Inequality, welfare and order statistics

Encarnación M. Parrado Gallardo

Dpto. de Economía Aplicada (Estadística y Econometría, 68). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Campus El Ejido s/n. 29013. Málaga, Spain. e-mail: mparrado@uma.es

Elena Bárcena Martín

Dpto. de Economía Aplicada (Estadística y Econometría, 68). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Campus El Ejido s/n. 29013. Málaga, Spain. e-mail: barcenae @uma.es

Luis José Imedio Olmedo

Dpto. de Economía Aplicada (Estadística y Econometría, 68). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Campus El Ejido s/n. 29013. Málaga, Spain. e-mail: imedio@uma.es

Corresponding author:

Elena Bárcena Martín

Dpto. de Economía Aplicada (Estadística y Econometría, 68). Facultad de Ciencias Económicas y Empresariales. Universidad de Málaga. Campus El Ejido s/n. 29013. Málaga, Spain. Telf.+34 952131191/ +34951952129. e-mail: barcenae@uma.es

Abstract

In this paper we use the distributions of order statistics to define functions with the appropriate properties and represent social preferences regarding income distributions. Following the approach of Yaari (1987, 1988), this allows constructing a set of social welfare functions from which the corresponding inequality indices are derived. The obtained measures incorporate diverse normative criteria, with different degrees of preference for equality. The generalized Gini coefficients and the family of indices proposed in Aaberge (2000) are obtained as particular cases. This approach shows that each of these families of indices characterizes the income distribution, but for a change of scale.

Key words: Lorenz curves, social preference function, order statistic, inequality aversion, transfers.

JEL code: C02, C10, D31, I30.

Inequality, welfare and order statistics

1. Introduction

When studying the relationship between inequality measures and social welfare functions (SWF) in the context of income distributions following the approach of Yaari (1987, 1988), the distribution functions of social preferences have an essential role. These functions incorporate the normative aspects or value judgments that are always present in the evaluation of both magnitudes, specifically welfare and inequality. The properties of the functions determine the degree of preference for equality[1] (or inequality aversion) of the measure, thus affecting the behavior of the measure when certain changes take place in the income distribution.

This paper shows that from the distributions of order statistics it is possible to define functions that meet the properties required to represent social preferences about the distribution of income between units in a population. This allows us to build a set of SWFs from which the corresponding inequality indices are derived. The obtained measures incorporate diverse normative criteria with different degrees of preference for equality, and hence a different response to progressive income transfers. The generalized Gini coefficients (Kakwani, 1980; Donaldson and Weymarck, 1980, 1983; Yitzhaki, 1983) and the family of indices proposed in Aaberge (2000), in addition to measures that, in the normative aspect, occupy an intermediate position between the two families are obtained as particular cases.

Although the use of order statistics in the analysis of welfare and of inequality is rare, there are some results in the literature in this regard. For instance, it is known that the SWFs of the generalized Gini coefficients have a simple interpretation from the first-order statistics (Lambert 2001, Ch. 5, pp. 125-126), while Kleiber and Kotz (2002) use the mean values of these statistics to show that this family of indices characterizes any income distribution with finite mean, except for a factor of proportionality. Our proposal extends and generalizes this analysis using certain statistical averages of the order statistics, from the minimum to the maximum.

The approach taken in this paper provides a constructive way to define new measures of welfare and inequality as well as some additional advantages. First, it permits obtaining an alternative characterization of distributions that are not determined by their potential moments due to the fact that they only allow a small number of moments. This is the case of empirical income distributions. Since these distributions usually have a heavy tail with a pronounced positive skew, the asymptotic convergence to the Pareto law (Mandelbrot 1960) is a condition usually required of the theoretical parametric models used to adjust the observed distributions. The models that satisfy

1 A SWF shows inequality aversion or preference for equality if it fulfils the Pigou-Dalton Principle of Transfers. That is, if a given income transfer takes place from a richer individual to a poorer one without changing the relative order between both (progressive transfer), the social welfare (inequality) increases (decreases).

this condition, such as the Pareto or Singh-Maddala (1976) models, cannot be characterized from their potential moments[2]. This justifies the interest in alternative procedures to characterize such distributions.

Moreover, in addition to providing diverse distributive criteria in assessing welfare and inequality, the procedure used in this paper allows a clear interpretation of each measure in terms of the statistics computed from a randomly selected sample drawn from the population, as well as identifying unbiased estimators of both the SWFs and their associated inequality indices.

The paper is structured as follows. In the second section, we present the analytical framework with special reference to the Lorenz curve and, in particular, to the inequality measures that weight the area between the curve and the line of equal distribution. We also examine the relationship between welfare and inequality, as well as issues related to order statistics. The main results are obtained in section three. In this section we define the distribution of preferences, the corresponding SWFs and their associated inequality indices, considering some particular cases. We also show the relationship between the preference functions and the functions that weight the Lorenz differences in the inequality indices, and their policy implications. This relationship is related with the behavior of the indices in reference to the fulfillment of principles of transfers that are more demanding than the Pigou-Dalton principle. The final section briefly summarizes the results and includes some comments.

2. Analytical framework. Previous considerations

Let us assume that the income distribution is represented by the random variable X, whose range is the positive real numbers $R_0^+ = [0, \infty)$, where $F(\cdot)$ is its distribution function, and $\mu = E(X) = \int_0^\infty x dF(x) < \infty$ is its mean income.

2.1. Inequality

A common and intuitive way to assess inequality in the income distribution is to weight the deviations between the income perceived by each individual and the mean income, $x - \mu$, or to weight deviations relative to the mean income, $(x-\mu)/\mu$, using a weight function, $\omega(\cdot): R_0^+ \to R$, which incorporates the value judgments when adding local inequality. This procedure yields inequality measures of the type

$$I = \frac{1}{\mu}\int_0^\infty (x-\mu)\omega(x)dF(x) = \frac{1}{\mu}\int_0^1 (F^{-1}(p)-\mu)\lambda(p)dp, \qquad [1]$$

[2] Even some models used to adjust income distributions, such as the lognormal model, cannot be characterized by the sequence of their moments despite the fact that all of them are finite (Heyde 1963).

where $F^{-1}(p) = \inf\{x : F(x) \geq p\}$, $0 \leq p \leq 1$ is the income of an individual in the pth percentile of the distribution and $\lambda(p) = \omega(F^{-1}(p))$.

The above measures allow a simple geometric interpretation from the Lorenz curve, L (·), which is associated with the distribution. This curve is defined as:

$$L : [0,1] \to [0,1] \text{ , } L(p) = \frac{1}{\mu}\int_0^x sdF(s) = \frac{1}{\mu}\int_0^p F^{-1}(t)dt \text{ , } 0 \leq p \leq 1 \text{ , } p=F(x). \quad [2]$$

For each p=F(x), L(p) is the proportion of total income volume accumulated by the set of units with an income lower than or equal to x. It is clear that for 0≤p≤1 we have L(p)≤p, where L(p)=p in the case of perfect equality and L(p)=0 for 0≤p<1, L(1)=1 if the concentration is maximum. For any distribution, X, the Lorenz curve is increasing and convex and given the mean income, determines the density function. Hereinafter, $\Lambda = \left\{ F : 0 < \int_0^\infty xdF(x) < \infty \right\}$ represents the set of distributions with finite mean income, thus supporting the Lorenz curve.

From [1] and [2], and integrating by parts, we get:

$$I = \int_0^1 (p - L(p))\pi(p)dp, \ \pi(p) = \lambda'(p), \ 0 < p < 1. \quad [3]$$

Each of these indices weights the area between the Lorenz curve and the line of perfect equality[3]. The weight used is what differentiates one index from another.

The best known measure of inequality, the Gini (1914) coefficient, is obtained when $\omega(x) = 2F(x)$, or $\pi(p) = 2$. Its expression is:

$$G = \frac{2}{\mu}\int_0^\infty (x - \mu)F(x)dF(x) = 2\int_0^1 (p - L(p))dp.$$

The generalized Gini coefficients with a positive integer parameter are obtained by weighting the Lorenz differences with $\pi(p) = n(n-1)(1-p)^{n-2}$, $n \geq 2$. Their expressions are:

$$G(n) = 1 - n(n-1)\int_0^1 (1-p)^{n-2} L(p)dp = n(n-1)\int_0^1 (p - L(p))(1-p)^{n-2} dp. \quad [4]$$

If $\pi(p) = np^{n-2}$, $n \geq 2$, we obtain the family of indices proposed by Aaberge (2000):

[3] Note that $p - L(p)$ is the difference between the share in the total income of the individuals with an income smaller or equal to $x = F^{-1}(p)$, in the case of equally distributed income, and the effective share in the total income in the distribution considered. The function $\pi(\cdot)$ provides a criterion to add this difference along the distribution.

$$A(n) = 1 - n\int_0^1 p^{n-2}L(p)dp = n\int_0^1 (p - L(p))p^{n-2}dp. \quad [5]$$

For $n = 2$, $G(2) = A(2) = G$, therefore the Gini coefficient belongs to both families.

The above indices are all compromise indices[4]. These are relative indices (invariant to changes in the income scale) that when multiplied by the mean income become absolute measures (invariant under changes of origin).

2.2. Normative aspects. Welfare and inequality

In order to establish the relationship between inequality and social welfare we follow the Yaari approach (1987, 1988). If $F(\cdot)$ is the income distribution and $\phi : [0,1] \to R$ is a distribution function[5] that represents social preferences, the Yaari social welfare function (YSWF) is given by

$$W_\phi(F) = \int_{R^+} x d\phi(F(x)) = \int_0^1 F^{-1}(p)d\phi(p) = \int_0^1 \phi'(p)F^{-1}(p)dp . \quad [6]$$

Thus, Wϕ is additive and linear in the incomes and weights them according to the rankings assigned to the individuals in the distribution. The weight attached to the income of an individual with rank p, 0<p<1, is $\phi'(p) \geq 0$. Yaari (1988) shows that $W_\phi(F)$ presents an aversion to inequality if and only if $\phi'(p)$ is decreasing, which is equivalent to the concavity of ϕ.

If μ is the mean of $F(\cdot)$ and L(p) is its Lorenz curve, the YSWF W_ϕ can be expressed as a social welfare function associated to a linear measure of inequality of the type defined in [3]. Then[6],

$$W_\phi(F) = \mu[1 - I_\phi(F)], \quad [7]$$

where

$$I_\phi(F) = \int_0^1 (p - L(p))\pi_\phi(p)dp, \quad \pi_\phi(p) = -\phi''(p) . \quad [8]$$

The above two expressions yield an explicit relationship between the YSWFs, $W_\phi(\cdot)$, and their associated inequality indices, $I_\phi(\cdot)$, thus relating the distribution of preferences, $\phi(\cdot)$, and the weighting scheme of the Lorenz differences, $\pi_\phi(\cdot)$.

---

[4] A relative index, I, is a compromise index if μI is an absolute index. An absolute index, J, is a compromise index if J/μ is a relative index (Blackorby and Donaldson, 1978).

[5] We assume it to be a class C^2 function, which is twice continuously derivable. When necessary, we will admit the existence of higher order derivatives in later results.

[6] For a detailed calculation see Imedio and Bárcena (2007)

According to the Blackorby and Donaldson approach (1978), expression [7] is the equally distributed equivalent income[7], in which case $\mu I_\phi(F)$ measures the loss of social welfare due to inequality.

As pointed out earlier, $W_\phi(\cdot)$ or, equivalently, $I_\phi(\cdot)$ satisfy the Pigou-Dalton Principle of Transfers (PDPT) if and only if ϕ is a concave function. When studying more demanding redistributive criteria by which the effect of a transfer is greater the lower the part of the distribution in which it takes place, Kolm (1976a b) and Mehran (1976) propose two alternative versions. According to the Principle of Diminishing Transfers (PDT), a progressive transfer between two individuals with a given difference in income implies that the lower the income of these individuals, the greater the reduction (increase) in the index (social welfare). According to the Principle of Positional Transfer Sensitivity (PPTS), when there is a given difference in ranks among the individuals for whom the transfer takes place, the effect of the transfer is greater when it occurs among individuals in the lower part of the distribution. Although both principles are analogous with regard to the transfers, the income difference between the donor and the recipient is relevant for the PDT, while the proportion of individuals located between both is relevant for the PPTS. The following result shows how both principles are satisfied.

Proposition 1. Let $F(\cdot)$ be an income distribution with a positive mean and $I_\phi(F)$ an inequality index whose preference distribution, ϕ, is concave. Then,

(i) (Mehran, 1976; Zoli, 1999) Index $I_\phi(F)$ satisfies the PPTS if and only if $\phi'''(p) > 0$.

(ii) (Aaberge, 2000) Index $I_\phi(F)$ satisfies the PDT if and only if $\phi''(F(x))F'(x)$ is strictly increasing for x>0. This is equivalent to the condition

$$-\frac{\phi'''(F(x))}{\phi''(F(x))} > \frac{F''(x)}{(F'(x))^2}, \quad x > 0 \qquad (28)$$

The above proposition proves that an inequality measure satisfies, or does not satisfy, the PPTS depending on the properties of its preference distribution, ϕ, irrespective of the income distribution to which it is applied. It is, therefore, a characteristic of the index. However, the same does not occur with the PDT. That $I_\phi(F)$ satisfies the PDT depends not only on the properties of its preference distribution, but also on the shape of the income distribution. That is, given ϕ, index $I_\phi(F)$ verifies the PDT only for a given class of income distributions whose extension depends on the degree of inequality aversion of ϕ.

[7] This refers to a level of income such that if it is equally attached to all the individuals of the population, it will provide an identical level of social welfare, according to the specified SWF, to that of the existing distribution. This concept is the basis of the AKS approach (Atkinson, 1970; Kolm, 1976a y b; Sen, 1973) for relating social welfare and inequality.

2. 3. Order statistics

Let $X_1, X_2, ..., X_n$ be a sample of size n, $n \in N$, $n \geq 2$, from a distribution $F(\cdot)$ and define the order statistics, $X_{k:n}$, $k \in \{1, 2, ..., n\}$, in the ascending order by:

$$X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n} .$$

That is, in each specific realization of the sample, once the values are ordered from lowest to highest, the variable $X_{k:n}$ assigns the value at position k-th to each sample.

The cumulative distribution function $F_{k:n}(\cdot)$ of $X_{k:n}$, can be written as:

$$F_{k:n}(x) = \sum_{j=k}^{n} \binom{n}{j} (F(x))^j (1 - F(x))^{n-j} . \qquad [10]$$

If the variable, X, is continuous and $f(\cdot)$ is its density function, $f(\cdot) = F'(\cdot)$, the density function of $X_{k:n}$ is:

$$f_{k:n}(x) = k \binom{n}{k} (F(x))^{k-1} (1 - F(x))^{n-k} f(x) .$$

Particularly, for the first order statistic we get:

$$F_{1:n}(x) = 1 - (1 - F(x))^n , \quad f_{1:n}(x) = n(1 - F(x))^{n-1} f(x) .$$

For the maximum, we obtain:

$$F_{n:n}(x) = (F(x))^n , \quad f_{n:n}(x) = n(F(x))^{n-1} f(x) .$$

The mean values of the order statistics are given by the expressions:

$$E(X_{k:n}) = \int_0^{\infty} x dF_{k:n}(x) = k \binom{n}{k} \int_0^1 F^{-1}(p) p^{k-1} (1 - p)^{n-k} dp . \qquad [11]$$

From the above expression it is evident that $E|X_{k:n}| \leq cE|X|$ for some $c > 0$. Consequently, if the distribution $F(\cdot)$ has a finite mean, the existence of the first moment of any order statistic is assured. This property is important because there are distributions, such as heavy-tailed income distributions, for which only a few potential moments exist, and therefore no characterization in terms of (ordinary) moments is feasible. In these cases, it is interesting to analyze whether the distribution can be characterized by the moments of the order statistics, $\{E(X_{k:n})\}_{(k,n) \in N \times N}$, or a subset of them. There is also a recurrence relation between the first moments of order statistics (David 1981, p. 46)

$$(n - k)E(X_{k:n}) + kE(X_{k+1:n}) = nE(X_{k:n-1}) ,$$

which allows knowing the whole array $\{E(X_{k:n})\}_{(k,n) \in N \times N}$ if we can have one moment (of first order) for each sample size.

The following result characterizes a distribution function with finite mean from the expectations of their order statistics.

Proposition 2[8]. Let X be a random variable with $E|X| < \infty$ and $k(n)$ a positive integer, $1 \le k(n) \le n$, then the distribution $F(\cdot)$ is uniquely determined by the sequence $\{E(X_{k(n):n})\}_{n \in N}$.

Based on the previous proposition, in the next section any element of Λ is characterized by certain families of SWFs or their corresponding inequality measures.

3. Welfare functions and inequality measures generated through mean values of order statistics

From the distribution functions of order statistics, we can obtain functions with the right properties to be considered distributions of social preferences. After specifying these functions, and applying the procedure and the results detailed in section 2.2., their corresponding SWFs and corresponding inequality measures are identified. These measures are of the type defined in [1] or [3] applicable to distributions $F \in \Lambda$.

3.1. First order statistics and generalized Gini coefficients

If we write $x = F^{-1}(p)$ in the distribution function of the first order statistic, $F_{1:n}(\cdot)$, we get the following function:

$$\phi_{1:n}(p) = F_{1:n}(F^{-1}(p)) = 1-(1-p)^n, \quad 0 \le p \le 1, \; n \ge 2. \qquad [12]$$

The properties of $\phi_{1:n}(\cdot)$, growth and concavity, allow interpreting it as a distribution of preferences that results in a SWF with aversion to inequality. Applying [6], [11] and [12], we get:

$$W_{1:n}(F) = \int_0^1 F^{-1}(p)d\phi_{1:n}(p) = E(X_{1:n}) = n(n-1)\mu\int_0^1 (1-p)^{n-2}L(p)dp, \quad n \ge 2. \qquad [13]$$

The inequality measure corresponding to the above SWF, from [7] and [8], is:

$$I_{1:n}(F) = 1 - \frac{W_{1:n}(F)}{\mu} = 1 - \frac{E(X_{1:n})}{E(X)}, \; n \ge 2, \qquad [14]$$

or equivalently:

$$I_{1:n}(F) = 1 - n(n-1)\int_0^1 (1-p)^{n-2}L(p)dp = n(n-1)\int_0^1 (p-L(p))(1-p)^{n-2}dp, \quad n \ge 2.$$

[8] Huang (1989) proves this result and makes a detailed revision of the literature.

$I_{1:n}(F) = G(n)$ is the generalized Gini coefficient of order n for the distribution $F(\cdot)$, expression [4]. When the sample size varies, we obtain the different indices of this family. Particularly, the traditional Gini coefficient is obtained for samples of size 2, $I_{1:2}(F) = G$.

The corresponding absolute indices, $\mu I_{1:n}(F)$, which evaluate the loss of social welfare due to inequality, are given by:

$$\mu I_{1:n}(F) = \mu - W_{1:n}(F) = E\left(\overline{X}_n - X_{1:n}\right), \qquad [15]$$

where $\overline{X}_n = (X_{1:n} + X_{2:n} + \cdots + X_{n:n})/n$ is the sample mean.

Expression [13] implies that if we take random samples of size n, n≥2, from the income distribution $F(\cdot)$ and the welfare associated to each sample is identified with the minimum income, the mean value that is obtained when considering all possible samples of the given size is the welfare that the underlying SWF assigns to the generalized Gini coefficient of parameter n. Moreover, as a consequence of Proposition 2, we can ensure that any distribution $F \in \Lambda$ is characterized by the succession of SWFs $\{W_{1:n}(F)\}_n \equiv \{E(X_{1:n})\}_n$. Taking into account expressions [14] and [15], this result is equivalent to saying that any $F \in \Lambda$ is characterized by the sequence of the generalized absolute Gini coefficients $\{\mu I_{1:n}(F)\}_n$, or (up to a scale) by the sequence of the generalized relative Gini coefficients $\{I_{1:n}(F)\}_n$ (Kleiber and Kotz , 2002)

3.2. General case

The result obtained from the distributions of preferences, $\{\phi_{1:n}(\cdot)\}_n \equiv \{F_{1:n}(F^{-1}(\cdot))\}_n$, might suggest that, in general, the functions $\{F_{k:n}(F^{-1}(\cdot))\}_{2\le k\le n}$ can be preference distributions. It is found that these functions are increasing in the range $(0,1)$, but not necessarily concave over the whole range[9]. This would result in SWFs and indices of inequality that would not meet the Pigou-Dalton Principle of Transfers (PDPT); a condition that is equivalent to the aversion to inequality of the measures.

However, if for fixed sample size n, $n \ge 2$, we calculate consecutively the arithmetic mean of the functions $\{F_{k:n}(F^{-1}(\cdot))\}_{1\le k\le n}$, we obtain a sequence of functions which have an appropriate behavior to be considered distributions of social preferences.

[9] For example, $F_{2:n}(F^{-1}(p)) = 1-(1-p)^n - np(1-p)^{n-1}$ is strictly increasing in $(0,1)$ and strictly convex in $(0, 1/(n-1))$.

Definition. For each (n, k), $n \geq 2$, $k = 1, 2, ..., n$, we consider the function $\phi_{k:n} : [0,1] \to R$ given by:

$$\phi_{k:n}(p) = \frac{1}{k}\sum_{i=1}^{k} F_{i:n}(F^{-1}(p)),\ 0 \leq p \leq 1,\ \ k = 1, 2, ..., n. \qquad [16]$$

It is:

$$\phi_{1:n}(p) = F_{1:n}(F^{-1}(p)),$$

$$\phi_{2:n}(p) = \frac{F_{1:n}(F^{-1}(p)) + F_{2:n}(F^{-1}(p))}{2},$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\phi_{n-1:n}(p) = \frac{F_{1:n}(F^{-1}(p)) + F_{2:n}(F^{-1}(p)) + \cdots + F_{n-1:n}(F^{-1}(p))}{n-1},$$

$$\phi_{n:n}(p) = \frac{F_{1:n}(F^{-1}(p)) + F_{2:n}(F^{-1}(p)) + \cdots + F_{n-1:n}(F^{-1}(p)) + F_{n:n}(F^{-1}(p))}{n} = p.$$

Proposition 3. Each of the functions $\{\phi_{k:n}(\cdot)\}$, $1 \leq k \leq n$, defined in the interval $[0,1]$ shows the properties required of a distribution of social preferences.

Proof. From [10] and [16] the function $\phi_{k:n}(\cdot)$ is expressed as:

$$\phi_{k:n}(p) = \frac{1}{k}\sum_{i=1}^{k}\left(\sum_{j=i}^{n}\binom{n}{j}p^{j}(1-p)^{n-j}\right).$$

It is verified that $\phi_{k:n}(0) = 0$, $\phi_{k:n}(1) = 1$. Each $\phi_{k:n}(\cdot)$ is strictly increasing and for $1 \leq k \leq n-1$ is strictly concave in the interval $(0,1)$, as its first two derivatives are:

$$\phi'_{k:n}(p) = \frac{1}{k}\sum_{i=1}^{k}\left(i\binom{n}{i}p^{i-1}(1-p)^{n-i}\right),$$

$$\phi''_{k:n}(p) = -(n-k)\binom{n}{k}p^{k-1}(1-p)^{n-k-1}. \qquad [17]$$

It is evident that $\phi'_{k:n}(p) > 0, 1 \leq k \leq n$, $\phi''_{k:n}(p) < 0,\ 1 \leq k \leq n-1$, $p \in (0,1)$. □

As a consequence of the previous result, and applying the methodology described in section 2.2., we get the corresponding YSWFs $W_{k:n}(\cdot)$ from the distributions of preferences $\phi_{k:n}(\cdot)$. Their expressions are:

$$W_{k:n}(F) = \int_0^1 F^{-1}(p)d\phi_{k:n}(p) = E\left(\frac{1}{k}\left(X_{1:n} + X_{2:n} + \cdots + X_{k:n}\right)\right), k = 1, 2, .., n. \qquad [18].$$

Therefore, $W_{k:n}(F)$, $n \geq 2$, $1 \leq k \leq n$, is the expectation of the arithmetic mean of the k first order statistics for samples of size n from the distribution $F(\cdot)$. That is, if the level of welfare assigned to any sample of n incomes from $F(\cdot)$ is identified with the mean of their k lower incomes, the welfare of the population is the expectation of those values when considering all possible samples of size n.

The inequality measures underlying the above YSWFs, given [7], [8] and [18] are:

$$I_{k:n}(F) = 1 - \frac{W_{k:n}(F)}{\mu} = 1 - \frac{1}{k\mu} E\left(X_{1:n} + X_{2:n} + \cdots + X_{k:n}\right), n \geq 2,$$

or:

$$I_{k:n}(F) = (n-k)\binom{n}{k}\int_0^1 (p - L(p))p^{k-1}(1-p)^{n-k-1}dp, \ n \geq 2. \qquad [19]$$

The indices $\{I_{k:n}(\cdot)\}$ are linear inequality measures of the type defined in [1] or [3]. They weight Lorenz differences by:

$$\pi_{k:n}(p) = -\phi''_{k:n}(p) = (n-k)\binom{n}{k}p^{k-1}(1-p)^{n-k-1}.$$

The welfare loss due to inequality is measured by the corresponding absolute indices, $\mu I_{k:n}(F)$:

$$\mu I_{k:n}(F) = \mu - W_{k:n}(F) = E\left(\overline{X}_n - \frac{1}{k}(X_{1:n} + X_{2:n} + \cdots + X_{k:n})\right),$$

where $\overline{X}_n$ is the sample mean of order n. The above equality indicates that $\overline{X}_n - \left(X_{1:n} + X_{2:n} + \cdots + X_{k:n}\right)/k$ is an unbiased estimator of $\mu I_{k:n}(F)$. In a particular sample of n incomes, the difference between the mean and the mean of the k lower incomes is a point estimate of the index $\mu I_{k:n}(F)$.

3.3. Particular cases

• The case $k = 1$ has been studied, which provides the SWFs corresponding to the family of the generalized Gini coefficient.

• For $k = n-1$, the distribution of preferences are given by:

$$\phi_{n-1:n}(p) = \frac{\sum_{i=1}^{n-1} F_{i:n}(F^{-1}(p))}{n-1} = \frac{np - F_{n:n}(F^{-1}(p))}{n-1} = \frac{np - p^n}{n-1}, \ n \geq 2, \ 0 \leq p \leq 1.$$

The corresponding SWFs are expressed by applying [18] as:

$$W_{n-1:n}(F) = \frac{\sum_{i=1}^{n-1} E(X_{i:n})}{n-1} = \frac{E(n\overline{X}_n - X_{n:n})}{n-1} = \frac{n}{n-1}\mu - \frac{1}{n-1}E(X_{n:n}) . \qquad [20]$$

Their associated inequality measures are:

$$I_{n\text{-}1:n}(F) = 1 - \frac{W_{n-1:n}(F)}{\mu} = 1 - n\int_0^1 L(p)p^{n-2}dp = n\int_0^1 (p - L(p))p^{n-2}dp, \quad n \geq 2 .$$

$\{I_{n-1:n}(F)\}_n \equiv \{A(n)\}_n$ coincides with the family of indices proposed by Aaberge (2000), expression [5]. For $n = 2$ we get the Gini coefficient $I_{1:2}(F) = A(2) = G.$ The welfare loss due to inequality is given by the corresponding absolute indices, $\mu I_{n-1:n}(F)$:

$$\mu I_{n-1:n}(F) = \mu - W_{n-1:n}(F) = E\left(\overline{X}_n - \frac{1}{n-1}\sum_{i=1}^{n-1} X_{i:n}\right). \qquad [21]$$

The expressions [20] and [21] imply that, in this case, the welfare associated with each set of n incomes is identified with the average obtained by excluding the highest income. The difference between this average and the one of the whole group, including the maximum income, is the welfare loss due to inequality. For the population, the welfare and the cost of inequality is the average that would be obtained by considering all possible sets of n incomes. Moreover, using Proposition 2 again, we can say that any distribution $F \in \Lambda$ is characterized by the sequence of SWFs $\{W_{n-1:n}(F)\}_n$ or by the absolute indices $\{\mu I_{n-1:n}(F)\}_n$. The family of relative indices $\{I_{n-1:n}(F)\}_n$ also determines the distribution F, except for a multiplicative factor.

• If $k = n,$ the distribution of preferences is:

$$\phi_{n:n}(p) = \frac{1}{n}\sum_{i=1}^{n} F_{i:n}(F^{-1}(p)) = p , \; n \geq 2 .$$

That is, $\phi_{n:n}(\cdot)$ is the identity in $[0,1]$, which is strictly increasing but not strictly concave. Hence the resulting SWF shows no aversion to inequality. This SWF identifies the welfare of each income distribution, $F(\cdot),$ with its average income. Indeed, applying [18] we get:

$$W_{n:n}(F) = \int_0^1 F^{-1}(p)dp = E\left(\frac{1}{n}\left(X_{1:n} + X_{2:n} + \cdots + X_{n:n}\right)\right) = E\left(\overline{X}_n\right) = \mu .$$

Consequently, the associated inequality index is zero for any distribution:

$$I_{n:n}(F) = 1 - \frac{W_{n:n}(F)}{\mu} = 0, \; n \geq 2 .$$

This does not imply the absence of inequality, but that both the SWF and its corresponding index are indifferent to inequality.

3.4. Some additional policy considerations

In general, when the value of n is fixed and k varies, the criteria used in the measurement of welfare and inequality are modified. When k increases, the preference distributions reduce their concavity, and the SWFs therefore show less inequality aversion, form the corresponding to the generalized Gini coefficient, $W_{1:n}(.)$, until indifference, $W_{n:n}(F) = \mu_F$. Consequently, when k increases, the associated inequality measures assign less weight to the inequality corresponding to low incomes and greater weight to the inequality corresponding to high incomes. Indeed, in the expressions of the indices that weight Lorenz differences, expression [19], these weights are:

$$\pi_{1:n}(p) = n(n-1)(1-p)^{n-2},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\pi_{k:n}(p) = (n-k)\binom{n}{k}p^{k-1}(1-p)^{n-k-1},$$

$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,$$

$$\pi_{n-1:n}(p) = np^{n-2},$$

$$\pi_{n:n}(p) = 0.$$

It is evident that if $n > 2$, $\pi_{1:n}(\cdot)$ is strictly decreasing in the interval $[0,1]$, $\pi_{k:n}(\cdot)$, $1 < k < n-1$, is increasing in $[0,(k-1)/(n-2)]$ and decreasing in the rest of the interval, while $\pi_{n-1:n}(\cdot)$ is strictly increasing. Moreover, $\pi_{n:n}(\cdot)$ assigns a weight of zero to all Lorenz differences.

Figure 1 displays, for $n = 5$, the distributions of preferences $\{\phi_{k:5}\}_{1\le k\le 5}$ for the SWFs $\{W_{k:5}\}_{1\le k\le 5}$ and the weights $\{\pi_{k:5}\}_{1\le k\le 5}$ of the Lorenz differences corresponding to the indices $\{I_{k:5}\}_{1\le k\le 5}$.

----Figure1. Distributions of preferences and weights of the Lorenz differences, n=5, $1 \le k \le 5$------

As k increases, the concavity of the distribution of preferences decreases (being linear for k=5), while the weight given to inequality in low income diminishes, with income inequality weighing more in intermediate and high incomes. Indeed, $\pi_{1:5}$ is strictly decreasing, $\pi_{2:5}$ is bell-shaped and maximum for $p = 1/3$, $\pi_{3:5}$ is bell-shaped and maximum for $p = 2/3$, and $\pi_{4:5}$ is strictly increasing. For $k = 5$, the SWF is indifferent to inequality and both the weighting and the index are zero.

The above considerations show that given an income distribution, $F \in \Lambda$, the family of SWF $\{W_{k:n}(F)\}_{(k,n)}$ results in a family of inequality indices $\{I_{k:n}(F)\}_{(k,n)}$, which includes the generalized Gini coefficients, $\{I_{1:n}(F)\}_n$, and the Aaberge (2000) indices as subfamilies. The weights of local inequality in both subfamilies of inequality measures are monotonic functions along the distribution so that the greater weight is assigned to one of its ends. However, the weights for the Lorenz differences for the indices of the family, $\{I_{k:n}(F)\}_{(k,n)}$, $1 < k < n-1$, are not monotonic. They can reach their maximum or minimum value at any percentile. This allows for measures with different attitudes in assessing inequality and welfare, as they pay more attention to different parts of the distribution.

For each inequality index, the expression [8] relates the weighting function of the cumulative local inequality up to each percentile of the income distribution with the distribution of preferences of the associated SWF. These functions, which are related to each other, determine the characteristics of the index. These include the degree of inequality aversion and the response to income transfers when the difference in rank, or the income difference, between the individuals involved in the transfer and their location in the distribution is considered. The behavior of the indices $\{I_{k:n}(F)\}_{(k,n)}$ regarding PPTS and DTP is obtained by applying Proposition 1 to their preference distributions.

Proposition 4

a) The indices of the family $\{I_{k:n}(F)\}_{(k,n)}$, $n \geq 2,\ 1 \leq k \leq n$, satisfy the PPTS if and only if[10]:

$$T(n,k,p) = [(n-2)p - (k-1)] > 0, \ \text{for any } p \in (0,1). \quad [22]$$

b) The index $I_{k:n}(F)$, which is applied over the distribution function F, satisfies the DTP if and only if:

$$\frac{(n-2)F(x)-(k-1)}{F(x)(1-F(x))} > \frac{F''(x)}{(F'(x))^2}, \ x > 0. \quad [23]$$

The above proposition proves that the Gini index, $G = I_{1:2}$, does not satisfy the PPTS because T(2,1,p)=0, 0<p<1. That is, given a rank difference, the effect over G of any progressive transfer is the same irrespective of the income distribution to which it is applied. However, the generalized Gini coefficient, $G(n) = I_{1:n}$ for n>2, does fulfill this principle since in this case $T(n,1,p) = (n-2)p > 0$, $0 < p < 1$. Other indices exhibit a behavior opposite to the PPTS. That is, a progressive transfer reduces the value of the index but given a difference of ranks, the reduction is greater the higher the individuals involved are in the income distribution. This is the case for the Aaberge indices, $A(n) = I_{n-1:n}$ for n> 2, because

[10] The sign of the third derivative of the preference distribution, which can be constant or not in (0,1) depending on the values of n and k, coincides with expression T(n,k,p).

$T(n, n-1, p) = (n-2)(p-1) < 0$, $0 < p < 1$. There are also indices $I_{k:n}$ whose behavior with respect to this principle is not uniform. For example, for $I_{2:5}$ it is $T(5,2,p) = 3p - 1$, so that this index satisfies the PPTS if p> 1/3 and behaves the opposite for 0 <p <1/3.

Regarding the DTP, if an index has aversion towards inequality ($\phi''(p)<0$) and its preference function has a non-negative positive third derivative $\phi'''(p) \geq 0$, it will satisfy the PDT for any concave income distribution ($F''(x)<0$), as condition [9] is satisfied. This is the case of the Gini coefficient. The concavity of F is a sufficient condition in these cases. However, the observed distribution functions do not present a uniform behavior throughout the income scales regarding concavity/convexity. They tend to be unimodal distributions with asymmetry to the right. In the observed distributions, in a given setting, there are income levels where the slope of the density function does not have a constant sign. Therefore the PDT will not be verified throughout the range of the income variable, but can be verified in specific intervals. In general, it can be shown that if an inequality measure satisfies the PDT in a certain range, any other measure with greater inequality aversion also verifies that principle on that interval and possibly on others of greater amplitude. In our case, for indices $I_{n:k}$ given $n \geq 2$, the smaller K and the greater the inequality aversion of the index, the wider the set of income distributions for which the PDT is satisfied.

4. Conclusions

The use of order statistics in the definition of SWFs and indices of inequality provides a joint treatment of measures that share common features, but differ from and complement each other from the normative standpoint.

Our approach leads to linear measures that are the result of weighting the differences between the incomes of the distribution and the average income. This is equivalent to weighting the Lorenz differences or inequality accumulated up to each income percentile. The different weighting schemes generate different attitudes in the assessment of inequality and welfare throughout the distribution, depending on the part of the distribution considered, and the degree of inequality aversion of the indices.

The approach adopted in this work not only provides a constructive procedure to define the measures under study, but also proves that, given the average income, certain families of indices characterize the income distribution, and provides a clear statistical interpretation to each FBS and its corresponding index of inequality.

In practice, the appropriate selection of various elements of the set $\{I_{k:n}\}_{(n,k)}$, or $\{W_{k:n}\}_{(n,k)}$, permits applying very different distributional judgments when comparing levels of inequality or welfare associated with different income distributions. Hence, the conclusion in a particular application may be interesting either when a robust result is obtained, or if the outcome is different depending on the index considered, as the properties of the different measures are taken into account.

Acknowledgements

Elena Bárcena acknowledges the financial support provided by the Spanish Ministry of Education through Grant ECO2012-33993. She also thanks the financial support provided by Fundación Ramón Areces, XI Concurso Nacional para la Adjudicación de Ayudas a la Investigación en Ciencias Sociales.

References

Aaberge, R. (2000). Characterizations of Lorenz curves and income distributions. *Social Choice and Welfare*, 17, 639-653.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of Economic Theory*, 2, 244-63.

Blackorby, C. & Donaldson, D. (1978). Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory*, 18, 59-80.

David, H. A. (1981). *Order Statistics.* (2nd edition). New York. Wiley.

Donaldson, D. & Weymark, J. A. (1980). A single parameter generalization of the Gini indices of inequality. *Journal of Economic Theory*, 22, 67-86.

Donaldson, D. & Weymark, J. A (1983). Ethically flexible indices for income distributions in the continuum. *Journal of Economic Theory*, 29, 353-358.

Gini, C. (1914). Sulla misura della concentrazione e della variabilità dei caratteri. *Atti del Reale Istituto Veneto di Scienze, Lettere ed Arti*, 73, 1203-1248.

Heyde, C. C. (1963). On a property of the lognormal distribution. *Journal of the Royal Statistical Society*, Series B, 25, 392-393.

Huang, J. S. (1989). Moment problem of order statistics. *International Statistical Review*, 57, 59-66.

Imedio Olmedo, L. J. & Bárcena Martín, E. (2007). Dos familias numerables de medidas de desigualdad. *Investigaciones Económicas*, XXX(1), 191-217.

Kakwani, N. C. (1980). On a class of poverty measures. *Econometrica*, 48, 437-446.

Kleiber C. & Kotz, S. (2002). A characterization of income distributions in terms of generalized Gini coefficients. *Social Choice and Welfare*, 19, 789-794.

Kolm, S. C. (1976a). Unequal inequalities, I. *Journal of Economic Theory*, 12, 416-442.

Kolm, S. C. (1976b). Unequal inequalities, II. *Journal of Economic Theory*, 13, 82-111.

Lambert, P. J. (2001). *The distribution and redistribution of income*. Manchester University Press.

Mandelbrot, B. (1960). The Pareto-Lévy law and the distribution of income. *International Economic Review*, 1, 79-106.

Mehran, F. (1976). Linear measures of inequality. *Econometrica*, 44, 805-809.

Sen, A. K. (1973). *On economic inequality.* Clarendon Press, Oxford.

Singh S. K. & Maddala, G. S. (1976). A function for the size distribution of incomes. Econometrica, 44, 963-970.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica*, 55, 99-115.

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

Yitzhaky, S. (1983). On an extension of the Gini index. *International Economic Review*, 24, 617-628.

Zoli, C. (1999). Intersecting generalized Lorenz curves and the Gini index. *Social Choice and Welfare*, 16, 183-196.

Figure 1. Preference distributions and weights of the Lorenz differences, n=5, $1 \le k \le 5$