterogeneity can be eliminated through labor supply choices, then the economy is de facto an Arrow Debreu setting in which the VP government can achieve the no redistribution objective. Second, VP governments will redistribute income across agents as long as there are differences in either the capital to ability ratios or in labor supply. Interestingly, redistribution with a VP government could be positive or negative depending on the type of initial heterogeneity and on the resulting labor supply behavior, as we illustrate below.

First, consider an economy in which agents have the same capital to ability ratios but preferences are such that one of the agents works more. In this case, the agent working harder will have a higher STP contribution with no taxes or transfers, and we shold expect redistribution from the VP government towards that agent in order to equalize the STP contributions. In contrast, a utilitarian government will redistribute towards the agent with the lowest income, even if he works less. We consider the fact that VP weights reward effort a desirable property.

Second, assume that the level of initial wealth is the same, $k_1 = k_2$, but individuals of type one have a higher ability, $\omega_1 > \omega_2$ so that $\frac{k_1}{\omega_1} < \frac{k_2}{\omega_2}$. If labor supply increases in ability so that $l_1 \geq l_2$ (for example because the substitution effect dominates or because higher skilled individuals have a lower disutility of labor), the STP contribution will be higher for the high ability agent and we should expect the VP government to redistribute towards high ability (rich) individuals. In contrast, in the more realistic case in which $\omega_1 \geq \omega_2$, $k_1 > k_2$ and $\frac{k_1}{\omega_1} > \frac{k_2}{\omega_2}$, reflecting the fact that inequality in initial wealth is bigger than inequality in income, as in the US data. In this case, if preferences are such them households exert the same level of effort, then the STP contribution will be higher for the low wealth agent and we should expect the VP government to redistribute towards low wealth (poor) agents, also consistent with the data.[33]

[32] Note also that, in the absence of initial wealth, $k_1 = k_2 = 0$, the proposition also implies that $\tau = R = 0$ is VP as long as agents work the same number of hours, regardless of their level of ability.

[33] Even if the high wealth agent were to supply a higher effort, he would have to work more than proportionally to the differences in capital to ability ratios to overturn this result.

In sum, VP weights tend to reward effort and they do not always redistribute towards the wealth rich. Moreover, in constrast to EW, which typically redistribute towards the poor, regardless of whether this is because of low wealth or low labor income, VP weights can do different types of redistribution depending on the reason for why an agent is poor.

Positive Government Spending. With positive government spending, $G > 0$, $\tau = R = 0$ is clearly not a solution to the government problem and less clear cut implications can be established regarding the type of redistribution the VP government will do, but we can still gain some intuition by comparing again the STP of the two agent types if no redistribution is allowed, namely, if social preferences are such that the government wants to set $R = 0$ so that $\tau > 0$ necessarily. As shown in the computational appendix, with $R = 0$, the tax effects simplify to:

$$\Phi_i = \tau \left[l_i + \frac{k_i}{\omega_i} + \frac{\partial l_i^*}{\partial \omega_i} + \frac{\partial l_i^*}{\partial T} T + \frac{\partial l_i^*}{\partial k_i} k_i \right] \tag{27}$$

and the STP contribution of an agent can be writen as:

$$SPT_i = \frac{g_i \left[l_i + \frac{k_i}{\omega_i} + T \right] + \frac{\tau}{(1-\tau)} \left[l_i + \frac{k_i}{\omega_i} + \frac{\partial l_i^*}{\partial \omega_i} + \frac{\partial l_i^*}{\partial T} T + \frac{\partial l_i^*}{\partial k_i} k_i \right]}{g_i \left(\frac{k_i}{\omega_i} + T \right)} \tag{28}$$

As reflected by the previous expression, the STP contribution of an individual consists of two additive terms. The first term, which we denote endowment term, is the same term that we have with no government spending and it reflects the direct effect of an individual's endowments to his social contribution. The second term, which we denote tax revenue term, reflects the social contribution of agents through the indirect effect of their endowments on tax revenues.

As before, the endowment term depends positively on effort and ability and negatively on capital (and the capital to ability ratio). Whereas the tax revenue term also depends on these two things, there are other factors that affect it: the elasticity of the labor supply with respect to a change in the endowments and the MGWW g_i. Agents with a higher elasticity to a change in the endowments and a lower MGWW will contribute more from a tax revenue perspective. Given these additional effects, the direction of redistribution is ambiguous, since the endowment and tax revenue effects could go in opposite directions.

In what follows, we present an illustrative numerical example that is calibrated to match the US differences in income and wealth inequality in which the endowment and tax revenue effects do not go in the same direction. For comparison, we also compute the solution with a utilitarian government. As stated above, we choose the ability and initial capital differences of the two types to match the US before tax income Gini coefficient of 0.499 and the US wealth Gini of 0.852 respectively. The individual utility function is assumed to be:

$$U(c_i, l_i) = \frac{\left[c_i^{\mu} (T - l_i)^{1-\mu} \right]^{1-\sigma}}{1-\sigma}$$

where $T = 1$, $\mu = 0.3$ so that labor supply is around 1/3 and $\sigma = 1$. Moreover, $G = 0.1$, which results in a government to income ratio of around 0.17 with VP weights. The numerical results for this case are displayed in Table 2 below.

Table 2: Allocations with $G > 0$, $\sigma = 1$ and $\frac{k_1}{\omega_1} > \frac{k_2}{\omega_2}$

	(λ_1, λ_2)	(τ, R)	(g_1, g_2)	(STP_1, STP_2)	ΔG_w
$R = 0$	$(4.9, 0.55)$	$(0.13, 0)$	$(0.51, 0.62)$	$(1.31, 1.37)$	0%
VP	$(4.80, 0.57)$	$(0.17, 0.02)$	$(0.51, 0.66)$	$(1.34, 1.34)$	-4.2%
U	$(1.00, 1.00)$	$(0.56, 0.12)$	$(0.18, 1.68)$	$(4.4, 0.98)$	-50%

The table displays the results for three cases. The first row corresponds to the case in which we exogenously constrain R to be zero. The next two rows corresponds to the case in which the government has VP and EW respectively. For each case, the table displays the Ramsey government policy (τ, R), the MGWW $g_i = \lambda_i u_c^i$ (not adjusted by the government multiplier Ω), the STP contributions of the two agents and the decrease of the income Gini in percentage terms, which is an indicator of the amount of redistribution the government exerts.

First, the fact that taxes are distortionary prevents the government from equalizing the MGWW, which are much lower for agent 1, who has both a higher ability and a higher wealth. This tends to increase the tax revenue effect and thus his STP contribution. At the same time, the capital to ability ratio is much higher for agent 1 (and he also works less), and this tends to decrease the endowment effect and thus his STP. If we look at the results with $R = 0$, we see that the tax endowment effect dominates the tax revenue effect, implying that the STP contribution of the rich agent of type 1 is lower. The government with VP weights will then equalize the STP contributions by redistributing away from him towards the low ability, low wealth agent. As in the US data, we do observe redistribution towards the poor but bigger welfare weights for the rich with a VP government, although the degree of redistribution, as measured by the change in the Gini of income before and after taxes, is considerably lower than in the data. Whereas the income Gini decreases by 23% in the data, it decreases only by 4.2% in the model.

For comparison, the last two rows of the table reflect that a utilitarian planner will also redistribute towards the low ability, low wealth agent, who is considerably poorer. However, the reason is that a utilitarian government would like to eliminate inequality due to differences in ability and wealth, whereas the VP government wants to equalize the social and private contributions, which are not too far apart with $R = 0$. Hence the relatively modest amount of redistribution with VP weights, compared to the much higher degree of redistribution with EW. As reflected by the table, the utilitarian government redistributes way too much income compared to the data, with a 50% decrease in the income Gini. Interestingly, this seems to suggest a compromise between utilitarian and libertarian governments in the data, as suggested by Weinzierl (2014). In what follows, we also compute the allocations with the equal sacrifice principle that he proposes (ES), showing that the implications of both the ES and the VP principles are indeed close.

Equal Sacrifice. As discussed earlier, an alternative principle to utilitarianism that has been used in the literature is the principle of equal sacrifice (ES), which essentially equates the "sacrifice" from paying taxes across agent types. In this section, we compare the ES solution to the one obtained with VP weights. Following Weinzierl (2014), we measure the sacrifice using utility levels and we assume that the reference point is the allocation with no government policy or government spending $\tau = R = G = 0$. If we denote the allocations with no goverment policy by $\{c_i^*, l_i^*\}_{i=1,2}$, the equal sacrifice principle can be stated as follows[34]:

$$u(c_1^*, T - l_1^*) - u(c_1, T - l_1) = u(c_2^*, T - l_2^*) - u(c_2, T - l_2) \quad (29)$$

This implies that public policy will be set up so as to equalize the utility loss across agents. In order to calculate the weights supporting ES, we can then use condition (29) instead of the VP condition.

The first thing we observe is that no government policy trivially satisfies condition (29) when $G = 0$. Therefore, we assume that $G > 0$ in what follows and analyze under what conditions a government policy imposing $\tau \in [0, 1]$ and $R = 0$ satisfies equation (29). To do this, we use the non separable utility in (**??**) and resort to numerical simulations. Before that, however, we can establish a couple of analytical results. First, if $R = 0$, $l_i = l_i^*$ and the optimal consumption is given by $c_i = (1 - \tau)(\omega_i l_i + k_i) = (1 - \tau)(\omega_i l_i^* + k_i) = (1 - \tau) c_i^*$. So, condition (29) simplifies to:

$$(\omega_1 l_1^* + k_1)^{\mu(1-\sigma)} = (\omega_2 l_1^* + k_2)^{\mu(1-\sigma)}$$

which is not true in general. In fact, if $\sigma \neq 1$, the individual sacrifice of an agent of a given type is decreasing in ω_i for $\sigma > 1$ but increasing for $\sigma < 1$. Given this, we should expect the government to choose $R > 0$ for $\sigma > 1$ and $R < 0$ for $\sigma < 1$. Table 3 below presents the results for for $\sigma = 1.5$, $\sigma = 1$ and $\sigma = 0.5$ and the same parameters as in the previous example with VP weights.

Table 3: ES Allocations with $G > 0$ and $\frac{k_1}{\omega_1} > \frac{k_2}{\omega_2}$

	(λ_1, λ_2)	(τ, R)	(g_1, g_2)	(STP_1, STP_2)	ΔG_w
VP	$(4.80, 0.57)$	$(0.17, 0.02)$	$(0.51, 0.66)$	$(1.34, 1.34)$	-4.2%
ES $(\sigma = 1.5)$	$(6.95, 0.33)$	$(0.16, 0.02)$	$(0.64, 0.81)$	$(1.33, 1.35)$	-3%
ES $(\sigma = 1.0)$	$(6.08, 0.43)$	$(0.13, 0.00)$	$(0.58, 0.71)$	$(1.31, 1.37)$	0%
ES $(\sigma = 0.5)$	$(5.11, 0.54)$	$(0.11, -0.01)$	$(0.51, 0.60)$	$(1.29, 1.40)$	3%

The first row displays the results with VP welfare weights, which are the same for all levels of σ, whereas the last 3 rows display the results with ES welfare weights for the three different levels of σ. Whereas both the VP and ES approaches are in the libertarian tradition and have in common very moderate levels of redistribution, one important observation that we have pointed out already before is that utility levels matter when they are used to compute the sacrifice. Consistent with this, we do observe positive redistribution if if $\sigma > 1$ and negative redistribution if $\sigma < 1$. In contrast, a different version of equal sacrifice that uses our contribution mechanism rather than utility levels to calculate the sacrifice, as discussed in Section 6, would be inmune to this problem.

[34] The equal sacrifice principle can also be stated in terms of ratios rather than differences.

8. Conclusion

This paper provides an axiomatic approach to determine the contribution of a bundle of goods (the value of a bundle) to any welfare function, regardless of whether it is a social or an individual welfare function. We then postulate a value preserving principle which consists in equating across agents the ratio of the value of the initial bundle of goods to social welfare to the value of the final bundle of goods to private welfare. This principle is used to choose amongst different welfare weights or among different constrained efficient allocations. We show that these weights, which we refer to as value preserving weights, can be thought of as an extension of the classic Negishi weights to non Arrow-Debreu economies. In several examples of optimal taxation, we point out the usefulness of our approach in disentangling redistribution motives for taxation from other considerations, such as efficiency or insurance and we illustrate different interesting implications that are in constrast to the ones that would arise with a utilitarian government.

Our approach could be used in other interesting applications. The first is the study of constrained efficient allocations under incomplete markets. Using an equal-weights approach, Davila, Hong, Krusell and Rios-Rull (2012) investigate the constrained efficient level of aggregate capital under incomplete markets. In particular, they investigate whether the equilibrium level of capital is too high or too low. For the most commonly used specification of risk, they find that the competitive equilibrium level of aggregate capital is too low. The extent to which this finding relies on the equal-weights assumption is not obvious. To put it differently, it remains unclear whether the constrained efficient level of capital is higher because this provides better insurance or more redistribution and the VP approach can shed some light on this question. Another potential application is the endogenous determination of the objective for the firm when markets are incomplete and shareholders disagree. In such a setting, one could aggregate the preferences of shareholders by maximizing a weighted average of their utilities, using value preserving weights that equate across shareholders the ratio of the contribution of the initial investment to the firm value and the contribution of the final allocation to private welfare. We leave these applications for further research.

Appendix A: Proofs

Proof of Theorem 1.

It is straightforward to verify that the per unit contribution mechanism $\widehat{C}_j(F,w) = \alpha\frac{\partial F}{\partial x_j}(w)$ for $1 \leq j \leq m$ defined for every $F \in \mathcal{F}^m$ and $w \in \mathbb{R}^m_+ \setminus \{0\}$ satisfies the three axioms. The proof of the other direction is less obvious. Suppose that $\widehat{C}_j(\cdot,\cdot)$ is a per unit contribution mechanism that satisfies the three axioms.

We first prove the theorem for linear welfare functions of the form $F(x) = \sum_{j=1}^m b_j x_j + c$. Namely, we have to show that $\widehat{C}_j(F,w) = \alpha b_j$, $1 \leq j \leq m$ for $\alpha = \widehat{C}(I,1)$, where $I(x) = x$ for all $x \in \mathbb{R}$. We first show that $\widehat{C}(I,1) = \widehat{C}(I,a)$ for all $a \in \mathbb{R}_{++}$. Let $F(x) = I(ax)$. By the rescaling axiom for the units of the commodity, $\widehat{C}(F,1) = a\widehat{C}(I,a)$. On the other hand, $F = aI$, hence $\widehat{C}(F,1) = \widehat{C}(aI,1) = a\widehat{C}(I,1)$ (by the rescaling axiom). These imply $\widehat{C}(I,a) = \widehat{C}(I,1)$. Next define $F^j(x_j) = b_j x_j + c$. Then $F = \sum_{j=1}^m F^j(x_j)$. By separability and the rescaling of welfare units,

$$\widehat{C}_j(F,w) = \widehat{C}\left(F^j, w_j\right) = \widehat{C}(b_j I, w_j) = b_j \widehat{C}(I, w_j)$$

Suppose first that $w_j > 0$. Then

$$\widehat{C}_j(F,w) = b_j \widehat{C}(I,1) = b_j \alpha.$$

Next suppose that $w_j = 0$. By the continuity axiom, for $\epsilon > 0$ there exists $\eta > 0$ sufficiently small s.t. $\left|\widehat{C}(I,0) - \widehat{C}(I,\eta)\right| < \epsilon$. But $\widehat{C}(I,\eta) = \widehat{C}(I,1) = \alpha$. Hence $\left|\widehat{C}(I,0) - \alpha\right| < \epsilon$. Since this is true for any $\epsilon > 0$, $\widehat{C}(I,0) = \alpha$ and $\widehat{C}_j(F,w) = b_j\alpha$.

Next we prove the general case. Let $F \in \mathcal{F}^m$ and $w \in \mathbb{R}^m_+ \setminus \{0\}$. Since F is continuously differentiable on $\mathbb{R}^m_+ \setminus \{0\}$

$$F(x) = \bigtriangledown F(w)\, x + F(w) - w \bigtriangledown F(w) + o(w - x) \tag{30}$$

Let $\epsilon > 0$. By the continuity axiom there exists $\eta_1 > 0$ and $\delta > 0$ s.t. for all $G \in \mathcal{F}^m$ with $\|F - G\|^1_{X(w,\eta_1)} < \delta$

$$\left|\widehat{C}_j(F,w) - \widehat{C}_j(G,w)\right| < \epsilon,\ 1 \leq j \leq m$$

Let $G(x) = \nabla F(w)\, x + c$, where $c = F(w) - w\nabla F(w)$. For the above δ there exists $\eta_2 > 0$ s.t. $\|o(w-x)\|^1_{X(w,\eta_2)} < \delta$.

Let $\eta = \min(\eta_1, \eta_2)$. By (30) and by the definition of the C^1 norm

$$\|F - G\|^1_{X(w,\eta)} \leq \|F - G\|^1_{X(w,\eta_1)} < \delta$$

Therefore,

$$\left|\widehat{C}_j(F,w) - \widehat{C}_j(G,w)\right| < \epsilon \tag{31}$$

By axiom 1,

$$\widehat{C}_j(G,w) = \widehat{C}_j(\nabla F(w)\, x, w)$$

Since $\nabla F(w)x$ is linear

$$\widehat{C}_j(G,w)=\frac{\partial F}{\partial x_j}(w)\alpha$$

Substituting this into (31)

$$\left|\widehat{C}_j(F,w)-\frac{\partial F}{\partial x_j}(w)\alpha\right|<\epsilon$$

Since the last inequality holds for all $\epsilon>0$, we conclude that $\widehat{C}_j(F,w)=\frac{\partial F}{\partial x_j}(w)\alpha$ and the proof of the theorem is complete.■

Proof of Proposition 1. We start by showing that if λ is a Negishi weight then it is value preserving. Let $\lambda\in\Delta^I$ be a Negishi weight. Using (5), (6) and applying the envelope theorem it is straightforward to show (see for example Mas-Colell et al (1995)) that

$$\bigtriangledown F_\lambda(w)=\frac{p}{K}=\lambda_i\nabla u_i\left(x^i_\lambda(w)\right)$$

where $K\equiv\sum_{i=1}^I\frac{1}{\mu_i}$ is a constant.[35] Multiplying by $x^i_\lambda(w)$ we obtain

$$\lambda_i\nabla u_i\left(x^i_\lambda(w)\right)\cdot x^i_\lambda(w)=\bigtriangledown F_\lambda(w)\cdot x^i_\lambda(w)=\frac{p}{K}\cdot x^i_\lambda(w)$$

Since the budget constraint in the competitive equilibrium problem must hold,

$$\begin{aligned}p\cdot x^i_\lambda(w)&=p\cdot\left(w^i+\sum_{j=1}^J\theta^i_j z^j_\lambda(w)\right)\\&=p\cdot w^i\end{aligned}$$

where the last equality follows from the fact that the f_j are homogeneous of degree 1. We have thus shown that

$$\lambda_i\nabla u_i\left(x^i_\lambda(w)\right)\cdot x^i_\lambda(w)=\bigtriangledown F_\lambda(w)\cdot w^i$$

which by Theorem 1 implies

$$C\left(\lambda_i u_i,x^i_\lambda(w)\right)=C\left(F_\lambda,\tilde{w}^i,w\right)$$

and the Negishi weight λ is value preserving.

We now prove the other direction, namely, we show that if λ is a VP weight then it is a Negishi weight for the initial endowments w. Let $\lambda\in\Delta^I$ be a value preserving weight and let $(\bar{x},\bar{z})$ be the corresponding VP allocations so that

$$\bigtriangledown F_\lambda(w)\cdot w^i=c\lambda_i\bigtriangledown u_i\left(\bar{x}^i\right)\cdot\bar{x}^i\tag{32}$$

for some $c\in\mathbb{R}_+$. Since $(\bar{x},\bar{z})$ is a Pareto optimal allocation, by the second welfare theorem there exists some $\bar{p}\in\mathbb{R}^L_+$ such that for every i, $\bar{x}^i$ maximizes $u_i\left(x^i\right)$ over $\left\{x^i|\bar{p}\cdot x^i\le\bar{p}\cdot\bar{x}^i\right\}$

[35] Note that, in this case, the value F_λ depends on the aggregate endowment $W=(W_1,...,W_L)\equiv\sum_{i=1}^I w^i\in\mathbb{R}^L_{++}$ only and not on the whole distribution w. The implication is that $\frac{\partial F_\lambda}{\partial w^i_l}=\frac{\partial F_\lambda}{\partial w^h_l}=\frac{\partial F_\lambda}{\partial W_l}$ for all $i,h\in I$ and the notation $\bigtriangledown F_\lambda(w)$ should be interpreted to mean $\left(\frac{\partial F_\lambda}{\partial W_1},...,\frac{\partial F_\lambda}{\partial W_L}\right)$

and $\bar{z}^j$ maximizes $\bar{p} \cdot z^j$ over $f_j\left(z^j\right) = 0$. Assuming an interior solution, there exists $\alpha > 0$ such that for every $i \in I$

$$\alpha \bar{p} = \lambda_i \bigtriangledown u_i\left(\bar{x}^i\right) = \bigtriangledown F_\lambda(w)$$

By (32) $\bar{p} \cdot w^i = c\bar{p} \cdot \bar{x}^i$, $i \in I$. Since $\bar{p} \cdot \bar{z}^j = 0$ and $\sum_{i \in I} \bar{x}^i = \sum_{i=1}^I w^i + \sum_{j \in J} \bar{z}^j$, we have that $c = 1$. Hence, for all $i \in I$

$$\bar{p} \cdot \bar{x}^i = \bar{p} \cdot w^i$$

and $\bar{x}^i$ maximizes $u_i\left(x^i\right)$ over $\left\{x^i | \bar{p} \cdot x^i \leq \bar{p} \cdot w^i\right\}$, implying that λ is a Negishi weight.■

Proof of Proposition 2: Let $Z_j = \left\{z^j | f_j\left(z^j\right) \leq 0\right\}$, $B(G) = \left\{\tau = \left(\tau^i\right)_{i \in I} | \tau^i \in \mathbb{R}_+^L,\ \sum_{i \in I} \tau^i = G\right\}$ and $w = \left(w^1, ..., w^I\right) \in \mathbb{R}_+^{LI}$. Let the set of feasible allocations with respect to $W - G$ be

$$\bar{L}(G) = \left\{(x, z) \mid \sum_{i \in I} x^i = W - G + \sum_{j \in J} z^j,\ f_j\left(z^j\right) \leq 0,\ j = 1, ..., J\right\}$$

and recall that the set of allocations implementable as a competitive equilibrium with lump sum taxes is

$$\widehat{L}(G) = \left\{(\widehat{x}, \widehat{z}) \,|\text{there exist } \widehat{\tau} \in \mathbb{R}^{IL} \text{ and } \widehat{p} \in \mathbb{R}_+^L \text{ s.t. } (\widehat{x}, \widehat{z}, \widehat{\tau}, \widehat{p}) \text{ satisfies (9)-(13)}\right\}$$

For any $\lambda \in \Delta_I$, let

$$\bar{F}_\lambda(w) = \max_{(x,z)} \sum_{i \in I} \lambda_i u_i\left(x^i\right) \text{ s.t. } (x, z) \in \bar{L}(G)$$

$$\widehat{F}_\lambda(w) = \max_{(x,z)} \sum_{i \in I} \lambda_i u_i\left(x^i\right) \text{ s.t. } (x, z) \in \widehat{L}(G)$$

Lemma 6. (x, z) is a maximizer of $\widehat{F}_\lambda(w)$ iff it is a maximizer of $\bar{F}_\lambda(w)$ and hence $\bar{F}_\lambda(w) = \widehat{F}_\lambda(w)$.

Proof: Since $\widehat{L}(G) \subseteq \bar{L}(G)$ it follows that $\widehat{F}_\lambda(w) \leq \bar{F}_\lambda(w)$. To prove the converse inequality, let $(\bar{x}, \bar{z})$ be a maximizer of $\bar{F}_\lambda(w)$. Since $(\bar{x}, \bar{z})$ is a Pareto optimal allocation with respect to the initial resources $\sum_{i \in I} w^i - G$, by the second welfare theorem, there exists a supporting price $\bar{p} \in \mathbb{R}_+^L$ s.t. for every $i \in I$, $\bar{x}^i$ maximizes $u_i\left(x^i\right)$ over $\left\{x^i \in \mathbb{R}_+^L | \bar{p}x^i \leq \bar{p}\bar{x}^i\right\}$ and for every $j \in J$, $\bar{z}^j$ maximizes $\bar{p}z^j$ over z^j s.t. $f_j\left(z^j\right) \leq 0$. Now let

$$\bar{\tau}^i = w^i - \bar{x}^i + \sum_{j \in J} \theta_j^i \bar{z}^j,\ i \in I$$

where $\theta_j \in \mathbb{R}_+^I$ are the initial shares of agent i in firm j so that $\sum_{i=1}^I \theta_j^i = 1$ for all j. Then, by (9),

$$\sum_{i \in I} \bar{\tau}^i = \sum_{i \in I} w^i - \sum_{i \in I} \bar{x}^i + \sum_{i \in I} \sum_{j \in J} \theta_j^i \bar{z}^j = G$$

Also, since $\bar{p}\bar{z}^j = 0$ for all $j \in J$ by the homotheticity of technologies, we know $\bar{p} \sum_{j \in J} \bar{z}^j = 0$ and therefore

$$\bar{p}\bar{\tau}^i = \bar{p}\left(w^i - \bar{x}^i\right)$$

Hence, $\bar{x}^i$ maximizes $u_i\left(x^i\right)$ over $\left\{x^i \in \mathbb{R}_+^I | \bar{p}x^i \leq \bar{p}\left(w^i - \bar{\tau}^i\right)\right\}$. Thus, $(\bar{x}, \bar{z}, \bar{\tau}, \bar{p})$ satisfies the five conditions (9)-(13) and $(\bar{x}, \bar{z}) \in \bar{L}(G)$. This implies that

$$\bar{F}_\lambda(w) = \sum_{i \in I} \lambda_i u_i\left(\bar{x}^i\right) \leq \widehat{F}_\lambda(w)$$

Since $\bar{F}_\lambda(w) \geq \widehat{F}_\lambda(w)$ also holds, we have that $\bar{F}_\lambda(w) = \widehat{F}_\lambda(w)$ and $(\bar{x}, \bar{z})$ is also a maximizer of $\widehat{F}_\lambda(w)$.■

Suppose next that λ is a VP weight. Then,

$$\frac{\bigtriangledown \widehat{F}_\lambda(w)\tilde{w}^i}{\bigtriangledown \widehat{F}_\lambda(w)\tilde{w}^h} = \frac{\lambda_i \bigtriangledown u_i\left(\widehat{x}_\lambda^i\right)\widehat{x}_\lambda^i}{\lambda_h \bigtriangledown u_h\left(\widehat{x}_\lambda^h\right)\widehat{x}_\lambda^h} \text{ for all } i, h \in I \tag{33}$$

where $(\widehat{x}_\lambda, \widehat{z}_\lambda)$ is a maximizer of $\widehat{F}_\lambda(w)$ over $\widehat{L}(G)$. By Lemma 6, $(\widehat{x}_\lambda, \widehat{z}_\lambda)$ is a maximizer of $\bar{F}_\lambda(w)$ and thus it is Pareto optimal with respect to $\sum_{i \in I} w^i - G$. By the second welfare theorem, there exists a supporting price $\widehat{p} \in \mathbb{R}_+^L$ s.t. $(\widehat{x}, \widehat{z}, \widehat{p})$ is a competitive equilibrium with transfers. Namely, $\widehat{z}^j = \arg\max \widehat{p}z^j$ over $z^j \in Z_j$ for all $j \in J$ and $\widehat{x}^i = \arg\max u_i\left(x^i\right)$ over $\left\{x^i \in \mathbb{R}_+^m | \widehat{p}x^i \leq \widehat{p}\widehat{x}^i\right\}$ for all $i \in I$. Hence, for some c,

$$c\widehat{p} = \lambda_i \bigtriangledown u_i\left(\widehat{x}^i\right) \text{ for all } i \in I \tag{34}$$

Following arguments similar to Proposition 1, prices equal the marginal value of the goods (see once again Mas-Colell et al (1995))

$$\frac{\partial \bar{F}_\lambda}{\partial w_l^i}(w) = c\widehat{p}_l \tag{35}$$

By (33), (34) and (35),

$$\frac{\widehat{p}\widehat{x}_\lambda^i}{\widehat{p}w^i} = \alpha \tag{36}$$

where α is a constant (that does not depend on i). By (36),

$$\widehat{p}\sum_{i \in I} \widehat{x}_\lambda^i = \alpha \widehat{p} \sum_{i \in I} w^i = \alpha \widehat{p} W$$

Since $\sum_{i \in I} \widehat{x}_\lambda^i = W - G + \sum_{j \in J} \widehat{z}^j$,

$$\alpha = 1 - \frac{\widehat{p}G}{\widehat{p}W} \tag{37}$$

By (36) and (37),

$$\widehat{p}\widehat{x}_\lambda^i = \widehat{p}w^i - \widehat{p}G\frac{\widehat{p}w^i}{\widehat{p}W} \tag{38}$$

Since

$$\widehat{p}\widehat{x}^i = \widehat{p}w^i - \widehat{p}\widehat{\tau}^i \tag{39}$$

we have

$$\widehat{p}\widehat{\tau}^i = \frac{\widehat{p}w^i}{\widehat{p}W}\widehat{p}G ■$$

Proof of Proposition 3.

As shown in the computational Appendix, VP condition for the model with proportional capital and labor taxes can be writen as:

$$\frac{\lambda_1 U_{c_{11}} p_1 \left[R_1 k_{11} + w_1 T_{11} + \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2} T_{12}\right]}{\lambda_2 U_{c_{21}} p_2 \left[R_1 k_{21} + w_1 T_{21} + \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2} T_{22}\right]} = \frac{\lambda_1 U_{c_{11}} p_1 \left[R_1 k_{11} + w_1 T_{11} + \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2} T_{12}\right] + \Phi_1}{\lambda_2 U_{c_{21}} p_2 \left[R_1 k_{21} + w_1 T_{21} + \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2} T_{22}\right] + \Phi_2} \tag{40}$$

where Φ_i represent the price and tax effects of a change in the endowments of agent i, namely:

$$\begin{aligned}
\Phi_i \quad = \quad & \left(\frac{\partial R_1}{\partial k_{i1}} k_{i1} + \frac{\partial R_1}{\partial T_{i1}} T_{i1} + \frac{\partial R_1}{\partial T_{i2}} T_{i2}\right) \sum_j \lambda_j p_j U_{c_{j1}} k_{j1} \\
& + \left(\frac{\partial w_1}{\partial k_{i1}} k_{i1} + \frac{\partial w_1}{\partial T_{i1}} T_{i1} + \frac{\partial w_1}{\partial T_{i2}} T_{i2}\right) \sum_j \lambda_j p_j U_{c_{j1}} n_{j1} \\
& + \left[k_{i1} \frac{\partial \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2}}{\partial k_{i1}} + T_{i1} \frac{\partial \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2}}{\partial T_{i1}} + T_{i2} \frac{\partial \frac{(1-\tau^n)w_2}{(1-\tau^k)R_2}}{\partial T_{i2}}\right] \sum_j \lambda_j p_j U_{c_{j1}} n_{j2} \\
& - \left[k_{i1} \frac{\partial \frac{1}{(1-\tau^k)R_2}}{\partial k_{i1}} + T_{i1} \frac{\partial \frac{1}{(1-\tau^k)R_2}}{\partial T_{i1}} + T_{i2} \frac{\partial \frac{1}{(1-\tau^k)R_2}}{\partial T_{i2}}\right] \sum_j \lambda_j p_j U_{c_{j1}} c_{j2}
\end{aligned}$$

As shown in the appenfix, the price and tax effects can be simplified to:

$$\begin{aligned}
\Phi_i \quad = \quad & \Omega \theta Y_2 \left[\theta \left(\tau_k - \tau_n\right) + \tau_n\right] \frac{p_i \left[R_1 k_{i1} + w_1 T_{i1}\right]}{R_1 K_1 + w_1 N_1} + \\
& (\theta - 1)\, \theta \left[\frac{\pi_i k_{i1}}{K_1} - \frac{\pi_i T_{i1}}{N_1}\right] \sum_j \lambda_j u_{cj1} \left[\left[\frac{\pi_j k_{j1}}{K_1} - \frac{\pi_j n_{j1}}{N_1}\right] Y_1 + \left[\frac{\pi_j k_{j2}}{K_2} - \frac{\pi_j n_{j2}}{N_2} \frac{(1-\tau^n)}{(1-\tau^k)}\right] K_2\right]
\end{aligned} \tag{41}$$

Assume that $G_2 = 0$, in which case $\tau_k = \tau_n = 0$ is a feasible solution to the government problem. We now show that $\Phi_i = 0$ for $i = 1, 2$ at zero taxes. First, it is clear that the last term of (**??**), corresponding to the effects on social welfare of a change in the endowments through changes in taxes, is equal to zero if $\tau_k = \tau_n = 0$. Second, since the allocation at zero taxes are Pareto optimal, we have that $\lambda_1 U_{c_{i1}} = \lambda_2 U_{c_{21}} = g_1$ at the Negishi weights, which are also the VP weights by Proposition 2. Consequently, the first term of (**??**), corresponding to the effects on social welfare of a change in the endowents through a change in prices becomes:

$$(\theta - 1)\, \theta \left[\frac{\pi_i k_{i1}}{K_1} - \frac{\pi_i T_{i1}}{N_1}\right] g_1 \sum_j \left[\left[\frac{\pi_j k_{j1}}{K_1} - \frac{\pi_j n_{j1}}{N_1}\right] Y_1 + \left[\frac{\pi_j k_{j2}}{K_2} - \frac{\pi_j n_{j2}}{N_2}\right] K_2\right] = 0$$

where the last equality follows directly from the definitions of K_1, K_2, N_1 and N_2. This implies that $\Phi_i = 0$ for $i = 1, 2$ at zero taxes. Third, if the price and tax effects are equal to zero, the left and right hand side of the VP condition equalize and zero taxes are zero taxes are therefore VP.■

References

[1] Abraham, A., Carceles-Poveda, E., 2010. Endogenous Trading Constraints with Incomplete Asset Markets. Journal of Economic Theory, 145, pp. 974-1004.

[2] Aiyagari, S.R., 1994. Uninsurable Idiosyncratic Risk and Aggregate Saving. Quarterly Journal of Economics, 109(3), pp. 659-684.

[3] Aiyagari, S.R., 1995. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints and Constant Discounting. Journal of Political Economy 103(6), pp. 1158-75.

[4] Anagnostopoulos, A., Carceles-Poveda, E., Lin, D., 2012. Dividend and Capital Gains Taxation under Incomplete Markets. Journal of Monetary Economics, vol. 59 (7), pp. 599-611.

[5] Aumann, R., 1975. Values of markets with a continuum of traders, Econometrica vol. 43 (4), pp. 611-646.

[6] Aumann, R., Shapley, L.S., 1974. Values of Non-Atomic Games, Princeton University Press, Princeton, New Jersey.

[7] Benabou, R., 2002. Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?, Econometrica, 7 (2), pp. 481-517.

[8] Chang B. H, Chang, Y., Kim S. B. Pareto Weights in Practice: A Quantitative Analysis Across 32 OECD Countrie. Conditionally accepted for Review of Economic Dynamics.

[9] Davila, J., Hong, J.H., Krusell, P., Ríos-Rull, J.V., 2012. Constrained Efficiency in the Neoclassical Growth Model with Uninsurable Idiosyncratic Shocks. Econometrica vol. 80 (6), pp. 2431-2467.

[10] Domeij, D., Heathcote, J., 2004. On the Distributional Effects of Reducing Capital Taxes. International Economic Review 45(2), pp. 523-554.

[11] Fleurbaey M., Maniquet, F., 2014. Optimal Taxation Theory and Principles of Fairness. Core Discussion Paper.

[12] Harsanyi, J. C, 1953. Cardinal Utility in Welfare Economics and in the Theory of Risk-Taking, Journal of Political Economy, vol. 61 (5), pp. 434-435.

[13] Harsanyi, J. C, 1955. Cardinal Welfare, Individualistic Ethics, and the Interpersonal Comparison of Utility, Journal of Political Economy, vol. 63 (4), pp. 309-321.

[14] Heathcote, J., Storesletten, K, Violante, G. L., 2014. Optimal Tax Progressivity: An Analytical Framework, working paper.

[15] Kehoe P., and Perri F., 2002. International Business Cycles With Endogenous Incomplete Markets, Econometrica, 70, vol. 3, pp. 907-928.

[16] Mankiw, G. N., 2010. Spreading the Wealth Around: Reflection Inspired by Joe the Plumber, Eastern Economic Journal, 36, pp. 285-298.

[17] Mas -Colell, A, Whinston, M. D., Green, J. R., 1995. Microeconomic Theory. Oxford University Press, New York, New York.

[18] Mirman,L., Tauman, Y., 1982. Demand compatible equitable cost sharing prices, Mathematics of Operations Research, vol. 7 (1), pp. 40-56.

[19] Negishi, T., 1960. Welfare Economics and Existence of an Equilibrium for a Competitive Economy, Metroeconomica, vol. 12 (2-3), pp. 92-97.

[20] Noszick, R., 1974. Anarchy, State and Utopia, Basic Books, New York.

[21] Rawls, J., 1971. A Theory of Justice, Belknap Press of Harvard University Press, Cambridge, MA.

[22] Saez, E., Stantcheva, S., 2015. Generalized Social Marginal Welfare Weights for Optimal Tax Theory, working paper.

[23] Samet, D., Tauman, Y., 1982. The determination of marginal cost prices under a set of axioms, Econometrica, vol. 50 (4), pp. 895-909.

[24] Shapley, L.S., 1953. "A Value for n-person Games". In Contributions to the Theory of Games, volume II, ed. by H.W. Kuhn and A.W. Tucker, Princeton, NJ: Princeton University Press, pp. 307–317.

[25] Shapley, L.S., 1969. Utility Comparison and the Theory of Games. In "La Decision Aggregation Ordres de Preference", Paris, Editions du Centre National de la Recherche Scientifique, 251-263.

[26] Weinstein, J., 2011. Fairness and Tax Policy: A Response to Mankiw's Proposed "Just Deserts", Eastern Economic Journal, 37, 313-320.

[27] Weinzierl, M., 2014. The Promise of Positive Optimal Taxation: Normative Diversity and a Role for Equal Sacrifice, Working Paper, Harvard Business School.

[28] Young, H. P., 1988. Distributive Justice in Taxation, Journal of Economic Theory, vol. 44 (2), pp. 321-335.

[29] Young, H. P., 1990. Progressive Taxation and Equal Sacrifice, American Economic Review, vol 80 (1), pp. 253-266.