KATHOLIEKE UNIVERSITEIT
LEUVEN

Faculty of Economics and Applied Economics

Department of Economics

On Comparing Heterogenous Populations: is there really a Conflict between the Pareto Criterion and Inequality Aversion?

by

Bart CAPEAU
Erwin OOGHE

Public Economics

Center for Economic Studies
Discussions Paper Series (DPS) 04.07
http://www.econ.kuleuven.be/ces/discussionpapers/default.htm

September 2004

DISCUSSION PAPER

On comparing heterogenous populations:
Is there really a conflict between the Pareto criterion and inequality aversion?*

Bart Capéau† and Erwin Ooghe‡

August 2004

Abstract

The incompatibility between the Pareto indifference criterion and a concern for greater equality in living standards of heterogenous populations (see, amongst others, Ebert, 1995, 1997, Ebert and Moyes, 2003 and Shorrocks, 1995) might come as a surprise, since both principles are reconcilable when people differ only in income (homogenous population). We present two families of welfare rankings –(i) single parameter extensions of the generalized Lorenz dominance rule and (ii) a subset of Weymark's (1981) generalized Ginis– and show how and why these rules resolve the paradox.

JEL classification: D31, D63, I31.

Keywords: heterogeneity, welfare comparisons.

1 Motivation

In the literature on heterogenous welfare comparisons, there seems to be a conflict between welfarism and a concern for greater equality in living standards (see amongst others

*We would like to thank André Decoster, Udo Ebert, Marc Fleurbaey, Luc Lauwers, Eugenio Peluso, Erik Schokkaert, Frans Spinnewyn, and Alain Trannoy for helpful comments, suggestions and stimulating discussions.

†Center for Economic Studies, Naamsestraat, 69, B-3000 Leuven, Belgium. I am thankful for the financial support of the K.U.Leuven research fund, grant 02/07. e-mail to bart.capeau@econ.kuleuven.ac.be.

‡Postdoctoral Fellow of the Fund for Scientific Research – Flanders. Center for Economic Studies, Naamsestraat, 69, B-3000 Leuven, Belgium. e-mail to erwin.ooghe@econ.kuleuven.ac.be.

Ebert, 1995, 1997, Ebert and Moyes, 2003 and Shorrocks, 1995). Living standards are indices which convert the income of people of different type, say hearing blind *versus* seeing deaf, into welfare measures, assumed to be comparable across individuals. The cited contributions claim that the welfarist Pareto indifference principle –which requires social indifference between two situations in which all individuals reach the same living standard– is incompatible with the between type Pigou-Dalton (BTPD) transfer principle –preferring mean-preserving income transfers which equalize living standards.

Following Shorrocks (1995), we will discuss the issue, in this note, in its simplest and most pure setting: ranking income distributions amongst *individuals*, when those individuals might differ not only with respect to the income they obtain, but in other aspects too. The essentials of the problem of making welfare comparisons among heterogenous populations are safeguarded in the individual setting, without having to treat simultaneously the problem of converting distributions among households into distributions among individuals. Moreover, the solutions we will propose can be adapted to tackle the household-individual conversion problem too, as we will argue in section 3.

To motivate this point of view, we provide two examples which illustrate, for the generalized Lorenz dominance (GLD) ranking (Shorrocks, 1983) and in a purely individual framework, the stated incompatibility between Pareto indifference and the between type Pigou-Dalton principle, which underlies the difficulty of making welfare comparisons among heterogenous populations.[1] In those examples we consider the case of Eve and Mary; Eve has a weaker metabolism than Mary and therefore she needs more food (and thus more income) to reach the same level of calorie intake as Mary. We use ratio scale equivalent income functions to convert monetary incomes into comparable welfare measures. More precisely, Eve obtains only two third's of Mary's welfare level for the same amount of income.

In our first example, we consider the case where Eve has not only a weaker metabolism, but also earns less income than Mary. Their nominal incomes are $(6, 18)$. Because of Eve's defective metabolism, she will reach a welfare level of only 4 units, whereas Mary's income is identical to her welfare and thus equals 18. Consider now a transfer of 6 income units from Mary to Eve. After the transfer, nominal incomes are equal: $(12, 12)$. The transfer also results in a more equal distribution of living standards (welfare levels): $(8, 12)$. The situation is summarized in table 1 below. Notice that we designed the transfer such that, *ex post*, Eve is still worse off in terms of equivalent income (welfare).

Nevertheless, according to the GLD ranking, applied to equivalent incomes, both distribu-

[1] The same problem occurs for other sum-type welfarist rankings (see Ebert, 1997).

tions cannot be compared, the reason being, that despite the more equal distribution of living standards after transfer, the mean living standard, which was equal to 11 before the transfer, has decreased to 10, while mean nominal income, per definition, remained constant and equals 12. More equality in living standards with a heterogenous population might imply giving up efficiency: weak persons are *defined to be* non-efficient welfare producers.

Example 1: a between type Pigou-Dalton transfer

Situation before

type	weak metabolism	strong metabolism
income	6	18
equivalence scale	1.5	1
equivalent income	4	18

↓ income transfer from welfare rich to welfare poor
which preserves rankings of equivalent incomes

Situation after

type	weak metabolism	strong metabolism
income	12	12
equivalence scale	1.5	1
equivalent income	8	12

Ebert (1997) suggests solving the problem by applying the GLD criterion (or other welfarist alternatives) to the equivalent income vector weighted by the equivalence scales. In the present context, the weighted mean of equivalent incomes in both cases equals $\frac{48}{5}$ and the after transfer distribution (always) dominates the original one according to the weighted GLD-criterion.

In our second example, we consider the case where a surgery can cure Eve's defective metabolism. However, the surgery costs 6 income units. After the surgery, Eve and Mary both have an equally strong metabolism: they can attain the same level of welfare with equal incomes. The situation is summarized in table 2 below.

The essential point now is that equivalent incomes before and after the surgery are the same. Pareto indifference requires that the social welfare ordering is indifferent between both situations in such a case. Ebert's weighted GLD criterion runs into problems now. Indeed, average income, and thus weighted average welfare (or living standard) before the surgery is higher: the cost of the surgery was too high to warrant the cure of the defect. Weighted GLD prefers, in this case, a society with unequal type and income distributions, over a more equal, but less affluent, type and income distribution, even if the welfare levels did not change at all. The standard (un-weighted) GLD-criterion applied to equivalent incomes, is indifferent between both situations, as the Pareto indifference criterion requires.

Example 2: changing types by means of a surgery

Situation before

type	strong metabolism	weak metabolism
income	6	18
equivalence scale	1	1.5
equivalent income	6	12

↓ by a surgery which costs 6 income units

Situation after

type	strong metabolism	strong metabolism
income	6	12
equivalence scale	1	1
equivalent income	6	12

Since Pareto indifference in homogenous societies (when people all are of the same type, and welfare levels are in accordance with income) is compatible with any degree of inequality aversion, the suggested conflict might come as a surprise. In this note, we will show how this apparent conflict can be overcome. In the next section (section 2) we present two sets of continuous welfare rankings: the r-extended GLD rankings and families of what we will call r-generalized Ginis. Both sets are defined with the aid of a parameter $r \in \mathbb{R}_+$. The r-extended GLD rankings are defined exclusively by this parameter (there is one ranking for each choice of r). There exists, on the other hand, for each r, a family of r-generalized Ginis, and each such family forms a subset of Weymark's (1981) generalized Ginis. Both sets of rules are linked: for any given r, the r-extended GLD ranking is equivalent with unanimity among the members of the associated family of r-generalized Ginis. In a companion paper (Capéau and Ooghe, 2004), we characterize these rankings.
When r equals zero, we obtain either the "standard" GLD quasi-ordering or the complete set of generalized Ginis; when r approaches infinity, we can come, in both cases, arbitrarily

close to the leximin rule, which is known to overcome the conflict between Pareto and inequality aversion, but binds in on continuity (*cf.* Ebert and Moyes, 2003). Suitably adapting the r-parameter provides *continuous* welfare rankings which satisfy both Pareto indifference and the between type Pigou-Dalton transfer principle. The conflict is due to other, background properties which the cited contributions impose silently on welfare rankings.

In section 3 we return to examples 1 and 2 and illustrate how the rules can be applied to the problems advanced in those examples. We show how and why these rules solve the paradoxical incompatibility between welfarism and a concern for greater equality. Finally, we indicate how the rules can cope with differences in household composition as a source of heterogeneity. However, the conflict between a concern for greater equality and efficiency seems to reappear when we want to be able to compare heterogenous populations of different size. We show in section 4 that none of the rules we proposed can combine a concern for efficiency and inequality aversion with satisfying replication invariance. But the assumption that replicating a society does not affect overall welfare, despite its common acceptance in the literature on welfare measurement, presupposes an answer to some deep and fundamental questions in population ethics: is a society with more people of the same type as those already existing really to be considered equally good as the present one, or is it better or worse? And these questions are, we feel, even more difficult to answer, when people are heterogenous. We therefore consider the reconcilability of efficiency and inequality aversion with less demanding, and in our view, less controversial, aggregation axioms, and get some positive results. Section 5 concludes.

2 Two families of welfare rankings

In the present section, we introduce a set of rankings[2] of distributions of outcomes, without specifying *a priori* which is the outcome of interest we will look at: in the terminology of the preceding section, it could be income, equivalent income or any other well-being measure. As such, our rules might therefore be applicable to classical welfare problems, trying to rank income distributions among homogenous populations. Such a distribution of interest will be denoted by a vector $\mathbf{u} = (u_1, \ldots, u_n) \in \mathbb{R}^n$. So, depending on the context, u_i might be

[2] A ranking, say R, is a reflexive and transitive binary relation on a set X, and will also be called a *quasi-ordering*. Reflexivity means: $\forall\, x \in X : xRx$. Transitivity holds when $\forall x, y, z \in X$: if xRy and yRz then xRz. If the ranking is moreover complete ($\forall\, x, y \in X$: either xRy or yRx), we say it is an *ordering*.

the income or the living standard (equivalent income) of an individual $i \in N = \{1, \ldots, n\}$. Since u_i does not exclusively mean income, but can also refer to welfare, we allow for negative entries in $\mathbf{u}$. In the present text, the welfare level attained by a person's income, given her type, is used as a synonym for equivalent income or living standard. Therefore, if $\mathbf{u}$ denotes living standards, the welfare ranking is said to be welfarist. We consider a population of individuals with fixed size n. Ranking heterogenous populations with different size and welfare, will be discussed in section 4.

For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, the notation $\mathbf{u}R\mathbf{v}$ means that distribution $\mathbf{u}$ is weakly preferred to distribution $\mathbf{v}$ according to the ranking R. Associated with a weak preference relation R, there is:

the *strict preference relation*, denoted by $\mathbf{u}P\mathbf{v}$, and defined as: $\mathbf{u}R\mathbf{v}$ and not $\mathbf{v}R\mathbf{u}$;

the *indifference relation*, denoted by $\mathbf{u}I\mathbf{v}$, defined as: $\mathbf{u}R\mathbf{v}$ and $\mathbf{v}R\mathbf{u}$.

All the rankings presented here satisfy *anonymity*: any permutation of a distribution $\mathbf{u}$ is considered to be equally good as that distribution. Anonymity allows to focus attention on the domain of ordered distributions, denoted by the set $\mathbb{D} \equiv \{\mathbf{u} \in \mathbb{R}^n \,| u_1 \leq \ldots \leq u_n\}$, in the sequel.

Definition 1 *Given any $r \in \mathbb{R}_+$, the r-extended* GLD *quasi-ordering $R(r)$ is defined as:*

$$\mathbf{u}R(r)\,\mathbf{v} \Leftrightarrow \sum_{i=1}^{k} (1+r)^{k-i} (u_i - v_i) \geq 0 \text{ for all } k \in N,$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$.

Increasing r, increases the discriminatory power of this rule, *i.e.* the rule is able to rank more distributions. Formally, this statement means:

Proposition 1 *For all $q, r \in \mathbb{R}_+$ such that $q \geq r$:*

(a) $I(r) \subseteq I(q)$

(b) $P(r) \subseteq P(q)$.

Proof: see appendix.

In order to define the *second* collection of rankings, we introduce a set of positive, non-decreasing, and normalized weight vectors $\mathbb{W} \equiv \{\mathbf{w} \in \mathbb{R}^n \,| w_1 \geq \ldots \geq w_n = 1\}$.

Definition 2 *Given any $r \in \mathbb{R}_+$ and a weight vector $\mathbf{w} \in \mathbb{W}$, an r-generalized Gini ordering, $R(r, \mathbf{w})$ is defined as:*

$$\mathbf{u} R(r, \mathbf{w}) \mathbf{v} \Leftrightarrow \sum_{i=1}^{n} (1+r)^{n-i} w_i (u_i - y_i) \geq 0,$$

for all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$.

We will denote the family of r-generalized Ginis by $\mathcal{G}(r) \equiv \{R(r, \mathbf{w}) | \mathbf{w} \in \mathbb{W}\}$. It is the subset of all generalized Ginis for which the relative weight of two consecutive rank positions is at least $1 + r$.
Increasing r decreases the size of the family of r-generalized Ginis. More precisely:

Proposition 2 *For all $q, r \in \mathbb{R}_+$, such that $q \geq r$: $\mathcal{G}(q) \subseteq \mathcal{G}(r)$.*

Proof: see appendix.

When $r = 0$, the r-extended GLD quasi-ordering corresponds with the standard GLD quasi-ordering, while the r generalized Ginis then correspond with the whole set of Weymark's (1981) generalized Ginis. When r approaches ∞, we can obtain in both cases the leximin ordering.[3]
The next proposition provides a link between both sets of rankings: it shows that the r-extended GLD quasi-ordering is equivalent with unanimity among the members of the family of the corresponding r-generalized Gini orderings. Formally:

Proposition 3 *For all $r \in \mathbb{R}_+$: $R(r) = \bigcap_{\mathbf{w} \in \mathbb{W}} R(r, \mathbf{w})$.*

Proof: see appendix.

All the rankings, presented here, satisfy the following properties.

Continuity.

For any sequence of distributions $(\mathbf{u}^m)_{m \in \mathbb{N}_0}$, if there exists an $M \in \mathbb{N}_0$ such that for some $\mathbf{v} \in \mathbb{R}^n$: $\mathbf{u}^m R \mathbf{v}$ for all $m \geq M$ then $\lim_{m \to \infty} \mathbf{u}^m \equiv \mathbf{u}^* R \mathbf{v}$; or, alternatively, if $\mathbf{v} R \mathbf{u}^m$ for all $m \geq M$ then $\mathbf{v} R \mathbf{u}^*$.

Strong Pareto (SP).

For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$: if $\mathbf{u} \geq \mathbf{v}$, then $\mathbf{u} R \mathbf{v}$; if, in addition, $\mathbf{u} \neq \mathbf{v}$, then $\mathbf{u} P \mathbf{v}$.

[3]As there are different possibilities to take the limit of a sequence of rankings, other limiting cases might be obtained, such as the maximin rule (see Hammond, 1975).

We introduce a set of new equality preference axioms, one for each scalar $r \geq 0$.

r-extended Pigou-Dalton transfer principle (PD (r)).

For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, such that for some $i, j \in N$, (i) $u_k = v_k$, for all $k \neq i, j$, (ii) $v_i < u_i \leq u_j < v_j$, and (iii) $\frac{v_j - u_j}{u_i - v_i} = 1 + r$, it follows that $\mathbf{u} R \mathbf{v}$.

For any $r \geq 0$, the r-GLD ranking and all members of the r-Gini family $\mathcal{G}(r)$, satisfy the corresponding PD(r)-principle.

The PD(r)-principles are modifications of the standard Pigou-Dalton transfer principle. They state that an equalizing transfer between two individuals –which does not change their relative positions– does not lower social welfare, whenever the loss to the donor $(v_j - u_j)$ is exactly equal to $1 + r$ times the gains of the receiver $(u_i - v_i)$. For later use, we call a transfer which satisfies conditions (i)-(iii) in the definition of the r-extended Pigou Dalton transfer principle a PD(r)-*transfer*. Under the domain restriction $\mathbb{D}$, associated with anonymity, it might occur that the after transfer vector $\mathbf{u}$ does not belong any more to $\mathbb{D}$. By a slight abuse of notation, and without loss of generality, we will always read $\mathbf{u}$ to be the after transfer rank-reordered vector belonging to $\mathbb{D}$.

Choosing $r = 0$, leads to the standard Pigou-Dalton transfer principle. Furthermore, if R satisfies the strong Pareto principle, then increasing r increases the strength of the transfer principle, i.e., PD$(r) \Rightarrow$ PD(q), for all q in $[0, r]$. This can be seen as follows: assume that we impose PD(r), and assume that $\mathbf{u}$ is obtained from $\mathbf{v}$ via a PD(q)-transfer of size $\delta > 0$, for $q \leq r$ such that $\mathbf{u} = (v_1, \ldots, v_{i-1}, v_i + \delta, v_{i+1}, \ldots, v_{j-1}, v_j - (1+q)\delta, v_{j+1}, \ldots, v_n)$. Then construct $\mathbf{u}'$ from $\mathbf{v}$, by means of a PD(r)-transfer of size $\epsilon = \delta \frac{1+q}{1+r} \leq \delta$. Then $u'_j = v_j - (1+r)\delta \frac{1+q}{1+r} = u_j$ and $u'_i = v_i + \delta \frac{1+q}{1+r} \leq u_i = v_i + \delta$. Hence, by SP, $\mathbf{u} R \mathbf{u}'$ and by PD(r), $\mathbf{u}' R \mathbf{v}$. So, by transitivity of R, we obtain $\mathbf{u} R \mathbf{v}$.

Thus, under the strong Pareto principle, PD(r) is a strengthening of the standard Pigou-Dalton transfer principle. Then, we can adopt the convention that PD(∞) means that PD(r) holds for all $r \geq 0$. If so, we obtain (a slightly stronger version of) the Hammond equity principle.[4]

3 Inequality aversion and efficiency reconciled again

The r-extended Pigou-Dalton transfer principle is at the heart of our new rules. It is also easy to understand why it allows welfarist rules to satisfy the BTPD transfer principle (a

[4]Hammond's (1976) equity principle states that $\mathbf{u} R \mathbf{v}$ if $\exists\, i, j \in N$ such that (i) $u_k = v_k$, for all $k \neq i, j$, (ii) $v_i < u_i < u_j < v_j$. Contrary to Hammond's equity principle, our principle also applies when $u_i = u_j$.

formal definition of that principle is provided below). Recall example 1: as Mary is a more efficient equivalent income generator than Eve, any income transfer from Mary to Eve leads to a loss in total equivalent income. Indeed, as we go from a(n individual equivalent income) distribution $(4, 18)$ to $(8, 12)$, 2 equivalent income units are lost. But such a leak during the transfer is precisely what is allowed for by the r-extended Pigou-Dalton transfer principle: it suffices to choose $r = \frac{1}{2}$ (or larger), to see that $(8, 12)$ can be derived from $(4, 18)$ via a PD$\left(\frac{1}{2}\right)$-transfer. We will clarify below how to choose r in order to guarantee satisfaction of the BTPD-transfer principle.

We turn now to the application of the rules presented in the previous section to the problem of welfare comparisons among heterogenous populations as defined in section 1. Individuals in such a setting do not only differ with respect to the level of income they obtain, say y_i, but also with respect to other characteristics. People who are identical in all (relevant) aspects, except possibly for the level of income they obtain, are said to be of the same type, say $\theta_i = \theta_j$, where $\theta_i, \theta_j \in \Theta$, the set of all possible types. Welfare economics for homogenous populations is concerned with ranking income vectors: in the terminology of the previous section this means $\mathbf{u} = (y_1, \ldots, y_i, \ldots, y_n)$. Comparing heterogenous populations amounts to looking for suitable rankings of the vectors $(\mathbf{y}, \boldsymbol{\theta}) \equiv ((y_1, \theta_1), \ldots, (y_i, \theta_i), \ldots, (y_n, \theta_n))$. We limit ourselves to a setting where this problem can be converted into one of ranking equivalent incomes or welfare levels. More specifically, we define equivalent incomes by means of ratio equivalence scales: the equivalent income vector is then $e(\mathbf{y}, \boldsymbol{\theta}) \equiv (e_1(\mathbf{y}, \boldsymbol{\theta}), \ldots e_i(\mathbf{y}, \boldsymbol{\theta}), \ldots, e_n(\mathbf{y}, \boldsymbol{\theta})) \equiv \left(\frac{y_1}{m(\theta_1)}, \ldots, \frac{y_i}{m(\theta_i)}, \ldots, \frac{y_n}{m(\theta_n)}\right)$, where $m(\theta_i)$ is type θ_i's equivalence scale. In the language of the previous section, the vectors we now rank are those equivalent income vectors: $\mathbf{u} = e(\mathbf{y}, \boldsymbol{\theta})$. It is therefore natural to limit the domain of the vectors $\mathbf{u}$ to the non-negative part of the n-dimensional real vector space, $\mathbb{R}^n_+$, as we will do from now on.

In this context, the *between type Pigou-Dalton principle* and *Pareto indifference* are properties of a ranking, say R, defined on income and type distributions belonging to $\mathbb{R}^n_+ \times \Theta^n$.

Pareto indifference (PI).
For all $(\mathbf{y}, \boldsymbol{\theta}), (\mathbf{x}, \boldsymbol{\zeta}) \in \mathbb{R}^n_+ \times \Theta^n$: if $e(\mathbf{y}, \boldsymbol{\theta}) = e(\mathbf{x}, \boldsymbol{\zeta})$ then $(\mathbf{y}, \boldsymbol{\theta}) I (\mathbf{x}, \boldsymbol{\zeta})$.

Between type Pigou-Dalton principle (BTPD).
For all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n_+$ and any $\boldsymbol{\theta} \in \Theta^n$: if (i) $x_k = y_k$, for all $k \neq i, j$, (ii) $e_i(\mathbf{y}, \boldsymbol{\theta}) < e_i(\mathbf{x}, \boldsymbol{\theta}) \leq e_j(\mathbf{x}, \boldsymbol{\theta}) < e_j(\mathbf{y}, \boldsymbol{\theta})$, and (iii) $x_i - y_i = y_j - x_j$, then $(\mathbf{x}, \boldsymbol{\theta}) R (\mathbf{y}, \boldsymbol{\theta})$.

We show first *how* it is possible to reconcile the Pareto indifference principle and the

BTPD transfer principle by means of the rules introduced in the previous section. Let us return to the motivating examples introduced in section 1. If we choose r sufficiently high, for example $r \geq 1/2$, and apply the r-extended GLD quasi-ordering to the individual equivalent income vectors of example 1 and 2, then both the BTPD transfer principle (example 1) as well as the Pareto indifference principle (example 2) are satisfied.

Proposition 4 *The r-extended* GLD *quasi-ordering (resp. any member of the r-generalized Gini orderings) applied to individual equivalent income distributions, allows to reconcile the* BTPD *transfer principle and the Pareto indifference principle, if and only if*

$$r \geq \frac{\max_{\theta \in \Theta} \{m(\theta)\}}{\min_{\theta \in \Theta} \{m(\theta)\}} - 1.$$

Proof: see appendix.

Why then is it possible to reconcile both principles, despite the claim in the literature that these principles are incompatible? The weights in the proposed rankings are power terms of $1 + r$, where the power depends on the (individual) equivalent income position (rank). Rank order weights are however excluded in (i) Ebert (1997), who imposes separability, (ii) Ebert and Moyes (2003), where the weights can only depend on the reference type and the own type, and (iii) Shorrocks (1995), who imposes differentiability of the welfare ranking. It is fair to mention that Shorrocks actually describes a two-person r-generalized Gini ordering as a possible solution (when relaxing differentiability), but he discards this possibility for practical purposes, without further explanation.

Equivalence scales are mostly applied when individuals differ (only) with respect to the size of the household to which they belong. For practical applications, this context may raise some complications if only household incomes are observed or if only transfers between household incomes are available to the government. Indeed, in that case we are asked to judge, for example transfers from households with few happy members to large scale families, populated however by non-efficient welfare producers. These problems are reminiscent of the deep ethical problems when judging populations of variable size, to which we return in the next section. For a fixed size population, many of the problems can however be circumvented by introducing the concept of a *per capita equivalence scale*. Given a set of individuals, N, the set of partitions of N (a partition of N is a set of nonempty and non-overlapping subsets of N, the union of which equals N), denoted by $\mathcal{H}(N)$, constitutes the set of possible household constellations for this society. A household constellation is a description of the way individuals decide to join together in households; these are, basically, income pooling units. For a given household constellation, say $\hbar \in \mathcal{H}(N)$, let $H(\hbar) = \{1, 2, \ldots, h, \ldots\}$ be the index set for the households in that constellation to which individuals belong. Individuals can then be indexed by the correspondence $\boldsymbol{i} : H(\hbar) \rightarrow N$, such that $i \in \boldsymbol{i}(h)$ means that individual i belongs to household h. Equivalence scales are now associated with a household h and provide a measure to convert *household income*, into the nominal income level needed by a reference household (usually a single) to obtain the same welfare level as the members of that household. All non-income information concerning the household can be recollected into the *household* type, which will be denoted by $\varphi_h \in \Phi$, where Φ is the set of all possible household types. The equivalence scale is henceforth dependent on the household's type information (and the reference type): $m(\varphi_h)$.[5] The size of a household equals the number of its members: $s(\varphi_h) = |\boldsymbol{i}(h)|$. The *per capita equivalence scale* equals: $\frac{m(\varphi_h)}{s(\varphi_h)}$. It gives a measure of the extent to which there are (dis)economies of scale of living in larger units. An (income,type)-distribution is then a vector $(\mathbf{y}, \boldsymbol{\varphi}) \in \Delta \equiv \left\{ (\mathbf{x}, \boldsymbol{\psi}) \in \bigcup_{1 \leq m \leq n} \mathbb{R}_+^m \times \Phi^m \,\middle|\, \sum_{h=1}^{m} s(\psi_h) = n \right\}$. We follow the standard assumption, in this literature, that individual welfare or living standards within a household are equally distributed: $e_i(\mathbf{y}, \boldsymbol{\varphi}) = \frac{y_h}{m(\varphi_h)}$ for all $i \in \boldsymbol{i}(h)$.[6] A reformulation of the Pareto indifference criterion and the between type Pigou-Dalton principle in this context is straightforward. Notice, however, that the between type Pigou Dalton principle considers a monetary transfer between households, such that the living standard of all its members is affected. This is embodied in the notation: (i) $x_g = y_g$, for all $g \in H(\hbar) : g \neq h, h'$, (ii) $e_i(\mathbf{y}, \boldsymbol{\varphi}) < e_i(\mathbf{x}, \boldsymbol{\varphi}) \leq e_j(\mathbf{x}, \boldsymbol{\varphi}) < e_j(\mathbf{y}, \boldsymbol{\varphi})$ for all $i \in \boldsymbol{i}(h), \;\; j \in \boldsymbol{i}(h')$, and (iii) $x_h - y_h = y_{h'} - x_{h'}$. We get the following, slightly weaker, adjoint result to proposition 4:

Corollary to proposition 4 *The r-extended* GLD *quasi-ordering (resp. any member of the r-generalized Gini orderings) applied to individual equivalent income distributions, allows to reconcile the* BTPD *transfer principle and the Pareto indifference principle, if*

$$r \geq \frac{\max\limits_{\varphi \in \Phi} \{m(\varphi)/s(\varphi)\}}{\min\limits_{\varphi \in \Phi} \{m(\varphi)/s(\varphi)\}} - 1.$$

Proof: see appendix.

[5]By a slight abuse of notation, we continue to denote equivalence scale functions by m.

[6]Alternatively, it could be assumed that nominal household income is equally distributed and equivalence scale depend solely on individual type information, converting per capita incomes into individual living standards (as f.e. in Shorrocks, 1995).

4 Variable population size

Finally, we turn to applications of the presented rules to the comparison of heterogenous populations of variable size. We therefore have to extend the domain of the rules to the set $\mathcal{R} \equiv \bigcup_{n \in \mathbb{N}_0} \mathbb{R}^n$. The r-generalized Ginis are readily extended to this domain by selecting for each population size $n \in \mathbb{N}_0$, a weight vector of length n, say $\mathbf{w}_n \in \mathbb{W}_n$ where $\mathbb{W}_n \equiv \{\mathbf{w}_n \in \mathbb{R}^n \,| w_{1,n} \geq \ldots \geq w_{i,n} \geq \ldots \geq w_{n,n} = 1\}$.

For making comparisons of vectors of different size, we will use replications of a vector $\mathbf{u}$. Given a natural number, say $m \in \mathbb{N}_0$, a m-replication of a n-vector $\mathbf{u}$, denoted by $\lambda_m(\mathbf{u})$, is defined as:

$$\lambda_m(\mathbf{u}) \equiv \left(\underbrace{u_1, \ldots, u_1}_{m \text{ times}}, \ldots, \underbrace{u_n, \ldots, u_n}_{m \text{ times}} \right).$$

The variable population size equivalent of the r-extended GLD ranking is defined as follows:

$$\forall \mathbf{u} \in \mathbb{D}_n, \mathbf{v} \in \mathbb{D}_m :$$

$$\mathbf{u}R(r)\mathbf{v} \Leftrightarrow \sum_{i=1}^{k} \frac{(1+r)^{(m \cdot n - i)}}{(1+r)^{m \cdot n} - 1} \left(\lambda_m(\mathbf{u})_i - \lambda_n(\mathbf{v})_i\right) \geq 0 \text{ for all } k \leq m \cdot n,$$

where $\mathbb{D}_n$, with $n \in \mathbb{N}_0$, is the set of rank re-ordered real vectors of length n.

A problem occurs if we require the ranking to be *replication invariant*, which means that social welfare must not change when replicating the distribution a number of times.

Replication invariance (RI)
For all $n \in \mathbb{N}_0$, for all $\mathbf{u} \in \mathbb{R}^n$, for all $m \in \mathbb{N}_0$: $\mathbf{u}I\lambda_m(\mathbf{u})$.

The next proposition states that replication invariance reinforces the r-extended Pigou-Dalton transfer principle:

Proposition 5 *For all $r > 0$,* SP*,* PD(r) *and* RI *imply* PD(q) *for all $q \in \mathbb{R}_+$.*

Proof: see appendix.

Recall that PD(r) for all $r \in \mathbb{R}_+$ corresponds with Hammond equity. As a consequence of the last proposition, given $r > 0$, there is only one ordering satisfying SP, PD(r) and RI: it is the leximin rule applied to (possibly replicated) distributions of equal length. Requiring in addition continuity, would lead to an impossibility result.

This result must not come as a surprise. The combination of the Pareto criterion with the PD(r)-principle leads to a specific trade-off between mean equivalent income and a more equal distribution of living standards: the amount of mean equivalent income one wants to give up increases with the number of people in society, even if inequality does not change. Replication invariance goes against this principle, but is not uncontroversial *per se*: it takes a specific stance in the discussion between average and total sum utilitarians (opting resolutely for the former).

Admittedly, it is even impossible for an anonymous, complete and continuous welfare ranking which satisfies the PD(r) principle and the strong Pareto criterion, to satisfy simultaneously the weaker population principle:

Population principle (PP)
For all $n \in \mathbb{N}_0$, for all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, for all $m \in \mathbb{N}_0$: $\mathbf{u} I \mathbf{v}$ if and only if $\lambda_m(\mathbf{u}) I \lambda_m(\mathbf{v})$.

The latter criterion does not impose a specific choice between average and total sum utilitarianism, but poses the questions involved in comparing societies with different population size in those terms.

Corollary to proposition 5 *For any $r > 0$, there is no member of the r-generalized Ginis, $\mathcal{G}(r)$, which satisfies* PP.

Proof: see appendix.

There are other consistency requirements between rankings of distributions among populations of different size, such as the restricted aggregation principles (see Ebert, 1988). Let $\mathbf{u}_k$ be a distribution in a population of size k: $\mathbf{u}_k \in \mathbb{R}^k$. For any $k \in \mathbb{N}_0 : k \geq 2$, and any vector $\mathbf{u}_k \in \mathbb{R}^k$, define the *equally distributed equivalent* $\xi(\mathbf{u}_k)$ to be the outcome which, if it were obtained by all individuals in society, would yield the same welfare as the stated distribution $\mathbf{u}_k$: $\lambda_k(\xi(\mathbf{u}_k)) I \mathbf{u}_k$.

Restricted aggregation from above (RAA)
For all $n \in \mathbb{N}_0 : n \geq 3$, for all $\mathbf{u} \in \mathbb{R}^n$, for all $k : 0 \leq k \leq n$:

$$\mathbf{u} I (u_1, \ldots, u_k, \lambda_{n-k}(\xi(u_{k+1}, \ldots, u_n))).$$

The latter vector should be read as $\lambda_n(\xi(\mathbf{u}))$ if $k = 0$.

Restricted aggregation from below (RAB)
For all $n \in \mathbb{N}_0 : n \geq 3$, for all $\mathbf{u} \in \mathbb{R}^n$, for all $k : 1 \leq k \leq n$:

$$\mathbf{u} I (\lambda_k(\xi(u_1, u_2, \ldots, u_k)), u_{k+1}, \ldots, u_n).$$

The latter vector should be read as $\lambda_n(\xi(\mathbf{u}))$ if $k = n$.

The following result is obtained:

Proposition 6 *For any $r > 0$, the r-generalized Ginis such that, for each $n \in \mathbb{N}_0 : n \geq 3$, $\mathbf{w}_n = (w_{1,n}, \ldots, w_{i,n}, \ldots, w_{n,n}) \in \mathbb{W}_n$ satisfies: $w_{i,n} = b^{n-i}$ for some scalar $b \geq 1$, satisfy* RAA *and* RAB*.*

Proof: see appendix.

5 Conclusion

In this note we showed how it is possible to circumvent the claimed incompatibility between efficiency (Pareto indifference) and inequality aversion (between type Pigou-Dalton transfer principle) when comparing heterogenous individuals. The intuition for our result is straightforward: the proposed r-extended GLD quasi-ordering can be interpreted as an *extension* of the standard GLD quasi-ordering. While the latter can only approve classical Pigou-Dalton transfers, the former also approves some transfers which might cause a decrease in average equivalent income. In this sense, we say that the r-extended GLD quasi-ordering is more inequality averse than its ordinary counterpart. Similarly, for each member of the class of r-generalized Ginis, the rate of substitution between two consecutive equivalent income positions is at least $1 + r$. By using rank dependent weights, we are able to define the minimal amount of inequality aversion a welfare ranking should exhibit in order to meet the between type Pigou-Dalton criterion, independent of the given equivalent income distribution and the size of transfer, while at the same retaining the anonymity criterion (abandoned by Ebert, 1997). In this way, we could determine a set of sufficiently inequality averse welfare rankings, without having to violate the Pareto criterion.

Appendix

Proof of proposition 1

(a) follows from the fact that for all $r \in \mathbb{R}_+$ and all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$: $\mathbf{u}I(r)\mathbf{v}$ if and only if $\mathbf{u} = \mathbf{v}$ (indifference sets associated with $R(r)$ defined on $\mathbb{D}$, are singletons).

To prove (b), suppose $\mathbf{u}P(r)\mathbf{v}$ holds, for some $r \in \mathbb{R}_+$ and for some $\mathbf{u}, \mathbf{v} \in \mathbb{D}$. Abbreviate $\Delta_i = u_i - v_i$ for all $i \in N$; we will prove, for all $q \geq r$, that

$$\sum_{i=1}^{k} (1+q)^{k-i} \Delta_i \geq \sum_{i=1}^{k} (1+r)^{k-i} \Delta_i, \text{ for all } k \in N, \tag{1}$$

which indeed leads to $\mathbf{u}P(q)\mathbf{v}$, because $\mathbf{u}P(r)\mathbf{v}$ implies that the right hand-side summation is non-negative for all k (and positive for at least one k).

Proof by induction. First, (1) is obvious for $k = 1$. Secondly, suppose (1) holds for some $1 \leq k < n$ (induction hypothesis), then it also holds for $k+1$. We rewrite

$$\sum_{i=1}^{k+1} (1+q)^{(k+1)-i} \Delta_i = \underbrace{(1+q)}_{A_1} \overbrace{\sum_{i=1}^{k} (1+q)^{k-i} \Delta_i}^{B_1} + \Delta_{k+1}.$$

Given (α) $q \geq r \geq 0$, (β) the induction hypothesis, and (γ) $\mathbf{u}P(r)\mathbf{v}$ we have

$$A_1 = 1 + q \overset{(\alpha)}{\geq} \underbrace{1+r}_{A_2} \overset{(\alpha)}{>} 0 \text{ and } B_1 = \sum_{i=1}^{k} (1+q)^{k-i} \Delta_i \overset{(\beta)}{\geq} \underbrace{\sum_{i=1}^{k} (1+r)^{k-i} \Delta_i}_{B_2} \overset{(\gamma)}{\geq} 0.$$

The desired result follows, because

$$A_1 B_1 + \Delta_{k+1} \geq A_2 B_2 + \Delta_{k+1} \Rightarrow \sum_{i=1}^{k+1} (1+q)^{(k+1)-i} \Delta_i \geq \sum_{i=1}^{k+1} (1+r)^{(k+1)-i} \Delta_i.$$

Proof of proposition 2

Consider $r, q \in \mathbb{R}_+$ with $q \geq r$. We show that for each rule $R(q, \mathbf{w})$ in $\mathcal{G}(q)$ there exists a weight vector $\mathbf{w}' \in \mathbb{W}$ such that $R(q, \mathbf{w}) = R(r, \mathbf{w}') \in \mathcal{G}(r)$. Abbreviate $\Delta_i = u_i - v_i$ for all $i \in N$; $R(q, \mathbf{w})$ is defined as follows:

for all distributions $\mathbf{u}, \mathbf{v} \in \mathbb{D}$,

$$\mathbf{u}R(q, \mathbf{w})\mathbf{v} \Leftrightarrow \sum_{i=1}^{n} (1+q)^{n-i} w_i \Delta_i \geq 0.$$

Define a weight vector $\mathbf{w}' = \left(\frac{(1+q)^{n-i}}{(1+r)^{n-i}} w_i; i \in N\right)$, which belongs to $\mathbb{W}$; this leads to the desired result.

Proof of proposition 3

Actually, the proposition says:
$\forall\, r \in \mathbb{R}_+, \forall\, \mathbf{u}, \mathbf{v} \in \mathbb{D}:$

$$\mathbf{u}R(r)\mathbf{v} \Leftrightarrow \mathbf{u}R(r,\mathbf{w})\mathbf{v} \;\; \forall\, \mathbf{w} \in \mathbb{W}.$$

We show first that all orderings in $\mathcal{G}(r)$ are consistent with the r-extended GLD quasi-ordering (sufficiency): for all $r \in \mathbb{R}_+$ and for all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$: $\mathbf{u}R(r)\mathbf{v}$ implies $\mathbf{u}R(r,\mathbf{w})\mathbf{v}$ for all $\mathbf{w} \in \mathbb{W}$. Let $\Delta_i \equiv u_i - v_i$ for all $i \in N$. $\mathbf{u}R(r)\mathbf{v}$ implies (start with $k = n$)

$$\begin{aligned}
\sum_{i=1}^{n} (1+r)^{n-i} \Delta_i &\geq 0 \\
\underbrace{(1+r)}_{>0} \overbrace{\sum_{i=1}^{n-1} (1+r)^{n-1-i} \Delta_i}^{\text{non-negative via } k=n-1} + \alpha_n \Delta_n &\geq 0, \text{ with } \alpha_n = 1 \\
(1+r)\,\alpha_{n-1} \sum_{i=1}^{n-1} (1+r)^{n-1-i} \Delta_i + \alpha_n \Delta_n &\geq 0, \; \forall \alpha_{n-1} \geq \alpha_n \\
\underbrace{(1+r)^2 \alpha_{n-1}}_{>0} \overbrace{\sum_{i=1}^{n-2} (1+r)^{n-2-i} \Delta_i}^{\text{non-negative via } k=n-2} + \sum_{i=n-1}^{n} (1+r)^{n-i} \left(\prod_{k=i}^{n} \alpha_k\right) \Delta_i &\geq 0, \forall \alpha_{n-1} \geq \alpha_n \\
(1+r)^2 \alpha_{n-2}\alpha_{n-1} \sum_{i=1}^{n-2} (1+r)^{n-2-i} \Delta_i + \sum_{i=n-1}^{n} (1+r)^{n-i} \left(\prod_{k=i}^{n} \alpha_k\right) \Delta_i &\geq 0, \forall \alpha_{n-2}, \alpha_{n-1} \geq \alpha_n \\
\underbrace{(1+r)^3 \alpha_{n-2}\alpha_{n-1}}_{\cdots} \overbrace{\sum_{i=1}^{n-3} (1+r)^{n-3-i} \Delta_i}^{\cdots} + \sum_{i=n-2}^{n} (1+r)^{n-i} \left(\prod_{k=i}^{n} \alpha_k\right) \Delta_i &\geq 0, \forall \alpha_{n-2}, \alpha_{n-1} \geq \alpha_n \\
\ldots & \\
\sum_{i=1}^{n} (1+r)^{n-i} \left(\prod_{k=i}^{n} \alpha_k\right) \Delta_i &\geq 0, \text{ for all } \alpha_1, \ldots, \alpha_{n-1} \geq \alpha_n = 1. \text{ Hence,} \\
\sum_{i=1}^{n} (1+r)^{n-i} w_i \Delta_i &\geq 0, \text{ for all } \mathbf{w} \in \mathbb{W}, \text{ as required.}
\end{aligned}$$

Conversely, for all $r \in \mathbb{R}_+$ and for all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$: $\mathbf{u}R(r,\mathbf{w})\mathbf{v}$ for all $\mathbf{w} \in \mathbb{W}$ implies $\mathbf{u}R(r)\mathbf{v}$. We prove this statement by contradiction. Suppose thus that there exists an $r \in \mathbb{R}_+$ and $\mathbf{u}, \mathbf{v} \in \mathbb{D}$ such that $\mathbf{u}R(r,\mathbf{w})\mathbf{v}$ for all $\mathbf{w} \in \mathbb{W}$, but not $\mathbf{u}R(r)\mathbf{v}$, *i.e.*, there must exist a $k \in N$ such that $\sum_{i=1}^{k} (1+r)^{k-i} \Delta_i < 0$. Choose a family of weight vectors

$$\mathbb{W}(k) = \{\mathbf{w} \in \mathbb{W} \,|\, w_1 = \ldots = w_k \geq w_{k+1} = \ldots = w_n = 1\}.$$

Notice that $\mathbb{W}(k) \subset \mathbb{W}$. Therefore, we must have

$$\begin{aligned}
\sum_{i=1}^{n} (1+r)^{n-i} w_i \Delta_i &\geq 0, \text{ for all } w \in \mathbb{W}(k) \\
\sum_{i=1}^{k} (1+r)^{n-i} w_1 \Delta_i + \sum_{i=k+1}^{n} (1+r)^{n-i} \Delta_i &\geq 0, \text{ for all } w_1 \geq 1 \\
\underbrace{w_1 (1+r)^{n-k}}_{>0} \overbrace{\sum_{i=1}^{k} (1+r)^{k-i} \Delta_i}^{<0} + \sum_{i=k+1}^{n} (1+r)^{n-i} \Delta_i &\geq 0, \text{ for all } w_1 \geq 1,
\end{aligned}$$

which is false (as w_1 can be chosen arbitrarily high to make the expression strictly negative).

Proof of proposition 4

Notice that Pareto indifference is satisfied, since the proposed quasi-orderings are defined on the domain of distributions of equivalent incomes. We show that the BTPD transfer principle is satisfied by an r-generalized Gini ordering, if and only if we choose r sufficiently large, more precisely:

$$r \geq \frac{\max_{\theta \in \Theta} \{m(\theta)\}}{\min_{\theta \in \Theta} \{m(\theta)\}} - 1.$$

Due to proposition 3, the result then also holds for the r-extended GLD quasi-ordering.

First, consider a "worst" case scenario ("worst" in terms of a decrease in total equivalent income) of transferring money between two persons: the decrease in total equivalent income is largest, if we transfer an amount of income $\epsilon > 0$ from a household with the minimal per-capita equivalence scale, denoted type $\underline{\theta} = \arg\min_{\theta \in \Theta} \{m(\theta)\}$ to a household with the maximal per-capita equivalence scale (and a lower equivalent income), denoted type $\overline{\theta} = \arg\max_{\theta \in \Theta} \{m(\theta)\}$. Assume that both persons occupy adjacent rank positions. According to an r-generalized Gini-ordering, such a transfer will not be welfare decreasing if

$$(1+r) \left(m\left(\overline{\theta}\right)\right)^{-1} \geq \left(m\left(\underline{\theta}\right)\right)^{-1}. \quad (2)$$

Secondly, any other transfer might, in general, change the equivalent income positions of individuals. But it suffices to notice that (i) each transfer can be decomposed into a finite number of welfare improving transfers between two persons, which do not change the equivalent income positions in the distribution, and (ii) whenever two persons have the same equivalent income, (equivalent income) positions can be attributed arbitrarily (without changing total welfare). This completes the proof.

Proof of the corollary to proposition 4

The same logic as for the previous proof is followed. So, Pareto indifference is satisfied, since the proposed quasi-orderings are defined on the domain of distributions of *individual* equivalent incomes. We show that the BTPD transfer principle is satisfied by an r-generalized Gini ordering, if we choose r sufficiently large, more precisely:

$$r \geq \frac{\max_{\varphi\in\Phi}\{m(\varphi)/s(\varphi)\}}{\min_{\varphi\in\Phi}\{m(\varphi)/s(\varphi)\}} - 1.$$

Due to proposition 3, the result then also holds for the r-extended GLD quasi-ordering.

First, consider the "worst" ("worst" in terms of a decrease in total equivalent income) possible transfer of money between two *households*, such that the number of households with an equivalent income strictly in between both households remains fixed, say $a \geq 0$. The decrease in total equivalent income is largest, if we transfer an amount of income $\epsilon > 0$ from a household with the minimal per-capita equivalence scale, denoted type $\underline{\varphi} = \arg\min_{\varphi\in\Phi}\{m(\varphi)/s(\varphi)\}$ to a household with the maximal per-capita equivalence scale (and a lower equivalent income), denoted type $\overline{\varphi} = \arg\max_{\varphi\in\Phi}\{m(\varphi)/s(\varphi)\}$. According to an r-generalized Gini-ordering, such a transfer will not be welfare decreasing if

$$(1+r)^a \frac{\sum_{i=1}^{s(\overline{\varphi})}(1+r)^{s(\underline{\varphi})+s(\overline{\varphi})-i}}{m(\overline{\varphi})} \geq \frac{\sum_{i=s(\overline{\varphi})+1}^{s(\underline{\varphi})+s(\overline{\varphi})}(1+r)^{s(\underline{\varphi})+s(\overline{\varphi})-i}}{m(\underline{\varphi})}. \tag{3}$$

Because

$$(1+r)^a \sum_{i=1}^{s(\overline{\varphi})}(1+r)^{s(\underline{\varphi})+s(\overline{\varphi})-i} \geq s(\overline{\varphi})(1+r)^{s(\underline{\varphi})} \ ; \ \sum_{i=s(\overline{\varphi})+1}^{s(\underline{\varphi})+s(\overline{\varphi})}(1+r)^{s(\underline{\varphi})+s(\overline{\varphi})-i} \leq s(\underline{\varphi})(1+r)^{s(\underline{\varphi})-1},$$

equation (3) will be true if

$$\frac{s(\overline{\varphi})(1+r)^{s(\underline{\varphi})}}{m(\overline{\varphi})} \geq \frac{s(\underline{\varphi})(1+r)^{s(\underline{\varphi})-1}}{m(\underline{\varphi})}.$$

Rearranging terms leads to the desired result.

Secondly, for transfers that change the equivalent income positions of households, similar remarks apply as for the proof of proposition 4.

Proof of proposition 5

We first prove the following lemma:

Lemma 1 *For any $n \in \mathbb{N}_0$, with $n \geq 2$, for any $r \in \mathbb{R}_{++}$, for any $a \in \mathbb{R}$ and for all $\varepsilon > 0$, consider the following sequence of* PD(r)*-transfers, starting from a vector* $\mathbf{v}^1 = \left(a - (1+r)\,\varepsilon, \underbrace{a, \ldots, a}_{n-1 \text{ times}}\right)$*:*

1. *Construct $\mathbf{v}^2$ from $\mathbf{v}^1$ by means of a* PD(r)*-transfer from individual (at position) 2 to 1, such that the incomes of 1 and 2 are equal.*

2. *Construct $\mathbf{v}^3$ from $\mathbf{y}^2$ by means of* PD(r)*-transfers from 3 to 2 and from 3 to 1, such that the incomes of 1,2, and 3 are equal.*

   ...

$n-1$. *Construct $\mathbf{v}^n$ from $\mathbf{v}^{n-1}$ by means of* PD(r)*-transfers from n to all other individuals $n-1, \ldots, 1$, such that all incomes are equal.*

Following this procedure, we will end up with a distribution $\mathbf{v}^n = \left(\underbrace{b, \ldots, b}_{n \text{ times}}\right)$ *with*
$b = a - (1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)i}{1+(1+r)i}$.

Obviously, $\mathbf{v}^n R \mathbf{v}^1$ holds for any (quasi-)ordering satisfying PD(r).

Proof by induction. For $n = 2$ this result is obvious. Starting from $\mathbf{v}^1 = (a - (1+r)\,\varepsilon, a)$ the PD(r)-transfer δ which equalizes incomes is the one which satisfies

$$a - (1+r)\,\varepsilon + \delta = a - (1+r)\,\delta,$$

and thus $\delta = \frac{(1+r)\varepsilon}{1+(1+r)}$. We end up with $\mathbf{v}^2 = (b, b)$ with

$$\begin{aligned} b &= a - (1+r)\,\varepsilon + \delta = a - (1+r)\,\varepsilon + \frac{(1+r)\,\varepsilon}{1+(1+r)} \\ &= a - (1+r)\,\varepsilon \left(\frac{1+r}{1+(1+r)}\right). \end{aligned}$$

Suppose it holds for n (induction hypothesis); we show that it also holds for $n+1$. Start with a vector $\mathbf{v}^1 = \left(a - (1+r)\,\varepsilon, \underbrace{a, \ldots, a}_{n \text{ times}}\right)$. As it holds for n, we obtain after $n-1$ steps a vector $\mathbf{v}^n = \left(\underbrace{c, \ldots, c}_{n \text{ times}}, a\right)$ with

$$c = a - (1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)\,i}{1+(1+r)\,i}.$$

In the final step we perform $\text{PD}(r)$-transfers from $n+1$ to all other individuals such that incomes become equal. This transfer δ can be calculated as:

$$c+\delta = a-(1+r)\,n\delta,$$

and thus

$$\delta = \frac{a-c}{(1+(1+r)\,n)} = \frac{(1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)i}{1+(1+r)i}}{(1+(1+r)\,n)}.$$

We end up with $\mathbf{v}^{n+1} = \left(\underbrace{b,\ldots,b}_{n+1 \text{ times}}\right)$ and

$$\begin{aligned} b &= c+\delta = a-(1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)\,i}{1+(1+r)\,i} + \frac{(1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)i}{1+(1+r)i}}{(1+(1+r)\,n)} \\ &= a-(1+r)\,\varepsilon \prod_{i=1}^{n-1} \frac{(1+r)\,i}{1+(1+r)\,i}\left(1-\frac{1}{1+(1+r)\,n}\right) \\ &= a-(1+r)\,\varepsilon \prod_{i=1}^{n} \frac{(1+r)\,i}{1+(1+r)\,i}, \text{ as required.} \end{aligned}$$

■

We prove now that $\text{PD}(r)$ and RI imply $\text{PD}(t\,(m))$, for all $m \in \mathbb{N}_0$ with $m \geq 2$, and

$$t\,(m) = -1+(1+r)\,m \prod_{i=1}^{m-1} \frac{(1+r)\,i}{1+(1+r)\,i} = -1+\prod_{i=1}^{m} \frac{(1+r)\,i}{1+(1+r)\,(i-1)},$$

and that $\lim_{m\to\infty} t\,(m) = \infty$ whenever $r>0$. This will complete our proof for the following reason. Recall that, when SP holds, $\text{PD}(t) \Rightarrow \text{PD}(q)$, for all q in $[0,t]$. As m can be chosen arbitrarily large and given that $\lim_{m\to\infty} t\,(m) = \infty$, we obtain the desired result, i.e., $\text{PD}(q)$ must hold for all $q \in \mathbb{R}_+$.

First, we show that the limit $\lim_{m\to\infty} t\,(m)$ diverges. Notice that the sequence with elements

$$1+t\,(m) = \prod_{i=1}^{m} \left(\frac{(1+r)\,i}{1+(1+r)\,(i-1)}\right) \tag{4}$$

diverges if the series

$$\ln\left(1+t\,(m)\right) = \sum_{i=1}^{m} \ln\left(\frac{(1+r)\,i}{1+(1+r)\,(i-1)}\right) = \sum_{i=1}^{m} \ln\left(1+\frac{r}{1+(1+r)\,(i-1)}\right), \tag{5}$$

diverges. We prove that the series composed by the sequence $\{a_i\}_{i=1}^{\infty}$, with $a_i = \ln\left(1 + \frac{r}{1+(1+r)(i-1)}\right)$, diverges for $r > 0$. As all $a_i > 0$, for $r > 0$, it then must march off to infinity. As $\ln(1+x) \geq x - \frac{x^2}{2}$ for all $x \in \mathbb{R}_+$, it suffices to prove that the series $\{a_i^*\}_{n=1}^{\infty}$, with

$$a_i^* = \frac{r}{1+(1+r)(i-1)} - \frac{\left(\frac{r}{1+(1+r)(i-1)}\right)^2}{2} = \frac{2r(1+(1+r)(i-1)) - r^2}{2(1+(1+r)(i-1))^2}$$

diverges for $r > 0$. We use the de Morgan and Bertrand test. Define ρ_i (implicitly) by

$$\frac{a_i^*}{a_{i+1}^*} = \frac{2r(1+(1+r)(i-1)) - r^2}{2(1+(1+r)(i-1))^2} \frac{2(1+(1+r)i)^2}{2r(1+(1+r)i) - r^2} = 1 + \frac{1}{i} + \frac{\rho_i}{i \ln i},$$

and thus

$$\left(\frac{2r(1+(1+r)(i-1)) - r^2}{2(1+(1+r)(i-1))^2} \frac{2(1+(1+r)i)^2}{2r(1+(1+r)i) - r^2} - 1 - \frac{1}{i}\right) i \ln i = \rho_i.$$

The de Morgan and Bertrand test tells us that the series diverges, whenever $\lim\limits_{i\to\infty} \rho_i < 1$. We have

$$\begin{aligned}
\lim_{i\to\infty} \rho_i &= \lim_{i\to\infty}\left(\left(\frac{2r(1+(1+r)(i-1)) - r^2}{2(1+(1+r)(i-1))^2} \frac{2(1+(1+r)i)^2}{2r(1+(1+r)i) - r^2} - 1 - \frac{1}{i}\right) i \ln i\right) \\
&= \lim_{i\to\infty}\left(\left(r\frac{-2r + r^2 + i + 2ri^2 + i^2 - 2ri + r^2 i^2 - 3r^2 i}{i(i+ri-r)^2(2+2i+2ri-r)}\right) i \ln i\right) \\
&= \underbrace{\lim_{i\to\infty}\left(\frac{ri\left(-2r + r^2 + i + 2ri^2 + i^2 - 2ri + r^2 i^2 - 3r^2 i\right)}{(i+ri-r)^2(2+2i+2ri-r)}\right)}_{\frac{r}{2(r+1)}} \underbrace{\lim_{i\to\infty}\left(\frac{\ln i}{i}\right)}_{0} \\
&= 0 \text{ (for } r > 0).
\end{aligned}$$

Secondly, PD(r) and RI imply PD$(t(m))$, for all $m \in \mathbb{N}_0$ with $m \geq 2$. To prove this result, consider $\mathbf{u}, \mathbf{v} \in \mathbb{D}$, $m \in \mathbb{N}_0$, with $m \geq 2$, and $i, j \in N$ with (i) $u_k = v_k$, for all $k \neq i, j$, (ii) $v_i < u_i \leq u_j < v_j$, and (iii) $\frac{v_j - u_j}{u_i - v_i} = (1+r)\, m \prod\limits_{i=1}^{m-1} \frac{(1+r)i}{1+(1+r)i}$. Using PD$(r)$ and RI, we prove that $\mathbf{u}R\mathbf{v}$ holds. Define

$$\mathbf{v}^0 = \lambda_m(\mathbf{v}) = \left(\underbrace{u_1, \ldots, u_1}_{v_1^0 \text{ to } v_m^0}, \ldots, \underbrace{v_i, \ldots, v_i}_{v_{m(i-1)+1}^0 \text{ to } v_{mi}^0}, \ldots, \underbrace{v_j, \ldots, v_j}_{v_{m(j-1)+1}^0 \text{ to } v_{mj}^0}, \ldots, \underbrace{u_n, \ldots, u_n}_{v_{m(n-1)+1}^0 \text{ to } v_{mn}^0}\right).$$

Let $\varepsilon = u_i - v_i > 0$. Consider the following sequence of PD(r) transfers starting from $\mathbf{v}^0$:

1. Choose $v_k^1 = v_k^0$, for all $k \neq m(i-1)+1, m(j-1)+1$, $v_{m(i-1)+1}^1 = v_{m(i-1)+1}^0 + \varepsilon = u_i$ and $v_{m(j-1)+1}^1 = v_{m(j-1)+1}^0 - (1+r)\varepsilon$. Via PD$(r)$, we have $\mathbf{v}^1 R \mathbf{v}^0$.

2. Choose $v_k^2 = v_k^1$ for all $k \neq m(j-1)+1, \ldots, mj$, and noticing that the sub-vector

$$\left(v^1_{m(j-1)+1}, \ldots, v^1_{mj}\right) = \left(v_j - (1+r)\,\varepsilon, \underbrace{v_j, \ldots, v_j}_{m-1 \text{ times}}\right),$$

we can apply a sequence of transfers described in lemma 1 to obtain $\left(v^2_{m(j-1)+1}, \ldots, v^2_{mj}\right) = \left(\underbrace{b_1, \ldots, b_1}_{m \text{ times}}\right)$ with

$$b_1 = v_j - (1+r)\,\varepsilon \prod_{i=1}^{m-1} \frac{(1+r)\,i}{1+(1+r)\,i},$$

and $\mathbf{v}^2 R \mathbf{v}^1$.

3. Choose $v_k^3 = v_k^2$, for all $k \neq m(i-1)+2, m(j-1)+1$, $v^3_{m(i-1)+2} = v^2_{m(i-1)+2} + \varepsilon = u_i$ and $v^3_{m(j-1)+1} = v^2_{m(j-1)+1} - (1+r)\,\varepsilon$. Via PD$(r)$, we have $\mathbf{v}^3 R \mathbf{v}^2$.

4. Choose $v_k^4 = v_k^3$ for all $k \neq m(j-1)+1, \ldots, mj$, and noticing that the sub-vector

$$\left(v^1_{m(j-1)+1}, \ldots, v^1_{mj}\right) = \left(b_1 - (1+r)\,\varepsilon, \underbrace{b_1, \ldots, b_1}_{m-1 \text{ times}}\right),$$

we can apply a sequence of transfers described in lemma 1 to obtain $\left(v^4_{m(j-1)+1}, \ldots, v^4_{mj}\right) = \left(\underbrace{b_2, \ldots, b_2}_{m \text{ times}}\right)$ with

$$b_2 = b_1 - (1+r)\,\varepsilon \prod_{i=1}^{m-1} \frac{(1+r)\,i}{1+(1+r)\,i} = v_j - 2\,(1+r)\,\varepsilon \prod_{i=1}^{m-1} \frac{(1+r)\,i}{1+(1+r)\,i},$$

and $\mathbf{v}^4 R \mathbf{v}^3$.

5. Proceeding this way, we end up with

$$\mathbf{v}^{2m} = \left(\underbrace{u_1, \ldots, u_1}_{m \text{ times}}, \ldots, \underbrace{v_i + \varepsilon, \ldots, v_i + \varepsilon}_{m \text{ times}}, \ldots, \underbrace{b_m, \ldots, b_m}_{m \text{ times}}, \ldots, \underbrace{u_n, \ldots, u_n}_{m \text{ times}}\right),$$

and $\mathbf{v}^{2m} R \mathbf{y}^0$. Applying RI to both sides gives:

$$(u_1, \ldots, u_{i-1}, v_i + \varepsilon, u_{i+1}, \ldots, u_{j-1}, b_m, u_{j+1}, \ldots, u_n)\, R\mathbf{v}.$$

Since

$$\varepsilon = u_i - v_i \text{ and } b_m = v_j - \underbrace{m\,(1+r)\,\varepsilon \prod_{i=1}^{m-1} \frac{(1+r)\,i}{1+(1+r)\,i}}_{=v_j-u_j \text{ via (iii)}},$$

we obtain $\mathbf{u}R\mathbf{v}$, as required.

Proof of the corollary to proposition 5

For any $r > 0$, for any $n \in \mathbb{N}_0$ and any distribution $\mathbf{u} \in \mathbb{D}_n$, the equally distributed equivalent of $\mathbf{u}$, say $\xi(\mathbf{u})$ for the ordering $R(r, \mathbf{w}_n)$, with $\mathbf{w}_n = \left(\underbrace{1, \ldots, 1}_{n \text{ times}}\right) \equiv \mathbf{1}_n$, is defined as:

$$\xi(\mathbf{u}) = r \sum_{i=1}^{n} \frac{(1+r)^{(n-i)}}{(1+r)^n - 1} u_i.$$

Similarly, for an m-replication of $\mathbf{u}$, we obtain:

$$\xi(\lambda_m(\mathbf{u})) = r \sum_{j=1}^{m \cdot n} \frac{(1+r)^{(m.n-j)}}{(1+r)^{m \cdot n} - 1} \left(\lambda_m(\mathbf{u})\right)_j = \sum_{i=1}^{n} \frac{(1+r)^{m \cdot n-(i-1)m} - (1+r)^{m \cdot (n-i)}}{(1+r)^{m \cdot n} - 1} u_i,$$

the equally distributed equivalent for the ordering $R(r, \mathbf{1}_{m \cdot n})$.
For a general weighting vector, $\mathbf{w}_n \in \mathbb{W}_n$, the equally distributed equivalent of $\mathbf{u}$ with respect to $R(r, \mathbf{w}_n)$ is:

$$\xi_{\mathbf{w}_n}(\mathbf{u}) = \sum_{i=1}^{n} \frac{(1+r)^{(n-i)} w_{i,n}}{\sum_{j=1}^{n} (1+r)^{(n-j)} w_{j,n}} u_i.$$

Notice that $\xi_{\mathbf{w}_n}(\mathbf{u}) \geq \xi(\mathbf{u})$ for all $\mathbf{u} \in \mathbb{D}_n$.
PP implies that for some $\mathbf{w}_n \in \mathbb{W}_n$, it holds that:

$$\mathbf{u} I(r, \mathbf{w}_n) \lambda_n \left(\xi_{\mathbf{w}_n}(\mathbf{u})\right) \iff \lambda_m(\mathbf{u}) I(r, \mathbf{w}_{m \cdot n}) \lambda_{m.n}(\xi_{\mathbf{w}_n}(\mathbf{u})) \quad \text{for some } \mathbf{w}_{m \cdot n} \in \mathbb{W}_{m \cdot n}.$$

Consequently,

$$\xi_{\mathbf{w}_n}(\mathbf{u}) = \xi_{\mathbf{w}_{m \cdot n}}(\lambda_m(\mathbf{u})).$$

Consequently,

$$\sum_{i=1}^{n} \frac{(1+r)^{(n-i)} w_{i,n}}{\sum_{j=1}^{n} (1+r)^{(n-j)} w_{j,n}} u_i \geq \sum_{i=1}^{n} \frac{(1+r)^{m \cdot n-(i-1)m} - (1+r)^{m \cdot (n-i)}}{(1+r)^{m \cdot n} - 1} u_i,$$

and this should hold for all $\mathbf{u} \in \mathbb{D}_n$.
The limit of the RHS of this expression for $m \to \infty$ converges to u_1. Consequently, it is impossible to find a weight vector in the domain $\mathbb{W}_n$ which could guarantee the last inequality.

Proof of proposition 6

Ebert (1988) showed that for the whole class of welfare orderings R which could be represented by functionals of the form:

$$W_n(\mathbf{u}_n) = \sum_{i=1}^{n} \alpha_{i,n}\nu(u_i) \quad \text{with} \quad \alpha_{i,n} \geq \alpha_{i+1,n} > 0, \forall i \in \mathbb{N}_0 : i < n, \ \sum_{i=1}^{n} \alpha_{i,n} = 1,$$
$$\nu \text{ an increasing concave transformation, } n \in \mathbb{N} : n \geq 3.$$

satisfying RAA and RAB simultaneously, requires that there exists a scalar $c \geq 1$ such that $\alpha_{i,n} = \frac{c^{(n-i)}}{\sum_{j=1}^{n} c^{n-j}}$.

For any $r > 0$, $\mathcal{G}(r)$ belongs to the class considered by Ebert (1988). Indeed, let ν be the identity function, and $\alpha_{i,n} = \frac{(1+r)^{(n-i)} w_{i,n}}{\sum_{j=1}^{n} (1+r)^{(n-j)} w_{j,n}}$. Choosing the weighting vectors so that for $w_{i,n} = b^{n-i}$, for some $b \geq 1$, and all $n \in \mathbb{N} : n \geq 3$, yields the result (with $c = (1+r)b$).

References

[1] Capéau, B. and Ooghe, E. (2004), On the reconciliation of efficiency and inequality aversion with heterogenous populations: characterization results, mimeo, CES KULeuven.

[2] Ebert, U. (1988), Measurement of inequality: an attempt at unification and generalization, in, Gaertner, W., and Pattanaik, P.K. (eds.), *Distributive Justice and inequality*, Berlin: Springer.

[3] Ebert, U. (1995), Income inequality and differences in household size, *Mathematical Social Sciences* 30, 37-55.

[4] Ebert, U. (1997), Social welfare when needs differ: An axiomatic approach, *Economica* 64, 233-244.

[5] Ebert, U. and Moyes, P. (2003), Equivalence scales reconsidered, *Econometrica* 71, 319-343.

[6] Hammond, P.J. (1975), A note on extreme inequality aversion, *Journal of Economic Theory* 11, 465-467.

[7] Hammond, P.J. (1976), Equity, Arrow's conditions, and Rawl's difference principle, *Econometrica* 44, 793-804.

[8] Shorrocks, A.F. (1983), Ranking income distributions, *Economica* (50), 3-17.

[9] Shorrocks, A.F. (1995), Inequality and welfare evaluation of heterogeneous income distributions, Discussion Paper 447, University of Essex.

[10] Weymark, J.A. (1981), Generalized Gini inequality indices, *Mathematical Social Sciences* 1, 409-30.