

---

---

# Value Preserving Welfare Weights for Social Optimization Problems

Working Paper 2018-08

---

ALEXIS ANAGNOSTOPOULOS, EVA CARCELES-POVEDA,  
YAIR TAUMAN

July 5, 2018



Stony Brook  
University

---

---

# Value Preserving Welfare Weights for Social Optimization Problems\*

ALEXIS ANAGNOSTOPOULOS, EVA CARCELES-POVEDA, YAIR TAUMAN

July 5, 2018

ABSTRACT. We propose a method to endogenize social welfare weights. We provide three axioms which uniquely determine the contribution of a bundle of goods to a welfare function. We define weights to be value preserving (VP) if the contribution of an individual to social welfare is proportional to the contribution of the final allocation to the individual's welfare. We show that VP weights coincide with Negishi weights in Arrow-Debreu economies. With wealth heterogeneity, VP weights prescribe higher than capital labor taxes even for extreme inequality levels. With ability differences, VP weights redistribute towards high ability, high effort and low wealth types.

*Keywords: Social Welfare Function, Heterogeneous Households, Redistribution, Optimal Taxation, Endogenous Welfare Weights*

*JEL Classification: D63, E62, H21*

---

\*We would like to thank Marina Azzimonti, Dirk Krueger, Francesc 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.

The macroeconomic literature has often sidestepped the issue of how to choose welfare weights by assuming equal weights (EW) for all agents.<sup>1</sup> This choice implicitly introduces a strong assumption about social preferences for redistribution which stands in contrast with both survey evidence on these preferences as well as with observed levels of redistribution through tax systems.<sup>2</sup> It also suffers from an even more fundamental problem, which is that the resulting allocations are not invariant to affine transformations of agents' utilities. We suggest an alternative to the EW approach that does not suffer from this problem and that has the potential to improve the model's connection with data. Specifically, we define value preserving (VP) welfare weights, which are endogenously computed to 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.

Our first contribution lies in axiomatically characterizing a mechanism that can be used 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 then use this mechanism to define 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 contribution) and the contribution of the individual's final allocation of goods to his individual welfare (private contribution). The VP weights are the ones that equate the ratio of social to private contributions across agents.

Our second contribution is to provide a characterization of the VP weights and allocations. When the fundamental welfare theorems hold, Negishi (1960) provides a method which

---

<sup>1</sup>Several exceptions are discussed below.

<sup>2</sup>Weinzierl (2014) provides survey evidence indicating that only a minority of individuals consider the EW approach and its implications for taxation as fair. Chang et al (2018) uncover the Pareto weights that justify as optimal current income tax systems and redistribution levels in different OECD countries and show that these deviate considerably from EW.

endogenously determines welfare weights by requiring that the resulting allocations coincide with the Walrasian 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, 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. Whereas the Negishi approach is not available in environments where the fundamental welfare theorems do not hold, VP weights and allocations can still be computed for economies with frictions and they can therefore be seen as a natural extension of the Negishi method to these types of economies.

Our proposed welfare weights could also be thought of as a formalization of Mankiw's (2010) Just Deserts Theory, but our paper does not provide new arguments on the discussion about fairness.<sup>3</sup> 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 and one of our aims is to analyze the implications of such assumptions.<sup>4</sup> We illustrate this with the characterization of optimal tax policy in a benchmark first best scenario where the government uses lump sum taxes. In this case, we show that 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, tax policy can also be used to correct market inefficiencies and this alternative tax motive will have to be weighed against the distributional motive implicit in the assumed social welfare function. An additional complication arises when non-distortionary tax instruments are not available, in which case the government also has to take tax distortion minimization into consideration. We illustrate these trade-offs with two examples that have been commonly used in the literature on optimal taxation.

First, we study a two period example with wealth inequality, a flat distortionary tax on capital income and a flat tax on labor income that is not distortionary. The efficient tax scheme is therefore to tax labor income only in order to finance government spending. As inequality increases, however, EW taxes become higher for capital even with moderate levels of inequality purely for redistribution reasons and despite the distortions introduced, whereas VP capital taxes remain lower than labor taxes even with extreme levels of inequality. In sum, the different 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. One implication of using VP weights that is consistent with the EW weights is that capital taxes do increase with inequality, but the reason is not to effect

---

<sup>3</sup>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 (2014) provide authoritative reviews of both classic and recent contributions to that literature.

<sup>4</sup>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.

redistribution but rather to avoid the negative redistribution inherent in the efficient tax scheme.

In a second example, we study an economy closer to the standard optimal taxation model in the public finance literature. The environment is static, but includes heterogeneity in ability in addition to wealth heterogeneity. When ability is included in the endowments that individuals bring to society, we find that, in the absence of redistribution, the ratio of social to private contribution is higher for high ability, high effort and low wealth individuals. In other words, a VP government tends to redistribute towards low wealth-to-ability types and also tends to reward effort. Interestingly, a VP government distinguishes between ability and wealth as the underlying reasons for overall income inequality, in contrast with the EW approach, which redistributes to the low overall income type, regardless of whether income is low due to low ability, effort or wealth. Moreover, when we calibrate this simple model to reflect correlation between ability and wealth and a larger amount of inequality in wealth than in wages, as in the data, the VP approach delivers welfare weights and a corresponding optimal tax policy that are qualitatively consistent with the findings in Chang et al (2018) and in Weinzierl (2014). Specifically, Pareto weights are larger on the rich and the optimal tax scheme generates some redistribution, while the level of redistribution is significantly lower than what a utilitarian welfare function would imply.

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<sup>5</sup>, but it is also related to the recent macroeconomic literature with interest in social welfare functions and their effect on optimal taxation. 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. Heathcote, Storesletten and Violante (2017) consider a family of social welfare functions with a parameter governing the degree of inequality aversion. Heathcote and Tsujiyama (2017) parameterize instead a welfare weight function to analyze the effects of changing weights on optimal policy. Chang, Chang and Kim (2018) construct Pareto weights that could justify the level of progressivity observed in several OECD countries. With respect to these two strands of literature, we obtain endogenous welfare weights using a new principle, the VP principle, and we 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 with EW<sup>6</sup>

Alternatives to the EW approach have also been considered recently in the public finance literature. Prominent examples include Weinzierl (2014), Saez and Stantcheva (2015) and Fleurbaey and Maniquet (2014). Weinzierl (2014) shows that a social welfare function which

---

<sup>5</sup>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.

<sup>6</sup>Endogenous and time varying Pareto weights are also present 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.

includes a utilitarian component but penalizes deviations from equal sacrifice can explain a number of features of US tax policy. The standard implementation of the equal sacrifice principle is not invariant to affine transformations of utility. We show how our contribution mechanism can be used to correct this and also compare the predictions of equal sacrifice to those under the value preserving principle. Saez and Stantcheva (2015) propose a general non-welfarist approach to optimal tax theory that can accommodate different justice principles by applying social marginal welfare weights directly to the earnings equivalent welfare change. In contrast, Fleurbaey and Maniquet (2014) show that one can incorporate different fairness principles into the standard social welfare framework to investigate 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 and focus on a specific principle that takes a stand on social preferences about redistribution. However, we argue that our methodology of endogenously solving for welfare weights can be used to determine social welfare weights that satisfy other social justice principles and the proposed contribution mechanism is a helpful tool in ensuring invariance to affine transformations of utilities.

Our VP principle is closely related to the literature on values of cooperative games, such as Shapley (1969) and Aumann (1975), but our approach is different. 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 accommodate social welfare functions instead of cost functions.<sup>7</sup> In this sense, the axiomatic approach we provide constitutes a theoretical contribution in itself.

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 applications of the theory to two optimal taxation examples and Section 7 summarizes and concludes.

## 2. THE ECONOMY

Consider an economy with  $L$  goods that are indexed by  $l \in \{1, \dots, L\}$ ,  $I$  agents indexed by  $i \in \{1, 2, \dots, I\}$  and  $J$  firms that are indexed by  $j \in \{1, 2, \dots, J\}$ .<sup>8</sup> 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 positive orthant of  $\mathbb{R}^L$ . We let  $w^i = (w_1^i, \dots, w_L^i) \in \mathbb{R}_+^L$  be the vector of initial good allocations (or endowments) of agent  $i \in I$  and  $w = (w^1, \dots, w^I) \in \mathbb{R}_+^{LI}$  be the vector of initial good allocations. Similarly,  $x^i = (x_1^i, \dots, x_L^i) \in \mathbb{R}_+^L$  represents the vector of final allocations of agent  $i$  and  $x = (x^1, \dots, x^I) \in \mathbb{R}_+^{LI}$  is the vector of final allocations.

---

<sup>7</sup>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.

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

A firm is identified with a production plan  $z^j \in \mathbb{R}^L$ . Let  $z = (z^1, \dots, z^J) \in \mathbb{R}^{LJ}$  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  $u_i(x^i)$  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  $F_\lambda(w) : \mathbb{R}_+^{LI} \rightarrow \mathbb{R}$  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,

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

where  $g_s : \mathbb{R}_+^{LI} \times \mathbb{R}_+^{LJ} \times \mathbb{R}_+^{LI} \rightarrow \mathbb{R}$ . 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  $\lambda$  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  $\tilde{w}^i = (0, \dots, 0, w^i, 0, \dots, 0) \in \mathbb{R}_+^{LI}$  and notice that  $\tilde{w}^i \leq w$  and  $\sum_i \tilde{w}^i = w$ . Let  $\lambda = (\lambda_i)_{i \in I}$  and let  $C(F_\lambda, \tilde{w}^i, w)$  be the contribution of the bundle  $\tilde{w}^i$  to the social welfare function  $F_\lambda$  generated by the weights  $\lambda$  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  $C(\lambda_i u_i, x^i, x^i)$ . In this case, we can simplify notation and write  $C(\lambda_i u_i, x^i)$ .<sup>9</sup> Given the 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.

<sup>9</sup>We have assumed that the individual welfare function depends only on  $x^i$  and not any  $x^h$ ,  $h \neq i$ . It 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  $C(\lambda_i u_i, \tilde{x}^i, x)$ .

**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  $x_\lambda$  are value preserving iff

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

The contribution of the initial allocation  $\tilde{w}^i$  of agent  $i \in I$  to social welfare is  $C(F_\lambda, \tilde{w}^i, w)$ . Similarly, the contribution of the final bundle  $x_\lambda^i$  to the private welfare function  $\lambda_i u_i$  is  $C(\lambda_i u_i, x_\lambda^i)$ . Thus, under VP weights, the ratio of the social to the private contribution is equalized across all individuals. To put it simply, the value an individual receives 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. In this sense, our approach is related 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. Whether this is consistent with notions of justice has long been a subject of debate.<sup>10</sup> 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'.<sup>11</sup> Our paper does not offer new arguments to this debate. 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  $\lambda$ . 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 Weizierl (2014) and, earlier by Young (1988, 1990) in a fundamental way: private

<sup>10</sup>See the introduction and, specifically, footnote 3.

<sup>11</sup>See Harsanyi (1953, 1955).

and social values are defined for *the same* economy. In contrast, the equal sacrifice principle 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. CONTRIBUTION MECHANISM

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  $w = (w_1, \dots, w_m) \in \mathbb{R}_+^m$  denote a bundle of these goods.<sup>12</sup> Let  $\mathcal{F}^m$  be the set of all functions  $F : \mathbb{R}_+^m \rightarrow \mathbb{R} \cup \{-\infty, +\infty\}$  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  $j^{\text{th}}$  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

$$F(x) = qG(r_1x_1, \dots, r_mx_m) + c$$

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  $w \in \mathbb{R}_+^m \setminus \{0\}$

$$\widehat{C}_j(F, w) = qr_j\widehat{C}_j(G, r_1w_1, \dots, r_mw_m).$$

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

---

<sup>12</sup>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  $m = LI$ .

value of a welfare function  $G$  is measured in dollars. Let  $F$  represent the same welfare as  $G$  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  $\{1, \dots, m\}$  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  $x \in \mathbb{R}_+^m$

$$F(x) = F^1(x^1) + \dots + F^k(x^k)$$

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  $X \subseteq \mathbb{R}^m$  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  $\tilde{w}$ ,  $w_j \leq \tilde{w}_j \leq w_j + \eta$ ,  $1 \leq j \leq m$ , and for all  $G \in \mathcal{F}^m$

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

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

**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  $w \in \mathbb{R}_+^m \setminus \{0\}$

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

<sup>13</sup>Since some of the coordinates of  $w$  may be zero, we only consider the right neighbourhood of  $w$ .

The (positive) constant  $\alpha$  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

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

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:

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

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

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

Note that the contribution mechanism  $C(F, w', w)$  represents the overall contribution of the bundle  $\tilde{w} \leq w$  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 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  $f_j(z^j) \leq 0$ , where  $f_j(z^j)$  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  $\nabla f_j(z^j) \gg 0$  for all  $j \in J$ . Each firm  $j$  maximizes profits and solves:

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

where  $p \in \mathbb{R}_+^L$  is the vector of commodity prices. Since  $f_j$  is continuous,  $f_j(z^j) = 0$  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  $u_i : \mathbb{R}_+^L \rightarrow \mathbb{R}$ . Household  $i$  takes  $p, w^i$  as well as  $\theta_j^i, z^j$  for all  $j \in J$  as given and solves:

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

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

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

where  $\mu_i$  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:

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

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

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

**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  $\Delta^I = \left\{ \lambda \in \mathbb{R}_+^I \mid \sum_{i=1}^I \lambda_i = 1 \right\}$ . Then,  $\lambda \in \Delta^I$  is a value preserving weight iff it is a Negishi weight for the initial endowments  $w = (w^1, \dots, w^I) \in \mathbb{R}_+^{LI} \setminus \{0\}$  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,

<sup>14</sup>Note that both the value  $F_\lambda$  and the allocations  $x_\lambda, z_\lambda$  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  $(z^j(w))_{j=1}^J$ , multiplied by the shares of individual  $i$  in each firm  $(\theta_j^i)_{j \in J}$  as part of that individual's initial endowments. To be precise, define the modified initial endowment of individual  $i$  as

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

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

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

With this modification, replacing  $\tilde{w}^i = (0, \dots, 0, w^i, 0, \dots, 0)$  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 macro-economists to take the equal weights approach, partly because the Negishi approach is not applicable.<sup>15</sup> One could compute the Negishi weights for the first-best version of the model and apply the same weights to a model which incorporates externalities, incomplete markets and distortionary taxes. However, this would be equivalent to the VP approach only 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 uses our contribution mechanism to determine the value of each good in the actual (distorted) economy as the marginal increase in social welfare arising from a unit increase in the endowment of the good in question. Typically, this value will not coincide with the frictionless price.

## 6. APPLICATIONS TO OPTIMAL TAXATION

In this section, we discuss applications to economies where the government chooses optimal taxes to maximize a social welfare function. Although the models are stylized, they are rich

---

<sup>15</sup>For important exceptions see the discussion of the literature in the introduction.

enough to allow the analysis of optimal tax policy as an instrument for revenue-raising, for redistribution and for insurance provision while under the restrictions imposed by the loss of efficiency from distortionary taxation. First, we characterize the optimal VP lump sum taxes in a first best environment. We subsequently discuss two settings that depart from the first best. In the first one, individuals differ in their initial level of wealth and are potentially subject to idiosyncratic labor income shocks.<sup>16</sup> In the second, individuals differ in both wealth and skill endowments, none of which are tradable.

**6.1. Lump Sum Taxation.** Consider an Arrow Debreu economy with lump sum taxation. Let  $G \in \mathbb{R}_+^L$  be the (exogenous) vector of goods which the government needs to consume. Let  $\hat{\tau} \in \mathbb{R}^{LI}$  denote a tax policy, where  $\hat{\tau}_l^i$  is the lump sum tax (or transfer) of good  $l$  for agent  $i$ . We first define the competitive equilibrium given a feasible tax policy  $\hat{\tau} \in \mathbb{R}^{LI}$  such that  $\sum_{i \in I} \hat{\tau}^i = G$ .

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

$$\sum_{i \in I} \hat{x}^i = \sum_{i \in I} w^i - G + \sum_{j=1}^J \hat{z}^j \quad (7)$$

$$f_j(\hat{z}^j) \leq 0, \text{ for all } j \in J \quad (8)$$

$$\sum_{i \in I} \hat{\tau}^i = G \quad (9)$$

$$\hat{z}^j \text{ maximizes } \hat{p}z^j \text{ over } \{z^j | f_j(z^j) \leq 0\}, j \in J \quad (10)$$

$$\hat{x}^i \text{ maximizes } u_i(x^i) \text{ over } \left\{ x^i \in \mathbb{R}_+^L | \hat{p}x^i \leq \hat{p} \left( w^i - \hat{\tau}^i + \sum_{j \in J} \theta_i^j \hat{z}^j \right) \right\}, i \in I \quad (11)$$

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

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  $\hat{L}(G)$  on allocations:

$$\hat{L}(G) = \{(\hat{x}, \hat{z}) | \text{there exist } \hat{\tau} \in \mathbb{R}^{LI} \text{ and } \hat{p} \in \mathbb{R}_+^L \text{ s.t. } (\hat{x}, \hat{z}, \hat{\tau}, \hat{p}) \text{ satisfies (7)-(11)}\}$$

The social welfare function of the Ramsey taxation problem can then be defined as<sup>17</sup>

$$\hat{F}_\lambda(w) = \max_{x, z} \sum_{i \in I} \lambda_i u_i(x^i) \text{ s.t. } (x, z) \in \hat{L}(G) \quad (12)$$

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

$$\frac{\nabla \hat{F}_\lambda(w) \tilde{w}^i}{\nabla \hat{F}_\lambda(w) \tilde{w}^h} = \frac{\lambda_i \nabla u_i(\hat{x}_\lambda^i) \hat{x}_\lambda^i}{\lambda_h \nabla u_h(\hat{x}_\lambda^h) \hat{x}_\lambda^h}, i, h \in I$$

<sup>16</sup>This environment has been used extensively in the macroeconomic literature following Bewley (1977, 1983) and Aiyagari (1994).

<sup>17</sup>This formulation follows the primal approach to the Ramsey taxation problem, where the government chooses allocations subject to implementability constraints.

where  $\hat{x}_\lambda$  is a maximizer of (12).

As in Proposition 1, we maintain the assumptions of smooth and concave utilities, convex technologies as well as technologies that are homothetic.<sup>18</sup>

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

The proof of Proposition 2 is provided 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  $\hat{p}w^i$ . The implication is that

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

which means that the before-tax and after-tax distributions of wealth are the same, i.e. the VP government does not redistribute. In the special case where the government does not need to raise revenue ( $\hat{p}G = 0$ ),  $\hat{p}\hat{\tau}^i = 0$  for all  $i$  is a VP tax.<sup>19</sup>

**6.2. 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  $\omega_t$  and uses them as inputs to produce a consumption good  $Y_t$  using a constant returns to scale technology

$$Y_t = K_t^\theta N_t^{1-\theta}, \quad t = 1, 2 \quad (13)$$

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, 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. 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  $G_2$ .<sup>20</sup> The government uses flat

<sup>18</sup>We can dispense with homotheticity by modifying the initial bundle, see the discussion after Proposition 1.

<sup>19</sup>Note that the VP tax  $\hat{\tau}$  is not uniquely determined in this case unless the tax is only on one good. Every  $\tau$  such that  $\hat{p}\hat{\tau}^i = \hat{p}\tau^i$  for all  $i \in I$  is a VP tax as well. Section 1 in the online appendix accompanying the paper provides an example with uniquely determined lump sum taxes in which the VP weights can be computed analytically.

<sup>20</sup>The analysis of the case with taxes in both periods can be provided by the authors upon request.

rate taxes  $\tau_k$  and  $\tau_n$  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:

$$\begin{aligned} & \max_{\{c_{i1}, c_{i2}, k_{i1}, n_{i1}, n_{i2}\}} U(c_{i1}, c_{i2}) \text{ s.t.} \\ c_{i1} + k_{i2} &= R_1 k_{i1} + \omega_1 n_{i1} \\ c_{i2} &= R_2 k_{i2} (1 - \tau_k) + \omega_2 n_{i2} (1 - \tau_n) \\ 0 &\leq n_{it} \leq T_{it}, c_{it} \geq 0, \text{ for } t = 1, 2 \end{aligned} \quad (14)$$

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:

$$\begin{aligned} V_\lambda(w) &= \max_{\tau_k, \tau_n} \sum_{i=1}^2 \lambda_i p_i U(c_{i1}^*(\tau, w), c_{i2}^*(\tau, w)) \\ \text{s.t. } G_2 &= \tau_k R_2 K_2 + \tau_n \omega_2 N_2 \end{aligned} \quad (15)$$

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 conditions of the government problem, the government budget constraint and the VP condition, which in this example is given by:

$$\frac{\frac{\partial V_\lambda(w)}{\partial k_{11}} k_{11} + \frac{\partial V_\lambda(w)}{\partial T_{11}} T_{11} + \frac{\partial V_\lambda(w)}{\partial T_{12}} T_{12}}{\lambda_1 p_1 [U_{c_{11}} c_{11} + \beta U_{c_{12}} c_{12}]} = \frac{\frac{\partial V_\lambda(w)}{\partial k_{21}} k_{21} + \frac{\partial V_\lambda(w)}{\partial T_{21}} T_{21} + \frac{\partial V_\lambda(w)}{\partial T_{22}} T_{22}}{\lambda_2 p_2 [U_{c_{21}} c_{21} + \beta U_{c_{22}} c_{22}]} \quad (16)$$

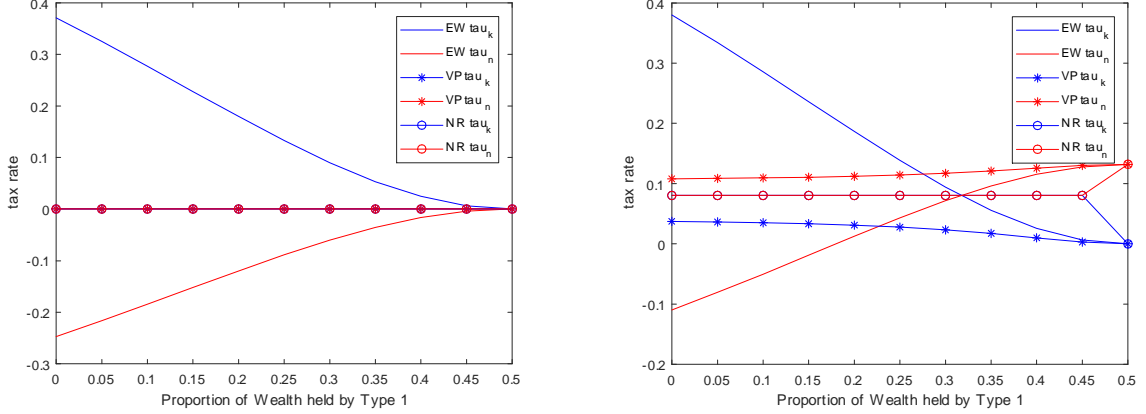
In what follows, we compare the prescribed optimal tax schemes under VP and EW weights. Note we have assumed agents do not derive utility from leisure, so labor supply is inelastic and the labor tax is non-distortionary. Therefore, the efficient tax scheme is to tax only labor income. However, the choice of welfare weights will determine the societal preference for equality and this can conflict with a pure efficiency objective. The EW objective builds in a preference for equalization of marginal utilities which, in this example, would tend to increase capital taxes and reduce labor taxes. The VP objective builds in a preference for no redistribution. In the case  $G_2 = 0$ , this coincides with the efficiency objective since zero taxes are both efficient and non-redistributive. This leads to the following result.

**Result 1.** If  $G_2 = 0$ , then  $\tau_k = \tau_n = 0$  is a VP tax scheme.

The proof is relegated to an online appendix accompanying the paper, but the intuition for this is straightforward. In the absence of taxes, this is an Arrow-Debreu economy and, hence, the no-tax allocations are VP allocations.

Adding the need to raise revenue  $G_2 > 0$  introduces a trade-off for the VP **government**, so we resort to a numerical example to illustrate the results. 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  $i = 1, 2$ . Initial aggregate capital  $K_1$  is normalized to one and we denote by  $\phi$  the fraction of initial capital owned by type 1 agents, i.e.  $\phi \equiv \frac{p_1 k_{11}}{K_1}$ . Type 2 agents own a fraction  $1 - \phi$ . The results for different levels of  $\phi$  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  $\phi$  is decreased towards zero, type 1 agents become progressively more wealth-poor relative to type 2 agents. In the two panels, we also display the EW taxes and the taxes that would achieve no redistribution (NR), in the sense of equalizing the ratio of the present value of tax liabilities to the ratio of the present value of endowments:

$$\frac{\frac{1}{R_2} (\tau_k R_2 k_{12} + \tau_n w_2 n_{12})}{R_1 k_{11} + w_1 n_{11} + \frac{w_2}{R_2} n_{12}} = \frac{\frac{1}{R_2} (\tau_k R_2 k_{22} + \tau_n w_2 n_{22})}{R_1 k_{21} + w_1 n_{21} + \frac{w_2}{R_2} n_{22}} \quad (17)$$

With the assumed utility, the no redistribution tax scheme requires  $\tau_n = \tau_k$ . Thus the VP government also faces a trade-off between avoiding capital taxes for efficiency reasons but needing capital taxes to prevent (negative) redistribution. The resolution of this trade-off is apparent in the right panel of figure 1. In the absence of any heterogeneity ( $\phi = 0.5$ ), distributional concerns are moot and the optimal tax scheme is the efficient one ( $\tau_k = 0$ ,  $\tau_n > 0$ ) regardless of the social welfare weights. As wealth inequality increases ( $\phi < 0.5$ ), both EW and VP taxes depart from efficiency and tax capital. The capital tax increases with inequality for both objectives, but does so more rapidly under the EW objective. For sufficient amounts of inequality, the prescription about capital and labor taxes is eventually reversed under the EW objective, with capital taxes becoming higher than labor taxes. The VP objective maintains the qualitative prescription of higher labor taxes dictated by efficiency even for extreme levels of inequality.<sup>21</sup>

Clearly, these features are specific to the stylized example used and most importantly the available tax instruments and the assumption of inelastic labor supply. The takeaway message is that optimal tax prescriptions based on a utilitarian objective alone can be entirely due to the assumed preference for equality. It is therefore important to consider alternative societal preferences, such as those inherent in the VP principle, in conjunction with utilitarianism to obtain tax prescriptions that are more robust to societal preferences.

In an economy with no frictions (other than the taxes) and only one dimension of inequality, it is not surprising that the VP taxes tend to favor the agents that start with a higher

<sup>21</sup>Introducing idiosyncratic labor income risk in the second period as in Aiyagari (1994, 1995) would strengthen this qualitative prediction since providing insurance in this model would require shifting the burden of taxation to labor income. The case of incomplete markets is discussed in the online appendix.

endowment of wealth. However, it is worth noting that the VP principle still prescribes capital taxes that are increasing in inequality. The following section introduces additional dimensions of heterogeneity and illustrates the ability of the VP approach to distinguish between heterogeneity in ability, in wealth and in preferences.

**6.3. Skill and Preference Heterogeneity.** We now consider a static framework with endogenous labor supply and additional sources of heterogeneity but with no trade and hence no endogenous price determination. The setup is standard in the public finance literature emanating from the seminal work of Mirrlees (1971) and more recently reviewed and updated in Piketty and Saez (2013).

The economy is still populated by two types 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 endowments of both wealth  $k_i$  and ability  $\omega_i$ . We also allow for potential heterogeneity in preferences. An individual's labor income is given by  $\omega_i l_i$ , where  $l_i$  is labor or effort and an individual's overall income is equal to  $\omega_i l_i + k_i$ . Government tax policy consists of a tax rate  $\tau$  that is proportional to income and a lump sum transfer  $R$ . Given the government's policy  $(\tau, R)$ , individuals choose consumption  $c_i$  and leisure  $T_i - l_i$  to maximize:

$$\begin{aligned} \max_{c_i, l_i} U^i(c_i, T - l_i) \quad \text{s.t.} \\ c_i = (1 - \tau)(\omega_i l_i + k_i) + R \end{aligned} \quad (18)$$

Denote the vector of endowments by  $w = (\omega_1, \omega_2, T_1, T_2, k_1, k_2)$  and denote the solution 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  $\lambda$ , the Ramsey planner chooses the pair  $(\tau, R)$  to maximize:

$$\begin{aligned} V_\lambda(w) &= \max_{\tau, R} \sum_i \lambda_i p_i [U^i(c_i^*, T - l_i^*)] \\ G + R &\leq \tau \sum_j p_j (\omega_j l_j^* + k_j) \end{aligned} \quad (19)$$

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  $\tau_\lambda$  and  $R_\lambda$  and the corresponding allocations by  $c_{\lambda, i}$  and  $l_{\lambda, i}$ . Tax policy  $(\tau, R)$  and the social welfare weights are determined by the government budget constraint, the first order conditions of the Ramsey problem and the VP condition, which in this model is given by

$$\frac{\frac{\partial V_\lambda(w)}{\partial \omega_1} \omega_1 + \frac{\partial V_\lambda(w)}{\partial T_1} T + \frac{\partial V_\lambda(w)}{\partial k_1} k_1}{\lambda_1 p_1 [U_{c_1}^1 c_{\lambda, 1} + U_{T-l_1}^1 (T_1 - l_{\lambda, 1})]} = \frac{\frac{\partial V_\lambda(w)}{\partial \omega_2} \omega_2 + \frac{\partial V_\lambda(w)}{\partial T_2} T + \frac{\partial V_\lambda(w)}{\partial k_2} k_2}{\lambda_2 p_2 [U_{c_2}^2 c_{\lambda, 2} + U_{T-l_2}^2 (T - l_{\lambda, 2})]} \quad (20)$$

It can be shown (see the online appendix) that the VP condition can be simplified to:

$$\frac{g_1 (1 - \tau) [\omega_1 T + k_1 + \omega_1 l_{\lambda, 1}] + \Phi_1}{g_1 [(1 - \tau) (\omega_1 T + k_1) + R]} = \frac{g_2 (1 - \tau) [\omega_2 T + k_2 + \omega_2 l_{\lambda, 2}] + \Phi_2}{g_2 [(1 - \tau) (\omega_2 T + k_2) + R]} \quad (21)$$

where the two sides of the VP condition represent the social to private (STP henceforth) contributions of each agent respectively. Here,  $g_i \equiv \lambda_i p_i U_{c_i}^i$  measures the value that society

puts on providing an additional dollar of consumption to any given individual. The terms  $\Phi_i$ , representing the indirect effects of a change in the endowments of agent  $i$  on the social welfare function through changes in tax policy, are given by:

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

Using the optimality of taxes from the Ramsey problem together with the government's budget constraint, these can be rewritten (see the online appendix) as:

$$\Phi_i = \Omega \tau p_i \left[ \omega_i \left( l_{\lambda, i} + \omega_i \frac{\partial l_i^*}{\partial \omega_i} \right) + \omega_i \frac{\partial l_i^*}{\partial T_i} T_i + \left( \omega_i \frac{\partial l_i^*}{\partial k_i} + 1 \right) k_i \right]$$

where  $\Omega$  is the social marginal cost of government revenues (i.e. the multiplier of the government budget constraint in the Ramsey problem). The bracket simply adds up the effects of a marginal increase in each endowment on the tax base  $\omega_i l_i + k_i$  times the corresponding endowment. Note that condition (21) can be re-written in terms of  $\frac{g_i}{\Omega}$ , which is what Piketty and Saez (2013) refer to as the social marginal welfare weight of agent  $i$ .

**No Government Spending.** Suppose first that  $G = 0$  and consider the possibility that  $\tau = R = 0$ . Using the VP condition, together with the fact that  $\Phi_i = 0$  when taxes are zero we can establish the following result.

**Result 2:** Assuming  $G = 0$ , then  $\tau = R = 0$  is a value preserving tax scheme iff:

$$\frac{l_1}{T + \frac{k_1}{\omega_1}} = \frac{l_2}{T + \frac{k_2}{\omega_2}} \quad (22)$$

The result follows directly from the VP condition in (21) evaluated at  $\tau = R = 0$ . Condition (22) involves exogenous endowments as well as the endogenous choice of effort  $l_i$ . Despite this, it can be used to provide sharp predictions regarding the direction of redistribution prescribed by VP weights.

First, notice that even if endowments are exactly equalized across the two types, no taxation is not a VP outcome unless effort is equalized. To be more specific, if agent  $i$  works more hours due to lower disutility of labor, then his STP is higher at the zero tax allocation. As a result, the VP government will increase his weight and redistribute income towards him. To put it simply, the VP government rewards effort.

Second, suppose for the sake of argument that effort is equalized across types (e.g. if labor supply is inelastic due to no preference for leisure). The STP is then increasing in ability and decreasing in wealth, indicating that the VP government would tend to redistribute towards the high ability and towards the low wealth agent. Thus the VP government distinguishes between ability and wealth as the underlying reasons for overall income inequality, whereas a utilitarian government would always redistribute to the type with low overall income  $\omega_i l_i + k_i$ .

Third, condition (22) implies an even sharper conclusion in the case of equal effort, namely, that the STP is decreasing in capital to ability ratios  $\frac{k}{\omega}$ . To the extent that wealth and ability (as measured by wages) are positively correlated and wealth is more unequally distributed than ability (which seems like the relevant case in the data), the VP government

would prescribe redistribution towards the low ability/low wealth type. This is a desirable feature since such redistribution is indeed the case in actual tax systems.

Whether the last two conclusions hold even when labor supply is allowed to endogenously vary across types will depend on the relative strength of income and substitution effects on labor supply. For example, if labor supply increases with ability, then the conclusion regarding VP redistributing towards high ability types is strengthened. In contrast, if labor supply increases with capital to ability ratios, this would introduce a countervailing force mitigating (or possibly reversing) the prescription of redistributing towards low capital to ability types.

In sum, VP weights tend to, *ceteris paribus*, favor high effort, high ability and low wealth types. The conclusion here is markedly different from our previous example where no redistribution was favored by the VP government. The fundamental difference is the presence of a non-traded endowment, namely ability, which is included in the social contribution of an agent.

**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 analytically regarding the type of redistribution the VP government will undertake. We can still gain some intuition by comparing the STP contribution of the two agent types in the case where no redistribution is allowed, i.e. assuming the tax scheme is restricted to  $R = 0$ . With  $G > 0$ , the additional terms  $\Phi_i$  have to be taken into account and these represent the social contribution of endowments through tax revenue effects. This additional component to the social contribution of type  $i$  depends on the level of endowments of  $i$  as well as on the elasticity of the tax base with respect to changes in those endowments. Importantly, since these tax revenue effects are valued at  $\Omega$ , whereas the direct effects are valued at  $g_i$ , the type with higher  $g_i$  will contribute proportionally less through tax revenues. We resort again to numerical computations to gauge the relative strength of the different components.

We assume type  $i = 1$  has higher wealth and higher ability and choose the ability and wealth differences between the two types to match a before-tax income Gini coefficient of 0.5 and a wealth Gini of 0.85 respectively, which are roughly the corresponding numbers in US data (see Chang et al (2018)). The individual utility function is assumed to be:

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

with  $\sigma = 1$ . Normalizing  $T = 1$ , we choose  $\mu = 0.3$  to deliver an average labor supply of approximately  $\frac{1}{3}$ . The level of government spending is chosen to be  $G = 0.1$ , which results in a government revenue to income ratio of approximately 0.17 in the economy with VP weights.

Table 2 below reports the numerical results for three economies, which differ in the societal preference for redistribution. Specifically we consider the economy with a VP government, with a government using equal weights EW and also an economy where the tax scheme is constrained to deliver no redistribution, i.e.  $R = 0$ .

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

	$(\lambda_1, \lambda_2)$	$(\tau, R)$	$(g_1, g_2)$	$(STP_1, STP_2)$	$\Delta G_w$
$R = 0$	(6.08, 0.43)	(0.13, 0)	(0.58, 0.71)	(1.31, 1.37)	0%
$VP$	(5.89, 0.45)	(0.17, 0.02)	(0.59, 0.75)	(1.34, 1.34)	-4.2%
$EW$	(1.00, 1.00)	(0.57, 0.12)	(0.19, 2.35)	(5.5, 0.97)	-54%

For each economy, the table displays the tax policy  $(\tau, R)$ , the social marginal welfare weights  $g_i = \lambda_i p_i U_{c_i}^i$  and the STP contributions of the two agents. The last column also displays the percentage decrease in the Gini coefficient of income, as an indicator of the amount of redistribution achieved by the tax scheme.

The fact that taxes are distortionary prevents the government from equalizing the  $g_i$  across agent types. In all cases, these are lower for type 1 agents, i.e. those with higher ability and higher wealth. As discussed earlier, this tends to increase their STP contribution through the tax revenue effect. At the same time, the capital to ability ratio is significantly higher for type 1 agents and this tends to decrease their STP contribution. At the no-redistribution allocation ( $R = 0$ ), the effect through the capital to ability ratio dominates the tax revenue effect, implying that the STP contribution of type 1 agents is lower. The government with VP weights, in an effort to equalize the STP contributions across agents, redistributes away from those agents and towards the low ability, low wealth agents.

This redistribution is consistent with actual tax schemes. Chang et al (2018) show that tax schemes across OECD countries effect some redistribution. For the case of the US, they show that the income Gini decreases by 23% through taxation. They also show that this level of redistribution can be made consistent with an optimal taxation model only by assuming larger weights for the wealthy. The VP approach captures these qualitative features by delivering larger welfare weights for the wealthy while, at the same time, delivering some level of redistribution. Quantitatively, the level of redistribution is more modest than in US data, decreasing the income Gini by only 4.2%. Notice, however, that the EW approach, which equalizes welfare weights across agents, delivers significantly more redistribution than in the data, reducing the income Gini by 54%. This is a point made both by Chang et al (2018) and by Weinzierl (2014). i.e. that the utilitarian social welfare function prescribes too much redistribution compared to what is observed in practice.

**Relation to the Equal Sacrifice Principle.** Another alternative to the VP principle that has been proposed in the literature is the equal sacrifice (ES) principle which requires the (utility) sacrifice resulting from taxes to be equated across individuals.<sup>22</sup> The sacrifice is defined relative to a benchmark economy with no taxes and can be computed in terms of the utility difference (absolute sacrifice) or utility ratios (relative sacrifice).

In the example above with log utility, the equal sacrifice principle prescribes a flat tax with no redistribution and corresponds to the case  $R = 0$  in Table 2. Weinzierl (2014) applies this principle and concludes, using survey data, that social preferences are best described by a mix of utilitarianism and equal sacrifice (no redistribution). The VP taxes reflect exactly such a mix, albeit with less weight on utilitarianism than implied by Weinzierl's calculations.

<sup>22</sup>See, for example, Young (1988, 1990) and Weinzierl (2014).

One way to rationalize additional redistribution within the VP principle framework would be to assign stronger preference for leisure to the high wealth/high ability type, thereby reducing their effort and social contribution.

One difficulty with the standard application of the ES principle is in defining the benchmark economy relative to which the sacrifice is computed. This could be a first best with no government (most commonly used) but could, equally reasonably, be the economy with frictions but no government. This ambiguity could be problematic since ES allocations and taxes will, in general, depend on the specification of the benchmark economy. One of the benefits of our VP principle relative to ES is that it naturally avoids having to make this choice of a reference economy altogether.

A second issue with the standard application of the equal sacrifice principle is that sacrifice (absolute or relative) is computed based on utility levels. This has the undesirable effect making the resulting ES allocations change when utilities are transformed in a non-uniform affine manner. The contribution mechanism we propose to measure the value of an allocation can be fruitfully used to implement the ES principle in a way that solves this problem.

To see how this could be implemented, assume that the benchmark point from which the sacrifice is calculated is the first best (no tax) allocation. The equal (relative) sacrifice condition can then be stated as follows:

$$\frac{C(\lambda_i u_i, x_\lambda^i)}{C(\lambda_i^N u_i, x_N^i)} = \frac{C(\lambda_h u_h, x_\lambda^h)}{C(\lambda_h^N u_h, x_N^h)} \text{ for all } i, h \in I \quad (23)$$

where  $\lambda^N = (\lambda_i^N)_{i \in I}$  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 (23) states that the ratio of the private value of the tax allocation to the private value of the first best allocation has to equalize across agents under the equal sacrifice weights. 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 the ratio of weighted utility levels  $\lambda_i u_i$  to compute the welfare weights.<sup>23</sup>

## 7. 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

---

<sup>23</sup>Similarly, our contribution mechanism could be used in a similar manner to implement other principles. For example, a Rawlsian principle could be set up as welfare weights determined by

$$\max_\lambda \min_i \left\{ C(\lambda_i u_i, x_\lambda^i) \right\} \quad (24)$$

where  $x_\lambda^i$  is the allocation with taxes for a given set of weights  $\lambda = (\lambda_i)_{i \in I}$  and  $C(\lambda_i u_i, x_\lambda^i)$  is the private value to the agent of this allocation. Using the contribution mechanism instead of utility levels ensures invariance to affine transformations of utility.

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. Applying the VP to standard optimal taxation problems, we find that VP taxes tend toward a no-redistribution benchmark but can still prescribe capital taxes that **increase** with the level of inequality. With ability taken into account when considering the social contribution of an agent, VP taxes favor ability and effort but do tend to redistribute wealth. Consistent with actual tax systems, the prescription involves some level of redistribution but to a much lower degree than the utilitarian principle.

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,

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

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.  $|\widehat{C}(I, 0) - \widehat{C}(I, \eta)| < \epsilon$ . But  $\widehat{C}(I, \eta) = \widehat{C}(I, 1) = \alpha$ . Hence  $|\widehat{C}(I, 0) - \alpha| < \epsilon$ . 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  $\mathbb{R}_+^m \setminus \{0\}$

$$F(x) = \nabla F(w)x + F(w) - w \nabla F(w) + o(w - x) \quad (25)$$

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

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

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

Let  $\eta = \min(\eta_1, \eta_2)$ . By (25) and by the definition of the  $C^1$  norm

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

Therefore,

$$|\widehat{C}_j(F, w) - \widehat{C}_j(G, w)| < \epsilon \quad (26)$$

By axiom 1,

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

Since  $\nabla F(w)$  is linear

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

Substituting this into (26)

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

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  $\lambda$  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

$$\nabla F_\lambda(w) = \frac{p}{K} = \lambda_i \nabla u_i(x_\lambda^i(w))$$

where  $K \equiv \sum_{i=1}^I \frac{1}{\mu_i}$  is a constant.<sup>24</sup> Multiplying by  $x_\lambda^i(w)$  we obtain

$$\lambda_i \nabla u_i(x_\lambda^i(w)) \cdot x_\lambda^i(w) = \nabla F_\lambda(w) \cdot x_\lambda^i(w) = \frac{p}{K} \cdot x_\lambda^i(w)$$

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

$$\begin{aligned} p \cdot x_\lambda^i(w) &= p \cdot \left( w^i + \sum_{j=1}^J \theta_j^i z_\lambda^j(w) \right) \\ &= p \cdot w^i \end{aligned}$$

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

$$\lambda_i \nabla u_i(x_\lambda^i(w)) \cdot x_\lambda^i(w) = \nabla F_\lambda(w) \cdot w^i$$

which by Theorem 1 implies

$$C(\lambda_i u_i, x_\lambda^i(w)) = C(F_\lambda, \tilde{w}^i, w)$$

and the Negishi weight  $\lambda$  is value preserving.

We now prove the other direction, namely, we show that if  $\lambda$  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

$$\nabla F_\lambda(w) \cdot w^i = c \lambda_i \nabla u_i(\bar{x}^i) \cdot \bar{x}^i \quad (27)$$

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  $u_i(x^i)$  over  $\{x^i | \bar{p} \cdot x^i \leq \bar{p} \cdot \bar{x}^i\}$

---

<sup>24</sup>Note that, in this case, the value  $F_\lambda$  depends on the aggregate endowment  $W = (W_1, \dots, W_L) \equiv \sum_{i=1}^I w^i \in \mathbb{R}_{++}^L$  only and not on the whole distribution  $w$ . The implication is that  $\frac{\partial F_\lambda}{\partial w_i^h} = \frac{\partial F_\lambda}{\partial w_i^h} = \frac{\partial F_\lambda}{\partial W_i}$  for all  $i, h \in I$  and the notation  $\nabla F_\lambda(w)$  should be interpreted to mean  $\left( \frac{\partial F_\lambda}{\partial W_1}, \dots, \frac{\partial F_\lambda}{\partial W_L} \right)$

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

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

By (27)  $\bar{p} \cdot w^i = c\bar{p} \cdot \bar{x}^i$ ,  $i \in I$ . 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  $u_i(x^i)$  over  $\{x^i | \bar{p} \cdot x^i \leq \bar{p} \cdot w^i\}$ , implying that  $\lambda$  is a Negishi weight. ■

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

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

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

$$\hat{L}(G) = \{(\hat{x}, \hat{z}) \mid \text{there exist } \hat{\tau} \in \mathbb{R}^{IL} \text{ and } \hat{p} \in \mathbb{R}_+^L \text{ s.t. } (\hat{x}, \hat{z}, \hat{\tau}, \hat{p}) \text{ satisfies (7)-(11)}\}$$

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

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

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

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

**Proof:** Since  $\hat{L}(G) \subseteq \bar{L}(G)$  it follows that  $\hat{F}_\lambda(w) \leq \bar{F}_\lambda(w)$ . 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  $u_i(x^i)$  over  $\{x^i \in \mathbb{R}_+^L | \bar{p}x^i \leq \bar{p}\bar{x}^i\}$  and for every  $j \in J$ ,  $\bar{z}^j$  maximizes  $\bar{p}z^j$  over  $z^j$  s.t.  $f_j(z^j) \leq 0$ . 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 (7),

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

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

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

Hence,  $\bar{x}^i$  maximizes  $u_i(x^i)$  over  $\{x^i \in \mathbb{R}_+^I | \bar{p}x^i \leq \bar{p}(w^i - \bar{\tau}^i)\}$ . Thus,  $(\bar{x}, \bar{z}, \bar{\tau}, \bar{p})$  satisfies the five conditions (7)-(11) and  $(\bar{x}, \bar{z}) \in \bar{L}(G)$ . This implies that

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

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  $\lambda$  is a VP weight. Then,

$$\frac{\nabla \widehat{F}_\lambda(w) \tilde{w}^i}{\nabla \widehat{F}_\lambda(w) \tilde{w}^h} = \frac{\lambda_i \nabla u_i(\widehat{x}_\lambda^i) \widehat{x}_\lambda^i}{\lambda_h \nabla u_h(\widehat{x}_\lambda^h) \widehat{x}_\lambda^h} \text{ for all } i, h \in I \quad (28)$$

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  $\widehat{x}^i = \arg \max u_i(x^i)$  over  $\{x^i \in \mathbb{R}_+^m | \widehat{p}x^i \leq \widehat{p}w^i\}$  for all  $i \in I$ . Hence, for some  $c$ ,

$$c\widehat{p} = \lambda_i \nabla u_i(\widehat{x}^i) \text{ for all } i \in I \quad (29)$$

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

$$\frac{\partial \bar{F}_\lambda}{\partial w_l^i}(w) = c\widehat{p}_l \quad (30)$$

By (28), (29) and (30),

$$\frac{\widehat{p}\widehat{x}_\lambda^i}{\widehat{p}w^i} = \alpha \quad (31)$$

where  $\alpha$  is a constant (that does not depend on  $i$ ). By (31),

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

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

$$\alpha = 1 - \frac{\widehat{p}G}{\widehat{p}W} \quad (32)$$

By (31) and (32),

$$\widehat{p}\widehat{x}_\lambda^i = \widehat{p}w^i - \widehat{p}G \frac{\widehat{p}w^i}{\widehat{p}W} \quad (33)$$

Since

$$\widehat{p}\widehat{\tau}^i = \widehat{p}w^i - \widehat{p}\tau^i \quad (34)$$

we have

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

## 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., 2018. Pareto Weights in Practice: A Quantitative Analysis Across 32 OECD Countries. *Review of Economic Dynamics*, 28, pp. 181 - 204.
- [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] Heathcote, J., Storesletten, K., Violante, G. L., 2017. *Optimal Tax Progressivity: An Analytical Framework*, *Quarterly Journal of Economics*, 132/4, pp. 1693-1754.

- [16] Heathcote, J., Tsujiyama, H., 2017. Optimal Income Taxations: Mirrlees meets Ramsey. Working Paper.
- [17] Kehoe P., and Perri F., 2002. International Business Cycles With Endogenous Incomplete Markets, *Econometrica*, 70, vol. 3, pp. 907-928.
- [18] Mankiw, G. N., 2010. Spreading the Wealth Around: Reflection Inspired by Joe the Plumber, *Eastern Economic Journal*, 36, pp. 285-298.
- [19] Mas -Colell, A., Whinston, M. D., Green, J. R., 1995. *Microeconomic Theory*. Oxford University Press, New York, New York.
- [20] Mirman,L., Tauman, Y., 1982. Demand compatible equitable cost sharing prices, *Mathematics of Operations Research*, vol. 7 (1), pp. 40-56.
- [21] Mirrlees, J., 1971. An Exploration in the Theory of Optimum Income Taxation, *The Review of Economic Studies*, vol. 38, no. 2, pp. 175 - 208.
- [22] Negishi, T., 1960. Welfare Economics and Existence of an Equilibrium for a Competitive Economy, *Metroeconomica*, vol. 12 (2-3), pp. 92-97.
- [23] Noszick, R., 1974. *Anarchy, State and Utopia*, Basic Books, New York.
- [24] Piketty, T., Saez, E., 2013. Optimal Labor Income Taxation. *Handbook of Public Economics*, Vol 5, Chapter 7, pp. 391 - 474.
- [25] Rawls, J., 1971. *A Theory of Justice*, Belknap Press of Harvard University Press, Cambridge, MA.
- [26] Saez, E., Stantcheva, S., 2015. Generalized Social Marginal Welfare Weights for Optimal Tax Theory, working paper.
- [27] Samet, D., Tauman, Y., 1982. The determination of marginal cost prices under a set of axioms, *Econometrica*, vol. 50 (4), pp. 895-909.
- [28] 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.
- [29] 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.
- [30] Weinzierl, M., 2014. The Promise of Positive Optimal Taxation: Normative Diversity and a Role for Equal Sacrifice, Working Paper, Harvard Business School.
- [31] Young, H. P., 1988. Distributive Justice in Taxation, *Journal of Economic Theory*, vol. 44 (2), pp. 321-335.

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