On the reconciliation of efficiency and inequality aversion with heterogeneous populations: characterization results*

Bart Capéau† and Erwin Ooghe‡

October 2004

Abstract

We characterize a family of r-extended generalized Lorenz dominance quasi-orderings and a family of r-Gini welfare orderings, on the basis of two allegedly "incompatible" axioms for heterogeneous welfare comparisons (Ebert, 1997, Ebert and Moyes, 2003, Shorrocks, 1995), but at the cost of either completeness or separability.

JEL classification: D31, D63, I31.

Keywords: heterogeneous welfare comparisons, equivalent income functions.

1 Motivation

Ebert (1997) and Shorrocks (1995) pinpoint some difficulties to reconcile the Pareto indifference principle with the between type Pigou-Dalton (BTPD) transfer principle. The first principle requires social indifference between two social states if each individual, given her type and income in the respective states, reaches the same living standard in both states. The latter principle says that money transfers which bring living standards closer together, are preferable. These difficulties originate from a paradox described in Pyatt (1990) and Glewwe (1991) and reappear in Ebert (1999) and Ebert and Moyes (2003). Furthermore, the BTPD transfer principle would imply to weight households in a social ranking by their equivalence scales, giving up the conventional wisdom of weighting households by the number of individuals (Ebert, 1997).

*We would like to thank André Decoster, Udo Ebert, Marc Fleurbaey, Peter Lambert Eugenio Peluso, Erik Schokkaert, Frans Spinnewyn, and Alain Trannoy for helpful comments and especially Luc Lauwers for help with proving proposition 3.

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

‡Erwin Ooghe is 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.be.

We characterize a family of r-extended generalized Lorenz dominance quasi-orderings and a family of r-Gini social welfare orderings, which (i) reconcile the Pareto indifference and the between type Pigou-Dalton transfer principle (at the cost of either completeness or separability) and (ii) are in line with the standard practice of weighting individuals equally.

2 Notation

A heterogeneous income distribution is denoted by $(\mathbf{x};\mathbf{a}) = (x_1,\ldots,x_H;a_1,\ldots,a_H)$, with $x_h \in \mathbb{R}_{++}$ the *income* of household h and $a_h = (n_h, c_h)$ its *type*, a description of all potentially relevant non-income characteristics, including household size $n_h \in \mathbb{N}_0$, as well as other characteristics c_h, e.g., the age of the household members, their handicaps and so on.[1] The set A collects all *possible* household types and is assumed to satisfy a *richness condition*: it contains at least three types with household size equal to one. Whether or not a social observer treats these types differently depends on her choice of the equivalent income function, to be defined later. The domain assumption plays the same role as in social choice theory: it allows to create imaginary societies which can be more easily compared according to our ethical intuitions, in order to derive principles for judging less transparent real world situations. Although the number of households H is variable, we keep the number of individuals $N = \sum_{h=1}^{H} n_h \geq 3$ fixed. In a companion paper (Capéau and Ooghe, 2004), we present some problems as well as a solution in case of a variable population size N. In this note, the set of heterogeneous income distributions equals:

$$\mathcal{D} = \left\{ (\mathbf{x};\mathbf{a}) \in \underset{1\leq H\leq N}{\cup} \mathbb{R}_{++}^{H} \times A^{H} \,\middle|\, \mathbf{a} \text{ satisfies } \sum_{h=1}^{H} n_h = N \right\}.$$

We use an equivalent income function $E : \mathbb{R}_+ \times A \to \mathbb{R}_+$, to compare living standards between individuals of different households. A particular single, say $a^* = (1, c^*)$, will be the reference type. The equivalent income $E(x,a)$ is then the income needed by this reference type to obtain the same living standard as the members of a household with income x and type a.[2] We consider equivalent income functions which belong to a general parameterized domain $\mathcal{E}(r)$, with $1 \leq r < \infty$, defined by the following assumptions:

E1. $E(0,a) = 0$, for all $a \in A$,

E2. $E(x,a^*) = x$, for all $x \in \mathbb{R}_+$, with $a^* = (1,c^*)$ the reference type, and,

E3. there exist $\alpha,\beta \in \mathbb{R}_{++}$, with $\frac{\beta}{\alpha} = r$, such that $n\alpha \leq \frac{\epsilon}{E(x+\epsilon,a)-E(x,a)} \leq n\beta$ holds, for all $x \in \mathbb{R}_+$, all $\epsilon \in \mathbb{R}_{++}^{N}$ and all $a \in A$.

[1] In the sequel, subscripts h will be dropped if we do not refer to a particular household.

[2] In this paper we assume perfect measurability and full comparability of equivalent incomes.

In words, we impose that, without household income, the equivalent income is zero for all types (E1), and, the equivalent income of the reference type equals its income level (E2).
Linear equivalent income functions ($E(x, a) = x/m(a)$, with $m(a) \in \mathbb{R}_{++}$ an income independent equivalence scale and $m(a^*) = 1$) satisfy conditions E1 and E2. In the linear case, the boundary condition E3 reduces to $\alpha \leq \frac{m(a)}{n} \leq \beta$ for all types $a \in A$. It puts restrictions on the *per capita* equivalence scales.
More generally, the term $\frac{\epsilon}{E(x+\epsilon,a)-E(x,a)}$ (locally) measures the additional income needed by a household of type a in order to increase its equivalent income by one unit.[3] Notice that the bounds on this needs measure grow with household size n, reflecting that larger households may have higher needs, everything else constant.
Finally, $\mathcal{E}(1)$ reduces to using *per capita* incomes –with a constant efficiency equal to one for all types– as the only possible living standard concept. Increasing r, expands the domain $\mathcal{E}(r)$, allowing for larger need differences. Letting $r \to \infty$, we end up with all increasing equivalent income functions satisfying E1 and E2.

3 Core axioms

We want to rank heterogeneous distributions $(\mathbf{x}; \mathbf{a})$ in $\mathcal{D}$ via a rule

$$f : \mathcal{E}(r) \to \mathcal{Q} : E \mapsto R_E = f(E),$$

with $\mathcal{Q}$ the set of quasi-orderings on $\mathcal{D}$.[4] P_E and I_E are the corresponding strict preference and indifference relations.
Strong monotonicity requires welfare to increase with household *income*.
Strong Monotonicity (SM)
For each $E \in \mathcal{E}(r)$, and for all $(\mathbf{x}; \mathbf{a}), (\mathbf{y}; \mathbf{a}) \in \mathcal{D}$: if $x_h \geq y_h$, for all $h = 1, \ldots, H$, then $(\mathbf{x}; \mathbf{a}) \, R_E \, (\mathbf{y}; \mathbf{a})$. If, in addition, $\mathbf{x} \neq \mathbf{y}$, then $(\mathbf{x}; \mathbf{a}) \, P_E \, (\mathbf{y}; \mathbf{a})$.

Define $e(\mathbf{x}; \mathbf{a}) \in \mathbb{R}^N_{++}$ as the vector of *individual* equivalent incomes, i.e.,

$$e(\mathbf{x}; \mathbf{a}) \equiv \left(\underbrace{E(x_1, a_1), \ldots, E(x_1, a_1)}_{n_1 \text{ times}}, \ldots, \underbrace{E(x_H, a_H), \ldots, E(x_H, a_H)}_{n_H \text{ times}} \right).$$

The Pareto indifference axiom requires social welfare to remain unchanged, whenever each *individual* reaches the same equivalent income in two heterogeneous distributions.
Pareto Indifference (PI).
For each $E \in \mathcal{E}(r)$, and for all $(\mathbf{x}; \mathbf{a}), (\mathbf{y}; \mathbf{b}) \in \mathcal{D}$: if $e(\mathbf{x}; \mathbf{a}) = e(\mathbf{y}; \mathbf{b})$, then $(\mathbf{x}; \mathbf{a}) \, I_E \, (\mathbf{y}; \mathbf{b})$.

[3] The lower bound on $\frac{E(x+\epsilon,a)-E(x,a)}{\epsilon}$ ($\frac{1}{n\beta} > 0$) implied by condition E3, guarantees that the equivalent income functions are strictly increasing, while the upper bound ($\frac{1}{n\alpha} < \infty$) guarantees Lipzitsch continuity.
[4] A quasi-ordering is a reflexive and transitive binary relation, and, an ordering is a complete quasi-ordering.

Due to Pareto indifference, we can convert every income distribution of H heterogeneous households into a welfare-equivalent income distribution of N homogeneous reference type singles, by replacing each household h by n_h reference type singles, each with an income equal to the equivalent income of household h.
The between type Pigou-Dalton transfer principle requires social welfare to increase whenever an income transfer between *households* makes their equivalent incomes closer together.
Between type Pigou-Dalton transfer principle (BTPD).
For each $E \in \mathcal{E}(r)$, for all $(\mathbf{x};\mathbf{a}), (\mathbf{y};\mathbf{a}) \in \mathcal{D}$, and for all $i, j \in \{1,\ldots,H\}$: if (i) $x_k = y_k$ for all $k \neq i, j$, (ii) $x_i - y_i = y_j - x_j$, and (iii) $E(y_i, a_i) < E(x_i, a_i) \leq E(x_j, a_j) < E(y_j, a_j)$, then $(\mathbf{x};\mathbf{a}) R_E (\mathbf{y};\mathbf{a})$.

Finally, consider two distributions where all households are of the reference type a^*. Homogeneity unanimity requires a unanimous ranking of such distributions for all equivalent income functions.
Homogeneity Unanimity (HU).
For all $E, E' \in \mathcal{E}(r)$, for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^N_{++}$: $(\mathbf{x}; a^*, \ldots, a^*) R_E (\mathbf{y}; a^*, \ldots, a^*)$ if and only if $(\mathbf{x}; a^*, \ldots, a^*) R_{E'} (\mathbf{y}; a^*, \ldots, a^*)$.

4 Intermediate results

Pareto indifference and homogeneity unanimity lead to welfarism.[5]

Proposition 1 *A rule $f : \mathcal{E}(r) \to \mathcal{Q}$ satisfies* PI *and* HU *if and only if there exists a unique quasi-ordering R defined on $\mathbb{R}^N_{++}$ such that, for all $E \in \mathcal{E}(r)$, and for all $(\mathbf{x};\mathbf{a}), (\mathbf{y};\mathbf{b}) \in \mathcal{D}$:*

$$(\mathbf{x};\mathbf{a}) R_E (\mathbf{y};\mathbf{b}) \iff e(\mathbf{x};\mathbf{a}) \; R \; e(\mathbf{y};\mathbf{b}).$$

Welfarism allows us to focus on quasi-orderings R defined over *individual* equivalent income distributions, denoted $\mathbf{u} = (u_1, \ldots, u_N) \in \mathbb{R}^N_{++}$. From now on, we look at individuals $i \in \{1,\ldots,N\}$, rather than households $h \in \{1,\ldots,H\}$. We define two additional axioms for R. According to the strong Pareto principle, higher living standards are better:
Strong Pareto (SP).
For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{++}$: if $u_i \geq v_i$ for all $i = 1,\ldots,N$, then $\mathbf{u}R\mathbf{v}$. If, in addition, $\mathbf{u} \neq \mathbf{v}$, then $\mathbf{u}P\mathbf{v}$.
The s-extended Pigou-Dalton transfer principle, with $0 < s < \infty$, requires social welfare to increase whenever the equivalent incomes of two individuals come closer together, and the loss in total equivalent income is restricted by s in a specific way:
s-extended Pigou-Dalton transfer principle (PD(s)).
For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{++}$, and for all $i, j \in \{1,\ldots,N\}$: if (i) $u_k = v_k$, for all $k \neq i, j$, (ii) $v_i < u_i \leq u_j < v_j$, and (iii) $\frac{v_j - u_j}{u_i - v_i} = s$, then $\mathbf{u}R\mathbf{v}$.

[5] All proofs are relegated to the appendix.

We call an equivalent income transfer which satisfies conditions (i)-(iii) a PD(s) transfer from j to i. Choosing $s = 1$, one recognizes the standard Pigou-Dalton transfer principle, here applied to equivalent incomes. Furthermore, if R satisfies strong Pareto, then, the higher s, the stronger the transfer principle:

Lemma 1 *Consider a quasi-ordering R on $\mathbb{R}^N_{++}$ which satisfies* SP*. For any given $s \in (0, \infty)$, if R satisfies the* PD(s) *transfer principle then R satisfies the* PD(t) *transfer principle for all t in $(0, s]$.*

If strong Pareto holds, then by letting $s \to \infty$, the PD(s) transfer principle approaches Hammond's equity principle: any equalizing transfer of individual equivalent incomes is an improvement, irrespective of the total loss in equivalent income.
The next proposition shows that, under welfarism, strong monotonicity and the BTPD transfer principle for a rule $f : \mathcal{E}(r) \to \mathcal{Q}$ are equivalent with strong Pareto and the PD(r) transfer principle for the corresponding quasi-ordering R (as defined in proposition 1). This result serves as a rationale for our endorsement of the PD(r) transfer principle, where the level of r depends on the maximal needs difference determined by the choice of the domain.

Proposition 2 *Consider a rule $f : \mathcal{E}(r) \to \mathcal{Q}$ which satisfies welfarism: f satisfies* SM *and the* BTPD *transfer principle if and only if the corresponding quasi-ordering R satisfies* SP *and the* PD(r) *transfer principle.*

5 Main characterizations

We consider first some additional properties for R.
Weak Continuity (CON).
For each convergent sequence of distributions $\left(\mathbf{v}^\ell\right)_{\ell \in \mathbb{N}_0}$, if there exists an $M \in \mathbb{N}_0$ such that for some $\mathbf{u} \in \mathbb{R}^N_{++}$: $\mathbf{v}^\ell R \mathbf{u}$ for all $\ell \geq M$ then, for $\mathbf{v}^* \equiv \lim_{\ell \to \infty} \mathbf{v}^\ell$, $\mathbf{v}^* R \mathbf{u}$; if $\mathbf{u} R \mathbf{v}^\ell$ for all $\ell \geq M$, then $\mathbf{u} R \mathbf{v}^*$.
Weak continuity ensures that small changes in distributions cannot cause large changes in social evaluation.[6]
Anonymity (A).
For each $\mathbf{u} \in \mathbb{R}^N_{++}$, and each permutation $\sigma : N \to N$: $\mathbf{u} I \sigma(\mathbf{u})$, with $\sigma(\mathbf{u}) = \left(u_{\sigma(1)}, \ldots, u_{\sigma(N)}\right)$.
Anonymity tells us that the names of the individuals do not matter. We can thus focus on rank- ordered distributions. We will consider the sub-domain $\mathbb{D} = \left\{\mathbf{u} \in \mathbb{R}^N_{++} \middle| u_1 \leq \ldots \leq u_N\right\}$.
Relative Invariance (RI).
For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{++}$ and for each $\lambda > 0$: if $\mathbf{u} I \mathbf{v}$, then $\lambda \mathbf{u} I \lambda \mathbf{v}$.

[6] For orderings it coincides with the ordinary continuity concept (closed weak better-than- and worse-than-sets), but for quasi-orderings, it does not always imply the strict better-than- and worse-than-sets to be open.

Absolute Invariance (AI).
For all $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{++}$ and for each $\lambda > 0$: if $\mathbf{u}I\mathbf{v}$ then $(\mathbf{u}+\lambda\mathbf{1})\, I\, (\mathbf{v}+\lambda\mathbf{1})$.
Relative (resp. absolute) invariance requires the indifference of two distributions to be unchanged, when multiplying both with the same constant (resp. adding the same constant vector to both).
Rank-ordered Separability (ROSE).
For all $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z} \in \mathbb{D}$, for all $M \subseteq \{1, \ldots, N\}$: if (i) $u_j = w_j$ and $v_j = z_j$ for all $j \in M$, and (ii) $u_i = v_i$ and $w_i = z_i$ for all $i \in N \setminus M$, then $\mathbf{u}R\mathbf{v}$ if and only if $\mathbf{w}R\mathbf{z}$.
Rank-ordered separability prohibits unconcerned "positions" to matter.
The r-extended GLD quasi-ordering $R(r)$ is defined, for all distributions $\mathbf{u}, \mathbf{v} \in \mathbb{D}$, as:

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

One recognizes the standard GLD quasi-ordering when $r = 1$. Increasing r, increases the weight for the lower equivalent incomes, and, when $r \to \infty$, we approach the leximin ordering. We obtain the following characterization:

Proposition 3 *For all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$, the following statements are equivalent:*

(a) $\mathbf{u}R(r)\,\mathbf{v}$,

(b) $\mathbf{u}$ *can be derived from* $\mathbf{v}$ *by a sequence of* PD(r) *transfers and income increments,*

(c) $\mathbf{u}R\mathbf{v}$ *holds, for all quasi-orderings R which satisfy* CON, SP, *and* PD(r).

Let $\mathbb{W} = \left\{\mathbf{w} \in \mathbb{R}^N_{++} \,|w_1 \geq \ldots \geq w_N = 1\right\}$ be a class of (strictly) positive, decreasing, and normalized weight vectors. An r-Gini ordering $R(r, \mathbf{w})$ is defined, for all $\mathbf{u}, \mathbf{v} \in \mathbb{D}$, as

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

Again, the extreme cases correspond with a standard generalized Gini ordering (when $r = 1$) and the leximin ordering (when $r \to \infty$). Our final proposition characterizes the r-Ginis.

Proposition 4 *Let $N \geq 3$. An ordering R satisfies* CON, SP, PD(r), RI, AI, *and* ROSE *if and only if R is an r-Gini ordering $R(r, \mathbf{w})$ for some $\mathbf{w} \in \mathbb{W}$.*

In a companion paper (Capéau and Ooghe, 2004), we show that the r-GLD criterion is the intersection of the corresponding family of r-Ginis: at the cost of completeness we regain separability.

6 Conclusion

In the present note we show that it is possible to reconcile Pareto indifference with the between type Pigou-Dalton transfer principle for comparing heterogeneous distributions. Given the difficulties to reconcile both principles, pinpointed in the literature (Ebert, 1997, Shorrocks, 1995), it should be no surprise that our results come at a cost. The r-GLD criterion gives up completeness, whereas all r-Ginis are complete, but separable only over rank-ordered distributions.

The leaky r-extended Pigou-Dalton transfer principle is at the heart of our results. It turns out to be equivalent with the between type Pigou-Dalton transfer principle if one is willing to accept welfarism.

Appendix

Proof of proposition 1

We show in the following auxiliary lemma equivalence between on the one hand, PI and HU, and strong neutrality on the other.
A rule $f : \mathcal{E}(r) \rightarrow \mathcal{Q}$ is said to be *strongly neutral* (SN), if for all $E, E' \in \mathcal{E}(r)$, for all $(\mathbf{x};\mathbf{a}), (\mathbf{y};\mathbf{b}), (\mathbf{x}';\mathbf{a}')(\mathbf{y}';\mathbf{b}') \in \mathcal{D}$ such that $e(\mathbf{x};\mathbf{a}) = e'(\mathbf{x}';\mathbf{a}')$ and $e(\mathbf{y};\mathbf{b}) = e'(\mathbf{y}';\mathbf{b}')$:

$$(\mathbf{x};\mathbf{a})\, R_E\, (\mathbf{y};\mathbf{b}) \iff \left(\mathbf{x}';\mathbf{a}'\right) R_{E'} \left(\mathbf{y}';\mathbf{b}'\right).$$

Lemma 2 *A rule $f : \mathcal{E}(r) \rightarrow \mathcal{Q}$ satisfies* PI *and* HU *if and only if it satisfies* SN.

Proof.
Sufficiency:

- SN implies HU: let $(\mathbf{x}_1;\mathbf{a}^*)$ play the role of $(\mathbf{x};\mathbf{a})$ and $(\mathbf{x}';\mathbf{a}')$ in the definition of SN, and similarly, let $(\mathbf{y}_1;\mathbf{a}^*) = (\mathbf{y};\mathbf{b}) = (\mathbf{y}';\mathbf{b}')$. By construction and property E2 of equivalent income functions, $\mathbf{x}_1 = e(\mathbf{x}_1;\mathbf{a}^*) = e(\mathbf{x};\mathbf{a}) = e'(\mathbf{x}';\mathbf{a}')$ and $\mathbf{y}_1 = e(\mathbf{y}_1;\mathbf{a}^*) = e(\mathbf{y};\mathbf{b}) = e'(\mathbf{y}';\mathbf{b}')$. By SN we have:

  $$(\mathbf{x}_1;\mathbf{a}^*)\, R_E\, (\mathbf{y}_1;\mathbf{a}^*) \iff (\mathbf{x}_1;\mathbf{a}^*)\, R_{E'}\, (\mathbf{y}_1;\mathbf{a}^*),$$

  which establishes the result.

- SN implies PI: let $E = E'$, let $(\mathbf{x}_1, \mathbf{a}_1)$ play the role of $(\mathbf{x};\mathbf{a})$, $(\mathbf{x}';\mathbf{a}')$ and $(\mathbf{y}';\mathbf{b}')$ in the definition of SN, and let $(\mathbf{x}_2;\mathbf{a}_2)$ play the role of $(\mathbf{y};\mathbf{b})$. Assume as in the antecedent of PI, that $e(\mathbf{x}_1;\mathbf{a}_1) = e(\mathbf{x}_2;\mathbf{a}_2)$. We also have $e(\mathbf{x}_1;\mathbf{a}_1) = e'(\mathbf{x}_1;\mathbf{a}_1)$ (because $E = E'$) and $e(\mathbf{x}_2;\mathbf{a}_2) = e'(\mathbf{x}_1;\mathbf{a}_1)$ (from $e(\mathbf{x}_1;\mathbf{a}_1) = e'(\mathbf{x}_1;\mathbf{a}_1)$ and $E = E'$). Applying SN gives:

  $$(\mathbf{x}_1;\mathbf{a}_1)\, R_E\, (\mathbf{x}_2;\mathbf{a}_2) \iff (\mathbf{x}_1;\mathbf{a}_1)\, R_E\, (\mathbf{x}_1;\mathbf{a}_1).$$

  By reflexivity of R_E we have established that $(\mathbf{x}_1;\mathbf{a}_1)\, R_E\, (\mathbf{x}_2;\mathbf{a}_2)$. Exchanging the roles of $(\mathbf{x}_1;\mathbf{a}_1)$ and $(\mathbf{x}_2;\mathbf{a}_2)$, we obtain in a similar fashion: $(\mathbf{x}_2;\mathbf{a}_2)\, R_E\, (\mathbf{x}_1;\mathbf{a}_1)$. Hence, by definition of I_E, we have: $(\mathbf{x}_1;\mathbf{a}_1)\, I_E\, (\mathbf{x}_2;\mathbf{a}_2)$.

Necessity:
Suppose the antecedent of SN is true. Recall that the reference type a^* is a single, thus $n^* = 1$. Construct distributions $(\mathbf{x}'';\mathbf{a}^*)$ and $(\mathbf{y}'';\mathbf{a}^*)$ as follows:

$$\begin{aligned}
\left(\mathbf{x}'';\mathbf{a}^*\right) &= \left(e\,(\mathbf{x};\mathbf{a})\,;\underbrace{a^*,\ldots,a^*}_{N\text{ times}}\right) = \left(e'\left(\mathbf{x}';\mathbf{a}'\right);\underbrace{a^*,\ldots,a^*}_{N\text{ times}}\right), \text{ and}\\
\left(\mathbf{y}'';\mathbf{a}^*\right) &= \left(e\,(\mathbf{y};\mathbf{b})\,;\underbrace{a^*,\ldots,a^*}_{N\text{ times}}\right) = \left(e'\left(\mathbf{y}';\mathbf{a}^*\right);\underbrace{a^*,\ldots,a^*}_{N\text{ times}}\right),
\end{aligned}$$

which all belong to $\mathcal{D}$. Via PI, transitivity and E2, we have:

$$\left(\mathbf{x};\mathbf{a}\right) R_E \left(\mathbf{y};\mathbf{b}\right) \;\Leftrightarrow\; \left(\mathbf{x}'';\mathbf{a}^*\right) R_E \left(\mathbf{y}'';\mathbf{a}^*\right), \tag{1}$$

$$\left(\mathbf{x}';\mathbf{a}'\right) R_{E'} \left(\mathbf{y}';\mathbf{b}'\right) \;\Leftrightarrow\; \left(\mathbf{x}'';\mathbf{a}^*\right) R_{E'} \left(\mathbf{y}'';\mathbf{a}^*\right). \tag{2}$$

Via HU, we get:

$$\left(\mathbf{x}'';\mathbf{a}^*\right) R_E \left(\mathbf{y}'';\mathbf{a}^*\right) \Leftrightarrow \left(\mathbf{x}'';\mathbf{a}^*\right) R_{E'} \left(\mathbf{y}'';\mathbf{a}^*\right). \tag{3}$$

Combining (1), (2), and (3) leads to the desired result. ■

We are now able to proof proposition 1. We show first that we can construct a binary relation R on $\mathbb{R}^N_{++}$ from any rule f that satisfies SN. For all $\mathbf{u},\mathbf{v} \in \mathbb{R}^N_{++}$, let $\mathbf{u}R\mathbf{v} \iff$

$$\exists E \in \mathcal{E}\left(r\right) \text{ and } \exists \left(\mathbf{x};\mathbf{a}\right), \left(\mathbf{y};\mathbf{b}\right) \in \mathcal{D} : \mathbf{u} = e\left(\mathbf{x};\mathbf{a}\right),\ \mathbf{v} = e\left(\mathbf{y};\mathbf{b}\right), \text{ and } \left(\mathbf{x};\mathbf{a}\right) R_E \left(\mathbf{y};\mathbf{b}\right).$$

From SN of f, it follows that for all $E \in \mathcal{E}\left(r\right)$ and for all $\left(\mathbf{x};\mathbf{a}\right), \left(\mathbf{y};\mathbf{b}\right) \in \mathcal{D}$ it holds that $\left(\mathbf{x};\mathbf{a}\right) R_E \left(\mathbf{y};\mathbf{b}\right)$ if and only if $e\left(\mathbf{x};\mathbf{a}\right)\ R\ e\left(\mathbf{y};\mathbf{b}\right)$, as required.

Notice that, for each vector $\mathbf{u} \in \mathbb{R}^N_{++}$ there exists an equivalent income function $E \in \mathcal{E}\left(r\right)$ and a distribution $\left(\mathbf{x};\mathbf{a}\right) \in \mathcal{D}$ such that $\mathbf{u} = e\left(\mathbf{x};\mathbf{a}\right)$. For example choose $\left(\mathbf{x};\mathbf{a}\right) = \left(\mathbf{u}; \underbrace{a^*,\ldots,a^*}_{N \text{ times}}\right)$.

Furthermore, strong neutrality ensures that the ranking of $\mathbf{u}$ and $\mathbf{v}$ does not depend on the chosen equivalent income function or the distributions which generate these equivalent incomes. Thus, R is a unique binary relation, which inherits (from f) reflexivity, transitivity, and possibly, completeness, in case the range of f is restricted to orderings (complete quasi-orderings). Conversely, construct a rule f from any quasi-ordering R on $\mathbb{R}^N_{++}$ as follows:

$$\left(\mathbf{x};\mathbf{a}\right) R_E \left(\mathbf{y};\mathbf{b}\right) \forall E \in \mathcal{E}\left(r\right) \text{ and } \forall \left(\mathbf{x};\mathbf{a}\right), \left(\mathbf{y};\mathbf{b}\right) \in \mathcal{D} \text{ with } \mathbf{u} = e\left(\mathbf{x};\mathbf{a}\right) \text{ and } \mathbf{v} = e\left(\mathbf{y};\mathbf{b}\right)$$
$$\iff \mathbf{u}R\mathbf{v}.$$

Now, f is well defined (for every $\mathbf{u},\mathbf{v} \in \mathbb{R}^N_{++}$, there exist $\left(\mathbf{x},\mathbf{a}\right), \left(\mathbf{y},\mathbf{b}\right) \in \mathcal{D}$: $\mathbf{u} = e(\mathbf{x},\mathbf{a})$ and $\mathbf{v} = e(\mathbf{y},\mathbf{b})$ for all $E \in \mathcal{E}(r)$, see above). Reflexivity and transitivity (and completeness, if R is an ordering) of R_E for each $E \in \mathcal{E}(r)$ are inherited from R, and f satisfies SN by construction.

Proof of lemma 1

Consider a quasi-ordering R which satisfies SP and the PD(s) transfer principle, for some $s \in (0,\infty)$. We show that R also satisfies the PD(t) transfer principle for all $t \in (0,s]$. Consider a distribution $\mathbf{u}$ obtained from $\mathbf{v}$ via a PD(t) transfer (with $t \in (0,s]$) of size $\delta > 0$:

$$\mathbf{u} = \left(v_1,\ldots,v_{i-1},v_i+\delta,v_{i+1},\ldots,v_{j-1},v_j-t\delta,v_{j+1},\ldots,v_N\right),$$

with $v_i + \delta \leq v_j - t\delta$.

Construct $\mathbf{u}'$ from $\mathbf{v}$ by means of a PD(s) transfer of size $\epsilon = \delta\frac{t}{s}$ (notice that $\epsilon \leq \delta$):

$$\mathbf{u}' = \left(v_1,\ldots,v_{i-1},v_i+\epsilon,v_{i+1},\ldots,v_{j-1},v_j - s\left(\delta\frac{t}{s}\right),v_{j+1},\ldots,v_N\right).$$

Using the PD(s) transfer principle, we must have $\mathbf{u}'R\mathbf{v}$. From SP and $\epsilon \leq \delta$, it follows that $\mathbf{u}R\mathbf{u}'$. By transitivity, we get $\mathbf{u}R\mathbf{v}$, as required.

Proof of proposition 2

Consider a rule f which satisfies welfarism, and denote by R the corresponding quasi-ordering defined over $\mathbb{R}^N_{++}$. It is obvious that (i) a welfarist rule f satisfies monotonicity if and only if R satisfies the strong Pareto principle. We only prove (in the remainder) that (ii) a welfarist rule f satisfies the BTPD transfer principle if and only if R satisfies the PD(s) transfer principle for all s in $\left[\frac{1}{r}, r\right]$. Statements (i) and (ii) together establish the required result, because lemma 1 tells us that a quasi-ordering R which satisfies SP, satisfies the PD(s) transfer principle for all $s \in \left[\frac{1}{r}, r\right]$ if and only if it satisfies the PD(r) transfer principle.

First, we show that if f satisfies the BTPD transfer principle, then R satisfies the PD(s) transfer principle for all s in $\left[\frac{1}{r}, r\right]$. Consider vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^N_{++}$, and individuals $i, j \in \{1, \ldots, N\}$, such that (i) $u_k = v_k$ for all $k \neq i, j$, (ii) $\frac{v_j - u_j}{u_i - v_i} = s$ for some $s \in \left[\frac{1}{r}, r\right]$ and (iii) $v_i < u_i \leq u_j < v_j$. We must prove that $\mathbf{u}R\mathbf{v}$ holds. Recall that A contains at least three possible types with household size equal to 1, the reference type a^* and two other types, say a_1 and a_2. Choose a linear equivalent income function E such that:

$$E(x, a) = \frac{x}{m(a)}, \text{ for all } x \in \mathbb{R}_+ \text{ and for all } a \in A, \text{ and } \frac{m(a_1)}{m(a_2)} = s.$$

Given $s \in \left[\frac{1}{r}, r\right]$, we can choose E in such a way that it belongs to $\mathcal{E}(r)$. Consider heterogeneous distributions $(\mathbf{x}; \mathbf{a})$ and $(\mathbf{y}; \mathbf{a})$ with

$$\begin{array}{lllllllllll}
\mathbf{x} & = & (& u_1, & \ldots, & u_{i-1}, & m(a_1)\, u_i, & \ldots, & u_{j-1}, & m(a_2)\, u_j, & u_{j+1}, & \ldots, & u_N &) \\
\mathbf{y} & = & (& v_1, & \ldots, & v_{i-1}, & m(a_1)\, v_i, & \ldots, & v_{j-1}, & m(a_2)\, v_j, & v_{j+1}, & \ldots, & v_N &) \\
\mathbf{a} & = & (& a^*, & \ldots, & a^*, & a_1, & \ldots, & a^*, & a_2, & a^*, & \ldots, & a^* &).
\end{array}$$

These distributions satisfy $e(\mathbf{x}; \mathbf{a}) = \mathbf{u}$, $e(\mathbf{y}; \mathbf{b}) = \mathbf{v}$, and $\mathbf{x}$ can be derived from $\mathbf{y}$ by transferring an amount of *income*, $\epsilon = m(a_2)(v_j - u_j) = m(a_1)(u_i - v_i)$ from individual/household j to individual/household i, without changing their mutual equivalent income position (from (iii)). Using the BTPD transfer principle, we have $(\mathbf{x}; \mathbf{a})\, R_E\, (\mathbf{y}; \mathbf{b})$ and using welfarism, we get $\mathbf{u}R\mathbf{v}$, as required.

Secondly, we show that if R satisfies the PD(s) principle for all $s \in \left[\frac{1}{r}, r\right]$, then f satisfies the BTPD principle. Consider an equivalent income function $E \in \mathcal{E}(r)$, distributions $(\mathbf{x}; \mathbf{a}), (\mathbf{y}; \mathbf{a}) \in \mathcal{D}$, and households $i, j \in \{1, \ldots, H\}$, such that (i) $x_k = y_k$ for all $k \neq i, j$, (ii) $\frac{y_j - x_j}{x_i - y_i} = 1$, and (iii) $E(y_i, a_i) < E(x_i, a_i) \leq E(x_j, a_j) < E(y_j, a_j)$. We must prove that

$(\mathbf{x};\mathbf{a})\, R_E\, (\mathbf{y};\mathbf{a})$ holds. Define $\epsilon = y_j - x_j = x_i - y_i$. Let

$$\mathbf{u} \equiv e(\mathbf{x};\mathbf{a}) = \left(\ldots, \underbrace{E(y_i+\epsilon, a_i), \ldots, E(y_i+\epsilon, a_i)}_{n_i \text{ times}}, \ldots, \underbrace{E(x_j, a_j), \ldots, E(x_j, a_j)}_{n_j \text{ times}}, \ldots \right),$$

$$\mathbf{v} \equiv e(\mathbf{y};\mathbf{a}) = \left(\ldots, \underbrace{E(y_i, a_i), \ldots, E(y_i, a_i)}_{n_i \text{ times}}, \ldots, \underbrace{E(x_j+\epsilon, a_j), \ldots, E(x_j+\epsilon, a_j)}_{n_j \text{ times}}, \ldots \right).$$

Now, we show that there exists an $s \in \left[\frac{1}{r}, r\right]$ such that $\mathbf{u}$ can be derived from $\mathbf{v}$ by a sequence of $n_i \times n_j$ $\text{PD}(s)$ transfers between individuals. Each individual from household j transfers an amount of $\frac{1}{n_i}\left(E(x_j+\epsilon, a_j) - E(x_j, a_j)\right)$ equivalent income units to each of the individuals belonging to household i, who only receive $\frac{1}{n_i s}\left(E(x_j+\epsilon, a_j) - E(x_j, a_j)\right)$ of it. We must have $\frac{n_j(E(x_j+\epsilon,a_j)-E(x_j,a_j))}{n_i s} = E(y_i+\epsilon, a_i) - E(y_i, a_i)$, or

$$s = \frac{n_j\left(E(x_j+\epsilon, a_j) - E(x_j, a_j)\right)}{n_i\left(E(y_i+\epsilon, a_i) - E(y_i, a_i)\right)},$$

which lies in the interval $\left[\frac{1}{r}, r\right]$, because our domain assumption E3 guarantees that the right-hand side lies in between $\frac{\alpha}{\beta} = \frac{1}{r}$ and $\frac{\beta}{\alpha} = r$.

Proof of proposition 3

Results (b) $\Rightarrow$ (c) and (c) $\Rightarrow$ (a) are obvious. We show (a) $\Rightarrow$ (b) for $r > 1$ (the case for $r = 1$ is well-known). Consider $r > 1$, and two distributions $\mathbf{u}, \mathbf{v} \in \mathbb{D}$ for which $\mathbf{u} R(r)\, \mathbf{v}$ holds. Consider the following algorithm, which leads to a sequence of distributions $\left(\mathbf{v}^{\ell}\right)_{\ell \in \mathbb{N}_0}$.

1. $\mathbf{v}^1 = \mathbf{v}$.

2. If $v_N^1 > u_N$, construct $\mathbf{v}^2$ by subtracting $\varepsilon_N = \min\{v_N^1 - u_N, \frac{r}{1+r}\left(v_N^1 - v_{N-1}^1\right)\}$ from v_N^1 and adding $\frac{1}{r}\varepsilon_N$ to v_{N-1}^1; this is the maximal $\text{PD}(r)$ transfer which ensures that $v_N^2 \geq u_N$. Otherwise, choose $\mathbf{v}^2 = \mathbf{v}^1$.

3. If $v_{N-1}^2 > u_{N-1}$, construct $\mathbf{v}^3$ by subtracting $\varepsilon_{N-1} = \min\{v_{N-1}^2 - u_{N-1}, \frac{r}{1+r}(v_{N-1}^2 - v_{N-2}^2)\}$ from v_{N-1}^2 and adding $\frac{1}{r}\varepsilon_{N-1}$ to v_{N-2}^2. Otherwise, choose $\mathbf{v}^3 = \mathbf{v}^1$.

   ...

N. If $v_2^{N-1} > u_2$, construct $\mathbf{v}^N$ by subtracting $\varepsilon_2 = \min\{v_2^{N-1} - u_2, \frac{r}{1+r}\left(v_2^{N-1} - v_1^{N-1}\right)\}$ from v_2^{N-1} and adding $\frac{1}{r}\varepsilon_2$ to v_1^{N-1}. Otherwise, choose $\mathbf{v}^N = \mathbf{v}^{N-1}$.

   Steps $k(N-1)+2$ to $(k+1)N+1-k$, with $k \in \mathbb{N}_0$: repeat steps 2 to N, starting from distribution $\mathbf{v}^{k(N-1)+1}$.

In the remainder of the proof we show that the sequence $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$ converges to a vector, say $\mathbf{v}^*$, which is Pareto dominated by $\mathbf{u}$ (lemma 4). If so, the sequence $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$ –based on $\mathrm{PD}(r)$ transfers– supplemented with an income increment $(\mathbf{u}-\mathbf{v}^*)$, allows to derive $\mathbf{u}$ from $\mathbf{v}=\mathbf{v}^1$, the desired result.
In order to proof convergence to a limit $\mathbf{v}^* \leq \mathbf{u}$, we first establish the following property of the sequence $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$.

Lemma 3 *For all $r \geq 1$, for all $\mathbf{u},\mathbf{v}\in\mathbb{D}$: $\mathbf{u}R(r)\,\mathbf{v}$ implies $\mathbf{u}R(r)\,\mathbf{v}^{\ell}$, for all $\ell\in\mathbb{N}_0$.*

Proof by induction. It holds for $\ell = 1$ (see step 1 of the algorithm). We show that, if it holds for some $\ell\in\mathbb{N}_0$ (induction hypothesis) then it also holds for $\ell+1$. Step $\ell+1$ (for some $j\in\{2,\dots,N\}$) is described as follows: if (a) $v_j^{\ell} > u_j$ then construct $\mathbf{v}^{\ell+1}$ by subtracting $\varepsilon_j = \min\left(v_j^{\ell}-u_j, \frac{r}{1+r}\left(v_j^{\ell}-v_{j-1}^{\ell}\right)\right)$ from v_j^{ℓ} and adding $\frac{1}{r}\varepsilon_j$ to v_{j-1}^{ℓ}. Otherwise (if (b) $v_j^{\ell}\leq u_j$) choose $\mathbf{v}^{\ell+1}=\mathbf{v}^{\ell}$. We focus on case (a); case (b) is straightforward. Notice that

$$\begin{aligned}
u_i - v_i^{\ell+1} &= u_i - v_i^{\ell} \text{ for all } i\notin\{j,j-1\}\,,\\
u_{j-1} - v_{j-1}^{\ell+1} &= u_{j-1} - \left(v_{j-1}^{\ell}+\frac{1}{r}\varepsilon_j\right),\text{ and,}\\
u_j - v_j^{\ell+1} &= u_j - \left(v_j^{\ell}-\varepsilon_j\right).
\end{aligned}$$

We obtain[7]

$$\begin{aligned}
\sum_{i=1}^{k} r^{k-i}\left(u_i - v_i^{\ell+1}\right) &= \underbrace{\sum_{i=1}^{k} r^{k-i}\left(u_i - v_i^{\ell}\right)}_{\geq 0 \text{ from } \mathbf{u}R(r)\mathbf{v}^{\ell}} \text{ for all } k\neq j-1, \text{ and}\\
\sum_{i=1}^{j-1} r^{j-1-i}\left(u_i - v_i^{\ell+1}\right) &= \sum_{i=1}^{j-2} r^{j-1-i}\left(u_i - v_i^{\ell}\right) + u_{j-1} - \left(v_{j-1}^{\ell}+\frac{1}{r}\varepsilon_j\right),\\
&= \frac{1}{r}\left(\sum_{i=1}^{j-1} r^{j-i}\left(u_i - v_i^{\ell}\right) - \varepsilon_j\right)\\
&\underbrace{\geq}_{\text{as } -\varepsilon_j \geq u_j - v_j^{\ell}} \frac{1}{r}\left(\sum_{i=1}^{j-1} r^{j-i}\left(u_i - v_i^{\ell}\right) + \left(u_j - v_j^{\ell}\right)\right)\\
&\geq \underbrace{\frac{1}{r}}_{\geq 0}\underbrace{\left(\sum_{i=1}^{j} r^{j-i}\left(u_i - v_i^{\ell}\right)\right)}_{\geq 0 \text{ from } \mathbf{u}R(r)\mathbf{v}^{\ell}},
\end{aligned}$$

and thus $\mathbf{u}R(r)\,\mathbf{v}^{\ell+1}$ must hold. ■
Notice that lemma 3 implies that $v_1^{\ell}\leq u_1$, for all $\ell\in\mathbb{N}_0$, by definition of $R(r)$.

[7] If $j=2$, the summation $\sum_{i=1}^{j-2}(\cdot)$ has to be interpreted as zero.

Lemma 4 *The sequence $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$ converges to a vector $\mathbf{v}^{*}\leq\mathbf{u}$.*

Proof.[8]

If $v_i \leq u_i$ for all $i \in \{1,\ldots,N\}$, the sequence is constant and thus converges. We consider now vectors for which $\exists k \in \{2,\ldots,N\} : v_k > u_k$. Notice that the range of the sequence $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$ is a compact subset of $\mathbb{R}^N_{++}$, bounded by the rectangle $[v_1, v_N]^N$. Hence there exist convergent subsequences. We now show that no two convergent subsequences can converge to different limits. If so, the sequence converges.

Consider the sequence of means of the vectors $\mathbf{v}^{\ell}$, say $\left(\mu^{\ell}\right)_{\ell\in\mathbb{N}_0}$, with $\mu^{\ell}\equiv\frac{\sum_{i=1}^{n}v_i^{\ell}}{n}$. This is a decreasing sequence bounded from below by v_1. Hence it converges to its greatest lower bound. As a consequence, the means of all convergent subsequences converge to the same limit.

Now, suppose there exist two subsequences of $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$ with different limits. Then the sequence oscillates. Oscillating is only possible via infinitely many (sets of successive) $\textsc{pd}(r)$ transfers . Each such a (set of) $\textsc{pd}(r)$ transfers decreases the mean with an amount at least as great as some positive constant. This would lead to divergent means, contradicting convergence of the mean. Hence, there can be no convergent subsequences with different limits. So, the sequence converges.

The sequence cannot converge to a point $\mathbf{v}^{*}$, in which there are still $\textsc{pd}(r)$ transfers possible. This can be shown as follows. Let $\left(\mathbf{w}^{k}\right)_{k\in\mathbb{N}}$ be a subsequence of $\left(\mathbf{v}^{\ell}\right)_{\ell\in\mathbb{N}_0}$, selecting the vectors $\mathbf{v}^{(k(N-1)+1)}$, for $k=\{0,1,2,\ldots\}$. Let $g:\mathbb{D}\to\mathbb{D}:\mathbf{w}^{k}\mapsto\mathbf{w}^{k+1}=g\left(\mathbf{w}^{k}\right)$. The function g collects $N-1$ subsequent steps of the algorithm that make $\textsc{pd}(r)$ transfers from richer individuals to their immediate successors, starting from the richest. If g is continuous, then $\lim\limits_{k\to\infty}\mathbf{w}^{k+1}=\mathbf{v}^{*}=g\left(\lim\limits_{k\to\infty}\mathbf{w}^{k}\right)=g\left(\mathbf{v}^{*}\right)$. If so, no further $\textsc{pd}(r)$ transfers can be made in the limit. Continuity of g can be shown as follows. Notice that g is composed of $N-1$ 'taxes' imposed on the richer of two subsequent individuals, pair-wisely associated with $N-1$ 'gifts' to the poorer of the two individuals. Denote the $N-1$ 'tax'-functions by $g_2^{\downarrow}, g_3^{\downarrow},\ldots,g_N^{\downarrow}$ and the $N-1$ associated 'gift'-functions by $g_1^{\uparrow}, g_2^{\uparrow},\ldots,g_{N-1}^{\uparrow}$. Then $g=g_1^{\uparrow}\circ g_2^{\downarrow}\circ g_2^{\uparrow}\circ\ldots\circ g_N^{\downarrow}$. We show, that for all $\epsilon>0$, for all $\mathbf{t},\mathbf{v}\in\mathbb{D}$: $\|\mathbf{t}-\mathbf{v}\|\equiv\sqrt{\sum_{i=1}^{n}\left(t_i-v_i\right)^2}<\delta\Rightarrow\|g\left(\mathbf{t}\right)-g\left(\mathbf{v}\right)\|<\epsilon$, provided that $0<\delta<\frac{\epsilon}{2^{N-1}}$. It can be checked that for any $\mathbf{r},\mathbf{s}\in\mathbb{D}$ and any $\delta>0$:

$$\|\mathbf{s}-\mathbf{r}\|<\delta \quad\Rightarrow\quad \left\|g_i^{\downarrow}\left(\mathbf{s}\right)-g_i^{\downarrow}\left(\mathbf{r}\right)\right\|\leq\sqrt{\left(\tfrac{r}{1+r}\right)^2\delta^2+\delta^2}<\sqrt{2}\delta \quad i=2,3,\ldots,N,$$

and

$$\|\mathbf{s}-\mathbf{r}\|<\delta \quad\Rightarrow\quad \left\|g_i^{\uparrow}\left(\mathbf{s}\right)-g_i^{\uparrow}\left(\mathbf{r}\right)\right\|\leq\sqrt{\left(\tfrac{1}{1+r}\right)^2\delta^2+\delta^2}<\sqrt{2}\delta \quad i=1,2,\ldots,N-1.$$

Because $g=g_1^{\uparrow}\circ g_2^{\downarrow}\circ g_2^{\uparrow}\circ\ldots\circ g_N^{\downarrow}$ we get $\|\mathbf{t}-\mathbf{v}\|<\delta\Rightarrow\|g\left(\mathbf{t}\right)-g\left(\mathbf{v}\right)\|\leq 2^{N-1}\delta$ and thus it suffices to choose $\delta<\frac{\epsilon}{2^{N-1}}$.

[8] We thank Luc Lauwers for help with proving convergence.

We conclude by showing that the limit of the sequence, say $\mathbf{v}^*$, is Pareto dominated by $\mathbf{u}$ (i.e. $\mathbf{v}^* \leq \mathbf{u}$). If not, there is a coordinate i such that $v_i^* > u_i$. By the previous property, it must hold that $v_i^* = v_{i-1}^* = \ldots = v_1^*$ (otherwise there still would be $\text{PD}(r)$ transfers possible) and thus $v_1^* > u_i \geq u_1$. But the latter is precluded by lemma 3. ■

Proof of proposition 4

Due to corollary 5 in Ebert (1988)[9] any ordering on $\mathbb{R}_{++}^N$ which satisfies CON, SP, ROSE, A, RI and AI has a linear representation over the rank-ordered subset $\mathbb{D}$, more precisely, for all $\mathbf{u} \in \mathbb{D}$, we have

$$W(\mathbf{u}) = \sum_{i=1}^{N} \alpha_i u_i \text{ with } \alpha_i > 0 \text{ and } \alpha_N = 1.$$

We show that, imposing the $\text{PD}(r)$ transfer principle for $r \geq 1$, leads to weights $\alpha_i = r^{N-i} w_i$ with $w_1 \geq \ldots \geq w_N = 1$. It is easy to verify that these weights guarantee the $\text{PD}(r)$ transfer principle. We show they are also necessary. The change in welfare due to a $\text{PD}(r)$ transfer of size $\delta > 0$ from individual j to i, with $j > i$, which does not change the position of *any* individual, has to be non-negative, or

$$\alpha_i (u_i + \delta) + \alpha_j (u_j - r\delta) - (\alpha_i u_i + \alpha_j u_j) \geq 0,$$

which is true if and only $r\alpha_j \leq \alpha_i$. As this should hold for any $i < j$ we obtain the desired result.

[9] See also Wakker (1993) for possible problems and a complete proof.

References

[1] Capéau, B. and Ooghe, E. (2004), On comparing heterogeneous populations: Is there really a conflict between the Pareto criterion and inequality aversion?, CES discussion paper series n° 04.07 (www.econ.kuleuven.ac.be/eng/ew/discussionpapers).

[2] Ebert, U. (1988), Measurement of inequality: an attempt at unification and generalization, *Social Choice and Welfare* 5, 147-169.

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

[4] Ebert, U. (1999), Using equivalent income of equivalent adults to rank income distributions, *Social Choice and Welfare* 16, 233-258.

[5] Ebert, U. and Moyes, P. (2003), Equivalence scales reconsidered, *Econometrica* 71 (1), 319-343.

[6] Glewwe, P. (1991), Household equivalence scales and the measurement of inequality: Transfers from the poor to the rich could decrease inequality, *Journal of Public Economics* 44, 211-216.

[7] Pyatt, G. (1990), Social evaluation criteria, in, C. Dagum and M. Zenga (eds.), *Income and Wealth Distribution, Inequality and Poverty*, Berlin: Springer-Verlag.

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

[9] Wakker, P. (1993), Additive representations on rank-ordered sets. II. The topological approach, *Journal of Mathematical Economics* 22, 1-26.