METRON

RIVISTA INTERNAZIONALE DI STATISTICA — REVUE INTERNATIONALE DE STATISTIQUE
INTERNATIONAL JOURNAL OF STATISTICS — INTERNATIONALE STATISTISCHE ZEITSCHRIFT
REVISTA INTERNACIONAL DE ESTADISTICA

FOUNDED IN 1920 BY
Corrado Gini

Vol. LVI n. 3-4 **1998**

KARL MOSLER - PIETRO MULIERE

Welfare means and equalizing transfers

DIPARTIMENTO DI STATISTICA, PROBABILITA' E STATISTICHE APPLICATE
UNIVERSITA' DEGLI STUDI DI ROMA «LA SAPIENZA»

KARL MOSLER (*) - PIETRO MULIERE (**)

Welfare means and equalizing transfers

CONTENTS:

1. INTRODUCTION

1.1. Means, welfare and disparity

The measurement of social welfare and of economic disparity are closely connected. A principal aspect of social welfare, besides per capita endowment, is the economic disparity in the society. Thus many social welfare indices are measures of income inequality among individuals or households. On the other hand, the measurement of economic disparity may be based on measures of social welfare. It was suggested in the pioneering paper of Dalton (1920) that every measure of economic disparity has an underlying social welfare function. Dalton proposes to measure economic disparity by the

> "ratio of the total economic welfare attainable under an equal distribution to the total economic welfare attained under the given distribution."

(*) Seminar für Wirtschafts- und Sozialstatistik, Universität zu Köln, D–50923 Köln, Germany

(**) Dipartimento di Economia Politica e Metodi Quantitativi, Università di Pavia, Via San Felice, 5, I–27100 Pavia, Italy

Dalton's approach was developed in Kolm (1969), Atkinson (1970), and Sen (1973).

Kolm (1969) considers a representative income which, if distributed equally, results in the same overall level of social welfare as the existing income distribution. Given n households and an income profile $X = (x_1, x_2, \ldots, x_n)$, he defines the *equally distributed equivalent income* of X to be that level of income which, if enjoyed by every household, would make the total welfare equal to the total welfare generated by X.

It seems to us that the natural way to characterize the equally distributed equivalent income is Chisini's functional approach regarding the mean (Chisini 1929):

> " ... la ricerca di una media ha lo scopo di semplificare una qualche nostra questione sostituendo, in essa, a due, o più, quantità date una quantità sola che valga a sintetizzarle, senza alterare la visione d'insieme del fenomeno considerato"(1)

As the equally distributed equivalent income is a mean, we call it the *welfare mean* of X. In contrast to other social evaluation functions, which are just ordinal, the welfare mean measures on a cardinal scale. The notion of equally distributed equivalent income corresponds to the notion of a "certainty equivalent" in decision theory under risk. Atkinson (1970) discloses this formal similarity between the measurement of income disparity in welfare evaluation and the measurement of risk in decision making under uncertainty. More precisely, he shows that the concavity of the utility function plays the same crucial role in the analysis of both classes of problems and that ranking probability distributions according to an expected utility index is formally similar to ranking income distributions according to an additive social welfare function. Machina (1982) pointed out that the classical measures of welfare and disparity, which are based on utilitarianism, might be modified according to the generalized theories of choice under risk.

An income distribution is said to be *more equal* than another if it is obtained as a result of an equalizing transfer from the other. This idea was first captured by Pigou (1912) and Dalton (1920). A pair

(1)"The search for a mean has the purpose of simplifying a given question by substituting a single summary variable for too many values and leaving the overall picture of the problem unchanged."

of income distributions is a Pigou-Dalton transfer if the second distribution is obtained from the first by a series of income transfers from richer to poorer households. This is tantamount to saying that the first distribution majorizes the second or that the Lorenz curve of the first lies below the Lorenz curve of the second. A welfare mean satisfies the Pigou-Dalton transfer principle if it is increasing with respect to such transfers, in other words, if it is a Schur–concave function. In utility theory, Rothschild and Stiglitz (1970) employed a similar idea. They introduced the notion of a mean–preserving increase in risk, which means that the distributions have equal expected values and the Lorenz curve of the first lies above that of the second distribution.

The Pigou-Dalton principle is the classic principle of transfers and so far nearly uniquely applied. But there exist different notions of transfers that model other important aspects of equalization. There is also some recent empirical evidence that Pigou-Dalton (PD) transfers are neither unambiguously seen as inequality reducing nor as welfare increasing operations; see Amiel and Cowell (1992), Harrison and Seidl (1994) and others. For example, the income profile (500, 640, 660, 800, 900) is obtained from (500, 600, 700, 800, 900) by a PD transfer. Harrison and Seidl (1994) performed a large experiment and report that, when asked about these two profiles, one quarter of the respondents replied that the second profile is "more equally distributed" than the first. Moreover, one third of the respondents regarded the second as "socially preferable" over the first. These replies are inconsistent with the PD principle of inequality measurement and the welfare measures based thereon. This and other empirical studies suggest that weaker versions of the PD principle should be considered. Such weakenings have been investigated in Eichhorn and Gehrig (1981), Castagnoli and Muliere (1990), Mosler and Muliere (1996). Several weakenings that restrict the class of admissible transfers are presented below in a general framework of equalizing transfers. They are related to threshold incomes which separate classes of "richer" from "poorer" people. Further, certain transitions of income that alterate the total income are introduced in Section 3.5. They draw on notions from reliability theory.

A central problem is whether a given welfare mean is consistent with a certain class of equalizing transfers. In a utilitarian setting, every welfare mean is quasilinear. A quasilinear welfare mean increases with a Pigou-Dalton transfer if and only if the elementary social evaluation function is increasing and concave. Similar results will be presented for various classes of means and transfers. Two basic postulates in measuring social welfare are that the measure should be anonymous, i.e. symmetric in the households, and population invariant, i.e. should not vary if the total population changes but the distribution of income remains the same. Then the welfare mean depends only on the empirical distribution of incomes.

In Part 2, we consider welfare means that are defined on arbitrary probability distributions. First, this approach allows the comparison of income profiles that differ in n. Besides that, a general probability distribution, which is not empirical, allows for the interpretation that the households are weighted according to their "importance" for society, e.g., the sizes of households or their contributions to social welfare. For another interpretation in terms of decision under risk we refer to Harsanyi's view (Harsanyi 1953). According to this view and in our setting, a subject evaluates the welfare of a given income profile by considering himself in a random position and calculating the expected utility of income. Then the certainty equivalent of the random income is taken as the welfare mean.

1.2. Preliminaries

We introduce some notations and basic definitions of disparity measurement.

Let $(\Omega, \mathcal{A}, P)$ be a probability space, and X be a nonnegative random variable defined on it that has finite positive expectation μ_X. Ω may be seen as a population of individuals or households and $X(\omega)$ as the income of the household $\omega \in \Omega$. To facilitate the presentation we speak of households and their incomes, but do not exclude other interpretations. X is called an *income profile*; its probability distribution function F_X is called an *income distribution*. Let I be a closed interval of positive length, bounded or not. $\mathcal{F}_I$ denotes the set of probability distributions F that have support in I and finite expectation $\mu_F = \int_I x dF(x)$, and $\mathcal{X}_I$ denotes the set of all income profiles having probability distribution in $\mathcal{F}_I$.

If $\Omega = \{\omega_1, \omega_2, \ldots, \omega_n\}$ and P gives equal mass $1/n$ to each ω_i, we write $x_i = X(\omega_i)$ and $X = (x_1, x_2, \ldots, x_n)$ for short. This is called the *n-point empirical case*. Note that the x_i are not necessarily ordered nor have different values. $\mathcal{X}_I^n$ and $\mathcal{F}_I^n$ are the sets of n-point empirical income profiles and distributions, respectively.

The disparity of an income profile is measured either by an index or by an ordering among profiles. A *disparity index* is a functional that assigns a real number to every income profile, while a *disparity order* is a preorder (transitive and reflexive) among income distributions. Of course, every disparity index induces a complete order of income profiles. For a given disparity order, it is important to know which indices are *consistent* with it, i.e. which indices increase in that order.

There exist three basic disparity orders, which are closely connected: *second degree stochastic dominance, majorization (= dilation = convex order)* and *Lorenz order*.

X dominates Y in *second degree stochastic dominance*, $X \geq_2 Y$, if

$$\int_I \psi(x) dF_X(x) \geq \int_I \psi(y) dF_Y(y) \tag{1}$$

for all increasing concave $\psi : I \to \mathbb{R}$, as far as the integrals exist.

Throughout the paper the terms increasing and decreasing are meant in the weak sense.

Second degree stochastic dominance, $X \geq_2 Y$, is defined by the following two equivalent conditions.

$$\int_0^z F_X(x)dx \leq \int_0^z F_Y(y)dy, \qquad \text{for all } z \in \mathbb{R}, \tag{2}$$

and

$$\int_0^t F_X^{-1}(s)ds \leq \int_0^t F_Y^{-1}(s)ds, \qquad \text{for all } t \in [0, 1]. \tag{3}$$

where F_X^{-1} denotes the usual quantile function of X, $F_X^{-1}(t) = \inf\{x : F_X(x) \geq t\}, 0 < t \leq 1$.

Higher degree stochastic dominance will be treated in Sections 3.3 and 3.4 below. X *majorizes* Y, $X \geq_M Y$ if (1) holds for all ψ that are convex. This is equivalent to saying that Y dominates X in second degree stochastic dominance and the means are equal.

To define the Lorenz order, consider the *Lorenz function* $L_X : [0, 1] \to [0, 1]$,

$$L_X(t) = \frac{1}{\mu_X} \int_0^t F_X^{-1}(s)ds, \qquad 0 \leq t \leq 1. \tag{4}$$

The graph of the Lorenz function is the *Lorenz curve*. X *dominates* Y *in the Lorenz order*, $X \geq_L Y$, if $L_X(t) \leq L_Y(t)$ holds for all $t \in [0, 1]$.

The Lorenz order is connected with the previous two orders by equivalences as follows.

$$X \geq_L Y \qquad \Leftrightarrow \qquad \frac{X}{\mu_X} \geq_M \frac{Y}{\mu_Y} \qquad \Leftrightarrow \qquad \frac{Y}{\mu_Y} \geq_2 \frac{X}{\mu_X}. \tag{5}$$

In the empirical case, the Lorenz function is piecewise linear. The Lorenz curve of an income profile $X = (x_1, \ldots, x_n)$ amounts to the points $(\frac{k}{n}, \frac{1}{n\bar{x}} \sum_{i=1}^{k} x_{(i)}), k = 0, \ldots, n$, and the straight lines connecting them. Here the $x_{(i)}$ denote the ordered components of X, $x_{(1)} \leq \ldots \leq x_{(n)}$.

Of two empirical income profiles, $X = (x_1, \ldots, x_n)$ and $Y = (y_1, \ldots, y_n)$, X majorizes Y if and only if $\sum_{i=1}^{n} x_{(i)} = \sum_{i=1}^{n} y_{(i)}$ and, for any k, $\sum_{i=1}^{k} x_{(i)} \leq \sum_{i=1}^{k} y_{(i)}$. This is the usual *vector majorization.*

A function $\mathbb{R}^n \to \mathbb{R}$ is *Schur–convex* if it increases with vector majorization.

In other words, a disparity index defined on $\mathcal{X}_I^n$ is consistent with majorization if and only if it is Schur–convex. A *Schur–concave* function is the negative of a Schur–convex function.

Similarly, a disparity index on $\mathcal{X}_I^n$ is consistent with second degree stochastic dominance if and only if it is Schur–concave and increasing.

From (5) we see that, on empirical profiles, an index is consistent with the Lorenz order if and only if it has the form

$$X = (x_1, \ldots, x_n) \mapsto S\left(\frac{x_1}{\bar{x}}, \ldots, \frac{x_n}{\bar{x}}\right)$$

with some Schur–convex S.

PROPOSITION 1.2.1. *Assume that $H(x) = \sum_{i=1}^{n} h(x_i)$. Then H is Schur–concave if and only if h is concave.*

Many special disparity indices have been proposed in the literature. For previous contributions that have surveyed the properties of inequality measures see Sen (1973), Kakwani (1980), Nygård and Sandström (1981), Chakravarty (1990), Jenkins (1991), Lambert(1993), Sen and Foster (1997), Cowell (1997) and others. The classic book on majorization is Marshall and Olkin (1979). For a survey of majorization in economic disparity measures, see Mosler (1994).

Note that – like almost every disparity order – the Lorenz order, majorization and second degree stochastic dominance depend on the distributions of income only. The same holds for the indices that are consistent with them. These orders and indices account for the relative frequency or weights of households that receive a certain level of income, but not for the total number nor for the "names" of the households. In the empirical case it means that the following two postulates are satisfied.

Anonymity

The evaluation of an income profile is not affected if two or more households exchange their incomes; the profile $(x_1, \ldots, x_n)$ *has the same value as any permutation* $(x_{\pi(1)}, \ldots, x_{\pi(n)})$ *of it.*

Anonymity is also called *impartiality* or *symmetry*. It implies that any n-variate social evaluation function is symmetric in its arguments.

Population invariance *A k-times replication of a given income profile is equally valuated as the given profile, for any k in* $\mathbb{N}$. *i.e.,* $(x_1, \ldots, x_n)$ *in* $\mathcal{X}_I^n$ *and* $(x_1, \ldots, x_1, \ldots, x_n, \ldots, x_n)$ *in* $\mathcal{X}_I^{n \cdot k}$ *have the same value.*

In most of Part 2 we shall assume that the welfare mean of a given income profile depends on the income distribution alone.

2. Welfare means

The problem we address in this part is to evaluate a given income profile X in $\mathcal{X}_I$ with respect to social welfare, i.e., to construct a real-valued function M that assigns social welfare $M(X)$ to X.

In Sections 2.1 to 2.6 we assume that $M(X)$ depends only on the distribution function of X. Especially, with an empirical distribution, the mean is population invariant and does not depend on the order of the x_i.

Assume that for every $F \in \mathcal{F}_I$ there is a number $m = M(F)$ such that the distribution F is equally valuated as the one-point income distribution H_m at m. Then the functional $F \mapsto m = M(F)$ is called a *welfare mean* or *representative income* (Kolm 1969, Atkinson 1970). H_m is named the *egalitarian distribution* at m. A welfare mean may be derived from different primitives. First we assume that an ordinal social evaluation function $\mathcal{J}(F)$, $F \in \mathcal{F}_I$, is given. If there exists a unique number $m \in I$ solving

$$\mathcal{J}(F) = \mathcal{J}(H_m) \qquad (6)$$

then $m = M^{\mathcal{J}}$ is a welfare mean and $\mathcal{J}$ the social evaluation function from which it is derived. While the social evaluation function measures on an ordinal scale, the welfare mean derived from it is a cardinal measure.

Second we presuppose a social equivalence relation $\sim$ in $\mathcal{F}_I$. If for every $F \in \mathcal{F}_I$ there exists a unique number m such that $F \sim H_m$, then $m = M(F)$ is the welfare mean derived from $\sim$. In other words, the value m of the welfare mean is chosen so that one is indifferent in order to evaluate the welfare between the given income distribution and a distribution concentrating all its mass at m. Obviously, a distribution concentrated at x_0 has welfare mean equal to x_0.

2.1. Quasilinear welfare means

Most measures of welfare and disparity have been defined ad hoc in the literature, followed by an investigation of their properties. Instead we will begin with a formulation of several postulates and then characterize classes of welfare means and disparity indices that are consistent with them.

For a characterization of the quasilinear welfare means we follow de Finetti (1931). In order to extend the Nagumo-Kolmogorov result (Nagumo 1930, Kolmogorov 1930) to this context, de Finetti considers a given function M that maps distribution functions into the reals. He employs three axioms on M, reflexivity, monotonicity and associativity.

D.1 Reflexivity

$$M(H_{x_0}) = x_0 \text{ for all } x_0 \in I.$$

D.2 Strict monotonicity *Let F and G be in $\mathcal{F}_I$. If $F \geq_1 G$ and $F \neq G$ then*

$$M(F_1) > M(F_2).$$

Here $\geq_1$ means usual stochastic order, which is also called *first degree stochastic dominance*: $F \geq_1 G$ if and only if $F(x) \leq G(x)$ holds for all x. In terms of welfare measurement, the first axiom D.1 says that, if everybody has the same income, the welfare mean equals this income. The second axiom D.2 is also known as "individualism" since the social welfare increases strictly with the individual incomes.

According to the second axiom, the welfare mean preserves usual stochastic order. If (after an eventual permutation of the incomes) every household has not less and some household has more income with G than with F, then the welfare mean increases.

Consider now a convex combination F^* of two distributions F and G, $F^* = \lambda F + (1-\lambda)G$ with some $\lambda \in]0, 1[$.

The property of associativity requires that the mean of F^* remains unchanged if the first component distribution is replaced by another one that has the same mean:

D.3 Associativity *For every $F_1, F_2, G \in \mathcal{F}_I$ and $\lambda \in]0, 1[$,*

$$M(F_1) = M(F_2)$$

implies

$$M(\lambda F_1 + (1-\lambda)G) = M(\lambda F_2 + (1-\lambda)G).$$

With other words, if we are indifferent in the evaluation of welfare between two income profiles of a subpopulation, say steel workers (or some part of the country), associativity means that we remain also indifferent if this subpopulation is merged with another subpopulation, say other workers (or the rest of the country), whose income profile remains unchanged.

Axiom D.3 is an axiom of mixture invariance. It is also called the *substitution* axiom. In de Finetti (1931) the three axioms are shown to assure a special functional form of M as follows. See also Hardy, Littlewood and Polya (1934).

PROPOSITION 2.1.1 (de Finetti 1931). *Let I be a bounded interval in $[0, \infty[$ and $M : \mathcal{F}_I \to \mathbb{R}$. M satisfies D.1, D.2 and D.3 if and only if there exists a function ψ, continuous and strictly monotone, such that for every $F \in \mathcal{F}_I$*

$$M(F) = \psi^{-1}\left(\int_I \psi(x)dF(x)\right). \tag{7}$$

Then ψ is unique up to positive affine transformations.

The proposition holds also for unbounded $I \subset [0, \infty[$ if the distribution is continuous or if two axioms are added (Chew 1983, p. 1073): compact support continuity and extension.

A welfare mean (7) with some $\psi : I \to \mathbb{R}$ continuous and strictly monotone is called a *quasilinear welfare mean* and denoted M_ψ. In the empirical case we write $M_\psi(x_1, x_2, \ldots, x_n) = M_\psi(F)$.

Observe that from the anonymity postulate follows that M is a symmetric function of its arguments $x_1, \ldots, x_n$, hence a *symmetric quasilinear mean.*

The function $\psi(x)$ is interpreted as an *elementary social evaluation function*, shortly ESEF. The function $\mathcal{J}$,

$$\mathcal{J}(F) = \int_I \psi(x) dF(x) = E(\psi(X)), \tag{8}$$

is called an *additive social evaluation function*. It is also called *linear* because it is linear in the proportions of the population.

Formula (8) goes back to Bonferroni (1924).

When $\psi(x) = x$ is chosen in (7), we obtain the *arithmetic mean*, i.e., the mathematical expectation μ_F of F. We obtain a *power mean* when $\psi(x) = x^k$ and $k \neq 0$, the *harmonic mean* when $\psi(x) = 1/x$, the *geometric mean* when $\psi(x) = \ln x$ and $x > 0$, the *exponential mean* when $\psi(x) = \exp(x)$. (7) corresponds to the certainty equivalent in decision under risk; ψ is the von Neumann–Morgenstern utility function, and $\int_I \psi(x) dF(x)$ the expected utility index. In this framework, Ramsey (1926) and von Neumann and Morgenstern (1947) provided different axiomatizations. For a survey see Muliere and Parmigiani (1993b). De Finetti (1952) pointed out the similarities between this characterization and the von Neumann and Morgenstern approach; see also Daboni (1984), Muliere and Parmigiani (1993b), Fishburn and Wakker (1995).

Under the usual continuity and monotonicity conditions for a preference relation over lotteries, a certainty equivalent exists for each lottery, and preferences and certainty equivalence functions are uniquely related to each other. Associativity is crucial for quasilinearity. In decision under risk, Axiom D.3 is equivalent to independence of the preference over mixtures.

2.2. *Disparity aversion*

There are close relations between the measurement of welfare and the measurement of economic disparity, some of which we explore in this section.

We proceed with another axiom for welfare means. $M : \mathcal{F}_I \to \mathbb{R}$ is called *disparity averse* if the following holds.

D.4 Disparity aversion

$$M(F) \leq \mu_F \textit{ holds for all } F \in \mathcal{F}_I.$$

Recall that H_μ denotes the egalitarian income distribution at μ. Since $\mu = M(H_\mu)$ by D.1, disparity aversion says that, for any $\mu \in \mathbb{R}$, among all income distributions having expectation μ, the egalitarian income distribution is socially most desirable.

Any real-valued index defined on $\mathcal{F}_I$ is called a *disparity index* if it satisfies disparity aversion D.4.

D.4 is characterized by majorizatión, stochastic dominance, and concavity of ψ.

PROPOSITION 2.2.1. *Let $M_\psi : \mathcal{F}_I \rightarrow \mathbb{R}$ be a quasilinear mean with ψ strictly increasing. Then the following four conditions are equivalent,* (i) *M_ψ is disparity averse,* (ii) *ψ is concave,* (iii) *M_ψ increases in second degree stochastic dominance.* (iv) *M_ψ decreases in majorization.*

Proof: Assume (i), which is equivalent to $\int_I \psi(x)dF(x) \leq \psi(\mu_F)$ for any F. For a two-point empirical distribution at x_1 and x_2, this means $\frac{1}{2}\psi(x_1) + \frac{1}{2}\psi(x_2) \leq \psi(\frac{1}{2}(x_1 + x_2))$, hence ψ is concave, (ii). The equivalence of (iii) to (ii) follows from a wellknown characterization of second degree stochastic dominance. (iv) follows immediately from (iii). Finally, note that any F majorizes H_{μ_F}. Therefore, (iv) implies (i). ■

PROPOSITION 2.2.2. *Let $\mathcal{J} : \mathcal{F}_I \rightarrow \mathbb{R}$ be an ordinal social evaluation function, and $M^{\mathcal{J}}$ the welfare mean defined by* (6). *Assume that $\mathcal{J}$ is strictly increasing on egalitarian distributions, i.e. $\mathcal{J}(H_a) < \mathcal{J}(H_b)$ if $a < b$. Then*

$M^{\mathcal{J}}$ is disparity averse if $\mathcal{J}$ is decreasing in majorization.

Proof: Let $F \in \mathcal{F}_I$. If $\mathcal{J}$ decreases in majorization, then $\mathcal{J}(F) \leq \mathcal{J}(\mu_F)$. Here, for an egalitarian distribution at a, we write shortly $\mathcal{J}(a)$ in place of $\mathcal{J}(H_a)$. It follows by (6) that $\mathcal{J}(M^{\mathcal{J}}(F)) = \mathcal{J}(F) \leq \mathcal{J}(\mu_F)$. As $\mathcal{J}$ is strictly increasing on egalitarian distributions, we conclude $M^{\mathcal{J}}(F) \leq \mu_F$. ■

The proof has shown that, under the statements of the Proposition 2.2.2, the following biimplication holds.

Corollary 2.2.1 $M^{\mathcal{J}}$ is disparity averse if and only if $\mathcal{J}(F) \leq \mathcal{J}(\mu_F)$ holds for any $F \in \mathcal{F}_I$.

Comparative disparity aversion.

Cooper (1927), Jessen (1931) and de Finetti (1931) compared quasilinear means obtained from different functionals on a given probability distribution. Let $M_{\psi_1}(F)$ and $M_{\psi_2}(F)$ be quasilinear means. If ψ_1 is increasing (decreasing) then

$$M_{\psi_1}(F) \geq M_{\psi_2}(F)$$

holds for every F if and only if $\psi_1 \circ \psi_2^{-1}$ is convex (resp. concave). We then say that the welfare mean M_{ψ_2} is *more disparity averse* than the welfare mean M_{ψ_1}. (For example, let $\psi_k(x) = x^{\gamma_k}$, $k = 1, 2$, with $\gamma_1 > \gamma_2$; then M_{ψ_2} is more disparity averse than M_{ψ_1} if either $\gamma_2 > 0$ or $\gamma_1 < 0$.)

This corresponds to the notion of being *more risk averse* in decision theory under risk. As Pratt (1964) and Arrow (1965) define, a household is more risk averse than another if his or her utility function is a concave transform of the other's utility function.

Dalton's approach. As Dalton (1920) pointed out, given a welfare mean $M(F)$ with ESEF ψ, there is a corresponding *index of economic disparity*, $D(F)$. Dalton proposes to use

$$D_D(F) = 1 - \frac{\psi(M(F))}{\psi(\mu_F)} \tag{9}$$

as a disparity index. Equation (9) says: It is the ratio between the social welfare of the given income distribution and the social welfare of a hypothetical egalitarian distribution of this income which makes up economic disparity. Dalton's ratio index equals zero if the given distribution is egalitarian, and it is bounded below by 1 if and only if the welfare mean is disparity averse. For a discussion of Dalton's approach see also Ferreri (1980), Benedetti (1980a), Giorgi(1984) and Muliere(1987).

Atkinson's approach. In place of (9), Atkinson (1970) introduces

$$D_A(F) = 1 - \frac{M(F)}{\mu_F}. \tag{10}$$

Here the disparity value of the egalitarian distribution is equal to zero, and greater than zero for every other distribution of income, provided the welfare is disparity averse. Note that Atkinson's $D_A(F)$ measures the disparity on a cardinal scale, viz. the equivalent income ratio, while Dalton's $D_D(F)$ does it only on an ordinal one. Equation (10) defines a large family of indices, depending on which quasilinear mean M, or elementary social evaluation function ψ, is used.

Let $I = [0, \infty[$. For $c > 0$ and $F \in \mathcal{F}_{[0,\infty[}$, we denote $F_c(x) = F(cx)$. Atkinson (1970) postulates that his disparity index be scale invariant: $D_A(F_c) = D_A(F)$ for all $c > 0$ and $F \in \mathcal{F}_{[0,\infty[}$. This is equivalent to the following postulate on the underlying mean.

D.5 Scale equivariance

$$M(F_c) = \frac{1}{c} M(F)$$

for every F and every $c > 0$.

D.5 says that M is homogeneous of degree one. The only quasilinear means (7) that are homogeneous of degree one are obtained by choosing for all x either

$$\begin{aligned} \psi(x) &= \alpha x^\gamma + \beta, \quad \text{with some } \gamma \neq 0, \alpha > 0, \beta \in \mathbb{R}, \quad \text{or} \\ \psi(x) &= \alpha \ln x + \beta, \quad \text{with some } \alpha > 0, \beta \in \mathbb{R}. \end{aligned} \tag{11}$$

In the second case the mean $M(F)$ is the *geometric mean*. When $\gamma = 2$, the mean $M(F)$ is the *root mean square*. Moreover if $\gamma \in]0, 1]$ then ψ is concave, and we obtain that M is disparity averse.

PROPOSITION 2.2.3. *Let $I \subset [0, \infty[$. Assume* D.1, D.2, D.3, D.5 *and* (10). *Then either $D_A(F)$ vanishes for all F, or equals*

$$D_{A\epsilon}(F) = 1 - \frac{1}{\mu_F} \left[\int_0^\infty x^{1-\epsilon} dF(x) \right]^{\frac{1}{1-\epsilon}} \tag{12}$$

with some $\epsilon < 1$, or equals

$$D_{A1}(F) = 1 - \frac{1}{\mu_F} \exp\left(\int_0^\infty \ln(x) dF(x) \right). \tag{13}$$

Moreover, (13) *satisfies* D.4 *(disparity aversion).* (12) *satisfies* D.4 *if and only if* $0 < \epsilon < 1$.

For $0 < \epsilon < 1$, $D_{A\epsilon}$ is *Atkinson's index*; D_{A1} is *Champernowne's index* (Champernowne 1953). Both are disparity averse welfare means, while the parameter ϵ in $D_{A\epsilon}$ reflects the degree of risk aversion. In the empirical case, the Atkinson index is a *power mean,*

$$M_\epsilon(x_1, \ldots, x_n) = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^{1-\epsilon}\right)^{\frac{1}{1-\epsilon}},$$

whose properties are: (i) $M_\epsilon(x_1, \ldots, x_n)$ decreases on ϵ, (ii) $\lim_{\epsilon\to\infty} M_\epsilon(x_1, \ldots, x_n) = \min_i x_i$, (iii) $\lim_{\epsilon\to-\infty} M_\epsilon(x_1, \ldots, x_n) = \max_i x_i$, (iv) $\lim_{\epsilon\to1} M_\epsilon(x_1, \ldots, x_n) = \prod_{i=1}^{n} x_i^{p_i}$.

For the properties of a power mean and their history and role in statistical theory, see Norris (1976) and Weerahandi and Zidek (1979).

Kolm's approach. Now let $I = \mathbb{R}$. For $a \in \mathbb{R}$ we denote $F_{+a}(x) = F(x-a)$. Kolm (1969) considers the "leftist" postulate,

$$D(F_{+a}) = D(F) \quad \text{for all } a \in \mathbb{R},\ F \in \mathcal{F}_I, \tag{14}$$

which means that adding a constant amount to every income does not affect economic disparity.

Kolm proposes an index that has the form

$$D_K(F) = \mu_F - M(F). \tag{15}$$

The disparity index (15) satisfies (14) if and only if M is *translative*, i.e. satisfies translation equivariance.

D.6 Translation equivariance

$$M(F_{+a}) = M(F) + a \quad \text{for all } F \text{ and } a.$$

A quasilinear mean is translative if and only if either for all x

$$\psi(x) = \alpha x + \beta \quad \text{with some } \alpha > 0, \beta \in \mathbb{R} \quad \text{or}$$

$$\psi(x) = \alpha \exp\{-\gamma x\} + \beta \quad \text{with some } \alpha > 0, \beta \in \mathbb{R}, \gamma \neq 0.$$

PROPOSITION 2.2.4. *Let* $I = \mathbb{R}$. *Assume* D.1, D.2, D.3, D.6 *and* (15).

(i) *Then either* $D_K(F)$ *vanishes for all* F *or, with some* $\gamma \in \mathbb{R} \setminus \{0\}$, *is given by*

$$D_{K\gamma}(F) = \frac{1}{\gamma}\ln\left(\int_{\mathbb{R}} e^{-\gamma(x-\mu_F)}\, dF(x)\right) \quad \text{for } F \in \mathcal{F}_{\mathbb{R}}. \tag{16}$$

(ii) (16) *satisfies* D.4 *(disparity aversion) if and only if* $\gamma > 0$.

$D_{K\gamma}$ is known as *Kolm's index*, $\gamma > 0$.

Proof: Plugging ψ from (16) into (15) yields $D_{K\epsilon}(F) = 0$ in the first case and, in the second,

$$D_{K\epsilon}(F) = \mu_F - \frac{1}{-\gamma} \ln\left(\int_{\mathbb{R}} e^{-\gamma x} dF(x)\right) = \frac{1}{\gamma} \ln\left(\int_{\mathbb{R}} e^{-\gamma(x-\mu_F)} dF(x)\right).$$

From D.4 follows that $\gamma > 0$. ■

2.3. Quasilinear means of the relative distribution

By many people, the social welfare of a given income profile is not perceived as depending from the realized absolute incomes but rather from the realized incomes relative to some location parameter of the income distribution, such as the arithmetic mean, the median or the minimum of all incomes. Often the social evaluation of a household income is not considered as just a function of the income level this household receives, but rather of the relation of its income to, say, total income in the population. In particular, a standard notion of poverty measurement says that a household is defined as "poor" if it receives an income that is below somefixed percentage of the arithmetic mean (or the median) of incomes.

Thus, if household incomes are evaluated in their relation to total income, we may consider welfare means that are based rather on the *relative income profile*

$$\frac{X}{\mu_X} \quad \text{or} \quad (r_1, \ldots, r_n) = \left(\frac{x_1}{\bar{x}}, \ldots, \frac{x_n}{\bar{x}}\right),$$

respectively.

Given $X \sim F$, let $\widetilde{F}$ denote the distribution of X/μ_X, which we call the *relative distribution* of X.

We define

$$\widetilde{M_\psi}(F) = M_\psi(\widetilde{F}) = \mu_X \cdot \psi^{-1}\left(\int_I \psi\left(\frac{x}{\mu_X}\right) dF(x)\right) \qquad (17)$$

as a welfare mean. Obviously, $\widetilde{M_\psi}$ satisfies the axioms D.1, D.2, D.3, and also the scale equivariance D.5. Further, D.4 (disparity aversion) holds if and only if ψ is concave.

For example, $\psi(x) = \alpha - \beta \exp(-\gamma x)$, $x \in I \subset [0, \infty[$, is concave when $\beta, \gamma > 0$ and $\alpha \in \mathbb{R}$. Then

$$\widetilde{M}_\psi(F) = -\frac{\mu_F}{\gamma} \ln\left(\int_I e^{-\gamma \frac{x}{\mu_F}} dF(x)\right). \tag{18}$$

2.4. *Welfare means with weight functions*

Despite the class of quasilinear means is large, several important welfare means are not included. Consider

$$P(x_1, \ldots, x_n) = \left(\frac{\sum_{i=1}^n x_i^{s+r}}{\sum_{i=1}^n x_i^s}\right)^{\frac{1}{r}}, \tag{19}$$

where $x_i \in I = [0, \infty[$ and s and r are positive numbers. (19) is a *power weighted power mean* and belongs to a more general class of means proposed by Gini (1938). In particular, with $s = 0$ we obtain the usual *power mean*,

$$P_0(x_1, \ldots, x_n) = \left(\frac{1}{n}\sum_{i=1}^n x_i^r\right)^{\frac{1}{r}}, \tag{20}$$

and with $s = r = 1$ the *contraharmonic mean*,

$$P_1(x_1, \ldots, x_n) = \frac{\sum_{i=1}^n x_i^2}{\sum_{i=1}^n x_i} \tag{21}$$

P_1 is connected to the coefficient of variation VC and the Herfindahl index $H(x_1, \ldots, x_n) = \sum(x_i/(n\bar{x}))^2$ by $P_1 = \bar{x}(VC^2 + 1) = n\bar{x}H$.

If $s \neq 0$, (19) is no quasilinear mean, but a *quasilinear weighted mean* of the form

$$M_{\alpha,\psi}(x_1, \ldots, x_n) = \psi^{-1}\left(\frac{\sum_{i=1}^n \alpha(x_i)\psi(x_i)}{\sum_{i=1}^n \alpha(x_i)}\right). \tag{22}$$

A mean defined by (22) is symmetric and includes quasilinearity as a special case (if α is constant). Here, weights are assigned to the households depending on their income levels. In (19) the elementary social evaluation function is $\psi(x) = x^r$ and the weight function is $\alpha(x) = x^s$. It can be shown that, for $s \neq 0$, (19) is neither monotone D.2 nor associative D.3.

The general form of a weighted mean is

$$M_{\alpha,\psi}(F) = \psi^{-1}\left(\frac{\int_I \alpha(x)\psi(x)dF(x)}{\int_I \alpha(x)dF(x)}\right). \tag{23}$$

with two continuous functions ψ and α, ψ strictly monotone and α positive.

Chew (1983) characterizes this class of means by axioms and proves a generalisation of the Theorem of de Finetti. In place of the above axioms D.2 and D.3, Chew employs the following weaker ones.

C.1 Betweenness. For all F and G in $\mathcal{F}_I$ and $M(F) < M(G)$. Then for every λ in$]0, 1[$,

$$M(F) < M(\lambda F + (1-\lambda)G) < M(G).$$

C.2 Weak substitution invariance. Let F_1, F_2, G be in $\mathcal{F}_I$ and λ_1 and λ_2 in $]0, 1[$ such that

$$M(F_1) = M(F_2) \neq M(G)$$

and

$$M(\lambda_1 F_1 + (1-\lambda_1)G) = M(\lambda_2 F_2 + (1-\lambda_2)G).$$

Then, for every H in $\mathcal{F}_I$ holds

$$M(\lambda_1 F_1 + (1-\lambda_1)H) = M(\lambda_2 F_2 + (1-\lambda_2)H).$$

Let $I \subset [0, \infty[$. If a function $M : \mathcal{F}_I \longrightarrow \mathbb{R}$ satisfies D.1, C.1, C.2 and is continuous with respect to weak convergence of measures, then there exist two continuous functions ψ and α, ψ being strictly monotone and α non-vanishing (except possibly at one endpoint x_1 of I at which $\psi(x_1) \neq 0$), such that $M = M_{\alpha,\psi}$ according to Equation (23) holds for all $F \in \mathcal{F}_I$. This has been shown by Chew (1983). Moreover, Chew has proved an if and only if characterization by employing two axioms of compact support continuity and extension in place of the above assumed continuity.

The interpretation of the two axioms C.1 and C.2 is obvious. Weak substitution invariance C.2 is a consequence of associativity D.3, but not vice versa. In fact, Chew's two axioms C.1 and C.2 imply the following: if $M(F_1) = M(F_2)$ and $\lambda_1 \in]0, 1[$ then there exists some $\lambda_2 \in]0, 1[$ such that for all G

$$M(\lambda_1 F_1 + (1-\lambda_1)G) = M(\lambda_2 F_2 + (1-\lambda_2)G).$$

Associativity requires the further restriction that $\lambda_1 = \lambda_2$.

Obviously, like M_ψ, the mean $M_{\alpha,\psi}$ can be used to generate measures of income inequality in the Atkinson and in the Kolm approach,

$$D_A(F) = 1 - \frac{M_{\alpha,\psi}(F)}{\mu_F} \quad \text{and} \tag{24}$$

$$D_K(F) = \mu_F - M_{\alpha,\psi}(F), \quad \text{respectively.} \tag{25}$$

PROPOSITION 2.4.1 (Chew 1983). *Let $M_{\alpha,\psi}$ be a weighted quasilinear welfare mean with elementary social evaluation function ψ and weight function α. Suppose that α and $\alpha \cdot \psi$ and their derivatives are continuous and bounded on I. $M_{\alpha,\psi}$ is decreasing with majorization if and only if for every $F \in \mathcal{F}_I$ the Gâteaux derivative at F is monotone in the opposite direction as ψ.*

Then, in particular, $M_{\alpha,\psi}$ is disparity averse.

Remark 2.4.1. Axiomatic treatments of means with a weight function have been also given in the literature on functional equations. See Bajractarevic (1958), Páles (1986, 1987). In the axiomatic treatment of information measures, several important functions, like Shannon entropy and Rényi entropy can be expressed as special cases of (22); see Aczél and Daróczy (1975). Also some measures of disparity are derived from (22) and (23); see Bürg and Gehrig (1978).

Example 2.4.1 (Beta-weighted mean). Let $I = [0, 1]$, $\alpha(x) = x^s(1-x)^t$, $\psi(x) = x^r$ with some positive r, s and t, hence

$$M_{\alpha,\psi}(x_1, \ldots, x_n) = \left(\frac{\sum_{i=1}^n x_i^{s+r}(1-x_i)^t}{\sum_{i=1}^n x_i^s(1-x_i)^t} \right)^{\frac{1}{r}}. \tag{26}$$

Next, we present three examples of weighted welfare means (22) in relative incomes $r_i = x_i/\bar{x}$, each of which is no quasilinear mean.

Example 2.4.2. With $\psi(r) = r^\gamma$ and $\alpha(r) = r/n$, we obtain

$$M_{\alpha,\psi}\left(\frac{x_1}{\bar{x}}, \ldots, \frac{x_n}{\bar{x}}\right) = \left\{ \frac{1}{n} \sum_{i=1}^n \frac{x_i}{\bar{x}} \left(\frac{x_i}{\bar{x}}\right)^\gamma \right\}^{\frac{1}{\gamma}} =: E_\gamma(x_1, \ldots, x_n).$$

$E_\gamma(x_1, \ldots, x_n)$ is the *generalized exponential mean* of order γ, $\gamma \neq 0$.

Example 2.4.3. We use $\psi(r) = \ln(r)$ and α as in the preceding example and obtain

$$E_0(X) = \exp\left\{\frac{1}{n}\sum_{i=1}^{n}\frac{x_i}{\bar{x}}\ln\left(\frac{x_i}{\bar{x}}\right)\right\},$$

the *generalized logarithmic mean.* There holds $E_0(X) = \lim_{\gamma\to 0} E_\gamma(X)$. The *Theil measure* of disparity (Theil 1967) is the logarithm of $E_0(x)$,

$$T_0(x_1, \ldots, x_n) = \sum_{i=1}^{n}\frac{x_i}{n\bar{x}}\ln(\frac{x_i}{\bar{x}}).$$

Example 2.4.4. Also the Herfindahl index H and the coefficient of variation VC can be expressed by a weighted quasilinear mean of relative incomes: Choose $\psi(r) = r$ and $\alpha(r) = r/n^2$.Then $M_{\alpha,\psi} = nH = VC^2 + 1$.

Implicit means. Let I be bounded and $v : I \times I \to \mathbb{R}$ be a skew-symmetric function, $v(y, x) = -v(x, y)$, continuous and strictly increasing in the first variable. The *implicit mean* of a distribution $F \in \mathcal{F}_I$ with respect to v is defined as the unique solution y of the equation

$$\int_I v(x, y)dF(x) = 0. \qquad (27)$$

Then D.1, D.2 and C.1 are satisfied, but not C.2; see Fishburn (1986) for a characterization by axioms. The above means are special cases: $v(x, y) = \psi(x) - \psi(y)$ yields a quasilinear mean, while $v(x, y) = [\psi(x) - \psi(y)]\alpha(x)\alpha(y)$ yields a weighted quasilinear mean.

PROPOSITION 2.4.2. The implicit mean (27) is disparity averse if and only if it its negative is consistent with majorization or if and only if $v(\cdot, y)$ is concave for all y.

The Proposition follows from Chew and Mao (1995, Theorem 2).

2.5. *Rank dependent quasilinear means*

It is well known that, when two Lorenz curves intersect, the use of different Schur–convex disparity measures can lead to conflicting results. Atkinson (1970) pointed out that we can always find Schur–convex measures of income disparity that rank income distributions in the reverse order of the Gini index. From Newbery (1970) follows that there exists no quasilinear welfare mean M_ψ which ranks all income distributions as the Gini index. In this section we consider the class of rank dependent quasilinear mean values (Chew 1990), which contains the Gini index and other useful indices. In utility theory, the rank dependent approach goes back to Quiggin (1982). Let $g : [0, 1] \to [0, 1]$ be continuous and nondecreasing, $g(0) = 0$, $g(1) = 1$, and $\psi : I \to \mathbb{R}$ be continuous, strictly increasing and bounded. Then

$$R_\psi^g(F) = \psi^{-1}\left(\int_I \psi(x)dg[F(x)]\right) \tag{28}$$

is called a *rank-dependent quasilinear mean* of F, $F \in \mathcal{F}_I$. In the empirical case, (28) reduces to

$$\begin{aligned} R_\psi^g(x_1, \ldots, x_n) &= \psi^{-1}\left(\sum_{i=1}^{n} \psi(x_{(i)})\left[g\left(\frac{i}{n}\right) - g\left(\frac{i-1}{n}\right)\right]\right) \\ &= \psi^{-1}\left(\psi(x_{(n)}) + \sum_{i=0}^{n-1} [\psi(x_{(i)}) - \psi(x_{(i+1)})]g\left(\frac{i}{n}\right)\right), \end{aligned} \tag{29}$$

with $\psi(x_0) = 0$. Every rank dependent quasilinear mean is continuous and satisfies D.1 and D.2. Obviously, when g is the identity function, we obtain a quasilinear mean.

The special case where ψ is the identity function,

$$R_\psi^{id} = \int_I x dg[F(x)] = \int_0^1 F^{-1}(t)dg(t), \tag{30}$$

respectively

$$R_\psi^{id}(x_1, \ldots, x_n) = \sum_{i=1}^{n} x_{(i)}\left[g\left(\frac{i}{n}\right) - g\left(\frac{i-1}{n}\right)\right],$$

has been explored by Yaari (1987),(1988); see the next section. Some examples follow.

Example 2.5.1. Suppose that $\psi(x) = x$ and $g(p) = 1 - (1-p)^2$. We obtain the *Gini arithmetic mean*

$$R^{Gmean}(F) = \int_I x d[1 - (1 - F(x))^2],$$

$$R^{Gmean}(x_1, \ldots, x_n) = \sum_{i=1}^{n} \frac{2(n-i)+1}{n^2} x_i.$$

Then the Gini index is $R(F) = 1 - R^{Gmean}(F)/\mu_F$. With $g(p) = 1-(1-p)^s$ we obtain the s-Gini mean; see Donaldson and Weymark (1980). For a bibliography on the Gini index see Giorgi (1990)

Example 2.5.2. If R^g_ψ is a rank dependent quasilinear welfare mean, as above, Atkinson's approach yields a disparity index $1 - {\mu_F}^{-1} R^g_\psi(F)$.

PROPOSITION 2.5.1 (Chew, Karni and Safra 1987). *Let ψ be differentiable. R^g_ψ is disparity averse if $-R^g_\psi$ is consistent with majorization if and only if ψ and g are both concave.*

In statistics, the rank-dependent mean corresponds to the class of L–estimators. For a characterisation of L–functionals inspired by the axioms in de Finetti's representation theorem for quasilinear means we refer to Giovagnoli and Regoli (1993). The next section discusses Yaari's special case of rank dependent quasilinear means.

2.6. *Measures linear in Lorenz deviation*

Many known measures of income disparity, such as the coefficient of variation, the relative mean deviation, the logarithmic variance, and the indices of Atkinson and Champernowne take the form

$$D(F) = \int_I \phi(x)dF(x) \tag{31}$$

with some real function ϕ. $D(F)$ is the expected value of some transformation $\phi(X)$ of an income profile X having distribution F. A disparity measure (31) is linear in the "probabilities", i.e. in the proportions of the population receiving certain incomes, but nonlinear in the income levels. Note that the Gini index is not of this form since it depends on ranks (Newbery 1970). Mehran (1976) suggests a class of disparity indices that are rank dependent quasilinear means and include the Gini index; see also Alzaid (1990). These indices have the form

$$D_h(F) = \frac{1}{\mu_F}\int_0^1 F^{-1}(t)h(t)dt. \tag{32}$$

with some function $h : [0,1] \to \mathbb{R}$. D_h is linear in the relative income levels $F^{-1}(t)/\mu_F$, which are weighted by $h(t)$ according to their ranks. Obviously, D_h ranges between 0 and 1 if $0 \le h(t) \le 1$

for all t. We assume $\int_0^1 h(t)dt = 0$ so that the index vanishes at egalitarian distributions, $D_h(H_\mu) = 0$. Then

$$D_h(F) = \frac{1}{\mu_F}\int_0^1 \left[F^{-1}(t) - \mu_F\right] h(t)dt. \tag{33}$$

If h is a monotone function, by partial Stieltjes integration (33) may be rewritten

$$D_h(F) = \int_0^1 [t - L_F(t)]dh(t), \tag{34}$$

where L_F denotes the *Lorenz function*, $L_F(t) = \int_0^t F^{-1}(s)ds/\mu_F$, $0 \le t \le 1$. The index D_h has a simple geometric interpretation in terms of the Lorenz curve: $D_h(F)$ is the weighted area between the Lorenz curve and the line of perfect equality, weighted by $dh(t)$. From this is easily seen:

PROPOSITION 2.6.1. *D_h is (strictly) consistent with the Lorenz order if and only if h is (strictly) increasing.*

Thus, with an increasing h, the index D_h is a disparity measure. It is called a *linear disparity measure* because it is linear in the deviation of the Lorenz curve from the line of equality. As the index is not linear in the usual sense, viz. in probabilities, we prefer to call it *linear in Lorenz deviation.*

With $h(t) = 2t-1$, which is strictly increasing, we obtain the Gini index. Next, consider $h(t) = -1$ if $t \le F(\mu_F)$,$h(t) = 1$ if $t > F(\mu_F)$. When plugging this h into (33) or (34) we get the mean deviation from the mean, which equals $D^{meand}(F) = 2[F(\mu_F) - L_F(F(\mu_F))] = \int_I |x - \mu_F| dF(x)$.

With other specifications of h we obtain known disparity indices due to Piesch, Bonferroni, de Vergottini and others; see also Giaccardi (1950), Piesch (1975), Mehran (1976), Benedetti (1980b), Buscemi (1980), Nygård and Sandström (1981).

A *welfare mean linear in Lorenz deviation*, R_h, is the negative of a disparity index (33),

$$R_h(F) = -D(F) = -\frac{1}{\mu_F}\int_0^1 \left[F^{-1}(t) - \mu_F\right] h(t)dt \tag{35}$$

with $h : [0, 1] \to \mathbb{R}$ and $\int_0^1 h(t)dt = 0$ as above.

It is immediate from the definition that $\mu_F R_h$ is a special rank dependent quasilinear mean, $\mu_F R_h = R_\psi^g$ with $\psi(x) = x$ and $g'(t) = -h(t)$.

PROPOSITION 2.6.2. *R_h is disparity averse if and only if R_h decreasing with majorization or if and only if h is an increasing function.*

We refer to sections 3.3 and 3.4 for higher degree transfers and their relation to welfare means that are linear in Lorenz deviation. In particular, such means are averse to downside inequality and consistent with mean-variance preserving transfers if and only if h is increasing and concave.

It is easy to see that for two given distributions F and G, if L_F crosses L_G at most once on $]0, 1[$ and from above, then every disparity measure D_h with h nondecreasing and concave ranks these distributions in the same order as the Gini index does.

2.7. *Quasi-means*

A *quasi-mean* is a function $Q_{\alpha,\psi,\phi} : \mathcal{X}_I^n \to \mathbb{R}$,

$$Q_{\alpha,\psi,\phi}(x_1, \ldots, x_n) = \phi^{-1}\left(\sum_{i=1}^{n} \alpha_i \psi(x_i)\right), \tag{36}$$

where ϕ and ψ are continuous, strictly monotone (increasing or decreasing) functions and $\alpha_i > 0$. In what follows we assume that ϕ and ψ are strictly increasing. In general the two evaluation functions ϕ and ψ are different and the weights $\alpha_1, \ldots, \alpha_n$ not constant. Therefore a quasi-mean is a nonsymmetric welfare mean, defined on empirical income profiles only. If $\phi = \psi$ and $\alpha_i = 1/n$, $i = 1, \ldots, n$, a symmetric quasilinear mean is obtained from (36). For a discussion of quasi-means and their characterization with the help of functional equations, see Muliere and Parmigiani (1993a). Obviously, a quasi-mean is Schur-concave if and only if $\alpha_1 = \ldots = \alpha_n$ and ψ is concave. In this case, the quasi-mean is disparity averse, i.e., for every $X = (x_1, \ldots, x_n) \in \mathcal{X}_I$

$$Q_{\alpha,\psi,\phi}(x_1, \ldots, x_n) \le Q_{\alpha,\psi,\phi}(\bar{x}, \ldots, \bar{x}).$$

If $I = \mathbb{R}$, also the reverse can be shown.

PROPOSITION 2.7.1. Let $I = \mathbb{R}$. *A quasi-mean* $Q_{\alpha,\psi,\phi}$ *defined on* $\mathcal{X}^n_{\mathbb{R}}$ *is disparity averse if and only if it is Schur–concave if and only if* $\alpha_1 = \ldots = \alpha_n$ *and* ψ *is concave.*

We proceed with two examples.

Example 2.7.1. Let $I = [0, \infty[$, $\phi(x) = x^\gamma$, $\gamma > 0$, and $\psi(x) = \ln(x+1)$, $x > -1$. Then the quasi-mean is

$$Q_{\alpha,\psi,\phi}(x_1, \ldots x_n) = \left[\sum_{i=1}^{n} \alpha_i \ln(x_i + 1)\right]^{\frac{1}{\gamma}}$$

For $\gamma = 1$, see Tinbergen (1991).

Example 2.7.2. Choose $\phi(x) = x$ and $\psi(x) = \ln x$. We obtain

$$Q_{\alpha,\psi,\phi}(x_1, \ldots x_n) = \sum_{i=1}^{n} \alpha_i \ln x_i,$$

the celebrated utility index due to Daniel Bernoulli (1738). Note that this is a quasi-mean but no quasilinear mean.

3. Equalizing transfers

In comparing two distributions of well-being, it is of interest to investigate and interprete the transformations by which one distribution is obtained from the other. Let X and Y be random variables (= income profiles) in $\mathcal{X}_I$, with distribution functions F and G and expectations μ_F and μ_G. In the sequel we assume that $I \subset [0, \infty[$ and suppress the index I. The pair (X, Y) is called a *transfer* if $\mu_F = \mu_G$. It is called a *growing (shrinking) transfer* if $\mu_F \leq (\geq)\mu_G$. We say that a real function ϕ is the *transition function* of (X, Y) if ϕ is nondecreasing and $Y = \phi(X)$. In the empirical case, with $X = (x_1, x_2, \ldots x_n)$ and $Y = (y_1, y_2, \ldots y_n)$, (X, Y) is a transfer if $\sum_{i=1}^n x_i = \sum_{i=1}^n y_i$. It is an growing (decreasing) transfer if $\sum_{i=1}^n x_i \leq (\geq) \sum_{i=1}^n y_i$. The function ϕ, given by $\phi(x_i) = y_i$, is the transition function of (X, Y). A transfer (X, Y) is interpreted as a change in income distribution from one period to the next (or, e.g., before and after taxation). The total income may increase, decrease or stay unchanged. The income profile Y is regarded to be more equal than the income profile X if Y is obtained from X by some kind of equalizing transfer.

3.1. Pigou-Dalton and related transfers

The classic approach to equalizing transfers goes back to Pigou (1912) and Dalton (1920). A transfer (X, Y) is a *Pigou–Dalton transfer*, shortly *PD transfer*, if X majorizes Y and $F_X \neq F_Y$. In particular, whether (X, Y) is a PD transfer depends only on the distributions of X and Y. A PD transfer consists of an exchange of incomes from rich to poor and, as it operates on the distributions only, a possible permutation of incomes among households. In the n-point empirical case, it depends only on the ordered income profiles, $X_{()} = (x_{(1)}, \dots x_{(n)})$ and $Y_{()} = (y_{(1)}, \dots y_{(n)})$. Then (X, Y) is a PD transfer if and only if the total income remains unchanged and, for every j, the sum of the j smallest incomes is, after the transfer, larger than before,

$$\sum_{i=1}^{j} x_{(i)} \leq \sum_{i=1}^{j} y_{(i)} \tag{37}$$

or, equivalently, for every j, the sum of the $n - j$ largest incomes decreases,

$$\sum_{i=j+1}^{n} x_{(i)} \geq \sum_{i=j+1}^{n} y_{(i)}. \tag{38}$$

(37) and (38) may be interpreted in terms of poverty lines: After a PD transfer the number of people below some poverty line is not larger than before, wherever the line is drawn. We denote the class of all PD transfers by $\mathcal{T}_{PD}$, and the PD transfers among n-point distributions by $\mathcal{T}_{PD}^{n}$. Another characterization of PD transfers between empirical distributions is by averages,

$$y = Tx \text{ for some doubly stochastic } T, \tag{39}$$

i.e., $y_i = \sum_{i=1}^{n} t_{ij} x_j$ with $t_{ij} \geq 0$ and $\sum_{i=1}^{n} t_{ij} = \sum_{j=1}^{n} t_{ij} = 1$. A PD transfer redistributes incomes in a way such that each post-transfer income is a weighted average of the pre-transfer incomes.

A function $H : \mathcal{X}^n \to \mathbb{R}$ satisfies the *Pigou–Dalton principle of transfers* if $H(X) \geq H(Y)$ whenever $(X, Y) \in \mathcal{T}_{PD}$. In the empirical case this means that H is Schur–concave. H satisfies the *strict Pigou–Dalton principle of transfers* if $H(X) > H(Y)$ holds for all $(X, Y) \in \mathcal{T}_{PD}$. Thus the class of (strictly) Schur–concave functions is the largest class of welfare indices that is consistent with the (strict) PD principle.

More general, for every $\mathcal{T} \subset \mathcal{T}_{PD}$, we say that H satisfies the $\mathcal{T}$-*principle of transfers* if

$$H(X) \geq H(Y) \qquad \text{whenever} \qquad (X, Y) \in \mathcal{T}. \tag{40}$$

The *strict $\mathcal{T}$-principle of transfers* is similarly defined with $>$ in place of $\geq$ in (40).

An *elementary PD transfer* is a PD transfer among empirical income profiles where only two incomes are changed. Let $\mathcal{T}_{ePD}$ denote the set of elementary PD transfers. It is well known (Hardy, Littlewood and Polya (1934, p. 47)) that every PD transfer among empirical income profiles decomposes into a finite number of elementary PD transfers. Therefore Condition (40) with general PD transfers $\mathcal{T} = \mathcal{T}_{PD}$ is equivalent to the same with elementary ones only, $\mathcal{T} = \mathcal{T}_{ePD}$. For general X, Y in $\mathcal{X}$, (X, Y) is called a *growing PD transfer* if $Y \geq_2 X$, i.e., in the empirical case, (37) for every j. This is a growing transfer. (X, Y) is a *shrinking PD transfer* if $(-X, -Y)$ is a growing PD transfer. An index $H : \mathcal{X}^n \to \mathbb{R}$ is consistent with growing (shrinking) PD transfers if and only if it is an increasing (resp. decreasing) Schur-concave function.

3.2. Transfers about a threshold

In the introduction we have given some reasons to weaken the PD principle. Here we will present two weakenings of it, which are defined for X, Y in $\mathcal{X}^n$: the principle of transfers about θ and the starshaped principle of transfers about θ. For more details, the reader is referred to Mosler and Muliere (1996).

Given $\theta \in I$ define the *set of transfers about* θ,

$$\mathcal{T}_\theta^n = \bigcup_{n \geq 2} \Big\{(X, Y) \in \mathcal{T}_{PD}^n : y_{(i)} \leq \theta \text{ if } y_{(i)} - x_{(i)} > 0,\ y_{(i)} \geq \theta \text{ if } y_{(i)} - x_{(i)} < 0\Big\}.$$

$\mathcal{T}_\theta^n$ refers to a specific income level θ that separates the "relatively rich" from the "relatively poor". It includes all PD transfers that, besides a permutation of households, give to people below θ and take from people above θ, while no income crosses the line. For example, ((100, 470, 500), (100, 480, 490)) is a transfer about $\theta = 490$ but not about $\theta = 470$. Note that $\mathcal{T}_{PD}^n = \bigcup_{\theta \in I} \mathcal{T}_\theta^n$.

A differentiable function H satisfies the principle of transfers about θ if and only if

$$\max_{x_i<\theta} \frac{\partial}{\partial x_i} H(x_1, \ldots x_n) \leq \min_{x_i>\theta} \frac{\partial}{\partial x_i} H(x_1, \ldots x_n) \text{ for all } x_1, \ldots x_n. \quad (41)$$

If

$$H(x_1, \ldots x_n) = \sum_{i=1}^{n} g(x_i) \text{ with some } g \in \mathcal{G}, \quad (42)$$

Condition (41) reads

$$g'(s) \leq g'(t) \text{ whenever } s < \theta < t. \quad (43)$$

For (43) we say that H has *increasing disparity weight about θ*.
Define further

$$\mathcal{T}^n_{next\theta} = \Big\{(X, Y) \in \mathcal{T}^n_{PD} : \text{there is some } k \text{ with } x_k \leq \theta \leq x_{k+1}, \quad (44)$$
$$y_i \geq x_i \text{ if } i \leq k,\ y_i \leq x_i \text{ if } i \geq k,\ x_k \leq y_i \leq x_{k+1} \text{ if } y_i \neq x_i \Big\}$$
$$\mathcal{T}^n_{star\theta} = \mathcal{T}^n_{next\theta} \cup \mathcal{T}_\theta. \quad (45)$$

$\mathcal{T}^n_{next\theta}$ is the set of *transfers next to θ*, which includes all PD transfers where only the incomes next to θ change. Regarding Proposition 3.2.1 below, $\mathcal{T}^n_{star\theta}$ is named the set of *starshaped transfers at θ*. E.g., ((100, 470, 500), (100, 480, 490)) is a transfer next to $\theta = 470$. In fact, it is an element of $\mathcal{T}^n_{next\theta}$ and $\mathcal{T}^n_{star\theta}$ for any $\theta \in [470, 500]$.

A function $f : I \to \mathbb{R}$ is starshaped above at θ and supported if

$$(f(s) - f(\theta))/(s - \theta) \text{ is increasing at all } s \in I - \{\theta\}.$$

Shortly, we say that such an f is *starshaped above at θ*. If f is differentiable, equivalently,

$$(f(s) - f(\theta))/(s - \theta) \geq f'(s) \quad \text{when} \quad s < \theta,$$
$$\text{respectively} \quad \leq f'(s) \quad \text{when} \quad s > \theta.$$

PROPOSITION 3.2.1 (Mosler and Muliere). *Let H be additive* (42) *with g in $\mathcal{G}$. Then H satisfies the starshaped principle of transfers at θ for all n if and only if g is starshaped above at θ.*

Castagnoli and Muliere (1990) introduce another principle of transfers with respect to θ, which they call the *strengthened PD principle of transfers*. The principle says that an index should follow the PD principle of transfers and, in addition, the strict principle of transfers about θ. So the index is sensitive against a transfer from a rich household to a poor one but possibly insensitive (though not decreasing) when income is transferred among rich households or poor households only.

$$\max_{x_i<\theta} \frac{\partial}{\partial x_i} H(x) < \min_{x_i>\theta} \frac{\partial}{\partial x_i} H(x) \tag{46}$$

is sufficient for H to satisfy the strict principle of transfers about θ. Therefore every disparity index H in

$$\Psi_1 = \{H : H \text{ is Schur–convex and (46) holds}\}$$

satisfies the strengthened PD principle of transfers.

3.3. *Transfers with intersecting Lorenz curves*

The Lorenz curves reveals immediately whether or not two distributions can be ranked by the PD principle of transfers. In Section 1.2 we have presented the relationship between Lorenz order, PD transfers (= inverse vector majorization) and second degree stochastic dominance. Here we shall discuss several finer orderings. First let us consider the n-th degree stochastic dominance and the n-th degree inverse stochastic dominance. For $F \in \mathcal{F}_{[0,\infty[}$, define

$$F_1(x) = F(x) \quad \text{and} \quad F_n(x) = \int_0^x F_{n-1}(s)d(s), \quad x \in [0, \infty[,$$

as far as the integrals exist. It is well known that

$$F_n(x) = \frac{1}{(n-1)!} \int_0^x (x-y)^{n-1} dF(y), \quad x \geq 0 \tag{47}$$

Formula (47) is used to define *stochastic dominance of degree n*, $n \in \mathbb{N}$, shortly $n-$SD. F dominates G with degree n, in symbols $F \geq_n G$, if $F_n(x) \leq G_n(x)$ for all x holds.

The sequence of n-th degree SD is a sequence of progessively finer partial orders; see Rolski (1976) and Fishburn (1980). Formula

(47) makes transparent the connections between stochastic dominance orders and the measurement of poverty, given a poverty line at x. See Foster and Shorrocks (1988). For $n = 1$ we obtain the head-count ratio, $F_1(x)$, and for $n = 2$ the total income gap, $F_2(x)$, of the poor. It is also possible to define a sequence of progressively finer partial orders of distributions called *n-th degree inverse stochastic dominance*. We introduce the notation

$$F_1^{-1}(x) = F^{-1}(x) \quad \text{and} \quad F_n^{-1}(x) = \int_0^x F_{n-1}^{-1}(s)ds, \quad 0 \leq x \leq 1$$

(as far as the integrals exist) and define *n-th degree inverse stochastic dominance* as follows. F is said to dominate G inversely with degree n, in symbols $F \geq_n^{-1} G$, if $F_n^{-1}(x) \geq G_n^{-1}(x)$ for all x.

The ordering $\geq_n^{-1}$ is a partial order on $\mathcal{F}_I$; it is shortly denoted as $n-$ISD. The orderings $\geq_n^{-1}$ form a sequence of progressively finer partial orders,

$$F \geq_n^{-1} G, \; k \geq n \quad \Rightarrow \quad F \geq_k^{-1} G. \tag{48}$$

The meaning of (48) is the following: $\geq_k^{-1}$ orders all pairs of distributions that are ordered by $\geq_n^{-1}$ and some more. Therefore, as we pass from $\geq_k^{-1}$ to $\geq_{k+1}^{-1}$, each of the previously performed comparisons between pairs of distributions holds true, and some more are added.

It can be shown that n-SD and n-ISD are equivalent when $n = 1$ or $n = 2$. For $n = 1$ the result is trivial; for $n = 2$ it is proved by a slight generalization of an argument due to Atkinson (1970). Moreover, if $\mu_F = \mu_G$, then

$$F \geq_2 G \Longleftrightarrow F \geq_2^{-1} G \Longleftrightarrow F \leq_L G.$$

When $n \geq 3$ the equivalence does not hold any more. Let X be an income profile with distribution function F, $\mu_F = a$. The following proposition summarizes some properties of n-th degree stochastic dominance $\geq_n$.

Let $\mathcal{V}_n$ be the set of all strictly increasing functions $\psi : [0, \infty) \to \mathbb{R}$ whose derivatives exist through order n and alternate in sign with $\psi^{(k)} \geq 0$ for $k = 1, 3, 5, \ldots$, and $\psi^{(k)} \leq 0$ for $k = 0, 2, 4, 6, \ldots$.

PROPOSITION 3.3.1 (Rolski 1976). *For all F and G in $\mathcal{F}_I$ and $n \in \mathbb{N}$*

$$F \geq_n G \iff M_\psi(F) \geq M_\psi(G) \text{ for every } \psi \in \mathcal{V}_n, \tag{49}$$

where M_ψ is the quasilinear mean generated by ψ.

See also Fishburn (1980) and Chew (1983). Proposition 3.3.1 means that, if $F \geq_n G$, then F has greater social welfare than G in terms of every individual welfare function ψ that has n derivatives alternating in sign. In particular, for $n = 3$,

$$F \geq_3 G \implies \int_I \psi(x)dF(x) \geq \int_I \psi(x)dG(x)$$

if the social evaluation function ψ is nondecreasing and concave with a nondecreasing second derivative. Fishburn and Willig (1984) associate $\geq_n$ to transfers and show some implications about ψ. We proceed with the characterisation of disparity measures that are coherent with $\geq_n^{-1}$. Resorting to the stochastic dominance results of Rolski (1976) and Fishburn (1980), for any positive integer n, it is possible to find a class of welfare means coherent with $\geq_n^{-1}$. We define

$$M^\psi(F) = \psi^{-1}\left(\int_0^1 \psi(x)dF^{-1}(x)\right) \tag{50}$$

PROPOSITION 3.3.2. *$F \geq_n^{-1} G$ if and only if $M^\psi(F) \geq M^\psi(G)$, where $\psi : [0, 1] \to \mathbb{R}$ is a function with*

$$\psi(x) = -\int_x^1 (s - x)^{n-1} dT(s) \tag{51}$$

and T is a distribution function on $[0, 1]$, $T(x) > 0$ if $x > 0$.

Proof: Muliere and Scarsini (1989) consider indices of the form $I^\psi(F) = \int_0^1 \psi(x)dF^{-1}(x)$. They show that $F \geq_n^{-1} G$ is equivalent to $I^\psi(F) \leq I^\psi(G)$ for every $\psi : [0, 1] \to \mathbb{R}$, $\psi(x) = -\int_x^1 (s - x)^{n-1} dT(s)$, where T is a distribution function on $[0, 1]$, $T(x) > 0$ if $x > 0$. As ψ in (51) is strictly increasing and $M^\psi(F) = \psi^{-1} I^\psi(F)$, the Proposition follows. ■

Muliere and Scarsini (1989) have proved that $F \geq_n^{-1} G$ implies

$$E(X_1 \wedge X_2 \wedge \ldots \wedge X_\kappa) \geq E(Y_1 \wedge Y_2 \wedge \ldots \wedge Y_\kappa)$$

if $\kappa \geq n-1$, $X_1, \ldots, X_\kappa$ i.i.d. from F and $Y_1, \ldots, Y_\kappa$ i.i.d. from G. Note that

$$x_1 \wedge x_2 \wedge \ldots \wedge x_\kappa = F_\kappa^{-1}\left(\frac{1}{\kappa}\right)$$

where F_κ^{-1} is the quantile function of the empirical distribution on $x_1, \ldots, x_\kappa$. This shows that, as n increases, lower incomes become more important for $\geq_n^{-1}$.

Now let us turn to the social evaluation function. We have seen that Atkinson (1970) provides the rationale for basing such a function on his index. His social evaluation function is

$$M(F) = \left(1 - D_{A\epsilon}(F)\right)\mu_F = \left(\int_0^\infty x^{1-\epsilon} dF(x)\right)^{\frac{1}{1-\epsilon}},$$

and obeys the Pigou–Dalton principle of transfers.

Consider next the Gini index. Here the social evaluation function is a rank dependent geometric mean,

$$M(F) = R^{Gmean}(F) = (1 - R(F))\mu_F = E(X_1 \wedge X_2).$$

Sen (1973) calls this social evaluation function the *pairwise maximin welfare criterion*, since the expected minimum income of any two households is considered. In Muliere and Scarsini (1989) the Sen approach is generalized to the case of the extended Gini coefficient. Let the welfare level of n income units be equal to the income of the poorest, for all $n \geq 1$. The average, over all n-tuples, welfare level gives the index

$$M(F) = E(X_1 \wedge X_2 \wedge \ldots \wedge X_n)$$

which has been named the *n-wise maximin welfare criterion* by Lambert (1993). The role played by n (the degree of inverse dominance) in distributional judgement is rendered particularly clear by this criterion. As $n \to \infty$, welfare becomes simply the income of the poorest unit. Finally we mention that the latter index resolves a paradox due to Newbery (1979). The paradox says that no von Neumann–Morgenstern expected utility ranks all distributions in the same order as the Gini index. However, this index does for any n.

Now we present a couple of examples.

Example 3.3.1. The Gini index is given by

$$\begin{aligned} R(F) &= 1 - 2\int_0^1 L(p)dp \\ &= 1 - \frac{2}{\mu_F}\int_0^1\int_0^p F^{-1}(t)dt\,dp \\ &= 1 - \frac{2}{\mu_F}F_3^{-1}(1) \end{aligned}$$

Therefore, by the definition of $\geq_3^{-1}$ we see that

$$F \geq_3^{-1} G,\ \mu_F = \mu_G \quad \Rightarrow \quad R(F) \leq R(G).$$

This means that the Gini index is coherent with 3-ISD when the distributions have the same expectation.

Example 3.3.2. Donaldson and Weymark (1980, 1983) propose two classes of single parameter indices (called S–Gini indices). For each $\gamma \geq 1$, they define an absolute index

$$Z_\gamma(F) = -\int_0^\infty x d(1 - F(x))^\gamma$$

and a relative index

$$I_\gamma(F) = 1 + \frac{\int_0^\infty x d(1 - F(x))^\gamma}{\mu_F}.$$

Note that the relative index is scale invariant, while the absolute index is translation equivariant. When $\gamma = 2$, we obtain the usual Gini index. When $\gamma = 1$, we get $Z_\gamma(F) = \mu_F$ and $I_\gamma(F) = 0$. Consider, for simplicity, $\gamma \in \mathbb{N}$. Then

$$Z_\gamma(F) = E(X_1 \wedge \ldots \wedge X_\gamma).$$

Therefore we obtain

$$F \geq_n^{-1} G \text{ implies } Z_\gamma(F) \geq Z_\gamma(G)$$

for $\gamma \geq n - 1$. If, furthermore, $\mu_F = \mu_G$ then

$$I_\gamma(F) \leq I_\gamma(G) \quad \text{for } \gamma \geq n - 1.$$

3.4. Third degree dominance: a transfer characterization

A PD transfer is also called a *progressive transfer* while the inverse of a PD transfer is called a *regressive transfer*. In the language of risk measurement, the regressive transfer is a "mean preserving spread", and the progressive transfer is a "mean preserving shrink". The Lorenz criterion assumes only that a transfer from richer to poorer reduces disparity, that is, it embodies the Pigou-Dalton principle of transfers. If the two Lorenz curves intersect then neither distribution is obtained from the other by a pure series of either progressive or regressive transfers. Either distribution can, however, be obtained from the other by a combination of progressive and regressive transfers. (5) points out the importance of the Lorenz order in risk and disparity measurements. In order to understand the role of transfers when the Lorenz curves intersect we need notions of transfers that combine regressive and progressive transfers in certain ways.

The third-degree stochastic dominance criterion imposes the normative requirement, often considered desirable, of transfer sensitivity. According to the weak notion of transfer sensitivity introduced by Shorrocks and Foster (1987) transfers made at lower income levels should be considered more important. A transfer sensitive inequality measure is then defined to be a measure that decreases under any mean-variance preserving transformation (favourable composite transfer).

A *mean–variance preserving transformation* (MVPT) is a combination of a regressive transfer and a progressive transfer with the following properties (Menezes, Geiss and Tressler, 1980):

(i) The progressive transfer occurs at lower income levels than the regressive transfer does.

(ii) The overall effect is to leave the variance unchanged.

The same transformation is referred to as a "favourable composite transfer" (FACT) by Shorrocks and Foster (1987). Many authors have suggested that for a fixed total income gap a transfer from a poorer to a richer person should be considered more disequalizing the lower it occures in the distribution; see Sen (1973), Kolm (1976), Shorrocks and Foster (1987), Davies and Hoy (1994). Following Davies and Hoy (1995) this viewpoint is called *aversion to downside inequality* (ADI).

Recently theorists have paid considerable attention to the consequences of assuming aversion to downside disparity (Shorrocks and

Foster (1987), Foster and Shorrocks (1988), Dardanoni and Lambert (1988), Muliere and Scarsini (1989)). This is the definition of Davies and Hoy (1995): (i) Let X and Y be two income profiles with he same arithmetic mean. X displays less downside disparity than Y if Y is obtained from X through a series of MVPTs. (ii) A disparity index $D(\cdot)$ satisfies aversion to downside disparity (ADI) if

$$D(F) \leq D(G)$$

whenever F displays less downside disparity than G. One reason the ADI restriction on disparity measurement has turned out to be fruitful is that, when added to the PD principle, it corresponds to 3-SD.

PROPOSITION 3.4.1 (Shorrocks and Foster 1987). *Let X and Y be two income distributions with equal arithmetic means. The following statements are equivalent.*

(i) *X is obtained from Y by a finite sequence of progressive tranfers and/or MVPT.*
(ii) $X \geq_3 Y$.
(iii) *$I(F) \leq I(G)$ for all indices I obeying ADI.*

Mehran (1976) introduces the *principle of diminishing transfers*. It says that the effect on inequality of a small tranfer between two households decreases as their incomes increase, holding constant their pre-transfer income gap. Roughly speaking the principle states that inequality among rich is less important than inequality among poor. Shorocks and Foster (1987) show that this property is equivalent to the concept of aversion to downside inequality. This is another equivalent condition to those in Proposition 3.4.1. A measure that is linear in Lorenz deviation satisfies the diminishing principle of transfers, and thus is averse to downside inequality, if and only if its score function h, in addition to being strictly increasing, has a strictly decreasing derivative (Mehran 1976).

Next, for characterizing 3–ISD we consider a particular class of transfers that combine a progressive and regressive transfer. Let,
(i) the progressive transfer take place lower down in the distribution.
(ii) the Gini index remain constant with transfer.
This is named a *favourable composite positional transfer* (FCPT) by Zoli (1998). Like an MVPT, an FCPT combines a progressive transfer

and a regressive transfer, the progressive transfer taking place lower down in the distribution.

PROPOSITION 3.4.2 (Zoli 1998). *Let X and Y be two income distributions with equal arithmetic means. The following statements are equivalent.*

(i) *X is obtained from Y by a finite sequence of progressive tranfers and/or FCPT.*
(ii) $X \geq_3^{-1} Y$

3.5. *Transfer principles related to reliability orders*

Several authors have observed that there exist formal connections between commonly used indices of economic inequality and some notions in reliability theory. See Chandra and Singhpurwalla (1981), Klefsjö (1984), Bergman and Klefsjö (1984), Bhattacharjee (1993), and others. In this section we want to introduce several new principles of transfers that are related to existing and widely used reliability orders.

Let X and Y be nonnegative random variables having distribution functions F and G. Define $\psi(y) = F^{-1} \circ G(y), y > 0$. Y is said to be smaller than X with respect to

(i) *increasing failure rate*, $G \leq_{IFR} F$, if ψ is convex,
(ii) *increasing failure rate average*, $G \leq_{IFRA} F$, if ψ is starshaped,
(iii) *new better than used*, $G \leq_{NBU} F$, if ψ is superadditive, $\psi(x_1 + x_2) \leq \psi(x_1) + \psi(x_2)$.

The random variable X may be rearranged, i.e. be replaced by another $\tilde{X}$ with the same distribution F, such that $X = \psi(Y)$ holds. Therefore each of the three notions says that, in a specific sense, X is more dispersed than Y. The three orderings are successively weaker, and all are

$$G \leq_{IFR} F \Rightarrow G \leq_{IFRA} F \Rightarrow G \leq_{NBU} F \Rightarrow G \leq_{L} F \qquad (52)$$

For this, see, e.g., Barlow and Proschan (1975). Here we take the function ψ as a transition function which defines the inverse $(Y, \psi(Y))$ of an equalizing transfer $(\psi(Y), Y)$. Obviously, ψ is nondecreasing. Moreover, if F and G are absolutely continuous, ψ maps the α quantile of X to that of Y, $0 < \alpha < 1$. Thus, as a transition function of incomes, ψ preserves the income order. We call $(\psi(Y), Y)$ an IFR

(IFRA, NBU) transfer if ψ satisfies one of the above three properties. There arise three transfer principles which, by (52), all are weaker than the PD principle. I.e., if a disparity index satisfies the PD principle, then it satisfies the NBU principle, then the IFRA principle, and then the IFR principle. Note that these transfers may change the arithmetic mean. However, often a redistribution of income in an economy will not leave the total income unchanged, as the redistribution may make the economy more efficient or less. The economic interpretation of each of the three new types of transfers is based on the following proposition.

PROPOSITION 3.5.1. *Let* $\psi = F^{-1} \circ G$. *Then*
(i) *if* ψ *defines an IFR transfer we obtain*

$$\psi\left(\frac{x_1 + x_2}{2}\right) \leq \frac{\psi(x_1) + \psi(x_2)}{2} \quad \textit{for all } x_1, x_2, \tag{53}$$

(ii) *if* ψ *defines an IFRA transfer we obtain*

$$\frac{x_2}{x_1} \leq \frac{\psi(x_2)}{\psi(x_1)} \quad \textit{for all } x_1 < x_2, \tag{54}$$

(iii) *if* ψ *defines an NBU transfer we obtain*

$$\psi(x_1 + x_2) \leq \psi(x_1) + \psi(x_2) \quad \textit{for all } x_1, x_2. \tag{55}$$

The proof is obvious.

An IFR transfer is characterized by the fact that two individuals who share their incomes are better off if they share them before the transfer rather than after. (Shortly: "Rather marry before an IFR transfer!") An IFRA transfer is given if and only if the ratio between any two (relatively) rich and poor is decreased by the transfer or, equivalently, if

$$\frac{x_2}{\psi(x_2)} \leq \frac{x_1}{\psi(x_1)} \quad \text{for all } x_1 < x_2, \tag{56}$$

i.e., the richer individual has a worse ratio between his posterior and prior incomes x_2 and $\psi(x_2)$, respectively, than the poorer one between x_1 and $\psi(x_1)$. Finally, which makes an NBU transfer is the fact that, before the transfer, the sum of any two incomes is less than the income that will be transferred to the sum of the two post-transfer incomes. It means that merging the incomes of two individuals to one gives a larger result after the transfer. (Shortly: "Don't be an heir before an NBU transfer!")

4. Outlook

So far we have measured the disparity *within* a distribution, by studying the discrepancy between a given distribution and the distribution assigning all mass to one point. Another important task is measuring the disparity *between* two given distributions of income profiles in the same population. This has applications to horizontal equity.

Cifarelli and Regazzini (1987) define and investigate the concentration function between arbitrary probability measures which are defined on the same space.

Recently the multivariate measurement of disparity and welfare has been investigated. In the case of many attributes of wellbeing, we may consider an equally distributed endowment that is equivalent to a given distribution of endowments. But, in general this "welfare mean" is no single point but a manifold in the space of endowments.

The concept of Pigou-Dalton transfers is easily extended to many attributes; see Marshall and Olkin (1969, ch. 15). However, several multivariate generalizations of the majorization order arise. We refer the reader to the discussion in Mosler (1994), to Koshevoy and Mosler (1996) for a multivariate extension of the Lorenz curve and to Koshevoy and Mosler (1997) for multivariate Gini indices. See also Tsui (1995) for applications of multivariate majorizations.

REFERENCES

Aczel, J. and Daróczy, Z. (1975) *On Measures of Information* and *Their Characterizations*, Academic Press: New York.

Alzaid, A.A. (1990) Lorenz ranking of income distributions, *Statistical Papers*, 31, 209–224.

Amiel, Y. and Cowell, F. (1992) Measurement of income inequality: Experimental test by questionnaire, *Journal of Public Economics*, 47, 3–26.

Arrow, K.J. (1965) *Aspects of the Theory of Risk Bearing*, Academic Bookstae: Helsinki.

Atkinson, A.B. (1970) On the measurement of Inequality, *Journal of Economic Theory*, 2, 244–263.

Bajraktarevic, M. (1958) Sur une Equation fonctionelle aux valeurs moyennes, *Glasnik. Mat.-Fiz. Astr.*, 13, 243–248.

Barlow, R.E. and Proschan, F. (1975) *Statistical Theory of Reliability* and *Life Testing. Probability Models*, Holt, Rinehart and Winston: New York.

BENEDETTI, C. (1980A) Di alcuni indici di disuguaglianza del benessere, *Statistica*, 40, 7–12.

BENEDETTI, C. (1980B) Contro–esempi in statistica e probabilità, in *Studi in onore di Paolo Fortunati*, vol.1, 23–29. Clueb: Bologna.

BERGMAN, B. and KLEFSJÖ, B. (1984) The total time on test concept and its use in reliability theory, *Operations Research*, 32, 596–606.

BERNOULLI, D. (1738) Specimen theoriae novae de mensura sortis, *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, 5, 175–192. [English translation in *Econometrica* 22(1954), 23–35.]

BHATTACHARJEE, M.C. (1993) How rich are the rich? Modeling affluence and inequality via reliability theory, *Sankhya*, 55, 1–26.

BONFERRONI, C.E. (1924) La media esponenziale in matematica finanziaria, *Annuano del Regio Istituto Superiore di Scienze Economiche e Commerciali di Bari*, AA 23/24, 1–14.

BÜRG, R. and GEHRIG, W. (1978) Indices of income inequality and societal income. An axiomatic approach, in *Theory* and *Applications of Economic Indices* , (Eichhorn, W. et al. eds.), 309–356. Physica: Wien.

BUSCEMI, S. (1980) A proposito di una espressione generale degli indici di concentrazione, *Metron*, 38, 177–186.

CASTAGNOLI, E. and MULIERE, P. (1990) A note on inequality measures and the Pigon-Dalton principle of transfers, in *Studies in Contemporary Economics*, (Dagum, C. and Zenga, M. eds.). Income and Wealth Distribution, Inequality and Poverty. Springer Verlag: Berlin.

CHAKRAVARTY, S. R. (1990) *Ethical Social Index Numbers*, Springer Verlag: Berlin.

CHANDRA, M. and SINGPURWALLA, N.D. (1981) Relationships between some notions which are common to reliability theory and economics, *Mathematics of Operations Research*, 6, 113–121.

CHAMPERNOWNE, D. (1953) A model of income distribution, *Economic Journal*, 63, 318–351.

CHEW, S.H. (1983) A generalization of the quasilinear mean with applications to the measurement of the income inequality and decision theory resolving the Allais Paradox, *Econometrica*, 51, 1065–1092.

CHEW, S.H. (1990) *The rank-dependent quasilinear mean*, Department of Economics Working Paper, UCI.

CHEW, S.H., KARNI, E. and SAFRA, Z. (1987) Risk aversion in the theory of expected utility with rank dependent probabilities, *Journal of Economic Theory*, 42, 370–381.

CHEW, S.H. and MAO, M.H. (1995) A Schur concave characterization of risk aversion for non-expected utility preferences, *Journal of Economic Theory*, 67, 402–435.

CHISINI, O. (1929) Sul concetto di media, *Periodico di Mathematiche*, 9, 106–116.

CIFARELLI, D.M. and REGAZZINI, E. (1987) On a general definition of concentration functions, *Sankhya*, B 49, 307–319.

COOPER, R. (1927) Notes on certain inequalities, I and II, *J. London. Math. Soc.*, 2, 17–21, 159–163.

COWELL, F.A.(1997) Measurement of Inequality, in A.B. Atkinson and F. Bourguignon (eds.) *Handbook of Income Distribution*, to appear.

DABONI, L. (1984) On the axiomatic treatment of the utility theory, *Metroeconomica*, 36, 203–209.

DALTON, H.(1920) The measurement of the inequality, *Economic Journal*, 348–361.

DARDANONI, V. and LAMBERT, P.J. (1988) Welfare rankings of income distributions: a role for the variance and some insights for tax reform, *Social Choice* and *Welfare*, 5, 1–17.

DAVIES, J. and HOY, M. (1994) The normative significance of using third-degree stochastic dominance in comparing income distributions, *Journal of Economic Theory*, 64, 520-530.

DAVIES,J. and HOY, M. (1995) Making Inequality comparison when Lorenz curves intersect, *American Economic Review*, 85, 980-986.

DE FINETTI, B. (1931) Sul concetto di media, *Giornale dell'Istituto Italiano degli Attuari*, 2, 369–396.

DE FINETTI, B. (1952) Sulla preferibilitá, *Giornale degli Economisti e Annali di Economia 11*, 685–709.

DONALDSON, D. and WEYMARK, J.A. (1980) A single-parameter generalization of the Gini indices of inequality, *Journal of Economic Theory*, 22, 67–86.

DONALDSON, D. and WEYMARK, J.A. (1983) Ethically flexible Gini indices for income distributions in the continuum, *Journal of Economic Theory*, 29, 353–358.

EICHHORN, W. and GEHRIG, W. (1981) Generalized convexity and the measurement of inequality, in *Generalized Concavity in Optimization* and *Economics*, (Schaible, S. and Ziemba, W.T. eds.), 637–642. Academic Press: New York.

FERRERI, C. (1980) Considerazioni sulla misurazione della concentrabilità in statistica, in *Studi in onore di Paolo Fortunati*, vol.1 319–364. Clueb: Bologna.

FISHBURN, P. (1980) Continua of stochastic dominance relations for unbounded probability distributions, *Journal of Mathematical Economics*, 7, 271–285.

FISHBURN, P.C. (1986) Implicit mean value and certainty equivalence, *Econometrica*, 54, 1197–1205.

FISHBURN, P. and WAKKER, P. (1995) The invention of the independence condition for preferences, *Management Science*, 41, 1130–1144.

FISHBURN, P.C. and WILLIG, R.D. (1984) Transfer principles in income distribution, *Journal of Public Economics*, 323–328.

FOSTER, J. and SHORROCKS, A.F. (1988) Poverty orderings, *Econometrica*, 56, 173–177.

GIACCARDI, F. (1950) Un criterio per la costruzione di indici di concentrazione, *Rivista Italiana di Economia, Demografia e Statistica*, 4, 527–538.

GINI, C. (1938) Di una formula comprensiva delle media, *Metron*, 13, 3–22.

GIORGI, G.M. (1984) A Methodological survey of recent studies for the measurement of inequality of economic welfare carried out by some Italian Statisticians, *Economic Notes*, 1, 146–158.

GIORGI, G.M. (1990) Bibligraphic portrait of the Gini concentration ratio, *Metron*, 48, 83–98.

GIOVAGNOLI, A. and REGOLI, G. (1993) Some results on the representation of measures of location and spread as L-functionals, *Statistics* and *Probability Letters*, 16, 269–278.

HARDY, G.H., LITTLEWOOD, J.E. and PÓLYA, G. (1934) *Inequalities*, Cambridge University Press.

HARRISON, E. and SEIDL, C. (1994) Perceptional inequality and preferential judgement: An empirical examination of distributional axioms, *Public Choice*, 79, 61–81.

HARSANYI, J.C. (1953) Cardinal utility in welfare economics and theory of risk taking, *Journal of Political Economy*, 61, 434–435.

JESSEN, B. (1931) Bemaerkninger om konvekse funktioner og uligheder imellem middleværdier, I and II. *Matematisk Tidsskrift B*, 17–28, 84–95.

JENKINS, S.P. (1991) The measurement of economic inequality, in L. Osberg (ed.) *Readings on Economic Inequality*, M.E. Sharpe, Armonk, N.Y.

KAKWANI, N.C. (1980) *Income, Inequality* and *Poverty: Methods of Estimation* and *Policy Applications*, Oxford University Press: Oxford.

KLEFSJÖ, B. (1984) Reliability interpretations of some concepts from economics, *Naval Research Logistic Quarterly*, 31, 301–308.

KOLM, S.C. (1969) The optimal production of social justice, in *Public Economics*, (Marjolis, I. and Guitton, H. eds.). London: MacMillan.

KOLM, S.C. (1976) Unequal Inequalities, I and II, *Journal of Economic Theory*, 12, 416–442, 13, 82–111.

KOLMOGOROV, A.N. (1930) Sur la notion de la moyenne, *Atti dell'Accademia Nazionale dei Lincei*, 9, 388–391.

KOSHEVOY, G.A. and MOSLER, K. (1996) The Lorenz zonoid of a multivariate distribution, *Journal of the American Statistical Association*, 91, 873–882.

KOSHEVOY, G.A. and MOSLER, K. (1997) Multivariate Gini indices, *Journal of Multivariate Analysis*, 60, 252–276.

LAMBERT, P.J. (1993) *The Distribution* and *Redistribution of Income: A Mathematical Analysis*, (Second ed.) Manchester UK: Manchester University Press.

MACHINA, M.J. (1982) Expected utility analysis without the independence axiom, *Econometrica*, 50, 277–323.

MARSHALL, A.W. and OLKIN, I. (1979) *Inequalities: Theory of Majorization* and *its Applications*, Academic Press: New York.

MEHRAN, F. (1976) Linear measures of income inequality, *Econometrica*, 44, 805–809.

MENEZES, C., GEISS, C. and TRESSLER, J. (1980) Increasing downside risk, *American Economic Review*, 70, 921–932.

MOSLER, K. (1994) Majorization in economic disparity measures, *Linear Algebra* and *Its Applications*, 199, 91–114.

MOSLER, K. and MULIERE, P. (1996) Inequality indices and the starshaped principle of transfers, *Statistical Papers*, 37, 343–364.

MULIERE, P. (1987) Misure della concentrazione: alcune osservazioni sull'impostazione assiomatica, in M. Zenga (ed.), *La distribuzione personale del reddito*, Vita e Pensiero, Milano, 190–213.

MULIERE, P. and PARMIGIANI, G. (1993A) On quasi-means, *Utilitas Mathematica*, 43, 79–87.

MULIERE, P. and PARMIGIANI, G. (1993B) Utility and means in the 1930s, *Statistical Science*, 8, 421–432.

MULIERE, P. and SCARSINI, M. (1989) A note on stochastic dominance and inequality measures, *Journal of Economic Theory*, 49, 314–329.

NAGUMO, M. (1930) Ober eine Klasse der Mittelwerte, *Japanese Journal of Mathematics*, 7, 72–79.

NEWBERY, D. (1970) A theorem on the measurement of inequality, *Journal of Economic Theory*, 2, 264–266.

NORRIS, N. (1976) General means and statistical theory, *American Statistician*, 30, 8–12.

NYGÅRD , F. and S andSTRÖM, A. (1981) *Measuring Income Inequality*, Almquist and Wiksell International: Stockholm.

PÁLES, ZS. (1986) On the generalization of quasiarithmetic means with weight functions, *Aequationes Mathematica*, 31, 324–325.

PÁLES, ZS. (1987) On the characterization of quasiarithmetic means with weight functions, *Aequationes Mathematicae*, 32, 171–194.

PIESCH, W. (1975) *Statistische Konzentrationsmaße*, J.C.B. Mohr (Paul Siebeck): Tübingen.

PIGOU, A.C. (1912) *Wealth* and *Welfare*, MacMillan: New York.

PRATT, J.W. (1964) Risk aversion in the small and in the large, *Econometrica*, 32, 122–136.

QUIGGIN, J. (1982) Anticipated utility theory, *Journal of Economic Behaviour* and *Organization*, 3, 323–343.

RAMSEY, F. (1926) Truth and probability, in *The Foundations of Mathematics*, 156–211. Routledge and Kegan Paul: London.

ROLSKI, T. (1976) Order relations in the set of probability distribution functions and their applications in queueing theory, *Dissertationes Mathematicae*, 132.

ROTHSCHILD, M. and STIGLITZ, J.E. (1970) Increasing risk: I. A definition, *Journal of Economic Theory*, 2, 225–243.

SEN, A. (1973) *On Economic Inequality*, Oxford: Clarendon Press.

SEN, A. and J.E. FOSTER (1997) *On Economic Inequality*, (Second ed.) Oxford: Clarendon Press.

SHORROCKS, A.F. and FOSTER, J. (1987) Transfer sensitive inequality measures, *Review of Economic Studies*, 54, 485–497.

THEIL, H. (1967) *Economics* and *Information Theory*, R and McNally: Chicago.

TINBERGEN, J. (1991) On the measurement of welfare, *Journal of Econometrics*, 50, 7–13.

TSUI, K.Y. (1995) Multidimensional generalizations of the relative and absolute inequality indices:The Atkinson-Kolm-Sen approach, *Journal of Economic Theory*, 2, 251–265.

VON NEUMANN, J. and MORGENSTERN, O. (1947) *Theory of Games* and *Economic Behavior*, 2nd Edition. J. Wiley: New York.

WEERAH andI, S. and ZIDEK, J.V. (1979) A characterization of the general mean, *Canadian J. of Statistics*, 7, 83–90.

Yaari, M.E. (1987) The dual theory of choice under risk, *Econometrica*, 55, 95–115.

Yaari, M.E. (1988) A controversial proposal concerning inequality measurement, *Journal of Economic Theory*, 44 (4), 381–397.

Zoli, C. (1998) Intersecting generalized Lorenz curves and the Gini index, *Social Choice and Welfare*, to appear.

Welfare means and equalizing transfers

Summary

Economic disparity and social welfare are measured by real-valued indices and orderings. Often a welfare index can be seen as a proper mean value, based on an equivalent equally distributed income, while a welfare ordering is constructed from inequality reducing transfers. This paper surveys welfare and disparity indices from the view of quasilinear means and their generalizations and gives an account of recent notions of equalizing transfers.

Le medie come indicatori del benessere sociale ed i trasferimenti redistributivi

Riassunto

La disuguaglianza in senso economico ed il benessere sociale vengono misurati mediante indicatori e ordinamenti. Sovente un indicatore del benessere può essere visto come una vera e propria media basata su un reddito equivalente equamente distribuito, mentre un ordinamento del benessere può essere costruito a partire da trasferimenti che riducono la disuguaglianza. In questo lavoro viene presentata una rassegna degli indicatori del benessere sociale e della disuguaglianza dal punto di vista delle medie quasi-lineari e della loro generalizzazione e vengono presi in esame le recenti nozioni di trasferimenti redistributivi

[Manuscript received February 1998; final version received May 1998.]