Udo Ebert*

Ethical inequality measures and the redistribution of income when needs differ

January 2005

* Address: Department of Economics, University of Oldenburg, D-26111 Oldenburg, Germany

Tel.: (+49)441-798-4113

Fax: (+49)441-798-4116

e-mail: ebert@uni-oldenburg.de

Abstract

The paper considers social welfare functions and ethical inequality measures for linear inequality concepts when households may differ in needs. The welfare functions are nested, i.e. separable in the level of welfare attained by the homogeneous subpopulations, and possess a homogeneity property depending on the inequality concept imposed. Several principles of transfers between different household types are introduced and systematically examined. Their implications for the form of welfare functions and inequality measures are derived. The corresponding classes are completely described.

Keywords: Welfare functions, inequality measures, inequality concepts, differences in needs, transfer principles, axiomatization.

JEL codes: D63, I31

1. Introduction[1]

The principle of progressive transfers postulates that the redistribution of (a small amount of) income from a richer individual to a poorer one increases the level of social welfare and decreases the degree of inequality. The idea goes back to Pigou and was generalized by Dalton (1920). The principle is without doubt the foundation stone of a normative theory of redistribution: Whenever one income distribution is more equal than another one (with the same average income) – measured by Lorenz dominance, it can be generated by a finite series of progressive transfers from the original income distribution. In this case, however, the individuals considered have to be identical in every respect but possibly income. Things become much more complicated if we assume that the economic units under consideration may also differ with respect to further attributes: For instance households can have different size or needs. Then it is not at all clear what kind of income transfer may improve welfare and equality since two variables – income and needs – have to be taken into account. Furthermore, the concept of inequality used has to be taken into consideration, as well.

The objective of this paper is to investigate some principles of redistribution in a heterogeneous framework systematically and to discuss their implications for the measurement of welfare. Moreover we translate these results to the measurement of inequality and derive the corresponding ethical inequality measures. We also examine various inequality concepts. The underlying framework can be described simply: a typical household can be characterized by its income and its type. (Both variables can in principle be observed!) We assume throughout the paper that household types can be ranked by needs.

Then a transfer principle has at first to compare the situation of the households involved: which one is 'richer', and which one 'poorer'? Here the notions of 'rich' and 'poor' have to take into consideration the living standards, i.e. the level of income *and* the household's needs. Thus the principles crucially depend on the comparability of these information. The weakest principle is only able to compare the level of income and needs separately, i.e. a household can be identified as the 'richer' one if its income is higher *and* if it is less needy. Thus the ranking of living standards is a dominance criterion and is incomplete. The strongest one is able to compare the situation of households with arbitrary incomes and needs. Here a complete ordering of the living standard of households is given. Between these extremes two further possibilities are discussed. Then in a second step, when the measuring of 'richer' and 'poorer' has been clarified, the transfer itself has to be defined.

[1] I would like to thank Patrick Moyes for helpful comments.

We examine a general class of social welfare orderings. The corresponding welfare functions are nested: at first the level of welfare of all households having the same type is determined. Then these welfare levels are aggregated to an overall welfare ordering in the second step. This class contains the set of separable welfare functions as proper subclass. Since ethical inequality measures are to be derived, the welfare orderings have to possess an appropriate homogeneity property depending on the inequality concept chosen. In the paper we consider linear inequality concepts (which are the only coherent inequality views, cf. Ebert (2004)), i.e. we derive measures according to the relative, absolute, intermediate, and reference-point inequality view.

The transfer principles introduced have different power. The weakest one only assumes that income and needs can be ranked. Another one allows to compare the income of different household types by using one-sided bounds. A third one is based on two-sided bounds. Finally, the strongest principle is formulated on the basis of an equivalent income function. It corresponds to the between-type transfer principle used in the literature. The transfer principles are imposed in two different ways: at first only transfers between two particular household types are considered. In a second step transfers between arbitrary subpopulations are admitted. The implications of these principles for the form of welfare orderings and inequality measures are derived for a fixed population. Furthermore, depending on the transfer principle imposed the corresponding classes are described completely. As expected, the weaker the transfer axiom the greater is the corresponding class of measures. We obtain some classes which have not yet been presented in the literature.

There are few papers dealing with this topic: Ebert (1995) introduces one generalized Pigou-Dalton principle for the measurement of relative inequality. Ebert (1997) is concerned with an analogous principle for absolute inequality. Shorrocks (1995) is primarily interested in some (incomplete) welfare and inequality orderings. In Ebert (2004) a different approach is used: here equivalent income functions and weights reflecting the type of household are given a priori and then a between-type transfer principle is imposed on a class of welfare and inequality orderings. To sum up, a systematic and general analysis of transfer principles for linear inequality concepts is still missing.

The paper is organized as follows: Section 2 introduces the notation. Section 3 presents the framework. In section 4 the transfer principles are defined and discussed. Their implications are investigated. Section 5 concludes.

2. Notation

We investigate a heterogeneous population. There are $n \geq 2$ types of household having different composition and/or needs. So we face n homogeneous subpopulations. We will assume that a household of type i comprises i adults, but this assumption is not necessary. Each subpopulation i, $i = 1,...,n$, consists of n_i households of type i. The total number of households is denoted by $N := \sum_{i=1}^{n} n_i$. It is important that the household types can be ranked by needs. They are numbered by *increasing* needs. We assume that the income X_j^i and the type i of a household j can be observed. Of course, households can vary in income. Each income has to be feasible, i.e. $X_j^i \in \Omega_d$, where $\Omega_d = (d, \infty)$ denotes the set of feasible incomes. For $d = 0$ we obtain $\Omega_0 = \mathbb{R}_{++}$ and for $d = -\infty$ the set $\Omega_{-\infty} = \mathbb{R}$. Incomes are either bounded from below or may be arbitrary. We can interpret d as minimum or reference income. If income is negative it is assumed that a household is able to survive by getting credit or by using savings. If $X_j^i \in \Omega_d$ the income $X_j^i - d$ is called normalized income for $d \in \mathbb{R}$. It is the amount of income of household j measured with respect to the minimum (reference) income.

Let $X^i = \left(X_1^i,...,X_{n_i}^i\right) \in \Omega_d^{n_i}$ denote an income vector of subpopulation i. A vector of incomes for the overall population is given by $X = \left(X^1,...,X^n\right) \in \Omega^N$. The average income of X is denoted by $\mu(X) = \Sigma\Sigma X_j^i / N$. Furthermore, let $\mathbf{1}_k$ denote a vector of k ones. An ordering defined on Ω_d^N is represented by $\succsim_d$, where $\sim_d$ and $\succ_d$ denote its symmetric and, respectively, asymmetric part.

3. Welfare and inequality

In the following we describe the framework we are interested in and prove a number of general results. They will provide the background for the analysis performed in the next section. At first we will introduce social welfare orderings. Then the class of coherent inequality concepts is presented and the derivation of ethical inequality measures is discussed. If an inequality measure has to be consistent with one of the inequality concepts examined, the underlying welfare ordering has to satisfy a homogeneity property. Therefore the class of feasible welfare orderings is determined. It turns out that the corresponding welfare functions are well-structured. They allow us to describe the corresponding class of ethical inequality orderings/measures precisely.

3.1 Welfare orderings

We will confine ourselves to a particular class of social welfare orderings $\succsim_d^W$ defined on Ω_d^N which can be represented by nested social welfare functions. Therefore in a first step we define a social welfare ordering $\succsim_d^{W_i}$ for each subpopulation i, $i = 1,...,n$. We assume that it can be represented by

$$\xi_i\left(X^i\right) = V_i^{-1}\left(\frac{1}{n_i}\sum_{j=1}^{n_i} V_i\left(X_j^i\right)\right) \tag{1}$$

where $V_i : \Omega_d \to \mathbb{R}$ is an increasing, strictly concave household utility function and where $V_i^{-1}(t)$ denotes its inverse function. It turns out that $\xi_i\left(X^i\right)$ is the equally distributed equivalent income (EDEI) of X^i: if given to each household in subpopulation *i* this income yields the same level of welfare as X^i. Because of the concavity of V_i the ordering $\succsim_d^{W_i}$ satisfies the (usual) principle of progressive transfers (the subpopulation is homogeneous!). In the second step we assume that the social welfare ordering $\succsim_d^W$ is defined on the vector $\left(\xi_1\left(X^1\right),...,\xi_n\left(X^n\right)\right)$ and that it can be represented by a welfare function

$$W(X) = V^{-1}\left(\sum_{i=1}^{n} \beta_i n_i V\left(\xi_i\left(X^i\right)\right)\right) \tag{2}$$

where V is again strictly increasing and concave, and where $\beta_1,...,\beta_n$ are strictly positive welfare weights. The basic idea underlying this construction can be described as follows: subpopulation *i*'s EDEI $\xi_i\left(X^i\right)$ is the representative income of a household of this type. Since there are n_i households in this subpopulation (we have to take into account the size of the subpopulation!) its total contribution to social welfare is $n_i V\left(\xi_i\left(X^i\right)\right)$. It is weighted by β_i which is to reflect the needs of household type *i*. Furthermore, the welfare function at this level is also separable. If we replace $\xi_i\left(X^i\right)$ by its proper definition we see that the welfare function $W(X)$ is a nested function having two levels. Since all functions involved are linear or concave $W(X)$ is strictly increasing and concave.[2]

Below, we require a social income function for the definition of ethical inequality orderings. The function represents the minimal amount of (aggregate) income that is just sufficient to yield the level of social welfare implied by a given income distribution (if total income is distributed optimally among households). We define it in analogy to an individual expenditure function in consumer theory by:

$$C(X) := \min_{\lambda_1^1,\dots,\lambda_{n_n}^n} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \lambda_j^i \tag{3a}$$

$$\text{s.t. } W\left(\lambda_1^1,\dots,\lambda_{n_n}^n\right) = W(X). \tag{3b}$$

Obviously there is a relationship to the concept of EDEI for homogeneous subpopulation. If there is only one subpopulation the value of the social income function is equal to the EDEI multiplied by the number of households. It is easy to see that the optimal incomes λ_j^i, for $j = 1,\dots,n_i$, have to be identical (since V_i is concave). Therefore $C(X)$ can also be represented by

$$C(X) = \min_{\lambda_1,\dots,\lambda_n} \sum_{i=1}^{n} n_i \lambda_i \tag{4a}$$

$$\text{s.t. } W\left(\lambda_1 \mathbf{1}_{n_1},\dots,\lambda_n \mathbf{1}_{n_n}\right) = W(X). \tag{4b}$$

3.2 Inequality concepts

Inequality measures (representing inequality orderings) are usually invariant with respect to a certain type of admissible transformation of incomes. Relative inequality measures do not change if all incomes are altered proportionally and absolute measures are invariant w.r.t. to adding the same amount to all incomes. In other words they are consistent with an inequality concept. We will examine the class of *coherent* inequality concepts (Ebert (2004a)). They have to satisfy two properties: path-independence and transfer-consistency. The first one requires that the composition of two admissible transformations is also admissible. Then

---

[2] Another possibility of defining an ordering $\succsim_d^W$ is to use a separable welfare function from the beginning and to consider $W(X) = V^{-1}\left(\Sigma\Sigma V_i\left(X_j^i\right)\right)$. Then there is only one level. It turns out that both forms coincide if and only if we define $V_i(t) = \beta_i V(t)$ for $i = 1,\dots,n$.

'having the same degree of inequality' is a transitive relation. The second one postulates that if a (sequence of) progressive transfers is needed in order to obtain one distribution from another one, then so it is the case after applying the same admissible transformation to both distributions. It requires that a change in the size of incomes (admissible transformation) and a redistribution of income are compatible with one another. This criterion implies in particular that Lorenz dominance (appropriately defined) is preserved if the distributions involved are transformed by admissible transformations.

It turns out that there is exactly one coherent inequality concept $\mathcal{J}_d$ for every Ω_d^N. The admissible transformations have to be *linear*. We obtain the functions

$$T_\lambda^d\left(t\right)=\lambda\left(t-d\right)+d \text{ for } t\in\Omega_d \text{ and } \lambda\in\mathbb{R}_{++} \text{ if } d\in\mathbb{R} \text{ and} \tag{5a}$$

$$T_\alpha^{-\infty}\left(t\right)=t+\alpha \text{ for } t\in\Omega_{-\infty} \text{ and } \alpha\in\mathbb{R} \text{ if } d=-\infty \tag{5b}$$

which define the transformations

$$T_\lambda^d\left(X\right):=\left(T_\lambda^d\left(X_1^1\right),\ldots,T_\lambda^d\left(X_{n_n}^n\right)\right) \text{ and } T_\alpha^{-\infty}\left(X\right):=\left(T_\alpha^{-\infty}\left(X_1^1\right),\ldots,T_\alpha^{-\infty}\left(X_{n_n}^n\right)\right).$$

For $d=0$ and $d=-\infty$ we get the relative and, respectively, absolute inequality concept, for $d<0$ the intermediate one and for $d>0$ reference-point inequality. They are discussed in more detail in Ebert (2004a).

3.3 Ethical inequality orderings

Ethical inequality orderings are derived from social welfare orderings. Suppose that $\succsim_d^W$ is a social welfare ordering for $d\in\mathbb{R}$. Then a corresponding inequality ordering $\succsim_d^I$ can be defined by means of the relative welfare loss due to inequality. It is calculated on the basis of the *normalized* income per household:

$$I_d\left(X\right):=\frac{\left(\mu\left(X\right)-d\right)-\left(C\left(X\right)/N-d\right)}{\mu\left(X\right)-d} \text{ for } d\in\mathbb{R}. \tag{6a}$$

Similarly, we consider the absolute welfare loss per household for $d=-\infty$

$$I_{-\infty}\left(X\right)=\mu\left(X\right)-C\left(X\right)/N. \tag{6b}$$

These indicators determine the corresponding inequality orderings uniquely.

But the orderings have to be consistent ($\mathcal{J}_d$-invariant) with the inequality concept $\mathcal{J}_d$, i.e.

$$X \sim_d^I T_\lambda^d(X) \text{ for all } X \in \Omega_d^N \text{ and } \lambda \in \mathbb{R}_{++} \text{ if } d \in \mathbb{R} \tag{7a}$$

and

$$X \sim_{-\infty}^I T_\alpha^d(X) \text{ for } X \in \Omega_d^N \text{ and } \alpha \in \mathbb{R}. \tag{7b}$$

We call a social welfare ordering $\mathcal{J}_d$-homogeneous with respect to the admissible transformations if for all $X, Y \in \Omega_d^N$:

$$X \sim_d^W Y \Leftrightarrow T_\lambda^d(X) \sim_d^W T_\lambda^d(Y) \text{ for all } \lambda \in \mathbb{R}_{++} \text{ if } d \in \mathbb{R} \tag{8a}$$

and, respectively,

$$X \sim_{-\infty}^W Y \Leftrightarrow T_\alpha^{-\infty}(X) \sim_{-\infty}^W T_\alpha^{-\infty}(Y) \text{ for all } \alpha \in \mathbb{R} \text{ if } d = -\infty \tag{8b}$$

and get[3]

Proposition 1

$\succsim_d^I$ is $\mathcal{J}_d$-invariant if and only if the corresponding ordering $\succsim_d^W$ is $\mathcal{J}_d$-homogeneous. Therefore $\mathcal{J}_d$-homogeneity is a crucial property of the welfare orderings we want to consider.

3.4 Homogeneous welfare orderings

We restrict ourselves to the class of $\mathcal{J}_d$-homogeneous welfare orderings which can be described precisely:

Proposition 2

a) *$\succsim_d^W$ is $\mathcal{J}_d$-homogeneous if and only if $\succsim_d^{W_i}$ and the overall ordering defined on $\left(\xi_1\left(X^1\right),\ldots,\xi_n\left(X^n\right)\right)$ are $\mathcal{J}_d$-homogeneous.*

b) *For $d \in \mathbb{R}$: $\succsim_d^W$ is $\mathcal{J}_d$-homogeneous if and only if there are $\varepsilon, \varepsilon_1, \ldots, \varepsilon_n \in (\infty, 1)$ and $\beta_1 = 1, \beta_2, \ldots, \beta_n \in \mathbb{R}_{++}$ such that $\succsim_d^W$ is represented by*[4]

$$W_d(X) = \left(\sum_{i=1}^{n} \frac{\beta_i n_i}{\Sigma \beta_h n_h} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(X_j^i - d \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon} + d \;\; for \;\; X \in \Omega_d^N. \tag{9a}$$

---

3 All proofs have been relegated to the Appendix.

4 If an exponent is equal to zero we have to use the corresponding geometric mean.

and

$$\tilde{W}_d(X) = \left(\sum_{i=1}^{n} \frac{m_i n_i}{\Sigma m_h n_h} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon} + d \;\; for \;\; X \in \Omega_d^N . \tag{9b}$$

where $m_i = \beta_i^{\frac{1}{1-\varepsilon}}$.

c) $\succsim_{-\infty}^W$ *is* $\mathcal{J}_{-\infty}$*-homogeneous if and only if there are* $\gamma, \gamma_1, ..., \gamma_n \in \mathbb{R}_{++}$ *and* $\beta_1 = 1, \beta_2, ..., \beta_n \in \mathbb{R}_{++}$ *such that* $\succsim_d^W$ *is represented by*

$$W_{-\infty}(X) = -\frac{1}{\gamma} \ln \left(\sum_{i=1}^{n} \frac{\beta_i n_i}{\Sigma \beta_h n_h} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i X_j^i} \right)^{\gamma/\gamma_i} \right) for \;\; X \in \Omega_{-\infty}^N . \tag{10a}$$

and

$$\tilde{W}_{-\infty}(X) = -\frac{1}{\gamma} \ln \left(\sum_{i=1}^{n} \frac{n_i}{N} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i \left(X_j^i - a_i \right)} \right)^{\gamma/\gamma_i} \right) for \;\; X \in \Omega_{-\infty}^N . \tag{10b}$$

where $a_i = \frac{1}{\gamma} \ln \beta_i$.

Thus homogeneity has to be satisfied at both levels. In particular we get for the EDEI $\xi_i\left(T_\lambda^d\left(X^i\right)\right) = T_\lambda^d\left(\xi_i\left(X^i\right)\right)$, for $i = 1,...,n$. The feasible welfare functions have to be generalized Atkinson and, respectively, Kolm-Pollak social welfare functions.[5] At this stage the parameters $(\varepsilon, \varepsilon_1, ..., \varepsilon_n$ and $\gamma, \gamma_1, ..., \gamma_n)$ may differ and may be chosen independently for the overall population and subpopulations. These parameters represent the respective inequality aversion.

Closer inspection of the welfare functions (9a) and (10a) and the discussion in Ebert (1995, 1997) demonstrate that homogeneous welfare orderings can also be represented in a different way. By introducing the constants m_i and a_i and rearranging we obtain (9b) and (10b). They are interpreted in the next subsection.

---

[5] If we use the orderings described in footnote 1 and impose $\mathcal{J}_d$-homogeneity the corresponding welfare functions also possess the form (9) and, respectively, (10), but we obtain $\varepsilon = \varepsilon_1 = ... = \varepsilon_n$ and $\gamma = \gamma_1 = ... = \gamma_n$ a priori.

3.5 Measurement of inequality

Given the representation in Proposition 2 we are now able to derive the corresponding social income functions. We get

Proposition 3

a) *For* $d \in \mathbb{R}$*: The social income function for (9) is given by*

$$C_d(X) = \left(\sum_{h=1}^{n} \beta_h^{\frac{1}{1-\varepsilon}} n_h\right)\left(\sum_{i=1}^{n} \frac{\beta_i n_i}{\Sigma \beta_h^{\frac{1}{1-\varepsilon}} n_h}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}\left(X_j^i - d\right)^{\varepsilon_i}\right)^{\varepsilon/\varepsilon_i}\right)^{1/\varepsilon} + Nd$$

$$= \left(\sum_{h=1}^{n} m_h n_h\right)\left(\sum_{i=1}^{n} \frac{m_i n_i}{\Sigma m_h n_h}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{X_j^i - d}{m_i}\right)^{\varepsilon_i}\right)^{\varepsilon/\varepsilon_i}\right)^{1/\varepsilon} + Nd. \tag{11}$$

b) *For* $d = -\infty$*: The social cost function for (10) is given by*

$$C_{-\infty}(X) = N\left[-\frac{1}{\gamma}\ln\left(\sum_{i=1}^{n} \frac{\beta_i n_i}{N}\left(\frac{1}{n_i}\sum_{j=1}^{n_i} e^{-\gamma_i X_j^i}\right)^{\gamma/\gamma_i}\right)\right] + \frac{1}{\gamma}\sum_{i=1}^{n} n_i \ln \beta_i$$

$$= N\left[-\frac{1}{\gamma}\ln\left(\sum_{i=1}^{n} \frac{n_i}{N}\left(\frac{1}{n_i}\sum_{j=1}^{n_i} e^{-\gamma_i\left(X_j^i - a_i\right)}\right)^{\gamma/\gamma_i}\right)\right] + \sum_{i=1}^{n} n_i a_i. \tag{12}$$

By the definition of the social income function $\lambda_1,\ldots,\lambda_n$ represent the optimal incomes of household types $1,\ldots,n$. The optimality conditions of this minimization process are given by

$$\frac{\lambda_i - d}{m_i} = \frac{\lambda_1 - d}{m_1} = \lambda_1 - d \text{ and } \quad \lambda_i - a_i = \lambda_1 - a_1 = \lambda_1 \text{ for } i = 1,\ldots,n. \tag{13}$$

They can be rearranged to

$$\lambda_1 = \frac{\lambda_i}{m_i} - \left(\frac{1}{m_i} - 1\right)d \text{ and } \lambda_1 = \lambda_i - a_i. \tag{14}$$

These equations can be interpreted as implicit equivalent income functions given single adults as reference type (see subsection 4.1 below and Ebert (2000b)) since they define the relationship between the optimal incomes of different household types in an *optimal* income distribution. Therefore the constants $m_1,\ldots,m_n$ and $a_1,\ldots,a_n$ can be interpreted as relative and, respectively, absolute implicit equivalence scales. Since the welfare functions are ordinal, relative scales are unique up to a scale factor; absolute scales can be changed by adding a constant. Therefore, without loss of generality we can choose household type 1 (single adults) as reference type by setting $m_1 = 1$ and $a_1 = 0$. $\left(X_j^i - d\right)\big/ m_i$ and, respectively, $X_j^i - a_i$ then represents the implicit equivalent income of a representative equivalent adult in household *j* of type *i*. For $d \in \mathbb{R},\quad d \neq 0$ we get a combination of relative and absolute scales.

Using the definition of inequality measures presented in subsection 3.3 and the result of Proposition 3 we are now able to describe the inequality orderings more precisely.

Proposition 4

The social welfare functions (9) and (10) imply

$$I_d\left(X\right) = 1 - \left(\sum_{i=1}^{n} \frac{m_i n_i}{\Sigma m_h n_h} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon} \Bigg/ \mu_d\left(X\right) \tag{15}$$

and

$$I_{-\infty}\left(X\right) = \mu_{-\infty}\left(X\right) + \frac{1}{\gamma} \ln \left(\sum_{i=1}^{n} \frac{n_i}{N} \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i \left(X_j^i - a_i\right)} \right)^{\gamma/\gamma_i} \right) \tag{16}$$

where

$$\mu_d\left(X\right) = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \frac{m_i}{\Sigma m_h n_h} \left(\frac{X_j^i - d}{m_i} \right) \quad \textit{and} \quad \mu_{-\infty}\left(X\right) = \frac{1}{N} \sum_{i=1}^{n} \sum_{j=1}^{n_i} \left(X_j^i - a_i \right), \qquad (17) + (18)$$

respectively.

We will come back to these representations of the inequality measures in section 5. Now we turn to a systematic discussion of transfer principles and of their implications.

4. Redistribution

The starting point of the following analysis is the Pigou-Dalton principle for homogeneous populations. It requires that a progressive transfer (i.e. a transfer from a richer to a poorer individual which does not reverse the ranking of incomes) improves social welfare (and decreases inequality). In our framework things are more complicated since the households belonging to different subpopulations have different needs. These differences in needs have to be taken into account, as well, when income is redistributed.

In this section we will introduce four different types of progressive transfers between different subpopulations. They will differ with respect to the assumption made on the comparability of living standards which depend on income and needs.

4.1 Transfer principles

The first kind of transfer takes into account only the ranking of income and needs. We introduce

WBT(d) *A transfer of income* $\delta > 0$ *changing* X_k^i *to* $X_k^i - \delta$ *and* X_l^{i+1} *to* $X_l^{i+1} + \delta$ *is a Weak Between-Type Progressive Transfer if* $X_k^i - \delta > X_l^{i+1} + \delta$.

The transfer redistributes income from a less needy and richer household to a needier and poorer household such that after redistributing income the first one is still richer than the second one. Here only information about the ranking of income and needs is required. Needs are not compared in a cardinal manner and it is clear that after the redistribution of income the household receiving the transfer is still worse off than the other one. The richer household is definitely better off than the poorer one, i.e. its income is higher and it is less needy. The definition can be employed for any $d \in \mathbb{R}$ and $d = -\infty$. Furthermore, the condition comparing incomes is equivalent to a formulation based on normalized incomes:

$$\left(X_k^i - \delta\right) - d > \left(X_l^{i+1} + \delta\right) - d \text{ for } d \in \mathbb{R}. \tag{19}$$

This kind of transfer is similar to Hammond's equity principle (which is applied to utility levels; cf. Hammond (1976)). It is identical with the transfer P3 used in Ebert (2000a) and also plays a role in Bourguignon (1989).

Sometimes one is willing to compare the living standards of households belonging to different subpopulations 'a little bit'; a(n upward) transfer from a household of type i to a household of type $i+1$ may even be desirable if the income of the recipient is a bit higher in the end than the income of the donor. The problem is to define the term 'a bit higher'. Formally we introduce

UBT(d) *Assume that* $r_i < 1$. *A transfer of income* $\delta > 0$ *changing* X_k^i *to* $X_k^i - \delta$ *and* X_l^{i+1} *to* $X_l^{i+1} + \delta$ *is an Upward Between-Type Progressive Transfer if*

$$\left(X_k^i - \delta\right) - d > r_i\left[\left(X_l^{i+1} + \delta\right) - d\right].$$

In this case the constant r_i determines the term ‘a bit’ precisely (in a relative way, i.e. by a ratio of incomes). The definition of this transfer is based on normalized income and is valid for $d \in \mathbb{R}$. For $d = -\infty$ we define

UBT($-\infty$) *Assume that* $b_i < 0$. *A transfer of income* $\delta > 0$ *changing* X_k^i *to* $X_k^i - \delta$ *and* X_l^{i+1} *to* $X_l^{i+1} + \delta$ *is an Upward Between-Type Progressive Transfer if*

$$\left(X_k^i - \delta\right) > \left(X_l^{i+1} + \delta\right) + b_i .$$

Here the constant b_i determines the bound.

If $r_i = 1$ or $b_i = 0$ the transfer UBT(d) coincides with WBT(d). Thus an upward between-type progressive transfer is an extension of the latter type of transfer. It is easy to see that constants r_i or b_i can really be chosen in practice. For instance two adults need a higher level of income than a single adult in order to be as well off. A constant $r_i = 10/11$ means that a transfer from the single to the couple is admissible as long as the couple’s (resulting) income is less than 110 % of the single person’s income.

If a needier household’s income is high enough we can also redistribute income to a less needy one. E.g. if a couple’s income is (more than) twice as high as a single adult’s income it seems to be better off. Therefore we introduce

DBT(d) *Assume that* $s_i > 1$. *A transfer of income* $\delta > 0$ *changing* X_l^{i+1} *to* $X_l^{i+1} - \delta$ *and* X_k^i *to* $X_k^i + \delta$ *is a Downward Between-Type Progressive Transfer if*

$$\left(X_l^{i+1} - \delta\right) - d > s_i\left[\left(X_k^i + \delta\right) - d\right]$$

for $d \in \mathbb{R}$ and

DBT($-\infty$) *Assume that* $c_i > 0$. *A transfer of income* $\delta > 0$ *changing* X_l^{i+1} *to* $X_l^{i+1} - \delta$ *and* X_k^i *to* $X_k^i + \delta$ *is a Downward Between-Type Progressive Transfer if*

$$\left(X_l^{i+1} - \delta\right) > \left(X_k^i + \delta\right) + c_i .$$

The interpretation of this kind of transfer is analogous.

Finally we assume that we are able to compare living standards completely. Suppose that an equivalent income function[6] E is explicitly defined in the following way: Choosing household type 1 (single adults) as reference type we introduce an equivalent income function E as a vector of n functions $E_i : \Omega_d \to \Omega_d$, where $E_i(t)$ denotes the income a single adult requires in order to be as well off as a household with i adults and household income t. (Here it is assumed implicitly that all members belonging to a household attain the same living standard.) The functions E_i have to satisfy some properties which allow us to make meaningful comparisons.

(i) $E_1(t) = t$

(ii) $E_i(t)$ is continuous and strictly increasing in t

(iii) $E_i(t)$ is strictly decreasing in i

(iv) $E_i(t)$ is an invertible function.

For an interpretation of these properties see e.g. Ebert (2000b).

$E_i(t)$ is called equivalent income. Then a household of type i is better off than a household of type h if and only if $E_i\left(X_j^i\right) > E_h\left(X_k^h\right)$ for any $h, i \in \{1,...,n\}$. If $E_i\left(X_j^i\right) = E_h\left(X_k^h\right)$ the living standards are the same. When an equivalent income function is given the living standards (equivalent incomes) of arbitrary households and household types can be compared. Then we define an appropriate transfer for $d \in \mathbb{R}$ or $d = -\infty$ by

SBT(d) *A transfer of income* $\delta > 0$ *changing* X_k^i *to* $X_k^i - \delta$ $\left[X_l^{i+1} \text{ to } X_l^{i+1} - \delta\right]$ *and* X_l^{i+1} *to* $X_l^{i+1} + \delta$ $\left[X_k^i \text{ to } X_k^i + \delta\right]$ *is a Strong Between-Type Progressive Transfer if* $E_i\left(X_k^i - \delta\right) > E_{i+1}\left(X_l^{i+1} + \delta\right)$ $\left[E_{i+1}\left(X_l^{i+1} - \delta\right) > E_i\left(X_k^i + \delta\right)\right]$.

It is the 'usual' Between-Type Progressive Transfer used in the literature (see e.g. Ebert (2004b)).

After having introduced four different kinds of transfers we define the corresponding **Principles of Transfers** which will be denoted by the *same* acronyms. They require that the respective transfer improves social welfare (and decreases inequality).

[6] See Donaldson and Pendakur (2004)

In the following we will always assume that the social welfare ordering $\succsim_d^W$ can be represented by a welfare function satisfying (9a) and, respectively, (10a). The implications for the representation (9b) and (10b) are discussed in section 5.

4.2 Two subpopulations

Since we want to analyze the implications of the transfer principles in detail we at first consider *two* subpopulations i and $i+1$. It is clear that the results depend on the principle chosen and the inequality concept considered. The latter determines a class of feasible social welfare orderings (see Proposition 2). Given this class of welfare orderings and the social welfare functions which represent them we can describe the implications of a transfer principle by two conditions: The parameters ($\varepsilon, \varepsilon_i, \varepsilon_{i+1}$ and, respectively, $\gamma, \gamma_i, \gamma_{i+1}$) have to be related and we obtain a condition on the weights β_i, β_{i+1}.

(a) Principle WBT

Starting with $d \in \mathbb{R}$ we obtain

Proposition 5a

$\succsim_d^W$ satisfies WBT(d) for $d \in \mathbb{R}$ if and only if

$$[\varepsilon_{i+1} \le \varepsilon \le \varepsilon_i \qquad \text{for } \varepsilon_{i+1} > 0, \tag{20a}$$

$$\varepsilon_{i+1} = \varepsilon \le 0 < \varepsilon_i \qquad \text{for } \varepsilon_{i+1} \le 0 \text{ and } \varepsilon_i > 0 \tag{20b}$$

$$\varepsilon_{i+1} = \varepsilon = \varepsilon_i \qquad \text{for } \varepsilon_{i+1} \le 0 \text{ and } \varepsilon_i \le 0, \tag{20c}$$

and

$$\beta_i \le \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i - \varepsilon}{\varepsilon_i}} \left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon - \varepsilon_{i+1}}{\varepsilon_{i+1}}} \beta_{i+1}].$$

In this case we have to distinguish several cases as far as the parameters are concerned: It turns out that the parameter (inequality aversion) of the receiver ε_{i+1} must not exceed (be higher than) the parameter (inequality aversion) of the total population ε which in turn must be less (higher) than or equal to the donor's parameter (inequality aversion) ε_i. In other words the inequality aversion of the needier subpopulation has to be weakly higher than the aversion of the less needy one. This is true as long as the parameters are strictly positive. If at least one is nonpositive, two or all three parameters have to be identical. Otherwise it is impossible to satisfy WBT(d). Furthermore we get a condition on the weights β_i and β_{i+1}. It may also depend on the size of the subpopulations involved. The condition requires that the weight of the needier subpopulation is greater than the other one.

For $d = -\infty$ we obtain a clearer result:

Proposition 5b

$\succsim_d^W$ *satisfies* $WBT(-\infty)$ *if and only if* $\gamma_i = \gamma = \gamma_{i+1}$ and $\beta_i \le \beta_{i+1}$.

In this case all parameters have to be identical and the coefficients have to be nondecreasing in needs.

(b) Principle UBT

We know that UBT is a more general transfer principle than WBT since here an additional bound is given. In this case the conditions on the weights β_i and β_{i+1} are stricter:

Proposition 6

a) $\succsim_d^W$ *satisfies UBT(d) for* $d \in \mathbb{R}$ *if and only if (20) and* $\beta_i \le \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i - \varepsilon}{\varepsilon_i}} \left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon - \varepsilon_{i+1}}{\varepsilon_{i+1}}} r_i^{1-\varepsilon} \beta_{i+1}$.

b) $\succsim_d^W$ *satisfies UBT(*$-\infty$*) if and only if* $\gamma_i = \gamma = \gamma_{i+1}$ *and* $\beta_i \le e^{\gamma_i b_i} \beta_{i+1}$.

Since $r_i < 1$ and $b_i < 0$ the increase of β_{i+1} measured with respect to β_i has to be even larger than above.

(c) Principle DBT

As one would expect the results for downward transfers are essentially analogous. We have to replace (20) by

$$\varepsilon_i \le \varepsilon \le \varepsilon_{i+1} \qquad \text{for } \varepsilon_i > 0 \tag{21a}$$

$$\varepsilon_i = \varepsilon \le 0 < \varepsilon_{i+1} \qquad \text{for } \varepsilon_i \le 0 \text{ and } \varepsilon_{i+1} > 0 \tag{21b}$$

$$\varepsilon_i = \varepsilon = \varepsilon_{i+1} \qquad \text{for } \varepsilon_i \le 0 \text{ and } \varepsilon_{i+1} \le 0. \tag{21c}$$

Then we obtain

Proposition 7

a) $\succsim_d^W$ *satisfies DBT(d) for* $d \in \mathbb{R}$ *if and only if (21) and* $\beta_i \geq \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i - \varepsilon}{\varepsilon_i}} \left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon - \varepsilon_{i+1}}{\varepsilon_{i+1}}} s_i^{\varepsilon-1} \beta_{i+1}$.

b) $\succsim_d^W$ *satisfies DBT(−∞) if and only if* $\gamma_i = \gamma = \gamma_{i+1}$ *and* $\beta_i \geq e^{-\gamma_i c_i} \beta_{i+1}$.

The results are similar to those above: The inequality aversion ε_i of the subpopulation receiving the transfer must not be lower than the donor's one. Furthermore, the ratio β_i / β_{i+1} now has a lower bound depending on s_i and c_i, respectively.

(d) Principles UBT and DBT

Both principles can be combined. Then we have to exclude 'circularity': it must not be possible to apply at first one principle and then the other one to the same households. This requires $r_i s_i \geq 1$ and $b_i + c_i \geq 0$. We obtain a combination of the results presented in subsection (b) and (c):

Proposition 8

a) *Assume that* $d \in \mathbb{R}$ *and* $r_i s_i \geq 1$. *Then*

$\succsim_d^W$ *satisfies UBT(d) and DBT(d) if and only if* $\varepsilon_i = \varepsilon = \varepsilon_{i+1}$ *and* $\beta_{i+1} \in \left[r_i^{\varepsilon-1} \beta_i, s_i^{1-\varepsilon} \beta_i \right]$

b) *Assume that* $b_i + c_i \geq 0$. *Then*

$\succsim_{-\infty}^W$ *satisfies UBT(−∞) and DBT(−∞) if and only if* $\gamma_i = \gamma = \gamma_{i+1}$ *and* $\beta_{i+1} \in \left[e^{-\gamma b_i} \beta_i, e^{\gamma c_i} \beta_i \right]$.

In this case even for $d \in \mathbb{R}$ the parameters have to be identical. We obtain precise restrictions for the weight β_{i+1}. The latter must belong to the interval $\left[r_i^{\varepsilon-1} \beta_i, s_i^{1-\varepsilon} \beta_i \right]$ and, respectively, $\left[e^{-\gamma b_i} \beta_i, e^{\gamma c_i} \beta_i \right]$. Thus the constants r_i, s_i and b_i, c_i impose bounds on the weights. The size of the subpopulations involved is no longer important since $\varepsilon = \varepsilon_i = \varepsilon_{i+1}$.

4.3 Many subpopulations

Finally we consider the implications if one of the transfer principles introduced above is imposed on *each* pair of subpopulations *i* and $i+1$ for $i = 1, ..., n-1$.

We establish

Proposition 6* [7*]

a) $\succsim_d^W$ *satisfies UBT(d) [DBT(d)] for* $i = 1,...,n-1$ *and* $d \in \mathbb{R}$ *if and only if*

$$\{(\varepsilon_n \le \varepsilon_{n-1} = ... = \varepsilon_2 = \varepsilon \le \varepsilon_1 \quad for\ \varepsilon_n > 0$$

or $\quad \varepsilon_n = \varepsilon_{n-1} = ... = \varepsilon_2 = \varepsilon \le \varepsilon_1 \quad$ *for* $\varepsilon_n \le 0$ *and* $\varepsilon_1 > 0$

or $\quad \varepsilon_1 = ... = \varepsilon_n = \varepsilon$ *otherwise)*

$$[\varepsilon_1 \le \varepsilon_2 = ... = \varepsilon_{n-1} = \varepsilon \le \varepsilon_n \quad for\ \varepsilon_1 > 0$$

or $\quad \varepsilon_1 = \varepsilon_2 = ... = \varepsilon_2 = \varepsilon \le \varepsilon_n \quad$ *for* $\varepsilon_1 \le 0$ *and* $\varepsilon_n > 0$

or $\quad \varepsilon_1 = ... = \varepsilon_n = \varepsilon$ *otherwise]*

and

$$(\beta_1 \le \left(\frac{1}{n_1}\right)^{\frac{\varepsilon_1-\varepsilon}{\varepsilon_1}} r_1^{1-\varepsilon}\,\beta_2, \quad \beta_i \le r_i^{1-\varepsilon}\,\beta_{i+1}\ for\ i = 2,...,n-2, \quad and\ \beta_{n-1} \le \left(\frac{1}{n_n}\right)^{\frac{\varepsilon-\varepsilon_n}{\varepsilon_n}} r_{n-1}^{1-\varepsilon}\,\beta_n)$$

$$[\beta_1 \ge \left(\frac{1}{n_1}\right)^{\frac{\varepsilon_1-\varepsilon}{\varepsilon_1}} s_1^{\varepsilon-1}\,\beta_2, \quad \beta_i \ge s_i^{\varepsilon-1}\,\beta_{i+1}\ for\ i = 2,...,n-2, \quad and\ \beta_{n-1} \ge \left(\frac{1}{n_n}\right)^{\frac{\varepsilon-\varepsilon_n}{\varepsilon_n}} s_{n-1}^{\varepsilon-1}\,\beta_n]\}.$$

b) $\succsim_{-\infty}^W$ *satisfies UBT($-\infty$) [DBT($-\infty$)] for* $i = 1,...,n-1$ *if and only if*

$$\gamma = \gamma_1 = ... = \gamma_n \ and$$

$$\beta_i \le e^{\gamma_i b_i}\,\beta_{i+1} \qquad for\ i = 1,...,n-1$$

$$[\beta_i \ge e^{-\gamma_i c_i}\,\beta_{i+1} \qquad for\ i = 1,...,n-1].$$

This proposition is easily proved by repeated application of the respective results derived in subsection 4.2. It is important to note that Proposition 6* also presents the outcome for an application of WBT(d) and WBT($-\infty$) (set $r_i \equiv 1$ and $b_i \equiv 0$). It turns out that the parameters ε_i have to be identical to ε for $i = 2,...,n-1$ (since these subpopulations may be the receiver *and* the donor of a transfer). The inequality aversion for the neediest and least needy subpopulation may differ from ε. If UBT(d) and DBT(d) are applied simultaneously (for $d \in \mathbb{R}$) this property vanishes. We obtain

Proposition 8*

a) *Assume that* $d \in \mathbb{R}$ *and* $r_i < 1, \quad s_i > 1, \quad r_i s_i \ge 1$ *for* $i = 1,...,n$.

$\succsim_d^W$ *satisfies UBT(d) and DBT(d) for* $i = 1,\dots,n-1$ *if and only if* $\varepsilon_1 = \dots = \varepsilon_n = \varepsilon$ *and* $\beta_{i+1} \in \left[r_i^{\varepsilon-1} \beta_i, s_i^{1-\varepsilon} \beta_i \right]$ *for* $i = 1,\dots,n-1$.

b) *Assume that* $b_i < 0, \quad c_i > 0, \quad b_i + c_i \geq 0$.

$\succsim_{-\infty}^W$ *satisfies UBT(−∞) and DBT(−∞) for* $i = 1,\dots,n-1$ *if and only if* $\gamma_1 = \dots = \gamma_n = \gamma$ *and* $\beta_{i+1} \in \left[e^{-\gamma b_i} \beta_i, e^{\gamma c_i} \beta_i \right]$ *for* $i = 1,\dots,n-1$.

Proposition 8* extends the results derived in Ebert (1995, 1997) to inequality concepts $d \neq 0$ and $d \neq -\infty$.

Finally we examine the Strong Between-Type Transfer Principle. In this case the living standard of arbitrary subpopulations can be compared. We get

Proposition 9

a) $\succsim_d^W$ *satisfies SBT(d) for* $i = 1,\dots,n-1$ *and* $d \in \mathbb{R}$ *if and only if* $\varepsilon_1 = \dots = \varepsilon_n = \varepsilon$, $1 = \beta_1 \leq \beta_2 \leq \dots \leq \beta_n$, *and* $E_i(t) = \beta_i^{\varepsilon-1}(t-d) + d$ *for* $i = 1,\dots,n$.

b) $\succsim_{-\infty}^W$ *satisfies SBT(−∞) for* $i = 1,\dots,n-1$ *if and only if* $\gamma_1 = \dots = \gamma_n = \gamma$, $1 \leq \beta_1 \leq \beta_2 \leq \dots \beta_n$, *and* $E_i(t) = t - \dfrac{1}{\gamma} \ln \beta_i$ *for* $i = 1,\dots,n$.

In this case the welfare ordering and the equivalent income function used in the transfer principle SBT are closely related. One has to employ relative and, respectively, absolute equivalence scales. These scales are already uniquely implied by the welfare weights β_i. Then these equivalence scales and the implicit scales (see subsection 3.5) are identical.

5. Discussion and conclusion

Section 4 has dealt with the characterization of some classes of social welfare orderings. Therefore the results are described in terms of the coefficients of the corresponding welfare functions (9a) and (10a). For a description of the inequality measures it seems to be easier to use the ‘implicit’ equivalence scales. Therefore we have to translate the restrictions imposed on the coefficients into corresponding restrictions for scales. Since $\beta_i = m_i^{1-\varepsilon}$ and, respectively, $\beta_i = e^{\gamma a_i}$, we obtain the following equivalences:

$$\beta_i \le \beta_{i+1} \Leftrightarrow m_i \le m_{i+1},$$

$$\beta_{i+1} \in \left[r_i^{\varepsilon-1}\beta_i, s_i^{\varepsilon-1}\beta_i \right] \Leftrightarrow m_{i+1} \in \left[m_i/r_i, s_i m_i \right] \text{ for } d \in \mathbb{R}$$

and

$$\beta_i \le \beta_{i+1} \Leftrightarrow a_i \le a_{i+1},$$

$$\beta_{i+1} \in \left[e^{-\gamma b_i}\beta_i, e^{\gamma c_i}\beta_i \right] \Leftrightarrow a_{i+1} \in \left[a_i - b_i, a_i + c_i \right] \text{ for } d = -\infty .$$

These conditions allow us to describe the implications of the transfer principles for the measurement of inequality.

The analysis presented above characterizes several (new) classes of welfare orderings and inequality orderings. Suppose, for example, that one wants to impose the transfer principle WBT and adheres to the relative inequality view. Then, for two different types, welfare functions of the following form satisfy this principle:

$$\left(\sum_{i=1}^{2} \beta_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(X_j^i \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon}$$

for $\beta_1 = 1 \le \beta_2$ and $\varepsilon_2 \le \varepsilon \le \varepsilon_1$.

Choosing $\varepsilon_1 = 3/4$, $\varepsilon = 1/2$, $\varepsilon_2 = 1/4$, and $\beta_2 = 2$ we get the inequality measure

$$I(X) = 1 - \left\{ \frac{1}{1+\sqrt{2}} \left[\frac{1}{n_1} \sum_{j=1}^{n_1} \left(X_j^1 \right)^{3/4} \right]^{2/3} + \frac{\sqrt{2}}{1+\sqrt{2}} \left[\frac{1}{n_2} \sum_{j=1}^{n_2} \left(\frac{X_j^2}{\sqrt{2}} \right)^{1/4} \right]^2 \right\}^2 \Bigg/ \mu_0(X)$$

which is a relative measure and allows us to treat the subpopulations differently. This form of inequality measure has not yet been discussed in the literature.

The objective of the paper was to characterize several classes of ethical inequality measures when households may differ in needs and household types can be ranked by needs. The starting point of the analysis is the distribution of household income. Whereas in practice the distribution of household income is often adjusted in order to take into account differences in needs (by introducing weights and equivalizing incomes), here social welfare orderings are directly defined on the income distribution observed. In this framework two-level welfare orderings have been considered: In a first step the level of welfare of all households having the same type is determined. In a second step these welfare levels are aggregated to an overall welfare ordering which can accordingly be represented by a two-level welfare function.

Since we are interested in the derivation of ethical inequality measures we had at first to choose the class of inequality concepts. In this paper the set of linear or coherent concepts is examined. Then one has to describe the way an inequality measure is derived from a welfare function. It is defined as the welfare loss per household due to inequality by means of the social income function. Since the ethical inequality measures have to be consistent with the respective inequality concept the corresponding welfare function has to satisfy an appropriate homogeneity property. Given the class of nested welfare orderings and an inequality concept, the corresponding subclass of feasible welfare functions and inequality measures has been characterized precisely.

Then four transfer principles (concerning the transfer of income between different household types) have been introduced and examined. They possess different power and define therefore different subclasses of welfare functions and inequality measures. As expected, the weaker the transfer the greater is the corresponding class of inequality measures.

To sum up, the systematic analysis of this paper allows us to make a reasonable choice among several transfer principles in a heterogeneous population (for linear inequality concepts). The functional structure of the corresponding inequality measures has been completely derived. New possibilities have opened up.

References

Aczél, J. (1966), Lectures on Functional Equations and Their Applications, Academic Press, New York.

Bourguignon, F. (1989), Family size and social utility. Income distribution dominance criteria, Journal of Econometrics 42, 67-80.

Dalton, H. (1920), The measurement of inequality of income, Economic Journal 30, 348-361.

Donaldson, D. and K. Pendakur (2004), Equivalent-expenditure functions and expenditure-dependent equivalence scales, Journal of Public Economics 88, 175-208.

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

Ebert, U. (1997), Absolute inequality indices and equivalence differences, in: S. Zandvikili (Ed.), Taxation and Inequality, Research on Economic Inequality, Vol. 7, JAI Press, 131-152.

Ebert, U. (2000a), Sequential generalized Lorenz dominance and transfer principles, Bulletin of Economic Research 52, 113-122.

Ebert, U. (2000b), Equivalizing incomes: a normative approach, International Tax and Public Finance 7, 619-640.

Ebert, U. (2004a), Coherent inequality views: linear invariant measures reconsidered, Mathematical Social Sciences 47, 1-20.

Ebert, U. (2004b), Social welfare, inequality, and poverty when needs differ, Social Choice and Welfare 23, 415-448.

Eichhorn, W. (1978), Functional Equations in Economics, Addison-Wesley, Reading.

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

Shorrocks, A.F. (1995), Inequality and welfare evaluation of heterogeneous income distribution. Discussion Paper, University of Essex.

APPENDIX

Proof of Proposition 1

We consider $I_d\left(T_\lambda^d(X)\right)$ for $d \in \mathbb{R}$. Then

$$I_d\left(T_\lambda^d(X)\right) = \frac{\left(\mu\left(\lambda\left(X-d\mathbf{1}\right)+d\mathbf{1}\right)-d\right)-\left(C_d\left(\lambda\left(X-d\mathbf{1}\right)+d\mathbf{1}\right)/N-d\right)}{\mu\left(\lambda\left(X-d\mathbf{1}\right)+d\mathbf{1}\right)-d}$$

$$= \frac{\lambda\left(\mu(X)-d\right)-\left(C_d\left(\lambda\left(X-d\mathbf{1}\right)+d\mathbf{1}\right)/N-d\right)}{\lambda\left(\mu(X)-d\right)}$$

and

$$I_d(X) = I_d\left(T_\lambda^d(X)\right)$$

$$\Leftrightarrow\; C_d\left(\lambda\left(X-d\mathbf{1}\right)+d\mathbf{1}\right)/N-d = \lambda\left(C_d(X)/N-d\right)$$

$$\Leftrightarrow\; C_d\left(T_\lambda^d(X)\right)/N = T_\lambda^d\left(C_d(X)/N\right) \;\Leftrightarrow\; \overline{C}_d\left(T_\lambda^d(X)\right) = T_\lambda^d\left(\overline{C}_d(X)\right)$$

where $\overline{C}_d(X) := C_d(X)/N$.

Now suppose that $X \sim_d^W Y$. Then $W_d(X) = W_d(Y)$ and therefore by definition $\overline{C}_d(X) = \overline{C}_d(Y)$ and thus $\overline{C}_d\left(T_\lambda^d(X)\right) = \overline{C}_d\left(T_\lambda^d(Y)\right)$. The latter equation implies that $T_\lambda^d(X) \sim_d^W T_\lambda^d(Y)$, i.e. homogeneity of $\succsim_d^W$ with respect to $\mathcal{J}_d$.

We obtain for $d = -\infty$

$$I_{-\infty}\left(T_\alpha^d(X)\right) = \mu\left(T_\alpha^d(X)\right) - C_{-\infty}\left(T_\alpha^d(X)\right)/N$$

$$= \mu(X) + \alpha - C_{-\infty}\left(T_\alpha^d(X)\right)/N.$$

Therefore again

$$I_{-\infty}(X) = I_{-\infty}\left(T_\alpha^d(X)\right) \Leftrightarrow \overline{C}_{-\infty}\left(T_\alpha^d(X)\right) = T_\alpha^d\left(\overline{C}_{-\infty}(X)\right).$$

The rest of the proof is the same as above. □

Proof of Proposition 2

a) Obvious

b) (i) We at first prove that

$$T_\lambda^d\left(\xi\left(X\right)\right)=\xi\left(T_\lambda^d\left(X\right)\right)$$

for a $\mathcal{J}_d$-homogeneous welfare ordering. By definition $X\sim_d^W \xi\left(X\right)\mathbf{1}$ and $T_\lambda^d\left(X\right)\sim_d^W \xi\left(T_\lambda^d\left(X\right)\right)\mathbf{1}$. $\mathcal{J}_d$-homogeneity implies that $T_\lambda^d\left(X\right)\sim_d^W T_\lambda^d\left(\xi\left(X\right)\right)\mathbf{1}$ which proves the claim.

(ii) We consider

$$W\left(X^i\right)=V_i^{-1}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(X_j^i\right)\right)=\xi_i\left(X^i\right).$$

$\mathcal{J}_d$-homogeneity (see (i)) implies that

$$V_i^{-1}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(T_\lambda^d\left(X_j^i\right)\right)\right)=T_\lambda^d\left(V_i^{-1}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(X_j^i\right)\right)\right)$$

and therefore

$$\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(T_\lambda^d\left(X_j^i\right)\right)=V_i\left(T_\lambda^d\left(V_i^{-1}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(X_j^i\right)\right)\right)\right).$$

Setting $t_j:=\frac{1}{n_i}V_i\left(X_j^i\right)$ we get

$$\frac{1}{n_i}\sum_{j=1}^{n_i}V_i\left(T_\lambda^d\left(V_i^{-1}\left(n_i t_j\right)\right)\right)=V_i\left(T_\lambda^d\left(V_i^{-1}\left(\sum_{j=1}^{n_i}t_j\right)\right)\right).$$

Theorem 1 and its Corollary in Aczel (1966), p. 142 imply that there are constants $a\left(\lambda\right)$ and $b\left(\lambda\right)$ such that

$$V_i\left(T_\lambda^d\left(V_i^{-1}\left(t\right)\right)\right)=a\left(\lambda\right)t+b\left(\lambda\right).$$

Now we replace *t* by $s:=V_i^{-1}\left(t\right)$ and obtain

$$V_i\left(\lambda\left(s-d\right)+d\right)=a\left(\lambda\right)V_i\left(s\right)+b\left(\lambda\right).$$

Define $r=s-d$ and $f\left(t\right):=V_i\left(t+d\right)$. Then

$$V_i\left(\lambda r+d\right)=a\left(\lambda\right)V_i\left(r+d\right)+b\left(\lambda\right) \text{ and } f\left(\lambda r\right)=a\left(\lambda\right)f\left(r\right)+b\left(\lambda\right).$$

The solution of this equation is given by Theorem 2.7.3 in Eichhorn (1978): There are $\rho \neq 0, \varepsilon \neq 0,$ and σ such that

$$f(t) = \rho t^{\varepsilon} + \sigma, \quad a(\lambda) = \lambda^{\varepsilon}, \quad b(\lambda) = \rho\left(1-\lambda^{\varepsilon}\right)$$

or $$f(t) = \rho \log t + \sigma, \quad a(\lambda) = 1, \quad b(\lambda) = \rho \log \lambda .$$

This implies the structural form of W^i, and analogously of W.

(iii) Now we define $m_i := \beta_i^{\frac{1}{1-\varepsilon}}$ and obtain $\beta_i = m_i\, m_i^{-\varepsilon}$. Using Proposition 2b we consider

$$\sum_{i=1}^{n} \beta_i n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(X_j^i - d \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} = \sum_{i=1}^{n} m_i m_i^{-\varepsilon} n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(X_j^i - d \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i}$$

$$= \sum_{i=1}^{n} m_i n_i \left(\frac{1}{n_i} m_i^{-\varepsilon(\varepsilon_i/\varepsilon)} \sum_{j=1}^{n_i} \left(X_j^i - d \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} = \sum_{i=1}^{n} m_i n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} .$$

Changing the normalization from $\Sigma \beta_h n_h$ to $\Sigma m_h n_h$ we obtain the result.

c) The proof of (10a) runs along the same lines. See also Proposition 2 in Ebert (1997). □

Now we define $a_i := \frac{1}{\gamma} \ln \beta_i$ and obtain $\beta_i = e^{\gamma a_i}$. Using Proposition 2c we consider

$$\sum_{i=1}^{n} \beta_i n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i X_j^i} \right)^{\gamma/\gamma_i} = \sum_{i=1}^{n} e^{\gamma a_i} n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i X_j^i} \right)^{\gamma/\gamma_i}$$

$$= \sum_{i=1}^{n} n_i \left(\frac{1}{n_i} e^{\gamma a_i(\gamma_i/\gamma)} \sum_{j=1}^{n_i} e^{-\gamma_i X_j^i} \right)^{\gamma/\gamma_i} = \sum_{i=1}^{n} n_i \left(\frac{1}{n_i} \sum_{j=1}^{n_i} e^{-\gamma_i\left(X_j^i - a_i\right)} \right)^{\gamma/\gamma_i}$$

which proves the result. □

Proof of Proposition 3

a) We consider

$$\min \sum_{i=1}^{n} n_i \lambda_i \text{ s.t. } W_d\left(\lambda_1 \mathbf{1}_{n_1}, \dots, \lambda_n \mathbf{1}_{n_n}\right) = W_d(X).$$

Introducing the Lagrange parameter κ and the Lagrange function $\mathcal{L}$ we obtain

$$\frac{\partial \mathcal{L}}{\partial \lambda_i} = n_i - \kappa \frac{1}{\varepsilon} (\dots)^{1/\varepsilon - 1} \frac{m_i n_i \varepsilon}{\Sigma m_h n_h} \left(\frac{\lambda_i - d}{m_i} \right)^{\varepsilon - 1} \frac{1}{m_i} = 0$$

and therefore

$$\frac{\lambda_i - d}{m_i} = \frac{\lambda_j - d}{m_j} \text{ for } i, j = 1,...,n \qquad \text{or} \qquad \frac{\lambda_i - d}{m_i} = \lambda_1 - d \ .$$

Then

$$W_d\left(\lambda_1 \mathbf{1}_{n_1},...,\lambda_n \mathbf{1}_{n_n}\right) = \left(\sum_{i=1}^{n} \frac{m_i n_i}{\Sigma m_h n_h}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\lambda_1 - d\right)^{\varepsilon_i}\right)^{\varepsilon/\varepsilon_i}\right)^{1/\varepsilon} + d$$

$$= \left(\lambda_1 - d\right) + d = \lambda_1$$

and by assumption

$$\lambda_1 = \left(\sum_{i=1}^{n} \frac{m_i n_i}{\Sigma m_h n_h}\left(\frac{1}{n_i}\sum_{j=1}^{n_i}\left(\frac{X_j^i - d}{m_i}\right)^{\varepsilon_i}\right)^{\varepsilon/\varepsilon_i}\right)^{1/\varepsilon} + d \ .$$

The FOC's imply that $\lambda_i = m_i\left(\lambda_1 - d\right) + d$ and therefore

$$C_d\left(X\right) = \sum_{i=1}^{n} n_i \lambda_i = \sum_{i=1}^{n} n_i\left(m_i\left(\lambda_1 - d\right) + d\right) = \sum_{i=1}^{n} m_i n_i\left(\lambda_1 - d\right) + Nd.$$

b) For $d = -\infty$ we similarly derive

$$\lambda_i - a_i = \lambda_j - a_j \text{ for } i, j = 1,...,n \ .$$

See also Ebert (1997). □

Proof of Proposition 5

Set $r_i = 1$ or $b_i = 0$ and use Proposition 6.

Proof of Proposition 6

a) Exponents

(i) UBT(d) is satisfied if and only if $\dfrac{\partial W_d}{\partial X_k^i} < \dfrac{\partial W_d}{\partial X_l^{i+1}}$ for $X_k^i > r_i \, X_l^{i+1}$.

The inequality is equivalent to

$$\frac{\beta_i}{\Sigma \beta_h n_h}\, \xi_i\left(X^i\right)^{\varepsilon-\varepsilon_i}\left(X_k^i\right)^{\varepsilon_i - 1} \ < \ \frac{\beta_{i+1}}{\Sigma \beta_h n_h}\, \xi_{i+1}\left(X^{i+1}\right)^{\varepsilon-\varepsilon_{i+1}}\left(X_l^{i+1}\right)^{\varepsilon_{i+1}-1} . \qquad (*)$$

Leaving X_k^i constant and letting $\xi_i\left(X^i\right)\to\infty$ implies that $\varepsilon-\varepsilon_i\le 0$. Similarly $\xi_{i+1}\left(X^{i+1}\right)\to\infty$ (for constant X_l^{i+1}) yields $\varepsilon-\varepsilon_{i+1}\ge 0$. Thus we obtain $\varepsilon_{i+1}\le\varepsilon\le\varepsilon_i$.

(ii) $\varepsilon_{i+1}\le 0$: In this case $\xi_{i+1}\left(X^{i+1}\right)\to 0$ if $X_j^{i+1}\to 0$ for all $j\ne l$. Therefore $\varepsilon-\varepsilon_{i+1}\le 0$. Then we get $\varepsilon=\varepsilon_{i+1}\le 0$.

(iii) $\varepsilon\le 0$: Then $\varepsilon_{i+1}\le\varepsilon\le 0$ and thus $\varepsilon=\varepsilon_{i+1}$ by (ii).

(iv) $\varepsilon_i\le 0$: By the same argument we obtain $\varepsilon-\varepsilon_i\ge 0$ and therefore $\varepsilon=\varepsilon_i$ and $\varepsilon_{i+1}=\varepsilon=\varepsilon_i$.

Weights

(i) Suppose that $\varepsilon_{i+1}>0$

Since $\varepsilon-\varepsilon_i\le 0$ the LHS of (*) is maximal (given X_k^i and X_l^{i+1}) if $X_j^i\to 0$ for all $j\ne k$. Then the LHS tends to $\dfrac{\beta_i}{\Sigma\beta_h n_h}\left[\left(\dfrac{1}{n_i}\right)^{1/\varepsilon_i}X_k^i\right]^{\varepsilon-\varepsilon_i}\left(X_k^i\right)^{\varepsilon_i-1}$.

Analogously, since $\varepsilon-\varepsilon_{i+1}\ge 0$ the RHS is minimal if $X_j^{i+1}\to 0$ for $j\ne l$ (given X_k^i). Then the following inequality has to be fulfilled:

$$\beta_i\left(\frac{1}{n_i}\right)^{\frac{\varepsilon-\varepsilon_i}{\varepsilon_i}}\left(X_k^i\right)^{\varepsilon-1}\ \le\ \beta_{i+1}\left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon-\varepsilon_{i+1}}{\varepsilon_{i+1}}}\left(X_l^{i+1}\right)^{\varepsilon-1}.$$

It is equivalent to

$$\beta_i/\beta_{i+1}\ \le\ \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i-\varepsilon}{\varepsilon_i}}\left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon-\varepsilon_{i+1}}{\varepsilon_{i+1}}}\left(X_l^{i+1}\right)^{\varepsilon-1}\Big/\left(X_k^i\right)^{\varepsilon-1}.$$

Since $\varepsilon-1<0$ the RHS is minimal for given X_l^{i+1} if $X_k^i\to r_i\,X_l^{i+1}$.

We obtain

$$\beta_i/\beta_{i+1}\ \le\ \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i-\varepsilon}{\varepsilon_i}}\left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon-\varepsilon_{i+1}}{\varepsilon_{i+1}}}r_i^{1-\varepsilon}.$$

(ii) $\varepsilon_{i+1}\le 0$ and $\varepsilon_i>0$: Then $\varepsilon=\varepsilon_{i+1}$ implies

$$\beta_i / \beta_{i+1} \ \le \ \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i - \varepsilon}{\varepsilon_i}} \ r_i^{1-\varepsilon} .$$

(iii) $\varepsilon_i \le 0$: Then $\beta_i / \beta_{i+1} \le r_i^{1-\varepsilon}$.

b) Exponents

UBT($-\infty$) is satisfied if and only if $\dfrac{\partial W}{\partial X_k^i} < \dfrac{\partial W}{\partial X_l^{i+1}}$ for $X_k^i > X_l^{i+1} + b_i$.

The inequality is equivalent to

$$\frac{\beta_i}{\Sigma \beta_h n_h} \, e^{-(\gamma-\gamma_i)\xi_i\left(X^i\right)} \ e^{-\gamma_i X_k^i} \ < \ \frac{\beta_{i+1}}{\Sigma \beta_h n_h} \, e^{-(\gamma-\gamma_{i+1})\xi_{i+1}\left(X^{i+1}\right)} \ e^{-\gamma_{i+1} X_l^{i+1}} . \qquad (**)$$

Leaving X_k^i constant and letting $\xi_i\left(X^i\right) \to \infty$ implies that $\gamma - \gamma_i \ge 0$. Similarly $\xi_{i+1}\left(X^{i+1}\right) \to \infty$ (for constant X_l^{i+1}) yields $\gamma - \gamma_{i+1} \le 0$. Therefore $\gamma_i \le \gamma \le \gamma_{i+1}$.

Then it is possible to consider $\xi_i\left(X^i\right) \to \infty$ and $\xi_{i+1}\left(X^{i+1}\right) \to -\infty$. We obtain $\gamma_{i+1} \le \gamma \le \gamma_i$ and therefore $\gamma_i = \gamma = \gamma_{i+1}$.

Weights

Then

$$\beta_i \, e^{-\gamma_i X_k^i} \ < \ \beta_{i+1} \, e^{-\gamma_{i+1} X_l^{i+1}} .$$

The RHS of (**) is maximal (for given X_l^{i+1}) if $X_k^i = X_l^{i+1} + b_i$. Thus

$$\beta_i \, e^{-\gamma_i\left(X_l^{i+1} + b_i\right)} \ \le \ \beta_{i+1} \, e^{-\gamma_{i+1} X_l^{i+1}} \quad \text{and} \quad \beta_i / \beta_{i+1} \le e^{\gamma_i b_i} \qquad \square$$

Proof of Proposition 7

This case can be handled by switching the indices and the inequality sign in the proof of Proposition 6. Then we obtain

a) Exponents

(i) General case: $\varepsilon_i \le \varepsilon \le \varepsilon_{i+1}$

(ii) $\varepsilon_i \le 0$: Then $\varepsilon = \varepsilon_i$

(iii) $\varepsilon \le 0$ implies $\varepsilon_i = \varepsilon$

(iv) $\varepsilon_{i+1} \le 0$ yields $\varepsilon_i = \varepsilon = \varepsilon_{i+1}$

Weights

Here we get

$$\beta_{i+1}/\beta_i \ \ge \ \left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon_{i+1}-\varepsilon}{\varepsilon_{i+1}}} \left(\frac{1}{n_i}\right)^{\frac{\varepsilon-\varepsilon_i}{\varepsilon_i}} s_i^{1-\varepsilon}$$

and thus

$$\beta_i/\beta_{i+1} \ \ge \ \left(\frac{1}{n_i}\right)^{\frac{\varepsilon_i-\varepsilon}{\varepsilon_i}} \left(\frac{1}{n_{i+1}}\right)^{\frac{\varepsilon-\varepsilon_{i+1}}{\varepsilon_{i+1}}} s_i^{\varepsilon-1} .$$

b) Exponents

$$\gamma_{i+1} = \gamma = \gamma_i$$

Weights

Then

$$\beta_{i+1}\, e^{-\gamma_{i+1} X_l^{i+1}} < \beta_i\, e^{-\gamma_i X_k^i} .$$

The LHS is maximal if $X_l^{i+1} = X_k^i + c_i$. Thus

$$\beta_{i+1}\, e^{-\gamma_{i+1}\left(X_k^i + c_i\right)} \le \beta_i\, e^{-\gamma_i X_k^i} \quad \text{and} \quad \beta_{i+1} \le \beta_i\, e^{\gamma c_i} . \qquad \square$$

Proof of Proposition 8

a) We obtain $\varepsilon_i = \varepsilon = \varepsilon_{i+1}$ and thus

$$\beta_i\, r_i^{\varepsilon-1} \ \le \ \beta_{i+1} \ \le \ \beta_i\, s_i^{1-\varepsilon} .$$

The constants r_i and s_i have to be chosen appropriately; i.e. it must not be possible that BOTH transfers are feasible. That would be the case if

$$X_k^i > r_i\, X_l^{i+1} > s_i\left(r_i\, X_k^i\right).$$

This condition cannot be met if $r_i s_i \ge 1$. In this case the interval $\left[r_i^{\varepsilon-1}\beta_i, s_i^{1-\varepsilon}\beta_i\right]$ is nonvoid.

We similarly obtain $\beta_{i+1} \in \left[e^{-\gamma b_i} \beta_i, e^{\gamma c_i} \beta_i \right]$ and $b_i + c_i \geq 0$.

Proof of Proposition 9

(i) Using the representation (11) we obtain

$$\frac{\partial \tilde{W}_d}{\partial X_k^i} = \frac{1}{\varepsilon} \left(\frac{\Sigma m_i n_i}{\Sigma m_h n_h} \left(\frac{1}{n_i} \Sigma \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon - 1} \cdot m_i n_i \frac{\varepsilon}{\varepsilon_i} \left(\frac{1}{n_i} \Sigma \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i - 1} \cdot \frac{1}{n_i} \varepsilon_i \left(\frac{X_k^i - d}{m_i} \right)^{\varepsilon_i - 1} \cdot \frac{1}{m_i}$$

$$= \left(\frac{\Sigma m_i n_i}{\Sigma m_h n_h} \left(\frac{1}{n_i} \Sigma \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i} \right)^{1/\varepsilon - 1} \cdot \left(\frac{1}{n_i} \Sigma \left(\frac{X_j^i - d}{m_i} \right)^{\varepsilon_i} \right)^{\varepsilon/\varepsilon_i - 1} \cdot \left(\frac{X_k^i - d}{m_i} \right)^{\varepsilon_i - 1}.$$

Now suppose that δ is redistributed from X_k^i to X_l^{i+1}.

Then

$$\frac{\partial \tilde{W}_d}{\partial X_k^i} < \frac{\partial \tilde{W}_d}{\partial X_l^{i+1}}$$

if and only if $\varepsilon = \varepsilon_i = \varepsilon_{i+1}$ and $\frac{X_k^i - d}{m_i} \leq \frac{X_l^{i+1} - d}{m_{i+1}}$ for $E_i\left(X_k^i\right) \geq E_{i+1}\left(X_l^{i+1}\right)$.

We get an analogue if $E_{i+1}\left(X_l^{i+1}\right) \geq E_i\left(X_k^i\right)$ and therefore

$$\frac{X_k^i - d}{m_i} = \frac{X_l^{i+1} - d}{m_{i+1}} \text{ for } E_i\left(X_k^i\right) = E_{i+1}\left(X_l^{i+1}\right).$$

This has to be true for $i = 1, ..., n-1$.

Therefore

$$E_i\left(X_k^i\right) = E_1\left(X_l^1\right) = X_l^1$$

and

$$\frac{X_k^i - d}{m_i} = \frac{E_i\left(X_k^i\right) - d}{m_1} = E_i\left(X_k^i\right) - d.$$

We obtain

$$E_i\left(X_k^i\right)=\frac{X_k^i}{m_i}+\left(d-\frac{d}{m_i}\right) \text{ and } E_i\left(X_k^i\right)=\frac{X_k^i}{\beta_i^{\frac{1}{1-\varepsilon}}}+\left(d-\frac{d}{\beta_i^{\frac{1}{1-\varepsilon}}}\right).$$

Property (iii) of the equivalent income function implies that $\beta_1 \le \beta_2 \le ... \le \beta_n$.

(ii) Using the representation (12) we obtain

$$\frac{\partial \tilde{W}_{-\infty}}{\partial X_k^i}=-\frac{1}{\gamma}\frac{1}{\Sigma\frac{n_i}{N}\left(\frac{1}{n_i}\Sigma e^{-\gamma_i\left(X_j^i-a_i\right)}\right)^{\gamma/\gamma_i}}\cdot\frac{n_i}{N}\frac{\gamma}{\gamma_i}\left(\frac{1}{n_i}\Sigma e^{-\gamma_i\left(X_j^i-a_i\right)}\right)^{\gamma/\gamma_i-1}\cdot\frac{1}{n_i}e^{-\gamma_i\left(X_k^i-a_i\right)}\cdot\left(-\gamma_i\right)$$

$$=\frac{1}{N}\frac{e^{-\gamma_i\left(X_k^i-a_i\right)}}{\Sigma\frac{n_i}{N}\left(\frac{1}{n_i}\Sigma e^{-\gamma_i\left(X_j^i-a_i\right)}\right)^{\gamma/\gamma_i}}\cdot\left(\frac{1}{n_i}\Sigma e^{-\gamma_i\left(X_j^i-a_i\right)}\right)^{\gamma/\gamma_i-1}.$$

Now suppose that δ is redistributed from X_k^i to X_l^{i+1}.

Then $\frac{\partial \tilde{W}_{-\infty}}{\partial X_k^i}<\frac{\partial \tilde{W}_{-\infty}}{\partial X_l^{i+1}}$ if and only if $\gamma=\gamma_i=\gamma_j$ and $e^{-\gamma_i\left(X_k^i-a_i\right)}<e^{-\gamma_{i+1}\left(X_l^{i+1}-a_{i+1}\right)}$

for $E_i\left(X_k^i\right)\ge E_{i+1}\left(X_l^{i+1}\right)$.

We get an analogous result if $E_{i+1}\left(X_l^{i+1}\right)\ge E_i\left(X_k^i\right)$ and therefore $X_k^i-a_i=X_l^{i+1}-a_{i+1}$ for $E_i\left(X_k^i\right)=E_{i+1}\left(X_l^{i+1}\right)$. This has to be true for $i=1,...,n-1$.

Therefore

$$E_i\left(X_k^i\right)=E_1\left(X_l^1\right)=X_l^1 \text{ and } X_k^i-a_i=E_i\left(X_k^i\right)-a_1=E_i\left(X_k^i\right).$$

Now observe that $a_i=\frac{1}{\gamma}\ln\beta_i$. Property (iii) of the equivalent income function again implies that $a_1\le a_2\le ...\le a_n$. □