Soc Choice Welfare (2004) 23: 415–448
DOI: 10.1007/s00355-003-0266-2

Social Choice
and Welfare

Social welfare, inequality, and poverty when needs differ

Udo Ebert

Department of Economics, University of Oldenburg, 26111 Oldenburg, Germany (e-mail: ebert@uni-oldenburg.de)

Received: 17 May 2001/Accepted: 23 June 2003

Abstract. The paper examines income distributions of a finite population consisting of households which may differ with respect to needs. Since observed incomes are not directly comparable, income distributions have to be adjusted. Incomes are transformed to equivalent incomes interpreted as living standards and measured for a reference type, and the latter are supplemented by weights depending on needs. A general class of social welfare orderings (being based on adjusted rank-ordered income distributions) is characterized by a set of properties. Severe limitations for the form of the adjustment process are implied. The consequences for the measurement of inequality and poverty are demonstrated, and corresponding orderings are derived.

1 Introduction[1]

The objective of the paper is to investigate the possibilities of measuring social welfare, inequality, and poverty for a heterogeneous population, i.e., when households may have different incomes and may differ in size, composition, and/or needs. These attributes determine the type of a household. In other words, households may also have different types. The latter represent the respective needs. The main difference between the evaluation of an income distribution for a homogeneous and a heterogeneous population is the

[1] The paper forms part of the research programme of the TMR network **Living Standards, Inequality and Taxation** [Contract No. ERBFMRXCT980248] of the European Communities whose financial support is gratefully acknowledged. I thank an Editor of this journal and two anonymous referees for most helpful critique, comments, and suggestions.The Editor and a referee helped me to avoid errors in an earlier draft.

dimensionality of the problem: When households may differ in type, we face a two-dimensional problem. A heterogeneous income distribution consists of a distribution of household income and of the corresponding household type. Then the incomes of different household types are no longer (directly) comparable, since they do not reflect the type and living standard attained. Any evaluation in terms of welfare, inequality, or poverty has to take into account income and type simultaneously.

The common practice[2] is to adjust a heterogeneous income distribution by transforming it to a homogeneous distribution for a hypothetical population. In the first step incomes are transformed into equivalent incomes for a reference type, and according to the type of households certain weights are attached to equivalent incomes. Usually single adults are chosen as the reference type. In this case a household's income is adjusted by a type-specific (relative) equivalence scale in order to define the corresponding equivalent income. The latter corresponds to the income a single adult requires to be as well off as the household under consideration. Furthermore the number of persons belonging to the household is used as a weight. Obviously the resulting income distribution (and hypothetical population) is homogeneous. In the second step the usual measures are employed, and the differing weights are included in the analysis.

The present paper examines the possibilities and limitations of this two-step process for the measurement of social welfare. Instead of using equivalence scales and the number of persons as weights, a general adjustment process is considered. Income and type is transformed by an equivalent income function and a weighting function, respectively. The former transforms income into equivalent income according to the respective household type. Equivalent income represents the living standard attained. The latter assigns a type-specific weight. It is assumed that both functions are given a priori and reflect a social decision maker's preferences and norms. Furthermore, a general form of social welfare orderings will be investigated. They are rank-dependent, i.e. depend on the rank of equivalent income and respect the weights attached.

Then the usual properties of social welfare orderings are introduced and appropriately generalized. Their implications are derived. It turns out that the equivalent income function must be affine and that the form of weights is restricted. It means in particular that one has to employ (relative) equivalence scales, if one confines oneself to strictly positive (equivalent) incomes. The weights have to be proportional to the respective scales. If incomes may be arbitrary, all weights have to be the same and absolute equivalence scales have to be used (incomes are translated by an amount depending on the respective type). Thus severe restrictions on the choice of the equivalent income function and weights are implied in this framework. The limitations

[2] See e.g., Buhmann et al.(1988), Coulter et al. (1992), and Burkhauser et al. (1996).

also hold for 'symmetric' welfare orderings, which do not depend on the rank of incomes, since they form a subset of rank-dependent orderings.

These results are applied to derive ethical inequality and poverty orderings/measures. Based on a given social welfare ordering the welfare loss due to inequality can be determined. This Atkinson-Kolm-Sen measure of inequality can be defined in the extended framework, too. By definition it reflects the inequality of the distribution of living standards. If additional invariance properties are imposed, one obtains a characterization of generalized Atkinson, Kolm-Pollak and linear measures (and orderings) for heterogeneous populations. In particular an appropriate generalization of the Gini coefficient is presented as an example. The generalized poverty orderings considered are based on a procedure first proposed by Clark et al. (1981). Here a censored (heterogeneous) income distribution is derived such that equivalent incomes of non-poor households are replaced by the poverty line. Then the representative poverty gap measured for this distribution defines the corresponding poverty ordering – given a social welfare ordering. In this way a class of rank-dependent poverty orderings can be characterized for heterogeneous populations.

The paper extends the literature in various ways. It does not add anything to the investigation of partial orderings for heterogeneous populations (Atkinson and Bourguignon 1987; Ebert 1999, 2000a). However, it generalizes and supplements work on social welfare orderings: Ebert (1995) and Ebert (1997a) derive partially symmetric social welfare functions for a heterogeneous population. In these papers the equivalent income function is characterized implicitly: One obtains relative or absolute equivalence scales. Below the equivalent income function and weights are assumed to be chosen by a decision maker from the beginning. Then only specific functional forms are compatible with further properties of the welfare ordering. Similarly, Ebert (1997b) investigates the relationship between a given set of weights and equivalence scales. Below the decision weights (or coefficients) used in a social welfare function are determined by the framework and depend on the rank of incomes.

Of course, there is a connection to the literature on rank-dependent expected utility functions (see e.g. Quiggin 1982; Segal 1989; Wakker 1989; for axiomatizations and related problems). Furthermore the work on rank-dependent social evaluation functions for homogeneous populations is extended (Mehran 1976; Donaldson and Weymark 1980; Weymark 1981; Yaari 1987, 1988; Ebert 1988). Put the other way around, if it is assumed that all households have the same type (the reference type) and are given the same weight, the results of this paper collapse to the corresponding ones for homogeneous populations. Above that, the representation of linear rank-dependent inequality measures for a heterogeneous population supplements recent work on this type of measure for a homogeneous population (cf. Ben-Porath and Gilboa 1994; Weymark 1994; Ben-Porath et al. 1997).

At first sight it seems that there might be a connection to the measurement of multidimensional inequality examined by Kolm (1977), Maasoumi (1986), and Tsui (1995, 1999). But this is not the case. These authors suppose that *all* variables considered are continuous. In this paper the variable 'household type' is a discrete one.

Summing up, the paper investigates the limitations for and the normative foundations of rank-dependent orderings and indicators for the measurement of social welfare, inequality, and poverty. It derives the restrictions which have to be imposed on the adjustment process, which is employed to make incomes of different household types comparable. It extends the literature on social indicators for homogeneous populations and relates its findings to the work on rank-dependent expected utility functions.

The organization of the paper is as follows. Section 2 considers rank-dependent orderings for homogeneous populations. Furthermore the implications of the principle of population are investigated, when the population is replicated. Section 3 introduces the framework and the notation for a heterogeneous population. The adjustment process of the income distribution leads to a distribution for a hypothetical homogeneous population. Given this background in Sect. 4 the redistribution of income is examined. Then the main result of the paper is derived. It describes the implications of the properties imposed on social welfare orderings: The possibilities and limitations are presented. Section 5 applies the result to the measurement of inequality and poverty. Section 6 concludes.

2 Income distribution for a homogeneous population

2.1 Framework and notation

We consider a population consisting of a number of individuals which form n groups. It is assumed that all individuals are the same with respect to type and needs. They may, however, differ with respect to the group, they belong, to and to their income. All individuals belonging to a group always possess identical incomes. The groups are numbered arbitrarily, but the numbering will be kept unchanged during the analysis. Group i ($i \in N := \{1, \ldots, n\}$) is described by its weight $w_i > 0$ and each member's income Y_i. The weight reflects the number of individuals belonging to that group and can be a strictly positive real number. Thus, the population comprises $\sum_{i=1}^{n} w_i$ identical individuals which form n groups. For ease of exposition we shall sometimes speak of w_i as if it were an integer, but nothing material will depend on this convenience. Incomes have to be feasible, i.e., $Y_i \in \Omega$, where the set of feasible incomes, Ω, is given by $\mathbb{R}$ or $\mathbb{R}_{++}$.

This model is an extension of the simple one used for welfare measurement: Whenever $w_i = 1$ for all $i \in N$ we get a homogeneous population having n individuals. Otherwise w_i can be interpreted as weight attributed to individual i. In the following sections a group corresponds to a household. w_i reflects its type; Y_i will be equal to its equivalent income. A homogeneous income distribution is given by a vector $(\mathbf{Y}, \mathbf{w})$, where $\mathbf{Y} = (Y_1, \ldots, Y_n) \in \Omega^n$

and $\mathbf{w} = (w_1, \dots, w_n) \in \mathbb{R}^n_{++}$ denote an income profile and the weight profile, respectively (see Ebert and Moyes (2002) for an investigation of symmetric social welfare functions in this framework).

Below further notation is required. For every income profile $\mathbf{Y}$ there is a permutation $\pi(\mathbf{Y}) = \pi(\cdot, \mathbf{Y})$ of N such that incomes are ordered decreasingly: $Y_{\pi(i,\mathbf{Y})} \geq Y_{\pi(i+1,\mathbf{Y})}$ for $i = 1, \dots, n-1$. The corresponding vector is denoted by $\mathbf{Y}_\downarrow := \left(Y_{[1]}, \dots, Y_{[n]}\right)$ and the corresponding weight profile by $\mathbf{w}_\downarrow = \mathbf{w}_\downarrow(\mathbf{Y}) := \left(w_{[1]}(\mathbf{Y}), \dots, w_{[n]}(\mathbf{Y})\right) := \left(w_{\pi(1,\mathbf{Y})}, \dots, w_{\pi(n,\mathbf{Y})}\right)$.

Here the argument $\mathbf{Y}$ is dropped whenever there is no ambiguity. If all incomes $Y_i, i \in N$, differ, the permutation $\pi(\mathbf{Y})$ is unique. Otherwise there are several permutations implying the same vector $\mathbf{Y}_\downarrow$. It turns out that it does not matter which permutation is chosen. The results derived later on are independent of this choice. $\pi(\mathbf{Y})$ is called the characteristic permutation of $\mathbf{Y}$. $Y_{[i]}$ is the income of an individual belonging to the group with rank i.

Two income profiles $\mathbf{Y}$ and $\mathbf{Y}'$ possessing the same characteristic permutation are called comonotonic if they are ordered in a similar way and satisfy the conditions $(Y_i - Y_j)(Y_i' - Y_j') > 0$ or $(Y_i \geq Y_j \ iff \ Y_i' \geq Y_j')$ for all $i, j \in N$ and $\mathbf{w}_\downarrow(\mathbf{Y}) = \mathbf{w}_\downarrow(\mathbf{Y}')$.

An income profile $\mathbf{Y}$ is sometimes modified: Let $\mathbf{Y}^{-j}\alpha$ be the income vector which is identical with $\mathbf{Y}$ in all components $i \neq j$, $i \in N$, and in which the j-th component is equal to α. More precisely, $\mathbf{Y}^{-j}\alpha := \mathbf{Y}'$, where $Y_i' = Y_i$ for all $i \neq j$ and $Y_j' = \alpha$.

Finally we introduce the abbreviation $w(A) := \sum_{i \in A} w_i$ for any nonempty subset $A \subset N$ and $w(\emptyset) := 0$. $w(N)$ denotes the total number of individuals or the population size. Often weights have to be normalized. We define $D(\mathbf{w}) := \{w(A)/w(N) \mid A \subset N\} \subset [0,1]$, a set containing all possible shares $w(A)/w(N)$ of the population, if groups are combined to a subpopulation A. Though weights are always strictly positive, it is assumed that $D(\mathbf{w})$ also contains 0 (the share of a void subpopulation). The reason for this extension will be clarified in the next section (see Footnote 3 there).

2.2 Social welfare ordering

Here it is our aim to consider homogeneous income distributions $(\mathbf{Y}, \mathbf{w}), (\mathbf{Y}', \mathbf{w})$ for a given population where $\mathbf{Y}, \mathbf{Y}' \in \Omega^n$. Therefore the weight profile $\mathbf{w}$ is always the same and reflects the formation and the respective size of groups. We introduce an ordering $\succsim_V$ on $\Omega^n \times \{\mathbf{w}\}$, i.e. a complete, reflexive, and transitive relation, which is to compare the social welfare inherent in $(\mathbf{Y}, \mathbf{w})$ and $(\mathbf{Y}', \mathbf{w})$; actually we evaluate social welfare of the income profiles $\mathbf{Y}$ and $\mathbf{Y}'$ for given $\mathbf{w}$. $\succsim_V$ can be interpreted as a preference relation of a social decision maker. The symmetric part is denoted by $\sim_V$ and the asymmetric part by $\succ_V$. The weights have, of course, to be taken into account.

The relation $\succsim_V$ should satisfy reasonable properties. We assume that $\succsim_V$ is represented by

$$\xi(\mathbf{Y},\mathbf{w}) = U^{-1}\left(\sum_{i=1}^{n} \alpha_i\big(\mathbf{w}_{\downarrow}(\mathbf{Y})\big)\, U\big(Y_{[i]}\big)\right), \tag{1}$$

where
(i) $U : \boldsymbol{\Omega} \to \mathbb{R}$ is continuous and strictly increasing,
(ii) there is a strictly increasing function $f : D(\mathbf{w}) \to [0, 1]$ with $f(0) = 0$ and $f(1) = 1$ such that[3]

$$\alpha_i\big(\mathbf{w}_{\downarrow}(\mathbf{Y})\big) = f\left(\sum_{j=1}^{i} w_{[j]}(\mathbf{Y})/w(N)\right) - f\left(\sum_{j=1}^{i-1} w_{[j]}(\mathbf{Y})/w(N)\right) > 0, \text{ and} \tag{2}$$

(iii) $\sum_{i=1}^{n} \alpha_i\big(\mathbf{w}_{\downarrow}(\mathbf{Y})\big) = 1$.

The social welfare ordering $\succsim_V$ is represented by social welfare functions defined on rank-ordered income profiles. Furthermore, the coefficients α_i depend on the weight of all groups whose income has rank i or a rank less than i. (Here the function f has also to be defined at 0, since for the group ranked highest the subpopulation being ranked even higher is void.) They reflect the position of $Y_{[i]}$ in the income scale. The form of the social welfare function is essentially identical with a (general) rank-dependent expected utility (RDEU) function. But the interpretation of these functions in both frameworks is different. In particular, in our model there is no risk or uncertainty. Formally, the weights w_i replace the probabilities, and the decision weights $\alpha_i\big(\mathbf{w}_{\downarrow}(\mathbf{Y})\big)$ (of the welfare function ξ) determine the contribution of the income level $Y_{[i]}$ having rank i to social welfare. They are generated by differences of transformed cumulated (relative) weights. The function f transforms the weights in a (generally) non-additive way. It corresponds to the decision maker's evaluation of the position of income in the income distribution. U can be interpreted as utility function. Therefore f and U are called transformation and utility function, respectively. (f, U) determine the underlying welfare ordering $\succsim_V$ uniquely.

The representation of $\succsim_V$ is a specific one: $\xi(\mathbf{Y},\mathbf{w})$ denotes the equally distributed equivalent income (EDEI) which is defined by $\xi(\mathbf{Y},\mathbf{w})\, \mathbf{1}_n \sim_V (\mathbf{Y},\mathbf{w})$, where $\mathbf{1}_n$ is a vector containing n ones. $\xi(\mathbf{Y},\mathbf{w})$ corresponds to the level of income which, if enjoyed by each individual (group) in society, yields the same level of social welfare as $(\mathbf{Y},\mathbf{w})$.

Social welfare orderings which are represented by (1) are characterized in Ebert (2003). They satisfy four properties:

Continuity. *The ordering* $\succsim_V$ *is continuous in income on* $\boldsymbol{\Omega}^n \times \{\mathbf{w}\}$.[4]

[3] If the set of feasible indices is void we define $\sum_{j=1}^{0} w_{[j]}(\mathbf{Y}) := 0$.
[4] The sets $\{\mathbf{Y}' \in \boldsymbol{\Omega}^n | (\mathbf{Y}',\mathbf{w}) \succsim_V (\mathbf{Y},\mathbf{w})\}$ and $\{\mathbf{Y}' \in \boldsymbol{\Omega}^n | (\mathbf{Y}',\mathbf{w}) \precsim_V (\mathbf{Y},\mathbf{w})\}$ are closed for all $\mathbf{Y} \in \boldsymbol{\Omega}^n$.

Independence.[5] *Choose any* $\mathbf{Y}, \tilde{\mathbf{Y}}, \mathbf{Y}', \tilde{\mathbf{Y}}' \in \Omega^n$ *,* $i, j \in N$ *and* $\alpha, \beta, \gamma, \delta \in \Omega$ *such that* $\mathbf{Y}, \mathbf{Y}', \mathbf{Y}^{-i}\alpha, \mathbf{Y}^{-i}\gamma, \mathbf{Y}'^{-i}\beta, \mathbf{Y}'^{-i}\delta$ *and* $\tilde{\mathbf{Y}}, \tilde{\mathbf{Y}}', \tilde{\mathbf{Y}}^{-j}\alpha, \tilde{\mathbf{Y}}^{-j}\gamma, \tilde{\mathbf{Y}}'^{-j}\beta, \tilde{\mathbf{Y}}'^{-j}\delta$*, respectively, are comonotonic.*

$$\textit{Then} \quad \left(\mathbf{Y}^{-i}\alpha, \mathbf{w}\right) \precsim_V \left(\mathbf{Y}'^{-i}\beta, \mathbf{w}\right), \tag{3}$$

$$\left(\mathbf{Y}^{-i}\gamma, \mathbf{w}\right) \succsim_V \left(\mathbf{Y}'^{-i}\delta, \mathbf{w}\right), \quad \textit{and} \tag{4}$$

$$\left(\tilde{\mathbf{Y}}^{-j}\alpha, \mathbf{w}\right) \succsim_V \left(\tilde{\mathbf{Y}}'^{-j}\beta, \mathbf{w}\right) \tag{5}$$

$$\textit{imply} \quad \left(\tilde{\mathbf{Y}}^{-j}\gamma, \mathbf{w}\right) \succsim_V \left(\tilde{\mathbf{Y}}'^{-j}\delta, \mathbf{w}\right). \tag{6}$$

Neutrality. *Suppose that there are* $\mathbf{Y}, \mathbf{Y}' \in \Omega^n$*,* $\alpha, \beta \in \Omega$ *and* $k, l \in N$ *such that*

$$\mathbf{Y}_\downarrow = (\alpha \mathbf{1}_k, \beta \mathbf{1}_{n-k}), \quad \mathbf{Y}'_\downarrow = (\alpha \mathbf{1}_l, \beta \mathbf{1}_{n-l}) \textit{and} \sum_{j=1}^{k} w_{[j]}(\mathbf{Y}) = \sum_{j=1}^{l} w_{[j]}(\mathbf{Y}').$$

Then $(\mathbf{Y}, \mathbf{w}) \sim_V (\mathbf{Y}', \mathbf{w})$.

Monotonicity. *Suppose that there are* $\mathbf{Y}, \mathbf{Y}' \in \Omega^n$*,* $\alpha, \beta \in \Omega$*, and* $k, l \in N$ *such that*

$$\mathbf{Y}_\downarrow = (\alpha \mathbf{1}_k, \beta \mathbf{1}_{n-k}), \quad \mathbf{Y}'_\downarrow = (\alpha \mathbf{1}_l, \beta \mathbf{1}_{n-l}) \quad \textit{and} \quad \sum_{j=1}^{k} w_{[j]}(\mathbf{Y}) > \sum_{j=1}^{l} w_{[j]}(\mathbf{Y}').$$

Then $(\mathbf{Y}, \mathbf{w}) \succ_V (\mathbf{Y}', \mathbf{w})$.

CONTINUITY implies that any continuous social welfare function representing $\succsim_V$ reacts to small changes of an income profile itself with a small change.

INDEPENDENCE considers in (3) and (4) two comonotonic income profiles. For group i the actual incomes are replaced by (α, β) and (γ, δ), respectively, in such a way that all profiles are still comonotonic. The substitution of α and β by γ and δ causes a preference reversal. The tradeoff between γ and δ (i.e. getting income γ instead of δ) leads to a stronger improvement than the tradeoff between α and β. This result can be applied to (5) and to a possibly different group j. Since here the profile containing α is already not welfare inferior to that containing β, substituting (α, β) by (γ, δ) leads to a further improvement. We (have to) expect (6). Rank separability of a social welfare function is implied.

[5] It is called 'comonotonic cardinal coordinate independence' (Com. CCI) by Wakker (1989).

NEUTRALITY is concerned with two income profiles in which only two income levels are involved. The share of the population which has income α and β, respectively, is the same. Then the composition of the respective subpopulations does not matter, since the sum of the weights represent overall size (cf., Yaari's 1987 Axiom A1).

MONOTONICITY requires that the income profile, in which the share of the population having the higher income α is greater, is welfare superior.

These four properties, which will be abbreviated by WELFARE, imply that a social welfare ordering $\succsim_V$ is given by (1) and (2), and that (f, U) satisfies the properties stated. The property INDEPENDENCE is stronger than the usually used additive separability condition (see e.g. Ebert 1988) for homogeneous populations. In the presence of the other axioms it implies the separation of the weights w_i from the common utility function U. Since the utility function is the same for all groups, all information regarding the group is captured in the decision weight attached to each group. The weight α_i depends on the rank of group i's (average) income Y_i. Groups having the same size, but different income levels may therefore end up with different decision weights: Lower incomes are in general assigned higher weights, i.e. they are regarded as 'more deserving'. But a specific case is also covered by (1): If $f(w) = w$, the decision weight $\alpha_i = w_{[i]}(\mathbf{Y})/w(N)$ depends only on the weight of group i. U is unique up to a positive affine transformation, i.e. cardinal, as always for a separable welfare function. The transformation function f is unique. Further interpretation and the proof can be found in Ebert (2003). It contains a generalization of Ebert (1988) to the case in which incomes may be weighted, and in which only one population is considered. Furthermore, no homogeneity property is required for the result. The proof is essentially based on Wakker (1989).

2.3 Principle of population

Up to now a population has been considered which has exactly n groups with fixed composition. In the following we introduce k-fold replications for $k \geq 2$. For each group i having weight w_i there are k groups with weight w_i in the replicated population. The k-fold replication of the weight profile $\mathbf{w}$ is $\mathbf{w}^{(k)}$, i.e. for instance $\mathbf{w}^{(2)} = (w_1, w_1, \dots, w_n, w_n)$. Analogously an income profile $\mathbf{Y}$ can be replicated to $\mathbf{Y}^{(k)}$. A social welfare ordering is then defined on the set[6] $\Omega^{kn} \times \{\mathbf{w}^{(k)}\}$ and is denoted by $\succsim_{V,k}$, where $\succsim_{V,1} :\equiv \succsim_V$. Let N_k be equal to $\{1, \dots, kn\}$ and D_k to $D(\mathbf{w}^{(k)})$. It should be obvious that the properties discussed – if imposed on every ordering $\succsim_{V,k}$ – are not sufficient to make $\succsim_{V,k}$ and $\succsim_{V,l}$ for $k \neq l$ consistent. Consistency can be attained by considering

[6] The set of k-fold replicated income distributions forms a proper subset of $\Omega^{kn} \times \{\mathbf{w}^{(k)}\}$.

replications of income distributions $(\mathbf{Y}, \mathbf{w}) \in \Omega^n \times \{\mathbf{w}\}$. Therefore we introduce the

Principle of Population. *For all $k \geq 2$ and $(\mathbf{Y}, \mathbf{w}), (\mathbf{Y}', \mathbf{w}) \in \Omega^n \times \{\mathbf{w}\}$ such that $(\mathbf{Y}, \mathbf{w}) \sim_{V,1} (\mathbf{Y}', \mathbf{w})$ we have $\left(\mathbf{Y}^{(k)}, \mathbf{w}^{(k)}\right) \sim_{V,k} \left(\mathbf{Y}'^{(k)}, \mathbf{w}^{(k)}\right)$.*

If two distributions are equivalent in terms of social welfare for the (original) population of size n, then the equivalence holds as well between the replicated income distributions in the k-fold replicated population. In other words, the equivalence of distributions depends on the (two-dimensional) distributions ($(\mathbf{Y}, \mathbf{w})$ and $\left(\mathbf{Y}^{(k)}, \mathbf{w}^{(k)}\right)$ are identical, if interpreted as two-dimensional distributions; i.e., their distribution functions are the same and therefore independent of the size of the actual population).

Furthermore we postulate that $\succsim_{V,k}$ satisfies WELFARE for every $k \geq 1$. Then the sequence of social welfare orderings $\succsim_{V,k}$ for $k \geq 1$ can be characterized by a sequence

$$\xi_k\left(\mathbf{Y}, \mathbf{w}^{(k)}\right) = U_k^{-1}\left(\sum_{i=1}^{k\,n} \alpha_i^k\left(\mathbf{w}_{\downarrow}(\mathbf{Y})\right) U_k\left(Y_{[i]}\right)\right) \text{ for } \mathbf{Y} \in \Omega^{k\,n},$$

where $U_k : \Omega \to \mathbb{R}$ is continuous and strictly increasing, and where there is a function $f_k : D\left(\mathbf{w}^{(k)}\right) \to [0, 1]$, strictly increasing with $f_k(0) = 0$ and $f_k(1) = 1$, such that

$$\alpha_i^k\left(\mathbf{w}_{\downarrow}(\mathbf{Y})\right) = f_k\left(\sum_{j=1}^{i} w_{[j]}(\mathbf{Y})/w(N_k)\right) - f_k\left(\sum_{j=1}^{i-1} w_{[j]}(\mathbf{Y})/w(N_k)\right)$$

and $\sum_{i=1}^{k\,n} \alpha_i^k\left(\mathbf{w}_{\downarrow}(\mathbf{Y})\right) = 1$.

In the following a countable and dense subset of the unit interval will be important: $D := \cup_{k \geq 1} D\left(\mathbf{w}^{(k)}\right)$. It contains 0 and all real numbers in $[0, 1]$ which can be generated by means of the weights for any replication of the population. We obtain

$$D = \{w(A_k)/w(N_k) \mid \exists\, k \geq 1, A_k \subset N_k\}.$$

Within this framework we investigate the consequences of the Principle of Population and establish

Proposition 1. *Let $\succsim_{V,k}$, fulfill WELFARE and be characterized by (f_k, U_k).*

Then the following two statements are equivalent.

(a) The welfare orderings $\succsim_{V,k}$, $k \geq 1$, satisfy the principle of POPULATION.

(b) (i) There is a continuous and strictly increasing function U (unique up to a positive affine transformation).

(ii) There is a unique strictly increasing function $f: D \to [0,1]$ *such that* $f(0)=0, f(1)=1$, *and* $f(w(A_k)/w(N_k)) = f_k(w(A_k)/w(N_k))$ *for every non empty* $A_k \subset N_{\mathrm{k}}$ *and* $k \geq 1$.

$$(iii) \quad \xi_k\left(\mathbf{Y}, \mathbf{w}^{(k)}\right) = U^{-1}\left(\sum_{i=1}^{k\,n}\left[f\left(\sum_{j=1}^{i} w_{[j]}(\mathbf{Y})/w(N_k)\right) - f\left(\sum_{j=1}^{i-1} w_{[j]}(\mathbf{Y})/w(N_k)\right)\right]U(Y_{[i]})\right)$$

represents $\succsim_{V,k}$ *for all* $\mathbf{Y} \in \Omega^{k\,n}$ *and* $k \geq 1$.

The proposition demonstrates the implications of the principle of population. The property requires that the social welfare orderings $\succsim_{V,k}$ are compatible with one another. They are completely defined by the transformation function f and the common utility function U. The restriction of f to $D\left(\mathbf{w}^{(k)}\right)$ determines the decision weights. Moreover there exists a relationship between the decision weights $\alpha_i^k(\mathbf{Y})$ for different populations:

$$\alpha_i^k\left(\mathbf{w}_{\downarrow}(\mathbf{Y})\right) = \sum_{j=(j-1)l+1}^{i\,l} \alpha_j^{k\,l}\left(\mathbf{w}_{\downarrow}\left(\mathbf{Y}^{(l)}\right)\right) \text{ for } l > 1 \tag{7}$$

It demonstrates that the orderings $\succsim_{V,k}$ and $\succsim_{V,kl}$ are compatible with one another; but they do not allow us to compare distributions for different population sizes.

At this stage it is not possible to prove that the transformation function f is continuous[7]. On the contrary, it may be discontinuous almost everywhere in D, as the following example[8] demonstrates:

Example 1. f is defined on the set D which is countable. Therefore it is possible to number the elements of D such that $D = \{w^i \mid i = 0, 1, 2, 3, \ldots\}$, where $w^0 := 0$. We choose a sequence of rational numbers $p_i := 2^{-i}$ for $i \geq 1$. Then $\sum_{i=1}^{\infty} p_i = \sum_{i=0}^{\infty} 2^{-i} - 1 = \frac{1}{1-1/2} - 1 = 1$.

Now define $f(0) := 0$ and $f(w) := \sum_{w^i \leqslant w} 2^{-i}$ for $w > 0$. f satisfies $f(0) = 0$ and $f(1) = 1$. Furthermore it is discontinuous at every $w \in D$, $w \neq 0$, since there is a jump at w: We define $f_-(w) := \sup_{\tilde{w} < w} f(\tilde{w})$ and $f_+(w) := \inf_{w \leq \tilde{w}} f(\tilde{w})$ and obtain that

$$f(w) = f_+(w) = f_-(w) + p_i \text{ if } w = w^i. \qquad \blacksquare$$

---

[7] f is defined on the discrete set D. Consider any converging sequence $w_i \in D$ with $w = \lim w_i \in D$. Then f is continuous in w if and only if $f(w) = \lim f(w_i)$.

[8] The construction is similar to that used in example 18 of Gelbaum and Olmsted (1964), p. 28.

This result is not really surprising, since CONTINUITY refers to continuity with respect to income and says nothing about the weights attached to the groups. Therefore one cannot expect that f is the restriction of a continuous function on $[0,1]$ to D. The only information one can extract is that f is strictly increasing and (a restriction of) an almost everywhere continuous function.

3 Income distribution for a heterogeneous population

Now we consider a heterogeneous population, which consists of $n \geq 2$ households. The numbering is arbitrary, but it is kept fixed. Households may differ in size, composition, and/or needs, which are reflected by a corresponding type. It is assumed that there are $n_H \geq 2$ different types. Without loss of generality types are characterized by a whole number[9] $m \in H := \{1, \cdots, n_H\}$. Here m can be interpreted as the number of persons belonging to the respective household; i.e. household types are numbered by increasing needs. It is supposed that $n_H \leq n$ and that for each $m \in H$ at least one household having type m belongs to the population. The set of feasible incomes Ω is given by $\mathbb{R}$ or $\mathbb{R}_{++}$. Each household $i \in N$, can be described or characterized by its total income $X_i \in \Omega$ and its type $m_i \in H$. A heterogeneous income profile is denoted by $\mathbf{X} = (X_1, \ldots, X_n) \in \Omega^n$, a type profile by $\mathbf{m} = (m_1, \ldots, m_n)$. $(\mathbf{X}, \mathbf{m})$ represents the corresponding heterogeneous income distribution. In order to distinguish between a heterogeneous and homogeneous income profile, now the symbol $\mathbf{X}$ is used instead of $\mathbf{Y}$ employed in Sect. 2. (X_i denotes household i's aggregate income, whereas Y_i represents the income of each individual belonging to group i.)

For an evaluation of a heterogeneous income distribution it is necessary to take into account the income and the type of households. In practice $(\mathbf{X}, \mathbf{m})$ is adjusted: Incomes are equivalized and types are translated into appropriate weights. We will proceed in the same way. Such a procedure allows us to take into consideration differences in size or needs and to compare the situation of households having different types. Comparisons of social welfare will be based on the adjusted income distribution. Furthermore, equivalent incomes reflect the situation of households by means of income measured for one household type – the reference type. Therefore the adjusted income distribution can be interpreted as a distribution of a hypothetical homogeneous population.

In order to compare living standards we introduce an equivalent income function. It is supposed that there is a reference household type $r \in H$. The basic idea is to measure the living standard of a type m household by the equivalent income of a type r household. The reference type r could be e.g., a

[9] It is possible to introduce a functional relationship $g(m)$ describing the needs or composition of a household having type m. Then the model becomes more complicated, but the extension does not contribute anything essential to the analysis.

single person, but in principle any type[10] can be chosen. An equivalent income function E is given by a vector of functions[11] $(E_1, \ldots, E_{n_H})$ such that

$$E_m : \Omega \to \Omega \text{ for } m = 1, \ldots, n_H.$$

$E_m(X)$ is equal to the income a household having the reference type r needs in order to be as well off as a type m household possessing income X. $E_m(X)$ is called equivalent income. E allows us to compare living standards. We define: household i is weakly better off than household j, if and only if $E_{m_i}(X_i) \geq E_{m_j}(X_j)$. Thus the living standard is measured by the respective equivalent income. Here it is implicitly assumed that there is no information about the distribution of income within households and, as a consequence, that all members belonging to a household attain the same standard of living. This assumption could be changed or weakened, if more detailed data were available. Then more elaborated models of household behaviour could be used (cf. e.g. Bourguignon 1989 or Chiappori 1992).

Equivalent income functions have to satisfy some conditions:

E(i) $E_m(X)$ is continuous on Ω for each $m \in H$.
E(ii) $E_r(X) = X$ for $X \in \Omega$.
E(iii) $E_m(X)$ is strictly increasing in $X \in \Omega$ for each $m \in H$.
E(iv) $E_m(X)$ is decreasing in $m \in H$ for all $X \in \Omega$.
E(v) $\mathrm{Im}(E_m(\cdot)) = \Omega$ for $m \in H$, where $\mathrm{Im}(E_m(\cdot))$ denotes the image of $E_m(X)$.

E(i) is a regularity condition. The income of a type r household has not to be transformed, since r is the reference type (E(ii)). E(iii) implies that the living standard (or the equivalent income) is strictly increasing in household income X. Condition E(iv) reflects the fact that needs are increasing in m (for all incomes) and requires that the ranking of equivalent incomes as a function of type does not depend on the level of household income. E(v) can be justified by its consequences: It implies that for any level of equivalent income X^* and any type m there is a feasible income X such that $X^* = E_m(X)$. Since this is true, arbitrary living standards (of any type) can always be compared. In particular it allows us to choose any type as reference type (which is important as there are no a priori grounds for choosing a particular type). Formally, E(iii) and E(v) imply that an equivalent income function is invertible.

In this framework relative and absolute equivalence scales are represented by an equivalent income function having the form

$$E_m(X) = X/a_m \text{ and } E_m(X) = X - b_m,$$

respectively, for $m = 1, \ldots, n_H$, where $0 < a_m < a_{m+1}$, $b_m < b_{m+1}$ for $m = 1, \ldots, n_H - 1$ and $a_r = 1$, $b_r = 0$. But, of course, the definition of E is

[10] As a referee suggests, the reference type could also be fictive and be introduced artificially by the social decision maker according to her preferences.
[11] See Ebert (2000b) for a discussion of equivalizing procedures/equivalent income functions.

more general. E transforms a heterogeneous income distribution $(\mathbf{X}, \mathbf{m})$ into a vector of equivalent income

$$\mathbf{X}^* = \left(X_1^*, \ldots, X_n^*\right) := E(\mathbf{X}, \mathbf{m}) := \left(E_{m_1}(X_1), \ldots, E_{m_n}(X_n)\right).$$

$\mathbf{X}^*$ reflects the distribution of equivalent income and therefore the distribution of living standard for a hypothetical population of type r households. Thus the differing needs of different types have been taken into account in the derivation of $\mathbf{X}^*$, but $\mathbf{X}^*$ itself does no longer provide any information about types. Therefore $\mathbf{X}^*$ has to be supplemented by additional information. In practice it is assumed that the type of a household is described by an appropriate weight. Thus we introduce a weighting function $w : H \to \mathbb{R}_{++}$, which assigns a weight $w(m)$ to each type $m \in H$. For instance $w(m)$ could be equal to the number of persons in a type m household or to the number of equivalent persons or to any other weight depending on the size, the composition, and the needs of a type m household. Of course, in general there will be a relationship between the choice of an equivalent income function E and a weighting function $w(m)$. But they might also be chosen separately or independently. Anyway they have to conform with and reflect the decision maker's value judgements.

Summing up, a heterogeneous income distribution $(\mathbf{X}, \mathbf{m})$ will be transformed to an adjusted (weighted) homogeneous distribution (of equivalent income) $(\mathbf{X}^*, \mathbf{w}^*) := (E(\mathbf{X}, \mathbf{m}), w(\mathbf{m}))$, where $\mathbf{w}^* = \left(w_1^*, \ldots, w_n^*\right) := w(\mathbf{m}) = (w(m_1), \ldots, w(m_n))$. Household i then corresponds to group i in the homogeneous framework. Each member of the group receives the equivalent income, i.e., $Y_i = X_i^* := E(X_i, m_i)$. The group's weight w_i is given by the household's weight $w(m_i)$.

Now suppose that the social decision maker possesses a social welfare ordering $\succsim_V$ defined on $\Omega^n \times \{w(\mathbf{m})\}$, i.e., on the set of adjusted homogeneous distributions of equivalent income for a given population of households. As above the population is kept fixed. It is represented by the type profile $\mathbf{m}$. Then it is obvious how to derive a social welfare ordering $\succsim_W$ for heterogeneous income distributions. We define

$$(\mathbf{X}, \mathbf{m}) \succsim_W (\mathbf{X}', \mathbf{m}) :\Leftrightarrow (E(\mathbf{X}, \mathbf{m}), w(\mathbf{m})) \succsim_V (E(\mathbf{X}', \mathbf{m}), w(\mathbf{m})) \tag{8}$$

for $(\mathbf{X}, \mathbf{m})$, $(\mathbf{X}', \mathbf{m}) \in \Omega^n \times \{\mathbf{m}\}$. Its symmetric part is denoted by $\sim_W$, its asymmetric one by $\succ_W$. The ordering $\succsim_W$ depends on the choice of the reference type r, on the equivalent income function E, the weighting function $w(m)$, and, of course, on the ordering $\succsim_V$. Such an ordering $\succsim_V$ has been examined in Sect. 2.

4 Social welfare for a heterogeneous population

In this section we now investigate a heterogeneous population and suppose that a social welfare ordering $\succsim_W$ is given; it is based on an ordering $\succsim_V$ satisfying WELFARE, a reference type r, an equivalent income function E, and a weighting function $w(m)$. $\succsim_W$ is represented by $\xi(\mathbf{X}, \mathbf{m})$, where $\xi(\mathbf{X}, \mathbf{m})$

represents the living standard (level of equivalent income) which, if enjoyed by each household in society, yields the same level of social welfare as $(\mathbf{X}^*, \mathbf{w}^*)$:

$$(\xi(\mathbf{X}, \mathbf{m})\mathbf{1}_n, \mathbf{w}^*) \sim_V (\mathbf{X}^*, \mathbf{w}^*).$$

Since in general needs differ, the corresponding household incomes depend on types. Set $\xi^{m_i} := E_{m_i}^{-1}(\xi(\mathbf{X}, \mathbf{m}))$, where $E_m^{-1}(\cdot)$ is the inverse of $E_m(\cdot)$, which is defined on Ω. Then we obtain

$$((\xi^{m_1}, \ldots, \xi^{m_n}), \mathbf{m}) \sim_W (\mathbf{X}, \mathbf{m});$$

i.e., the heterogeneous income distribution $((\xi^{m_1}, \ldots, \xi^{m_n}), \mathbf{m})$ leads to the same living standard for all households and yields the same level of social welfare as $(\mathbf{X}, \mathbf{m})$.

4.1 Principle of progressive transfers

Leaving aside the principle of population for a moment, we want to examine the redistribution of income. It should be clear that it is not possible to redistribute equivalent income. Thus we have to consider household income and the ordering $\succsim_W$, since it is defined on income vectors, whereas $\succsim_V$ is defined on vectors of equivalent income (whenever we examine a heterogeneous population). We introduce

Definition. $(\mathbf{X}, \mathbf{m})$ *is obtained from* $(\mathbf{X}', \mathbf{m})$ *by a progressive transfer, if a small amount of income is redistributed from a richer to a poorer household – measured by equivalent income – without changing the ranking of equivalent income, i.e., if there are* $i, j \in N$ *and* $\varepsilon > 0$ *such that* $X_h = X'_h$ *for* $h \neq i$, $h \neq j$, $X_i = X'_i - \varepsilon$, $X_j = X'_j + \varepsilon$, $E_{m_i}(X_i) \geq E_{m_i}(X_j)$, *and* $\pi(\mathbf{X}^*) = \pi(\mathbf{X}'^*)$.

Obviously the household type is not directly relevant, when income is redistributed among households: A(n order-preserving) progressive transfer is based on a comparison of equivalent income. It diminishes the difference between the equivalent incomes of household i and j, i.e., between the respective living standards. But household income is redistributed. Using the definition we formulate the

Principle of progressive transfers. *Whenever* $(\mathbf{X}, \mathbf{m})$ *is derived from* $(\mathbf{X}', \mathbf{m})$ *by a progressive transfer,* $(\mathbf{X}, \mathbf{m})$ *is weakly preferred to* $(\mathbf{X}', \mathbf{m})$: $(\mathbf{X}, \mathbf{m}) \succsim_W (\mathbf{X}', \mathbf{m})$.

Given the representation of $\succsim_W$ we expect some implications of this property for the transformation function f and/or utility function U. For a homogeneous population, in which $w_i = 1$ for $i = 1, \ldots, n$, one can prove that the decision weights α_i are nondecreasing in rank i (f is convex) and that the utility function U has to be concave (see Proposition 6 in Ebert 1988 when homotheticity or translatability is imposed). Moreover, it is well known that there is a connection between the principle of progressive transfers for social welfare orderings and (strong) risk aversion within the rank-dependent expected utility model. Risk aversion implies convexity of f and concavity of U (see Cohen 1995 for a survey and also Chateauneuf and Cohen 1994). Therefore we expect similar results in our framework. But neither property is necessary in our model as can be demonstrated by examples.

Thus the principle of progressive transfers seems to be weaker in the heterogeneous framework than it is in the homogeneous one. There are two reasons for this outcome. First, since the weights attached to types can be different in general, the domain of the transformation function does not consist of equidistant numbers. It is therefore not always possible to impose a condition on the values of the transformation function at neighbouring points of the domain by means of a progressive transfer. Then only monotonicity, but neither concavity of the utility function nor convexity of the transformation function is implied. Second, the equivalent income function E makes the model more complicated, but yields further degrees of freedom allowing the utility function to be convex.

These outcomes are driven by the choice of a finite and fixed population (characterized by $\mathbf{m}$). When also replications are admitted, the picture changes as the next subsection shows.

4.2 Main result

In the following we will impose the principle of progressive transfers and the principle of population simultaneously. (It is obvious how to define the principle of progressive transfers for a replicated population.) For every welfare ordering $\succsim_{V,k}$, $k \geq 1$, we define the corresponding welfare ordering $\succsim_{W,k}$ by

$$\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \succsim_{W,k} \left(\mathbf{X}', \mathbf{m}^{(k)}\right) \;:\Leftrightarrow$$
$$\left(E\left(\mathbf{X}, \mathbf{m}^{(k)}\right), w\left(\mathbf{m}^{(k)}\right)\right) \succsim_{V,k} \left(E\left(\mathbf{X}', \mathbf{m}^{(k)}\right), w\left(\mathbf{m}^{(k)}\right)\right)$$

for $\mathbf{X}, \mathbf{X}' \in \Omega^{kn}$ where $\mathbf{m}^{(k)}$ denotes the k-fold replication of $\mathbf{m}$.

Now we establish

Theorem 2. *Assume that the welfare orderings $\succsim_{V,k}$, $k \geq 1$, satisfy WELFARE and POPULATION and are represented by (f, U), where f is unique and U is cardinal. The welfare orderings $\succsim_{W,k}$, $k \geq 1$, are derived from $\succsim_{V,k}$ by means of the equivalent income function E, a reference type r, and a weighting function $w(m)$.*

Then the following statements are equivalent:
(a) $\succsim_{W,k}$ satisfies TRANSFER for $k \geq 1$.
(b) (i) f is continuous on $D \cap [0, 1)$ and convex[12] *on D.*

[12] f is called convex on D if and only if $f(\lambda w + (1-\lambda)\bar{w}) \leq \lambda f(w) + (1-\lambda) f(\bar{w})$ for all $w, \bar{w}, \lambda w + (1-\lambda)\bar{w} \in D, \lambda \in [0, 1]$ (cf. Footnote 17 in Donaldson and Weymark 1980).

(ii) $U(X) = U(E_r(X))$ *and* $U(E_m(X))$ *are concave in* $X \in \Omega$ *for* $m = 1, \ldots, n_H,\ m \neq r$.

(iii) *For* $\Omega = \mathbb{R}_{++}$: $E_m(X) = \frac{w(r)}{w(m)} X$ *and* $w(m) < w(m+1)$ *for* $m = 1, \ldots,\ n_H - 1$; *for* $\Omega = \mathbb{R}$: *the weights are identical* $(w(m) = w(r),\ m = 1, \ldots, n_H)$ *and* $E_m(X) = X - b_m$ *for* $m = 1, \ldots, n_H$ *where* $b_m < b_{m+1}$ *for* $m = 1, \ldots, n_H - 1$ *and* $b_r = 0$.

The combination of the principles of progressive transfers and of population yields the properties expected: The transformation function $f(w)$ has to be convex, and the utility functions $U(X)$ and $U(E_m(X))$ have to be concave. It is interesting to note that these results do not require any explicit differentiability assumption as e.g., employed in Chew et al. (1987) for the RDEU model. Convexity of the transformation function implies that – given an income distribution – the importance attributed to lower incomes is greater than the importance attributed to higher incomes. Similarly, the concavity of the utility function guarantees that a given increase in income is more appreciated for lower incomes.

Furthermore the functions $E_m(X)$ have to be proportional and translating functions, respectively, depending on the domain Ω. If incomes are restricted to be strictly positive, one has to apply relative equivalence scales being proportional to the corresponding weights; i.e., equivalence scales and weights can not be chosen independently. If arbitrary incomes are feasible, weights have to be equal. The respective functions $E_m(X)$ only translate incomes. In the situation considered there is not much room left for choosing an equivalent income function. $E_m(\cdot)$ must be affine, a priori. But linearity guarantees that the utility functions $U(E_m(X))$ possess the same degree of concavity (since it is not changed by an affine transformation).

Theorem 2 presents the main result of this paper. It characterizes the social welfare orderings possessing a small set of properties necessary for the measurement of inequality and poverty. At the same time it demonstrates that there are severe limitations to the choice of weights and equivalent income functions in this framework. The same restrictions are implied in Ebert and Moyes (2003) in which (generalized) Lorenz dominance is investigated for heterogeneous populations.

At this point the reader is reminded that the class of welfare orderings defined on rank-ordered income profiles also contains the class of orderings in which the rank of income does not play a role. In other words the result of Theorem 2 is more general than it looks like at first sight. Choose the transformation function $f(w) = w$ for $w \in D$. Then we obtain

$$\alpha_i\big(\mathbf{w}_{\downarrow}(\mathbf{X}^*)\big) = w_{[i]}(\mathbf{X}^*)/w(N) \text{ for } i = 1, \ldots, n.$$

The weights are no longer 'distorted' by f. Depending on the domain the social welfare ordering $\succsim_W$ characterized in Theorem 2 is represented by

$$\xi(\mathbf{X}, \mathbf{m}) = U^{-1}\left(\sum_{i=1}^{n} \frac{w(m_i)}{\Sigma\, w(m_k)}\, U\left(\frac{X_i}{w(m_i)}\right)\right) \text{ if } \Omega = \mathbb{R}_{++} \text{ and } w(r) = 1$$

and by

$$\xi(\mathbf{X}, \mathbf{m}) = U^{-1}\left(\frac{1}{n}\sum_{i=1}^{n} U\left(X_i + b_{m_i}\right)\right) \text{ if } \Omega = \mathbb{R} \text{ with } b_r = 0.$$

In the first case the decision weights are equal to the respective relative equivalence scales, which are appropriately normalized, in the second one all weights are equal. The form of these welfare functions is well known (see Ebert 1997a, 1997b, 1999). Obviously this kind of weighting is related to the generalization of the principle of progressive transfers (TRANSFER).

So far no properties like homotheticity or translatability[13] have been used. Thus we have to conclude that for the ethical assessment of heterogeneous income distributions we always have to employ relative or absolute equivalence scales and the specific weights discussed above. There is no escape from this insight, whenever the assumptions discussed above are made. In this sense Theorem 2 provides an impossibility result. If one wants to use social welfare orderings whose representation is different from that characterized in Theorem 2, one has to dispense with at least one of the properties imposed.

5 Measurement of inequality and poverty

5.1 Inequality

Social welfare orderings can be employed to define corresponding (ethical) inequality orderings. The basic idea[14] dates back to Atkinson (1970), Kolm (1969), and Sen (1973): they propose to determine and to use the welfare loss due to inequality. For a homogeneous population it is measured by the difference between average income and the equally distributed equivalent income (EDEI) of an income distribution, where, of course, the EDEI depends on the welfare ordering. Accordingly this idea has to be translated to our framework.

Suppose for the rest of this section that assumption WELFARE is satisfied. From the beginning we consider heterogeneous income distributions. However, since household incomes are not directly comparable, the evaluation of welfare has to be based on equivalent incomes. For an income distribution $\left(\mathbf{X}, \mathbf{m}^{(\mathbf{k})}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ we define

[13] A function W is homothetic if $W(\lambda\mathbf{X}) = \lambda W(\mathbf{X})$ and translatable if $W(\mathbf{X} + \kappa\mathbf{1}) = W(\mathbf{X}) + \kappa$ for $\mathbf{X} \in \Omega^n, \lambda > 0, \kappa \in \mathbb{R}$ and $\lambda\mathbf{X} \in \Omega^n$, $\mathbf{X} + \kappa\mathbf{1} \in \Omega^n$.

[14] The general relation between welfare and inequality is examined in Ebert (1987).

$$L\left(\mathbf{X},\mathbf{m}^{(k)}\right) := \mu\left(\mathbf{X}^*, w\left(\mathbf{m}^{(k)}\right)\right) - \xi_k\left(\mathbf{X}^*, w\left(\mathbf{m}^{(k)}\right)\right),$$

where

$$\mu\left(\mathbf{X}^*, w\left(\mathbf{m}^{(k)}\right)\right) = \sum_{i=1}^{kn} \frac{w\left(m_i^{(k)}\right)}{w(N_k)} X_i^*$$

denotes the weighted average (equivalent) income and ξ_k the EDEI of $(\mathbf{X}^*, \mathbf{w}^*)$.

L reflects the (representative) welfare loss per type r household (measured in equivalent income) due to the inequality of equivalent incomes, which represent the living standard attained by households. The different sizes and needs of households are taken into account by the (implicitly used) equivalent income function and the respective weights.

We introduce for $\left(\mathbf{X},\mathbf{m}^{(k)}\right)$, $\left(\mathbf{X}',\mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ inequality orderings by

$$\left(\mathbf{X},\mathbf{m}^{(k)}\right) \succsim_{I,k}^{a} \left(\mathbf{X}',\mathbf{m}^{(k)}\right) \;:\Leftrightarrow\; L\left(\mathbf{X},\mathbf{m}^{(k)}\right) \leq L\left(\mathbf{X}',\mathbf{m}^{(k)}\right), \text{ if } \Omega = \mathbb{R},$$

and

$$\left(\mathbf{X},\mathbf{m}^{(k)}\right) \succsim_{I,k}^{r} \left(\mathbf{X}',\mathbf{m}^{(k)}\right) \;:\Leftrightarrow\; L\left(\mathbf{X},\mathbf{m}^{(k)}\right) / \mu\left(\mathbf{X}^*, w\left(\mathbf{m}^{(k)}\right)\right)$$
$$\leq L\left(\mathbf{X}',\mathbf{m}^{(k)}\right) / \mu\left(\mathbf{X}'^*, w\left(\mathbf{m}^{(k)}\right)\right), \quad \text{if } \Omega = \mathbb{R}_{++}.$$

The definition of these measures is reasonable as the following proposition demonstrates.

Proposition 3. *Assume that $\succsim_{V,k}$, $k \geq 1$, satisfy WELFARE and POPULATION.*

The welfare orderings $\succsim_{W,k}$, $k \geq 1$, satisfy TRANSFER if and only if $L\left(\mathbf{X},\mathbf{m}^{(k)}\right)$ is nonnegative for all $\left(\mathbf{X},\mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ and weakly decreases for every progressive transfer, $k \geq 1$.

Thus the welfare loss is well defined and possesses the basic properties expected: it is always nonnegative and is weakly diminished by a rank-preserving progressive transfer of income from a richer to a poorer household.

In general one is interested in relative or absolute inequality measures. They are defined by

$$I(\lambda\mathbf{X},\mathbf{m}) = I(\mathbf{X},\mathbf{m}) \;\text{ for } \lambda > 0, \; \mathbf{X} \in \Omega^n, \; \lambda\mathbf{X} \in \Omega^n$$

and

$$I(\mathbf{X} + \kappa\mathbf{1}_n, \mathbf{m}) = I(\mathbf{X},\mathbf{m}) \text{ for all } \kappa \in \mathbb{R}, \; \mathbf{X} \in \Omega^n, \; \mathbf{X} + \kappa\mathbf{1}_n \in \Omega^n,$$

respectively. They can be generated in this framework, if the welfare orderings possess further properties. We define for the social welfare ordering $\succsim_V$.

Scale invariance. *For all* $\mathbf{Y}, \mathbf{Y}' \in \mathbb{R}^n$: $(\mathbf{Y}, \mathbf{w}) \sim_V (\mathbf{Y}', \mathbf{w})$ *implies* $(\lambda\mathbf{Y}, \mathbf{w}) \sim_V (\lambda\mathbf{Y}', \mathbf{w})$ *for all* $\lambda \in \Omega = \mathbb{R}_{++}$.

Translation invariance. *For all* $\mathbf{Y}, \mathbf{Y}' \in \Omega^n$:$(\mathbf{Y}, \mathbf{w}) \sim_V (\mathbf{Y}', \mathbf{w})$ *implies* $(\mathbf{Y}+\kappa\mathbf{1}_n, \mathbf{w}) \sim_V (\mathbf{Y}' + \kappa\mathbf{1}_n, \mathbf{w})$ *for all* $\kappa \in \mathbb{R}$ *if* $\mathbf{Y} + \kappa\mathbf{1}_n \in \Omega^n$ *and* $\mathbf{Y}' + \kappa\mathbf{1}_n \in \Omega^n$.

The properties consider specific modifications of income distributions being equivalent. For SCALE invariance the equivalence is preserved if all incomes are changed in the same proportion (they are multiplied by the same factor). TRANSLATION invariance requires that the simultaneous addition of the same amount κ to all incomes does not affect the equivalence in terms of social welfare.

Since we are now primarily interested in application, we will concentrate on deriving inequality measures and will not spell out the corresponding properties of the inequality orderings explicitly. We are able to establish

Proposition 4. *Assume that* $\succsim_{V,k}$, $k \geq 1$, *satisfy WELFARE and POPULATION and that the orderings* $\succsim_{W,k}$, $k \geq 1$, *satisfy TRANSFER.*

(a) $\succsim_V \equiv \succsim_{V,1}$, *defined on* $\mathbb{R}^n_{++} \times \{w(\mathbf{m})\}$, *satisfies SCALE invariance if and only if there is* $\varepsilon \leq 1$ *such that* $\succsim^r_{I,k}$ *is represented by*

$$I^\varepsilon_r\left(\mathbf{X}, \mathbf{m}^{(k)}\right) = \begin{cases} 1 - \left(\sum\limits_{i=1}^{kn} \alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right) \left(\frac{X^*_{[i]}}{\mu(\mathbf{X}^*,\mathbf{w}^*)}\right)^\varepsilon\right)^{1/\varepsilon} & \text{for } \varepsilon \leq 1,\ \varepsilon \neq 0 \\ 1 - \prod\limits_{i=1}^{kn} \left(\frac{X^*_{[i]}}{\mu(\mathbf{X}^*,\mathbf{w}^*)}\right)^{\alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right)} & \text{for } \varepsilon = 0 \end{cases}$$

for all $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \mathbb{R}^{kn}_{++} \times \left\{\mathbf{m}^{(k)}\right\}$ *and* $k \geq 1$.
Then I^ε_r *is a relative inequality measure, since* $X^* = E_m(X) = \frac{w(r)}{w(m)}X$.

(b) $\succsim_V \equiv \succsim_{V,1}$, *defined on* $\mathbb{R}^n_{++} \times \{w(\mathbf{m})\}$, *satisfies TRANSLATION invariance if and only if there is* $\gamma \leq 0$ *such that* $\succsim^a_{I,k}$ *is represented by*

$$I^\gamma_a\left(\mathbf{X}, \mathbf{m}^{(k)}\right) = \begin{cases} \mu(\mathbf{X}^*, \mathbf{w}^*) - \frac{1}{\gamma}\ln\left(\sum\limits_{i=1}^{kn} \alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right) e^{\gamma X^*_{[i]}}\right) & \text{for } \gamma < 0 \\ \mu(\mathbf{X}^*, \mathbf{w}^*) - \sum\limits_{i=1}^{kn} \alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right) X^*_{[i]} & \text{for } \gamma = 0 \end{cases}$$

for all $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \mathbb{R}^{kn}_{++} \times \left\{\mathbf{m}^{(k)}\right\}$ *and* $k \geq 1$.
Then I^γ_a *is an absolute inequality measure, since* $X^* = E_m(X) = X - b_m$.

(c) $\succsim_V \equiv \succsim_{V,1}$, *defined on* $\mathbb{R}^n_{++} \times \{w(\mathbf{m})\}$, *satisfies SCALE and TRANSLATION invariance if and only if* $\succsim^r_{I,k}$ *is represented by*

$$I^l_r\left(\mathbf{X}, \mathbf{m}^{(k)}\right) = 1 - \sum_{i=1}^{n} \alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right) X^*_{[i]} / \mu(\mathbf{X}^*, \mathbf{w}^*)$$

for all $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \mathbb{R}^{kn}_{++} \times \left\{\mathbf{m}^{(k)}\right\}$ *and* $k \geq 1$.
Then I^l_r *is a relative inequality measure, since* $X^* = E_m(X) = \frac{w(r)}{w(m)}X$.

In summary, we obtain rank-dependent relative and absolute inequality measures, a generalization of Atkinson-, Kolm-Pollak, and linear measures to the heterogeneous framework.[15] In addition to assumption WELFARE two factors determine this outcome: The invariance property imposed on the ordering $\succsim_V$ is, of course, necessary. But the form of the equivalent income function is important, as well. The property TRANSFER implies that the latter is linear. Furthermore, it has to fit to the domain. Therefore for SCALE invariance (part (a) and (c)) the equivalent income function is proportional. Furthermore, the (relative) equivalence scales have to be proportional to the respective weights. Similarly, for TRANSLATION invariance incomes have to be changed by an absolute equivalence scale. All weights have to be the same in this case (cf. Theorem 2).

In the framework considered an analogue to the Gini coefficient can be derived, too (see also Pyatt 1985). We present

Example 2. For a homogeneous population with identical individuals (if $w_i \equiv 1$ for $i = 1, \ldots, n$) the Gini coefficient is given[16] by

$$G(\mathbf{Y}, \mathbf{1}_n) = 1 - \sum_{i=1}^{n} \frac{2i-1}{n^2} \, \frac{Y_{[i]}}{\mu(\mathbf{Y}, \mathbf{1}_n)}.$$

It is well known (and can be seen directly) that $G(\mathbf{Y}, \mathbf{1}_n)$ is based on the linear rank-dependent social welfare function

$$V(\mathbf{Y}, \mathbf{1}_n) = \sum_{i=1}^{n} \frac{2i-1}{n^2} Y_{[i]} = \sum_{i=1}^{n} \frac{i^2 - (i-1)^2}{n^2} Y_{[i]}$$

(see Donaldson and Weymark 1980). It can be generalized to a homogeneous population of (not necessarily) identical groups by characterizing a general rank-dependent welfare function by the utility function $U(Y) = Y$ and the transformation function $f(w) = w^2$ (cf. Ebert 2003). Then we obtain

$$V(\mathbf{Y}, \mathbf{w}) = \sum_{i=1}^{n} \frac{\left(\sum_{j=1}^{i-1} w_{[j]} + w_{[i]}\right)^2 - \left(\sum_{j=1}^{i-1} w_{[j]}\right)^2}{w(N)^2} Y_{[i]} = \sum_{i=1}^{n} \frac{w_{[i]} \left(2 \sum_{j=1}^{i} w_{[j]} - w_{[i]}\right)}{w(N)^2} Y_{[i]}$$

and the corresponding Gini coefficient

$$G(\mathbf{X}, \mathbf{m}) = 1 - \sum_{i=1}^{n} \frac{w_{[i]} \left(2 \sum_{j=1}^{i} w_{[j]} - w_{[i]}\right)}{w(N)^2} \, \frac{X^*_{[i]}}{\mu(\mathbf{X}^*, \mathbf{w}^*)}$$

[15] Proposition 4 can be interpreted as a generalization of results derived in Ebert (1988) for homogeneous populations and in Ebert (1995,1997a) for heterogeneous populations.

[16] We confine ourselves to $k = 1$.

for a heterogeneous population. As the ordinary Gini index, $G(\mathbf{X}, \mathbf{m})$ is equal to the area enclosed between the Lorenz curve defined for $(\mathbf{X}^*, \mathbf{w}^*)$ (see Ebert 1999) and the egalitarian line, multiplied by 2. ■

5.2 Poverty

There are several approaches of deriving poverty measures from a social welfare function or ordering (see e.g. Chakravarty 1990). We employ a procedure first proposed by Clark, Hemming, and Ulph (1981) and generalized by Chakravarty (1983). Since household incomes cannot be compared directly, and the living standard attained is the relevant variable, poverty is (also) measured in terms of equivalent income – like inequality above. Let $Z^* > 0$ denote the living standard which separates poor households from nonpoor. Z^* is chosen exogenously and is interpreted as poverty line for the reference type.

It allows us to define a censored income profile in which all equivalent incomes above the poverty line are replaced by Z^*. More precisely, for every $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ we define

$$\text{the income profile } \mathbf{X}^p := \left(X_1^p, \ldots, X_{kn}^p\right) \text{ by } X_i^p := \begin{cases} X_i & \text{if } X_i^* \leq Z^* \\ E^{-1}_{m_i^{(k)}}(Z^*) & \text{if } X_i^* > Z^* \end{cases}.$$

Then we obtain the censored heterogeneous income distribution $\left(\mathbf{X}^p, \mathbf{m}^{(k)}\right)$ leading to the censored distribution of equivalent income

$$\mathbf{X}^{p*} = \left(\min\{X_1^*, Z^*\}, \ldots, \min\{X_{kn}^*, Z^*\}\right).$$

Given the welfare orderings $\succsim_{W,k}$, $k \geq 1$, we introduce the poverty measure

$$P\left(\mathbf{X}, \mathbf{m}^{(k)}; Z^*\right) := \frac{Z^* - \xi_k\left(\mathbf{X}^{p*}, w\left(\mathbf{m}^{(k)}\right)\right)}{Z^*}$$

for every $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ and $k \geq 1$. It measures the ratio between the (representative) poverty gap per type r household, $Z^* - \xi_k$, and the poverty line Z^*. The corresponding poverty orderings $\succsim_{P,k}, k \geq 1$, are defined by

$$\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \succsim_{P,k} \left(\mathbf{X}', \mathbf{m}^{(k)}\right) :\Leftrightarrow P\left(\mathbf{X}, \mathbf{m}^{(k)}; Z^*\right) \leq P\left(\mathbf{X}', \mathbf{m}^{(k)}; Z^*\right)$$

for all $\left(\mathbf{X}, \mathbf{m}^{(k)}\right), \left(\mathbf{X}', \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$. Both concepts are distribution-sensitive, as Sen (1976) postulates; we get an analogue to Proposition 3:

Corollary 5. *Assume that* $\succsim_{V,k}$, $k \geq 1$, *satisfy WELFARE and POPULATION.*

The poverty orderings $\succsim_{P,k}, k \geq 1$, *satisfy TRANSFER if and only if* $P\left(\mathbf{X}, \mathbf{m}^{(k)}; Z^*\right)$ *is nonnegative for all* $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ *and weakly decreases for every progressive transfer among the poor,* $k \geq 1$.

Obviously, a poverty measure is well-defined, and a progressive transfer weakly diminishes poverty. In practice often relative poverty measures ($P(\lambda \mathbf{X}, \mathbf{m}; \lambda Z^*) = P(\mathbf{X}, \mathbf{m}; Z^*)$ for $\lambda > 0$ and $\Omega = \mathbb{R}_{++}$) are in use: They do not change, if all incomes and the poverty line are changed in the same proportion[17]. These orderings are characterized by

Proposition 6. *Assume that $\succsim_{V,k}$, $k \geq 1$, satisfy WELFARE and POPULATION, that the orderings $\succsim_{W,k}$, $k \geq 1$, satisfy TRANSFER, and that $\Omega = \mathbb{R}_{++}$.*

$\succsim_{V,k}$, $k \geq 1$, satisfies SCALE invariance if and only if there is $\varepsilon \leq 1$ such that $\succsim_{P,k}$ is represented by

$$P^{\varepsilon}\left(\mathbf{X}, \mathbf{m}^{(k)}; Z^*\right) = \begin{cases} 1 - \left(\sum_{i=1}^{kn} \alpha_i\left(\mathbf{w}_{\downarrow}(\mathbf{X}^{p*})\right)\left(\frac{X_{[i]}^{p*}}{Z^*}\right)^{\varepsilon}\right)^{1/\varepsilon} & \text{for } \varepsilon \leq 1,\ \varepsilon \neq 0 \\ 1 - \prod_{i=1}^{kn}\left(\frac{X_{[i]}^{p*}}{Z^*}\right)^{\alpha_i\left(\mathbf{w}_{\downarrow}(\mathbf{X}^{p*})\right)} & \text{for } \varepsilon = 0 \end{cases}$$

for all $\left(\mathbf{X}, \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ and $k \geq 1$.

Then P^{ε} is a relative poverty measure, since $X^* = E_m(X) = \frac{w(r)}{w(m)} X$.

These measures are a generalization of the poverty indices derived by Clark et al. (1981).

There is a connection between the poverty measures P^{ε} and the relative inequality measures I_r^{ε} described in Proposition 4:

$$P^{\varepsilon}\left(\mathbf{X}, \mathbf{m}^{(k)}; Z^*\right) = 1 - \mu\left(\mathbf{X}^{p*}, \mathbf{w}^*\right)\left(1 - I_r^{\varepsilon}\left(\mathbf{X}^{p*}, \mathbf{m}^{(k)}\right)\right)/Z^*.$$

In other words, the measure P^{ε} is related to the inequality among the poor measured by the inequality index I_r^{ε}, which in turn is evaluated for the respective censored income distribution. This proves again that these measures are distribution-sensitive.

Example 3. If $\varepsilon = 1$ the class of poverty measures characterized contains those measures which are based on linear social welfare functions (and orderings). Using the Gini social welfare ordering we obtain in particular

$$P^G(\mathbf{X}, \mathbf{m}; Z^*) = 1 - \sum_{i=1}^{n} \frac{w_{[i]}\left(2\sum_{j=1}^{i} w_{[j]} - w_{[i]}\right)}{w(N)^2} \frac{X_{[i]}^{P*}}{Z^*} \text{ for } (\mathbf{X}, \mathbf{m}) \in \mathbb{R}_{++}^n \times \{\mathbf{m}\}.$$

■

The analysis of this section demonstrates that most inequality orderings and some poverty orderings well-known for homogeneous populations can be extended to take into account differences in needs.

[17] Similarly it is possible to define absolute poverty measures and to derive an analogue to Proposition 6.

6 Conclusion

The paper has determined the possibilities and limitations of defining a social welfare ordering for a heterogeneous population, when a two-step procedure is applied: At first the income distribution is adjusted; i.e., incomes are equivalized and weights reflecting needs are attached. Then, secondly, the adjusted income distribution of equivalent income is interpreted as an income distribution of a hypothetical homogeneous population, and a usual welfare ordering is applied. Therefore, we introduced the concept of an equivalent income function, which permits a lot of different ways of computing equivalent income (see Ebert 2000b). Furthermore, weights were determined by a (general) weighting function. Only one population was considered in this framework. Finally, the basic properties postulated for social welfare orderings were few: An ordering has to be continuous and increasing in income, has to satisfy an independence condition, and is defined on rank-ordered income vectors. Additionally the principles of population and of progressive transfers have been imposed. They guarantee that welfare depends only on the distribution (in the statistical sense) and that a redistribution of income from richer to poorer households weakly increases welfare. These principles seem to be necessary for the measurement of social welfare.

It turned out that severe restrictions on the adjustment process are implied, depending on the domain chosen. If incomes have to be strictly positive, one has to employ relative equivalence scales a priori. Then one can choose either the set of weights for the different household types (the weighting function) or the equivalence scales, since weights and equivalence scales have to be proportional. The factor of proportionality does not play a role. If arbitrary incomes are feasible, all weights have to be the same (and independent of household type). Again the magnitude of the common weight has no implications. Moreover absolute equivalence scales have to be employed: incomes have to be translated by an amount depending on household type. Thus in both cases not much room is left for the adjustment process, although a priori any sensible adjustment process (i.e. equivalent income function and weighting function) is permitted. The implications of the properties imposed are strong. It must be stressed that these limitations hold generally, not only for orderings defined on rank-ordered vectors, since welfare orderings for which the ranking of incomes does not play a role are specific forms of the former ones.

The welfare orderings derived were applied to the measurement of inequality and poverty. Then we obtained generalizations of Atkinson, Kolm-Pollak, and linear inequality measures and of the corresponding classes of poverty measures, respectively, to heterogeneous populations, if the invariance properties SCALE and TRANSLATION are imposed separately or jointly. Proceeding in this way allows us to reveal the value judgements underlying these indicators.

Appendix

Proof of Proposition 1

Consider $k = 1$:

$$(\mathbf{Y}, \mathbf{w}) \sim_{V,1} (\xi_1(\mathbf{Y}, \mathbf{w})\, \mathbf{1}_n, \mathbf{w}).$$

The principle of POPULATION and the definition of the EDEI imply

$$\left(\mathbf{Y}^{(k)}, \mathbf{w}^{(k)}\right) \sim_{V,k} \left(\xi_1(\mathbf{Y}, \mathbf{w})\mathbf{1}_{kn}, \mathbf{w}^{(k)}\right) \sim_{V,k} \left(\xi_k\left(\mathbf{Y}^{(k)}, \mathbf{w}^{(k)}\right)\mathbf{1}_{kn}, \mathbf{w}^{(k)}\right).$$

Thus

$$\xi_1(\mathbf{Y}, \mathbf{w}) = U_1^{-1}\left(\sum_{i=1}^{n} \alpha_i^1\left(\mathbf{w}_\downarrow(\mathbf{Y})\right) U_1\left(Y_{[i]}\right)\right)$$
$$= U_k^{-1}\left(\sum_{i=1}^{n}\left[\sum_{j=(i-1)k+1}^{ik} \alpha_j^k\left(\mathbf{w}_\downarrow\left(\mathbf{Y}^{(k)}\right)\right)\right] U_k\left(Y_{[i]}^k\right)\right) = \xi_k\left(\mathbf{Y}^{(k)}, \mathbf{w}^{(k)}\right).$$

Applying Theorem 2 of Aczel (1966), p. 290, we obtain

$$\alpha_i^1\left(\mathbf{w}_\downarrow(\mathbf{Y})\right) = \sum_{j=(i-1)k+1}^{ik} \alpha_j^k(\mathbf{w}(\mathbf{Y}))$$

and $U_k(Y) = c\, U_1(Y) + d$, where $c > 0$ because of monotonicity of U_1 and U_k. We have

$$\alpha_1^1\left(\mathbf{w}_\downarrow(\mathbf{Y})\right) = f_1\left(w_{[1]}(\mathbf{Y})/w(N)\right) - f_1(0) = f_1\left(w_{[1]}(\mathbf{Y})/w(N)\right)$$
$$= \sum_{j=1}^{k} \alpha_j^k\left(\mathbf{w}_\downarrow\left(\mathbf{Y}^{(k)}\right)\right) = f_k\left(kw_{[1]}(\mathbf{Y}^{(k)})/(kw(N))\right) - f_k(0) = f_k\left(w_{[1]}\left(\mathbf{Y}^{(k)}\right)/w(N)\right).$$

Therefore by induction $f_1(w) = f_k(w)$ for $w \in D(\mathbf{w})$, and, analogously, $f_k(w) = f_{kl}(w)$ for $w \in D\left(\mathbf{w}^{(k)}\right)$, $w \neq 0$ and $k, l > 1$.

Thus the function $f : \bigcup_{k \geq 1} D_k \to [0, 1]$, defined by $f(w) := f_k(w)$ for $w \in D$ and $k \geq 1$, is well-defined. It is strictly increasing and $f(0) = 0, f(1) = 1$ by construction.

The converse is obvious. ∎

Proof of Theorem 2

"(a) $\Rightarrow$ (b)":

Outline of the proof: The basic idea is to construct appropriate progressive transfers for the population N_k and to consider their implications when k tends to infinity. Step 1 presents a condition implied by the principle of progressive transfers, which is used later on repeatedly. Step 2 proves continuity of f. Step 3 derives a technical result; it considers a sequence of ratios of two decision weights and shows that a well-defined limit exists. Employing this result step 4 proves the concavity of U and $U(E_m(\cdot))$ and Step 5 the linearity of E_m. In Step 6 the implications of the domain Ω for the equivalent income functions are derived. Step 7 proves convexity of f.

Step 1. Progressive transfer

We consider an arbitrary $\left(\hat{\mathbf{X}}^k, \mathbf{m}^{(k)}\right) \in \Omega^{kn} \times \left\{\mathbf{m}^{(k)}\right\}$ with $\hat{X}^{k*}_{[i]} > \hat{X}^{k*}_{[j]}$ for $i < j$ and $k \geq 1$. Transferring the amount $\varepsilon > 0$ (ε appropriate) from the household with rank i to the household having rank j we obtain $\mathbf{X}^k \in \Omega^{kn}$. Since progressive transfers are rank-preserving, we get[18]:

$$\alpha^k_h\left(w_\downarrow\left(\mathbf{X}^{k*}\right)\right) = \alpha^k_h\left(w_\downarrow\left(\hat{\mathbf{X}}^{k*}\right)\right) \;=:\; \alpha^k_h \quad \textit{for } 1 \leq h \leq kn,$$

furthermore by the principle of progressive transfers and after rearrangement

$$\alpha^k_j\left[U\left(X^{k*}_{[j]}\right) - U\left(\hat{X}^{k*}_{[j]}\right)\right] \geq \alpha^k_i\left[U\left(\hat{X}^{k*}_{[i]}\right) - U\left(X^{k*}_{[i]}\right)\right] \tag{A1}$$

Step 2. f is continuous on $D \cap [0, 1)$

Proof. Suppose that f is discontinuous at $w \in D \cap (0, 1)$. Then there is a jump of f at w:

(i) $f(w) < \inf\limits_{\substack{w<\tilde{w}\\ \tilde{w}\in D}} f(\tilde{w})$ or (ii) $\sup\limits_{\substack{\tilde{w}<w\\ \tilde{w}\in D}} f(\tilde{w}) < f(w)$

We consider (i) and construct a sequence of progressive transfers 'at the jump of f' which will lead to a contradiction.

Define $k_0 = \min\limits_{w\in D_k} k$ and $k(l) := \; 2^l k_0$. We know that $w \in D_{k(l)}$ for all $l \geq 1$ and start by setting $l = 1$. There exists $A \subset N_{k(l)}$ such that $w = w(A)/w\left(N_{k(l)}\right)$. Then we define $i(l) := |A| + 1 < k(l)n$ and choose $j(l)$ such that $i(l) < j(l) \leq k(l)n$.

Now we want to construct a distribution $\hat{\mathbf{X}}^{k(l)} \in \Omega^{k(l)n}$ such that the households belonging to A are the $i(l) - 1$ richest and to perform a progressive transfer[19] from the household with rank $i(l)$ to that with rank $j(l)$. Therefore we choose any $\bar{X}^*$ and $\delta > 0$, define $\hat{\mathbf{X}}^{k(l)}$ s.t.

$$\hat{X}^{k(l)*}_{[h]} := \begin{cases} \bar{X}^* + \delta & \text{for } \; h = 1, \ldots, i(l) \\ \bar{X}^* + \delta/2 & \text{for } \; i(l) < h < j(l) \\ \bar{X}^* & \text{for } \; j(l) \leq \; h \leq \; k(l)n \end{cases} \tag{A2}$$

and require that $\left\{\pi\left(h, \hat{\mathbf{X}}^{k(l)*}\right) | \, h = 1, \ldots, i(l) - 1\right\} = A$.

Next we perform the progressive transfer of an amount $\varepsilon, \quad 0 < \varepsilon < \delta/2$. Then (A1) is satisfied for $k = k(l), \quad i = i(l), \quad j = j(l)$. Furthermore

[18] There is a slight abuse of notation because $w_\downarrow(\mathbf{X}^{k*})$ also depends on $\mathbf{m}^{(k)}$.

[19] Although at this point in the argument $l = 1$, the transfer is defined for an arbitrary l as the general formulation is needed for the subsequent steps in the induction process.

$$\alpha_{i(l)}^{k(l)} = f\big(w + w_{[i(l)]}/w\big(N_{k(l)}\big)\big) - f(w) \geq \inf_{\substack{w<\tilde{w}\\ \tilde{w}\in D}} f(\tilde{w}) - f(w) > 0 \tag{A3}$$

Now we duplicate the population. In order to utilize (A2) and (A3), we relabel the populations by letting the old population be indexed $l-1$ and the new one by l. We construct a transfer analogous to the one above by defining $i(l) := 2i(l-1) - 1$. Equation (7) yields $\alpha_{j(l-1)}^{k(l-1)} = \alpha_{2j(l-1)-1}^{k(l)} + \alpha_{2j(l-1)}^{k(l)}$; i.e., at least one decision weight of the RHS of this equation does not exceed $\alpha_{j(l-1)}^{k(l-1)}/2$. Define

$$j(l) := \begin{cases} 2j(l-1) - 1 & \text{if } \alpha_{2j(l)-1}^{k(l)} \leq \alpha_{j(l-1)}^{k(l-1)}/2 \\ 2j(l-1) & \text{otherwise} \end{cases}$$

Using the same $\delta, \bar{X}^*$, and ε we define $\hat{\mathbf{X}}^{k(l)}$ as in (A2) and perform an analogous transfer. Then again (A1) and (A3) are fulfilled.

When l tends to infinity, the terms in brackets in (A1) do not change and are strictly positive. On the other hand $\alpha_{j(l)}^{k(l)}$ goes to zero. Since (A3) is satisfied, the RHS of (A1) is always strictly positive. We get a contradiction.

The same construction is also possible for $w = 0$, but impossible for $w = 1$, since there is no $j(l) > i(l) = k(l)n$. The proof for (ii) runs along similar lines.

Step 3. Technical lemma

Lemma. *Define* $\hat{w} := \sum_{h=1}^{n_H} w(h)/w(N_2) \in D_2$ *and* $k(l) := 2^l$, *choose* $m, \tilde{m} \in H$, *and assume that for all* $l \geq 1$, $\mathbf{X}^{k(l)} \in \Omega^{k(l)n}$ *is chosen so that the ranks* $i(l)$ and $j(l) = i(l) + 1$ *are defined such that*

$$\alpha_{i(l)}^{k(l)} = f\big(\hat{w} + w(m)/w\big(N_{k(l)}\big)\big) - f(\hat{w}) \tag{A4}$$

$$\alpha_{j(l)}^{k(l)} = f\big(\hat{w} + (w(m) + w(\tilde{m}))/w\big(N_{k(l)}\big)\big) - f\big(\hat{w} + w(m)/w\big(N_{k(l)}\big)\big) \tag{A5}$$

Then

$$\lim_{l\to\infty} \alpha_{i(l)}^{k(l)}/\alpha_{j(l)}^{k(l)} = w(m)/w(\tilde{m}).$$

The proof consists of five parts. We suppose throughout this step that $\hat{w}, k(l)$ have been defined and $m, \tilde{m}$ are fixed. Note that $0 < \hat{w} \leq 1/2$. The latter inequality follows because the numerator of $\hat{w}$ does not exceed $w(N)$ (because $n_H \leq n$) and because $w(N_2) = 2w(N)$. $\hat{w}$ will also be employed in later steps.

Part a. Definition of a 'regular grid' $D(m) \subset D$ around $\hat{w}$ for $m \in H$

By assumption the population contains at least one household of each type. We duplicate the population. Then $\hat{w} = \sum_{h=1}^{n_H} w(h)/w(N_2) \in D_2$, $\hat{w} = w(A_2)/w(N_2)$ for $A_2 \subset N_2$ and there is at least one household of each type in A_2 and $N_2 - A_2$.

We define $\hat{w}^{-}(m) \ := \ \hat{w} - w(m)/w(N_2)$ and

$$D(m) := \left\{\hat{w}^{-}(m) + \frac{h}{k}\,\frac{w(m)}{w(N_2)} \mid 0 \le h \le 2k,\ \ h \in \mathbb{N}_0,\ \ k = 1, 2, \ldots\right\}.$$

$D(m)$ corresponds to a 'regular grid' around $\hat{w}$. The difference between $w' \in D(m)$ and $w'' \in D(m)$ can be represented as a multiple of $w(m)/(kw(N_2)) = w(m)/w(N_{2k})$ for an appropriate k. The regularity will be helpful in the rest of the proof, where we will confine us to income distributions which 'fit to $D(m)$'. We define $\Omega^{2kn}(m) \times \{\mathbf{m}^{(2k)}\}$ for $k \ge 1$, where $\Omega^{2kn}(m)$ contains all distributions $\mathbf{X} \in \Omega^{2kn}$ such that $\hat{w}^{-}(m) = \sum_{h=1}^{k(n_H-1)} w_{[h]}(\mathbf{X}^*)/w(N_{2k})$, and where the households with rank $i = k\ (n_H - 1) + 1, \ldots, k(n_H + 1)$ have type m.

We obtain

$$\begin{aligned}\alpha_i^{2k}(\mathbf{X}^*) &= f(\hat{w}^{-}(m) + (i - k(n_H - 1))w(m)/w(N_{2k})) \\ &\qquad - f(\hat{w}^{-}(m) + (i - 1 - k(n_H - 1))w(m)/w(N_{2k})) \ =: \alpha_i^{2k} \qquad \text{(A6)}\end{aligned}$$

for $i = k(n_H - 1) + 1, \ldots, k(n_H + 1)$ and $\mathbf{X} \in \Omega^{2kn}(m)$.

Thus these weights do not depend on the particular distribution $\mathbf{X} \in \Omega^{2kn}(m)$.

Part b. $\alpha_i^{2k} \le \alpha_{i+1}^{2k}$ for $\alpha_i^{2k}, \alpha_{i+1}^{2k}$ defined by (A6) and $i = k(n_H - 1)\ +1, \ldots, k(n_H + 1) - 1$ for all $k \ge 1$.

This result can be interpreted by saying that f is convex on $D(m)$.

Proof[20]. For given i and k we choose an $\hat{\mathbf{X}}^{2k} \in \Omega^{2kn}(m)$ with $\hat{X}_{[i]}^{2k*} > \hat{X}_{[i+1]}^{2k*}$ and appropriate $\varepsilon > 0$. Then we construct a progressive transfer of size ε from the household having rank i to the household with rank $i + 1$ and get $\mathbf{X}^{2k}$. (A1) implies

$$\frac{\left[U\left(X_{[i+1]}^{2k*}\right) - U\left(\hat{X}_{[i+1]}^{2k*}\right)\right]/\varepsilon}{\left[U\left(\hat{X}_{[i]}^{2k*}\right) - U\left(X_{[i]}^{2k*}\right)\right]/\varepsilon} \ge \frac{\alpha_i^{2k}}{\alpha_{i+1}^{2k}} \qquad \text{(A7)}$$

The RHS of (A7) is independent of $\hat{\mathbf{X}}^{2k}, \mathbf{X}^{2k}, \varepsilon$. We have to distinguish two cases.

(i) There is $\varepsilon > 0$ and $\hat{\mathbf{X}}^{2k} \in \Omega^{2kn}(m)$ such that the LHS of (A7) is less than or equal to unity. Then part b follows immediately.

(ii) The LHS of (A7) is always strictly greater than unity. Then $U(E_m(\cdot))$ is strictly concave and strictly increasing on Ω and therefore differentiable almost everywhere. Its derivative is strictly positive everywhere where it exists. Choosing $\hat{\mathbf{X}}^{2k}$ appropriately and letting $\varepsilon \to 0$ we obtain

[20] I owe the following arguments to a referee.

$$\frac{dU\left(E_m\left(\hat{X}^{2k}_{\pi\left(i+1,\hat{\mathbf{X}}^*\right)}\right)\right)/d\hat{X}^{2k}_{\pi\left(i+1,\hat{\mathbf{X}}^*\right)}}{dU\left(E_m\left(\hat{X}^{2k}_{\pi\left(i,\hat{\mathbf{X}}^*\right)}\right)\right)/d\hat{X}^{2k}_{\pi\left(i,\hat{\mathbf{X}}^*\right)}} \geq \frac{\alpha_i^{2k}}{\alpha_{i+1}^{2k}},$$

where households i and $i+1$ are of the same type.

Letting $\hat{X}^{2k}_{\pi\left(i,\hat{\mathbf{X}}^*\right)}$ tend to $\hat{X}^{2k}_{\pi\left(i+1,\hat{\mathbf{X}}^*\right)}$ we get $\alpha_{i+1}^{2k} \geq \alpha_i^{2k}$ (since $U(E_m(\cdot))$ is differentiable almost everywhere the incomes $\hat{X}^{2k}_{\pi\left(i,\hat{\mathbf{X}}^*\right)}$ and $\hat{X}^{2k}_{\pi\left(i+1,\hat{\mathbf{X}}^*\right)}$ can always be chosen such that the derivatives exist and are strictly positive. Then the LHS of the last equation tends to unity.)

Part c. Existence of a limit

Assume that $w^l := \frac{h(l)}{k(l)}\frac{w(m)}{w(N_2)}$ with a natural number $h(l)$, $0 < h(l) \leq k(l)$, $w^{l+1} < w^l$ for all l, and that $w^l \to 0$ for $l \to \infty$.

Then $\lim\limits_{l\to\infty} \frac{f\left(\hat{w}+w^l\right)-f(\hat{w})}{w^l} =: L(\hat{w}, m)$ exists and is strictly positive.

Proof. In view of the convexity of f on $D(m)$ (part b) we obtain

$$0 < \frac{f(\hat{w}) - f(\hat{w}^-(m))}{\hat{w} - \hat{w}^-(m)} \leq \frac{f\left(\hat{w} + w^{l+1}\right) - f(\hat{w})}{w^{l+1}} \leq \frac{f\left(\hat{w} + w^l\right) - f(\hat{w})}{w^l},$$

i.e., a decreasing sequence bounded from below. This proves part c.

Part d. Convergence of a specific sequence

We define

$$B^l := \frac{f\left(\hat{w} + (w(m) + w(\tilde{m}))/w\left(N_{k(l)}\right)\right) - f(\hat{w})}{(w(m) + w(\tilde{m}))/w\left(N_{k(l)}\right)}$$

We want to show that then $\lim\limits_{l\to\infty} B^l = L(\hat{w}, m)$.

Proof. If $m = \tilde{m}$ or $w(m) = w(\tilde{m})$ apply part c directly. Otherwise it is possible that $\hat{w} + (w(m) + w(\tilde{m}))/w\left(N_{k(l)}\right) \notin D(m)$. Then there exists a minimal $l_0 \geq 1$ such that $w(m)/w\left(N_{k(l_0)}\right) < (w(m) + w(\tilde{m}))/w\left(N_{k(l_0)}\right) \leq w(m)/w(N_2)$, i.e., $\hat{w} + (w(m) + w(\tilde{m}))/w\left(N_{k(l_0)}\right)$ belongs to $D(m)$ or is in the closure of $D(m)$.

We want to construct two specific sequences in order to approximate B^l from below and above for $l \geq l_0$. Thus we define $w^l := (w(m) + w(\tilde{m}))/\ w\left(N_{k(l)}\right)$, $q(s,l) := 2^s k(l) = k(l+s)$, and a positive integer $p(s) \geq 1$ such that $p(s)w(m)/2^s < w(m) + w(\tilde{m}) \leq (p(s)+1)w(m)/2^s$. Then $p(s)w(m)/w\left(N_{q(s,l)}\right) < w^l \leq (p(s)+1)w(m)/w\left(N_{q(s,l)}\right)$ since $w\left(N_{q(s,l)}\right) = 2^s k(l)w(N)$ and $w\left(N_{k(l)}\right) = k(l)w(N)$. Now we consider $\underline{w}_s^l := p(s)w(m)/w\left(N_{q(l,s)}\right)$ and $\bar{w}_s^l := (p(s)+1)w(m)/w\left(N_{q(l,s)}\right)$ and $\underline{A}_s^l := \left(f\left(\hat{w} + \underline{w}_s^l\right) - f(\hat{w})\right)/\underline{w}_s^l$ and $\bar{A}_s^l := \left(f\left(\hat{w} + \bar{w}_s^l\right) - f(\hat{w})\right)/\bar{w}_s^l$.

Then by construction $\hat{w} + \underline{w}_s^l,\ \hat{w} + \bar{w}_s^l \in D(m)$, $\underline{w}_s^l \leq \underline{w}_{s+1}^l \leq w^l \leq \bar{w}_{s+1}^l \leq \bar{w}_s^l$, and $\lim\limits_{s\to\infty} \underline{w}_s^l = w^l = \lim\limits_{s\to\infty} \bar{w}_s^l$.

Because of monotonicity of f, we obtain

$$\frac{\underline{w}_s^l}{w^l}\underline{A}_s^l \leq B^l \leq \frac{\bar{w}_s^l}{w^l}\bar{A}_s^l. \tag{A8}$$

Now suppose that $s = l$. Next we investigate the sequences $\underline{w}_l^l,\ \bar{w}_l^l,\ w^l$ and the corresponding $\underline{A}_l^l,\ \bar{A}_l^l,\ B^l$. Since $q(l+1, l+1) = k(2l+2) = 4k(2l) = 4q(l, l)$ and $p(s+1) \leq 2p(s) + 1$ we get $\underline{w}_{l+1}^{l+1} < \underline{w}_l^l$ and $\bar{w}_{l+1}^{l+1} < \bar{w}_l^l$ and therefore by part c: $\lim_{l\to\infty} \underline{A}_l^l = \lim_{l\to\infty} \bar{A}_l^l = L(\hat{w}, m)$. On the other hand $\lim_{l\to\infty} \underline{w}_l^l/w^l = \lim_{l\to\infty} \bar{w}_l^l/w^l = 1$. Thus $\lim_{l\to\infty} B^l = L(\hat{w}, m)$ because of (A8).

Part e. Combining results

Define $A^l := \left[f\left(\hat{w} + w(m)/w\left(N_{k(l)}\right)\right) - f(\hat{w})\right]/\left[w(m)/w\left(N_{k(l)}\right)\right]$ and B^l as above and assume that $i(l),\ j(l),\ k(l)$ are given as described in the Lemma. Then

$$\frac{\alpha_{i(l)}^{k(l)}}{\alpha_{j(l)}^{k(l)}} = \frac{A^l}{\frac{w(m)+w(\tilde{m})}{w(\tilde{m})}B^l - \frac{w(m)}{w(\tilde{m})}A^l} \cdot \frac{w(m)}{w(\tilde{m})}. \tag{A9}$$

Part c and part d imply that $\lim_{l\to\infty} A^l = \lim_{l\to\infty} B^l = L(\hat{w}, m)$. This proves the result.

Step 4. For $m \in H$: $U(E_m(\cdot))$ is concave and absolutely continuous. Therefore its derivative exists a.e. and is strictly positive a.e.

Proof. We choose $\bar{X}$ and $\varepsilon > 0$ and will prove that

$$U(E_m(\bar{X} + \varepsilon)) - U(E_m(\bar{X})) \leq U(E_m(\bar{X}) - U(E_m(\bar{X} - \varepsilon)). \tag{A10}$$

Let $m \in H$ be given and set $k(l) := 2^l$ and $l := 2$. Then $\hat{w},\ \hat{w} + w(m)/w(N_{k(l)}), \hat{w} + 2w(m)/w(N_{k(l)}) \in D(m)$. We define $i(2) := 2n_H + 1$, $j(2) := i(2) + 1$ and construct $\hat{\mathbf{X}}^{k(l)} \in \Omega^{k(l)n}(m)$ such that $\hat{X}^{k(l)}_{\pi\left(i(l), \hat{\mathbf{X}}^{k(l)*}\right)} = \bar{X} + \varepsilon_{i(l)}$ $\hat{X}^{k(l)}_{\pi(j(l), \hat{\mathbf{X}}^{k(l)*})} = \bar{X} - \varepsilon$, $\sum_{h=1}^{i(l)} w_{[h]}/w(N_{k(l)}) = \hat{w} + w(m)/w\left(N_{k(l)}\right)$, and $\sum_{h=1}^{i(l)} w_{[h]}/w(N_{k(l)}) = \hat{w} + 2w\ (m)/w(N_{k(l)})$.

Now we perform a progressive transfer of an amount ε from the household with rank $i(l)$ to the household having rank $j(l)$. Then (A1) is satisfied for $i = i(l),\ j = j(l),\ k = k(l)$, and we get

$$\begin{aligned}\frac{\alpha_{i(l)}^{k(l)}}{\alpha_{j(l)}^{k(l)}} &= \frac{f\left(\hat{w} + w(m)/(N_{k(l)})\right) - f(\hat{w})}{f\left(\hat{w} + 2w(m)/w(N_{k(l)})\right) - f\left(\hat{w} + w(m)/(N_{k(l)})\right)} \\ &\leq \frac{U\left(E_m(\bar{X})\right) - U\left(E_m(\bar{X} - \varepsilon)\right)}{U\left(E_m(\bar{X} + \varepsilon)\right) - U\left(E_m(\bar{X})\right)}\end{aligned} \tag{A11}$$

After increasing $l \to l+1$ we set $i(l) := 2i(l-1) - 1, \quad j(l) \;:=\; i(l) + 1$ and perform an analogous transfer. If l goes to infinity, the ratio (A11) converges to 1 (apply Step 3 for $m = \tilde{m}$). This proves (A10), which is equivalent to midpoint-concavity. Continuity and midpoint-concavity imply that $U(E_m(\cdot))$ is concave and absolutely continuous and that its derivative exists a.e. The latter is strictly positive a.e. because of strict monotonicity (cf. Kuczma 1985, Chapt. 7).

Step 5. $E_m(X) = \frac{w(r)}{w(m)} X - b_m$ for $m \in H$

Proof. We choose $m, \tilde{m} \in H$, define $k(l) := 2^l$, and consider $\hat{w}, \quad \hat{w} + w(m)/w(N_{k(l)}) \in D(m)$, and $\hat{w} + (w(m) + w(\tilde{m}))/w(N_{k(l)}) \in D$. We define $i(l) := k(l-1)n_H + 1, \quad j(l) \;:=\; i(l) + 1$ for $l \geq 1$. After choosing $X, \tilde{X}$ with $X^* \;:=\; E_m(X) > E_{\tilde{m}}(\tilde{X}) \;=:\; \tilde{X}^*$, we construct $\hat{\mathbf{X}}^{k(l)} \in \Omega^{k(l)n}$ such that $\sum_{h=1}^{i(l)} w_{[h]}/w(N_{k(l)}) = \hat{w} + w(m)/w(N_{k(l)}), \quad w_{[i(l)]} = w(m), \quad w_{[j(l)]} = w(\tilde{m})$, and $X^* = \hat{X}^{k(l)*}_{[i(l)]} > \hat{X}^{k(l)*}_{[j(l)]} = \tilde{X}^*$.

Performing a progressive transfer of a small amount $\varepsilon > 0$ we obtain (A1) and by definition

$$\frac{\alpha^{k(l)}_{i(l)}}{\alpha^{k(l)}_{j(l)}} = \frac{f\left(\hat{w} + w(m)/w(N_{k(l)})\right) - f(\hat{w})}{f\left(\hat{w} + (w(m) + w(\tilde{m}))/w(N_{k(l)})\right) - f\left(\hat{w} + w(m)/w(N_{k(l)})\right)}$$

By Step 3 this ratio converges to $w(m)/w(\tilde{m})$, when this step is repeated analogously and l goes to infinity. Thus we obtain from (A1)

$$w(\tilde{m})\left[U\left(E_{\tilde{m}}(\tilde{X} + \varepsilon)\right) - U\left(E_{\tilde{m}}(\tilde{X})\right)\right]/\varepsilon \geq w(m)[U(E_m(X)) - U(E_m(X - \varepsilon))]/\varepsilon$$

and for $\varepsilon \to 0$ (since $U(E_m(\cdot))$ is concave, it is differentiable a.e. with positive derivative)

$$w(\tilde{m}) \frac{d\, U\left(E_{\tilde{m}}(\tilde{X})\right)}{d\tilde{X}} \geq w(m) \frac{d\, U(E_m(X))}{dX} \text{ with } E_m(X) > E_{\tilde{m}}(\tilde{X}).$$

Now we change $\tilde{X}$ such that $E_{\tilde{m}}(\tilde{X})$ tends to $E_m(X)$. Then we get

$$w(\tilde{m}) \frac{d\, U(X^*)}{dX^*} E'_{\tilde{m}}(\tilde{X}) \geq w(m) \frac{d\, U(X^*)}{dX^*} E'_m(X).$$

Since U is absolutely continuous and strictly increasing, we obtain $w(\tilde{m})E'_{\tilde{m}}(\tilde{X}) \geq w(m)E'_m(X)$ and the converse by switching the roles of m and $\tilde{m}$. Therefore $w(\tilde{m})E'_{\tilde{m}}(\tilde{X}) = w(m)E'_m(X)$, if $E_{\tilde{m}}(\tilde{X}) = E_m(X)$. Since $E_r(X) = X$, we have $E'_r(X) = 1$ and $E'_m(X) = w(r)/w(m)$, which proves the claim.

Step 6. $E_m(X) = \frac{w(r)}{w(m)} X$ for $m \in H$ and $w(m) < w(m+1)$ for $m = 1, \ldots, n_H - 1$ if $\Omega = \mathbb{R}_{++}$, and $E_m(X) = X - b_m, \quad w(m) = w(r)$ for $m \in H$ and $b_m < b_{m+1}, \; m = 1, \ldots, n_H - 1, \; b_r = 0$ if $\Omega = \mathbb{R}$.

Proof. If $\Omega = \mathbb{R}_{++}$, the constants b_m have to be equal zero, since E_m is one-to-one. $w(m) < w(m+1)$ is implied by E(iv).

If $\Omega = \mathbb{R}$, the weights $w(m)$ have to be identical to $w(r)$, since otherwise E(iv) is violated ($E_r(X)$ and $E_m(X)$ would cross for $m \neq r$). $b_m < b_{m+1}$ is implied by E(iv) and $b_r = 0$ by E(v).

Step 7. For every k, i and $\mathbf{X} \in \Omega^{kn}$ we have

$$\frac{\alpha_i^k\left(w_\downarrow(X^*)\right)}{w_{[i]}(\mathbf{X}^*)/w(N_k)} \leq \frac{\alpha_{i+1}^k\left(w_\downarrow(\mathbf{X}^*)\right)}{w_{[i+1]}(\mathbf{X}^*)/w(N_k)}. \tag{A12}$$

Proof. We construct an appropriate $\hat{\mathbf{X}}^k \in \Omega^{kn}$ and progressive transfer implying (A1) for $j = i+1$. Dividing both sides by ε we obtain

$$\alpha_{i+1}^k \frac{d\,U(\mathbf{X}^*)}{d\,X^*} E'_{\tilde{m}}\left(X^k_{\pi(i+1,X^{k*})}\right) \geq \alpha_i^k \frac{d\,U(X^*)}{d\,X^*} E'_m\left(X^k_{\pi(i,X^{k*})}\right)$$

for $\varepsilon \to 0$, $X^{k*}_{[i+1]} \to X^{k*}_{[i]} =: X^*$, $m := m^{(k)}_{\pi\left(i,\mathbf{X}^{k*}\right)}$ and $\tilde{m} := m^{(k)}_{\pi\left(i+1,\mathbf{X}^{k*}\right)}$. Step 5 proves the result. ■

This allows us to prove convexity of f also generally: (A12) holds for given $\mathbf{X}$ and i, but simple computations demonstrate that the intervals can be 'enlarged' for given $\mathbf{X}$; i.e.

$$\frac{f(w_2) - f(w_1)}{w_2 - w_1} \leq \frac{f(w_3) - f(w_2)}{w_3 - w_2} \leq \frac{f(w_4) - f(w_3)}{w_4 - w_3}$$

for $w_1 < w_2 < w_3 < w_4$ also implies that

$$\frac{f(w_2) - f(w_1)}{w_2 - w_1} \leq \frac{f(w_4) - f(w_2)}{w_4 - w_2} \text{ and } \frac{f(w_3) - f(w_1)}{w_3 - w_1} \leq \frac{f(w_4) - f(w_3)}{w_4 - w_3}.$$

For arbitrary $w_1, w_2, w_3 \in D$ with $w_1 < w_2 < w_3$ we chose k and $\mathbf{X} \in \Omega^{kn}$ s.t. $w_2 = w_{[i]}(\mathbf{X}^*)/w(N_k)$ for an appropriate i and define $\underline{w}^k$ to be the minimal $w_{[j]}(\mathbf{X}^*)/w(N_k) \geq w_1$ and $\bar{w}_k$ to be the maximal $w_{[j]}(\mathbf{X}^*)/w(N_k) \leq w_3$. Then $f(w_1) \leq f\left(\underline{w}^k\right) < f(w_2) < f\left(\bar{w}^k\right) \leq f(w_3)$. Letting k tend to infinity we obtain

$$\frac{f(w_2) - f(w_1)}{w_2 - w_1} \leq \frac{f(w_3) - f(w_2)}{w_3 - w_2} \tag{A13}$$

in the limit, since the RHS and LHS of (A13) is approximated.

Proof of Proposition 3

Without loss of generality we consider the case $k = 1$.

Theorem 2 implies the linearity of the equivalent income function $E_{m_i}(X_i) = \frac{w(r)}{w(m_i)} X_i$ or $E_{m_i}(X_i) = X_i - b_{m_i}$. Therefore

$$\mu(\mathbf{X}^*,\mathbf{w}^*)=\sum_{i=1}^{n}\frac{w(m_i)}{w(N)}X_i^*=\begin{cases}\frac{w(r)}{w(N)}\sum\limits_{i=1}^{n}X_i & \text{for } \Omega=\mathbb{R}_{++}\\ \frac{1}{n}\left(\sum\limits_{i=1}^{n}X_i-\sum\limits_{i=1}^{n}b_{m_i}\right) & \text{for } \Omega=\mathbb{R}\end{cases}$$

which is unchanged by a progressive transfer (if $\Omega=\mathbb{R}$ then $w(m)=w(r)$ for $m=1,\ldots,n_H$ and therefore $w(N)=n\,w(r)$). Because of the concavity of U and of Jensen's inequality we obtain

$$\begin{aligned}\xi\left(\mathbf{X}^*,\mathbf{w}^*\right)&=U^{-1}\left(\sum_{i=1}^{n}\alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right)U\left(X^*_{[i]}\right)\right)\\ &\leq\sum_{i=1}^{n}\alpha_i\left(\mathbf{w}_\downarrow(\mathbf{X}^*)\right)X^*_{[i]}=\sum_{i=1}^{n-1}f\left(\frac{\sum\limits_{j=1}^{i}w_{[j]}(\mathbf{X}^*)}{w(N)}\right)\left(X^*_{[i]}-X^*_{[i+1]}\right)+X^*_{[n]}\\ &\leq\sum_{i=1}^{n-1}\left(\sum_{j=1}^{i}w_{[j]}(\mathbf{X}^*)/w(N)\right)\left(X^*_{[i]}-X^*_{[i+1]}\right)+X^*_{[n]}\\ &=\sum_{i=1}^{n}\frac{w_{[i]}(\mathbf{X}^*)}{w(N)}X^*_{[i]}=\mu\left(\mathbf{X}^*,\mathbf{w}^*\right),\end{aligned}$$

where the last inequality holds since $f(w)\leq w$. Thus $\mu(\mathbf{X}^*,\mathbf{w}^*)-\xi(\mathbf{X}^*,\mathbf{w}^*)\geq 0$. On the other hand a progressive transfer weakly increases $\xi(\mathbf{X}^*,\mathbf{w}^*)$.
The converse is obvious. ■

Proof of Proposition 4

(a) SCALE invariance implies that $\xi(\lambda\mathbf{X}^*,\mathbf{w}^*)=\lambda\,\xi(\mathbf{X}^*,\mathbf{w}^*)$. Observing the functional structure we obtain a functional equation, which can be solved. It implies that $U(X)=\gamma X^\varepsilon+\delta$ or $U(X)=\gamma\log X+\delta$, where $\gamma\neq 0$, $\varepsilon\neq 0$, and δ are real constants (see Proposition 2 in Ebert (2003) for details). Theorem 2 yields that $X^*=E_m(X)=\frac{w(r)}{w(m)}X$. Thus we obtain the result.
(b) Analogously.
(c) It is a consequence of (a) and (b). ■

Proof of Corollary 5

Observe that $\xi_k\left(\mathbf{X}^p,\mathbf{m}^{(k)}\right)\leq\max\left\{X_i^{p^*}\mid i\in N_k\right\}\leq Z^*$ and employ Proposition 3. ■

Proof of Proposition 6

Similar to the proof of Proposition 4. ■

References

Aczel J (1966) Lectures on functional equations and their applications. Academic Press, New York

Atkinson AB, Bourguignon F (1987) Income distribution and differences in needs. In: Feiwel GR (ed) Arrow and the foundations of the theory of economic policy. Macmillan, London, pp 350–370

Atkinson AB (1970) On the measurement of inequality. J Econ Theory 2: 244–263

Ben-Porath E, Gilboa I (1994) Linear measures, the Gini index, and the income-equality trade-off. J Econ Theory 64: 443–467

Ben-Porath E, Gilboa I, Schmeidler D (1997) On the measurement of inequality under uncertainty. J Econ Theory 75: 194–204

Bourguignon F (1989) Family size and social utility. Income distribution dominance criteria. J Econometrics 42: 67–80

Buhmann B, Rainwater L, Schmaus O, Smeeding TM (1988) Equivalence scales, well-being, inequality, and poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) database. Rev Income Wealth 34: 115–142

Burkhauser RV, Smeeding TM, Merz J (1996) Relative inequality and poverty in Germany and the United States using alternative equivalence scales. Rev Income Wealth 42: 381–400

Chakravarty SR (1983) Ethically flexible measures of poverty. Can J Econ 41: 74–85

Chateauneuf A, Cohen W (1994) Risk seeking with diminishing marginal utility in a non-expected utility model. J Risk Uncertainty 9: 77–91

Chew SH, Karni E, Safra Z (1987) Risk aversion in the theory of expected utility with rank dependent probabilities. J Econ Theory 42: 370–381

Chiappori P-A (1992) Collective labor supply and welfare. J Polit Econ 100: 437–467

Clark S, Hemming R, Ulph D (1981) On indices for the measurement of poverty. Econ J 91: 518–526

Cohen MD (1995) Risk-aversion concepts in expected- and non-expected utility models. Geneva Papers Risk Insur Theory 20: 73–91

Coulter FAS, Cowell FA, Jenkins SB (1992) Equivalent scale relativities and the extent of inequality and poverty. Econ J 102: 1067–1082

Donaldson D, Weymark JA (1980) A single-parameter generalization of the Gini indices of inequality. J Econ Theory 22: 67–86

Ebert U (1987) Size and distribution of incomes as determinants of social welfare. J Econ Theory 41: 23–33

Ebert U (1988) Measurement of inequality: an attempt at unification and generalization. Soc Choice Welfare 5: 147–169

Ebert U (1995) Income inequality and differences in household size. Math Soc Sci 30: 37–55

Ebert U (1997a) Absolute inequality indices and equivalence differences. In: Zandvikili S (ed) Taxation and inequality (Research on Economic Inequality, vol. 7). JAI Press, pp 131–152

Ebert U (1997b) Social welfare when needs differ: An axiomatic approach. Economica 64: 233–244

Ebert U (1999) Using equivalent income of equivalent adults to rank income distributions. Soc Choice Welfare 16: 233–258

Ebert U (2000a) Sequential generalized Lorenz dominance and transfer principles. Bull Econ Res 52: 113–122

Ebert U (2000b) Equivalizing incomes: a normative approach. Int Tax Publ Finance 7: 619–640

Ebert U (2003) Inequality measurement for homogeneous groups. In: Amiel Y, Bishop JA (eds) Inequality, welfare and poverty: Theory and measurement (Research on Economic Inequality, vol. 9). JAI Press, pp 47–67

Ebert U, Moyes P (2002) Welfare, inequality and the transformation of weighted income distributions. The case of weighted income distributions. In: Moyes P, Seidl C, Shorrocks AF (eds) Inequalities: Theory, measurement and applications. J Econ, (Supplement) 9: 9–50
Ebert U, Moyes P (2003) Equivalence scales reconsidered. Econometrica 71: 319–343
Gelbaum BR, Olmsted JMH (1964) Counterexamples in analysis. Holden-Day, Inc., San Francisco
Kolm SC (1969) The optimal production of social justice. In: Margolis J, Guitton H (eds) Public economics. Macmillan, London, pp 145–200
Kolm SC (1977) Multidimensional egalitarianism. Q J Econ 91: 1–13
Kuczma M (1985) An introduction to the theory of functional equations and inequalities. Państwowe Wydawnictwo Naukowe, Warszaw
Maasoumi E (1986) The measurement and decomposition of multidimensional inequality. Econometrica 54: 991–997
Mehran F (1976) Linear measures of income inequality. Econometrica 44: 805–809
Pyatt G (1985) An axiomatic approach to the Gini coefficient and the measurement of welfare. In: Basmann RL, Rhodes GF, Jr (eds) Advances in econometrics, vol. 4. JAI Press, Greenwich, Connecticut, pp 87–109
Quiggin J (1982) A theory of anticipated utility. J Econ Behav Organiz 3: 323–343
Segal U (1989) Anticipated utility: A measure representation approach. Ann Op Res 19: 359–373
Sen AK (1973) On economic inequality. Clarendon Press, Oxford
Sen AK (1976) Poverty: An ordinal approach to measurement. Econometrica 44: 219–231
Tsui KY (1995) Multidimensional generalizations of the relative and absolute inequality indices: the Atkinson-Kolm-Sen approach. J Econ Theory 67: 251–265
Tsui KY (1999) Multidimensional inequality and multidimensional generalized entropy measures: An axiomatic derivation. Soc Choice Welfare 16: 145–157
Wakker PP (1989) Continuous subjective expected utility with non-additive probabilities. J Math Econ 18: 1–27
Weymark JA (1981) Generalized Gini inequality indices. Math Soc Sci 1: 409–430
Weymark JA (1994) An alternative proof of Ben-Porath and Gilboa's generalized Gini characterization theorem. J Econ Theory 66: 573–580
Yaari ME (1987) The dual theory of choice under risk. Econometrica 55: 95–115
Yaari ME (1988) A controversial proposal concerning inequality measurement. J Econ Theory 44: 381–397