**First Meeting of The Society For The Study of Economic Inequality (ECINEQ)**

# THE EVOLUTION OF ECONOMIC INEQUALITY IN THE EU COUNTRIES DURING THE NINETIES[1]

**Juana Domínguez-Domínguez (juana.dominguez@uah.es)**

**José Javier Núñez-Velázquez (josej.nunez@uah.es)**

Departamento de Estadística, Estructura Económica y O.E.I. Universidad de Alcalá. Plaza de la Victoria, 2. Alcalá de Henares, 28802 Madrid
Phone: 918854276. Fax: 918854201

**ABSTRACT**

In recent years, there have been several studies comparing the evolution of economic inequality among different territorial units. In these cases, an inequality indicator was chosen, and then its value was calculated from sample data to give the information of inequality in each period of time considered. Thus, the problem turned out to be the selection of the inequality indicator.

This paper aims to prove that there is no need for a selection of a single inequality indicator. A whole set of inequality indicators are going to be considered and calculated for the European Countries, using income data from European Community Household Panel (ECHP). The information they provide is then resumed into a composite inequality indicator, through an adaptation of Principal Component Analysis (PCA). We analyze the conditions needed to make longitudinal comparisons possible.

Results obtained with this composite indicator are used to compare and analyze the trends followed by economic inequality in European Countries during the nineties and to classify them according to their evolution.

**KEY WORDS:** Economic Inequality, European Union, Multivariate Analysis, Economic Indicators.

---

[1] The authors gratefully acknowledge financial support from University of Alcalá (Research Project PI-UAH2004/034). ECHP data are used under permission of EUROSTAT contract ECHP/15/00, held with the University of Alcalá. The authors acknowledge the collaboration of Dr. Rivera-Galicia in previous versions of this paper. The usual disclaimer applies.

## 1. INTRODUCTION

Economic inequality has always been one of the most recurrent research fields in economics, because its presence is related to several important tasks, not only in economics but also in social, political and in many other ones. In such a sense, Sen (1973) pointed out how inequality can be connected to many causes of uneasiness, including social rebellions indeed. Also, its implications on other interesting economic concepts should be recognized, such as convergence, poverty, welfare and so on. Nevertheless, the last decades have seen an appreciable researchers' interest increase in several aspects related to economic inequality. Probably, this new increasing interest began with Atkinson (1970) and Sen (1973). Both have been considered as seminal studies, focused on basic aspects such as quantitative measuring, economic theoretical grounds or inequality comparisons.

Perhaps, overall agreement about analysing inequality can be placed on Lorenz curve. Since its presentation in Lorenz (1905), these curves remain as useful and comprehensive tools for comparing the accumulated percents of perceiving units and perceived resources, giving as a result the extent of inequality in the distribution [2]. But, in order to make comparisons between income distributions, their Lorenz curves will only produce valid results if they are completely nested and then the Lorenz curve closer to the egalitarian line is said to represent less inequality. Further, this so-called *Lorenz dominance criterion* allows only a quasi-order relationship among the set of income distributions, because intersections between Lorenz curves occur very often. Shorrocks (1983) proposed a generalized curve using the income mean to rescale the Lorenz curve ordinates, defining a *generalized Lorenz dominance criterion* to compare income distributions in a similar way. However, intersections may occur generating also a quasi-order relationship, and the underlying concept under these curves isn't yet inequality exactly, including income-welfare aspects. These elements constitute an active research field, with connections to stochastic dominance concepts[3].

---

[2] General details can be found in Kendall and Stuart (1977), for example.
[3] See Bishop, Formby and Sakano (1995), Davies and Hoy (1994) or Muliere and Scarsini (1989), for example. A revision of these elements can be found in Núñez (2002).

The natural way to overcome the difficulties associated with partial orders consists of inequality measures construction. These indicators will resume all the income inequality content in a single number, making possible a total order relationship over the income distribution space. Obviously, it seems clear that these indicators must be compatible with Lorenz criterion, which can be characterized by four well-known properties (Foster, 1985). Nevertheless, these restrictions result to be weak because there are a great number of inequality indexes fulfilling them. So, it is difficult to reach an agreement about the selection of a better inequality measure, resulting a set of them for use in current practice. There are several research fields related to the study of reasonable restrictions looking for isolating some index among the above-mentioned set. So, some authors are trying to restrict the so-called *Pigou-Dalton Transfers Principle* in an economic-based suitable way[4]. Other research field consists of imposing additional properties to narrow the set of available indicators; among these properties, it must be remarkable the use of the *additive decomposition properties* (Bourguignon, 1979) in order to separate the *between-groups inequality* and the *intra-groups inequality* in an additive manner, when the population is divided into subgroups. Another selection task consists of considering social welfare functions defined on economic theoretical grounds as an underlying support of inequality measures, but this research field presents hard controversies too[5].

Despite the above discussion, difficulties in choosing one better inequality measure still remain. The underlying problem arises because different measures may lead to different inequality orderings. Essentially, inequality measures we are discussing about, hide weighting schemes defined over the Lorenz curve ordinates, depending on different ideas behind their construction. So, in this paper the use of a whole set of admissible inequality measures is proposed in order to extract their common information, which will be inequality, necessarily. Beyond this idea, our proposal of a synthetic inequality indicator will be able to study dynamic trends too, after the necessary methodological adjustments.

Nevertheless, inequality trends studies are not new, but all of them use selected simple inequality measures or partial orders derived of domination relationships

[4] Further details can be seen in Shorrocks and Foster (1987) or Davies and Hoy (1995).
[5] See Atkinson (1970) and Dagum (1990), for example.

schemes such as Lorenz curve, generalized Lorenz curve or stochastic dominance based. Some examples referred to different countries are Lovell (1998), using Lorenz dominance and several inequality measures; Jenkins (1995), Achdut (1996) and Frick and Grabka (2003), using several inequality measures and decomposition property of some of them or Bishop, Formby and Smith (1991), using Lorenz dominance.

The structure of this paper is as follows. In section 2, general methodology and modifications introduced are presented. In sections 3 and 4, data used and empirical results are analysed. Last section will summarize the main conclusions of this paper.

## 2. METHODOLOGY

Through the following paragraphs, we are going to discuss the different decisions that must be made in order to define the final synthetic indicator proposed. First of all, we need to construct the space of income distributions as a useful background for next developments, keeping in mind that household economic position is going to be set through its global income[6]. Our formal framework follows the guidelines exposed in Ruiz-Castillo (1987), and further details can be found there.

Let *X* be a non-negative vector of incomes, defined in the usual way, with dimension determined by population size. Thus, the income space can be defined using the following set:

$$D = \bigcup_{N=2}^{\infty} D_N \ ,$$

where:

$$D_N = \left\{ (x_1, \dots, x_N) : x_i \geq 0, i = 1 \dots N; \sum_{i=1}^{N} x_i > 0 \right\}.$$

Obviously, usual definitions about inequality measures must be understood over the above set and they will be real-valued.

[6] The subsequent construction would be valid if the household economic position measurement is changed, using any other option, like expenditures, earnings or disposable incomes.

## 2.1. Selection of an inequality indicators set

There are a great number of inequality measures proposed in the literature (Foster and Sen, 1997; Nygard and Sandstrom, 1981, for example) and there is no agreement about which one could perform better. However, it is usual the establishment of a minimal set of properties to limit their scope. Let us consider the four axioms that characterize Lorenz dominance compatibility: *anonymity or symmetry, scale invariance, Dalton's Population Principle and the weak version of the Pigou-Dalton Transfers Principle* (Foster, 1985). We add the *Normalization Axiom* (inequality measures are either zero when all recipients have the same income or one if concentration attains its maximum). In such a case, the selection process could lead to the following simple inequality indicators[7], whose expressions are given in a descriptive mode over a general vector of incomes, $X \in D$:

1. Atkinson inequality index, with parameter 0.5[8]:

$$\text{ATKIN0.5} = 1 - \frac{1}{\mu} \cdot \left( \frac{1}{N} \sum_{i=1}^{N} \sqrt{x_i} \right)^2 ,$$

where $\mu$ is the income arithmetic mean.

2. Atkinson inequality index, with parameter 1:

$$\text{ATKIN1} = 1 - \prod_{i=1}^{N} \left( \frac{x_i}{\mu} \right)^{1/N} .$$

3. Atkinson inequality index, with parameter 2:

$$\text{ATKIN2} = 1 - (\mu_H / \mu),$$

where $\mu_H$ is the income harmonic mean.

[7] See Pena, Callealta, Casas, Merediz and Núñez (1996) and García, Núñez, Rivera and Zamora (2002), for further details.

[8] The family of Atkinson Index is obtained through the following equation: $A = 1 - \frac{1}{\mu} \cdot \left( \frac{1}{N} \sum_{i=1}^{N} x_i^{1-\varepsilon} \right)^{\frac{1}{1-\varepsilon}}$,

where $\varepsilon$ is a parameter of aversion to inequality. The sensitivity of the Atkinson index to different shares of the distribution depends on the value attributed to this parameter. The greater the level of $\varepsilon$, the greater the aversion to inequality.

4. Normalized Squared Coefficient of Variation [9]:

$$CV2.NORM = CV^2 / (1+CV^2),$$

where *CV* represents the coefficient of variation of the income distribution.

5. Gini index:

$$GINI = \frac{1}{2N^2\mu} \sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j| \quad .$$

6. Pietra or Schutz index:

$$PIETRA = \frac{1}{2N\mu} \cdot \sum_{i=1}^{N}|x_i - \mu| \, .$$

7. Normalized Theil index, with parameter 1:

$$TH1.NORM = 1 - \exp(-THEIL1),$$

where $THEIL1 = \frac{1}{N\mu}\sum_{i=1}^{N} x_i \cdot \log\left(\frac{x_i}{\mu}\right)$.

### 2.2. Construction of the cross-section synthetic inequality indicators

Once we have selected a group of good inequality measures, we would need the selection of a unique indicator to proceed. However, as long as no argument can be found to choose one of them, our option will be the use of the whole set as a battery of indicators. This latter approximation has a precedent in Sen (1973), who proposed the use of a battery of indicators to compare income vectors using his *intersection relationship*, giving as a result a quasi-order structure defined over the income set (*D*), like Lorenz domination does. However, our proposal will generate a complete order, by extracting the common inequality information they contain.

Let us expose the structure of data where methodology is going to be applied. Let $(I_1, I_2, ..., I_p)$ be a set of *p* simple indicators, which can be seen as a p-dimensional variable defined over the income space, whose values have been measured in each case

[9] We prefer the use of normalizing functions instead of the current option, which use the maximum value to divide. This last practice could produce Dalton Population Principle failure (Pena, Callealta, Casas, Merediz and Núñez, 1996).

of study (European countries in this paper), and let $T = \{t_0, t_1, ..., t_k\}$ be the set of different periods of time considered, when this set of simple indicators is calculated. For each $t \in T$, we compute these $p$ simple indicators over the income distribution of each territorial unit considered, thus having a ($n(t)$ x $p$)-dimensional matrix $\boldsymbol{I(t)}$, where $n(t)$ is the number of territorial units at moment $t$.

The above discussion suggests the possibility of considering a data matrices classification, where groups have been defined by the elements of the temporary set $T$. So, we can perform multivariate techniques on the data matrix defined over each point in time, generating a cross-section result. But the set of indicators are measuring inequality all of them and, thus, their content must be determined using such a fact. This conclusion leads us to think of Principal Components Analysis as a useful technique to extract the common information the battery of indicators offers. Particularly useful must be the First Principal Component if the accumulated variance is big enough, as we can expect.

The formal construction of such a cross-section indicator follows the guidelines exposed in García, Núñez, Rivera and Zamora (2002), when time is not taken into account. Let $(Y_1(t), Y_2(t), ..., Y_p(t))$ be the p-dimensional variable defined using the former variables under standardization along the corresponding cases in $t \in T$. Thus, data matrix in $t \in T$ will be $\boldsymbol{Y(t)}$, whose elements are defined by:

$$Y_{ij}(t) = Y_j(x_i(t)) = \frac{I_j(x_i(t)) - \mu_j(t)}{s_j(t)}, \quad i = 1,2,...,n(t); \quad j = 1,2,..., p; \quad t \in T \qquad (1)$$

where $x_i(t) \in D$ stands for the ith territorial unit vector of incomes, measured at moment $t$, $\mu_j(t)$ is the mean of the indicator $I_j$ calculated over all cases in $t$ and $s_j(t)$ is the corresponding standard deviation. In such circumstances, let $\boldsymbol{R(t)}$ be the associated variance-covariance matrix from **Y(t)**[10] and let $u_1(t), u_2(t), ..., u_p(t)$ be the eigenvectors extracted from **R(t)**, associated to its eigenvalues ordered from the greatest to the smallest one.

[10] As the variables have been standardized, this variance-covariance matrix is equivalent to the correlation matrix of the original variables.

The first principal component can be expressed as follows:

$$Z_1(t) = Z_1(x(t)) = u_1(t)\cdot(Y_1(x(t)),\ldots,Y_p(x(t)))' = \sum_{j=1}^{p} u_{1j}(t).Y_j(x(t)) \qquad (2)$$

with $x(t)\in D, \quad t\in T$. This becomes to be the optimal linear predictor when minimum squared error is used (Peña, 2002, 168-170). Furthermore, when the explained variability by first principal component becomes bigger, the obtained error is smaller.

After elementary algebraic manipulations, we have:

$$Z_1(x(t)) + K(t) = \sum_{j=1}^{p} \frac{u_{1j}(t)}{s_j(t)} \cdot I_j(x(t)),$$

where $K(t)$ is a value depending on $u_1(t)$, $\mu(t)$ and $s(t)$, but not on $x(t)$, except through the vectors expressed. Obviously, $\mu(t)$ and $s(t)$ are vectors compounded by the indicators means and standard deviations, respectively.

Finally, the proposed cross-sectional synthetic indicator can be expressed in the following way:

$$Z(x(t)) = \frac{Z_1(x(t)) + K(t)}{\sum_{h=1}^{p}(u_{1h}(t)/s_h(t))} = \sum_{j=1}^{p} a_j^*(t)\cdot I_j(x(t)), \quad x(t)\in D, \quad t\in T, \qquad (3)$$

with:

$$a_j^*(t) = \frac{u_{1j}(t)/s_j(t)}{\sum_{h=1}^{p}(u_{1h}(t)/s_h(t))}, \quad j = 1,2,\ldots,p,$$

and we have the synthetic longitudinal indicator as a convex linear combination of the initial simple indicators in the selected battery[11].

As it can be easily proved, this indicator is compatible with Lorenz domination relationship, and it is a normalized index. Furthermore, Z(t) constitutes an inequality indicator because it has been constructed using a battery of inequality ones, and this would be the primary content of the first principal component.

### 2.3. A dynamic synthetic indicator of inequality

Unfortunately, the synthetic indicator proposed in (3) will give us different functions on each point in time, because the first eigenvector of **R(t)** could change depending on *t*. To avoid this problem, we remind that data come from samples of households and, thus, correlation matrices are only estimations of the population ones. If we could admit that all these matrices were the same, then equality among first eigenvectors involved would be considered. In such a case, we might use a pooled estimate of the common variance-covariance matrix in order to obtain a unique eigenvector, which will be time-independent, providing a valid indicator for all periods in *T*.

So, as a first option, we propose the use of a test to contrast the hypothesis of a stable variance-covariance structure (correlation in our case). The selected test will be an adaptation of Box M, whose basic details can be found in Rencher (1995), for example[12].

If the same variance-covariance structure is accepted, then joint consideration of simple indicators is proposed, independently of their temporary period of reference, obtaining the pooled correlation matrix, ***R***. So, we might use only the first eigenvector, $u_1$, over the whole time period, and the proposed *global principal component synthetic indicator* can be written as:

---

[11] By construction, the elements of the eigenvector $u_1(t)$ must be non-negative because it was derived from the matrix $R(t)$.

[12] Further analytical details related to this process can be found in Domínguez, Núñez and Rivera (2004).

$$Z_{GPC}(x(t)) = \sum_{j=1}^{p} a_j^{*}.I_j(x(t)) = \sum_{j=1}^{p}\left(\frac{u_{ij}/S_j}{\sum_{i=1}^{p} u_{ij}/S_j}\right) I_j(x(t)), \quad t \in T \qquad (4)$$

As it can be observed, the convex linear combination coefficients are now constant across time. So, the incidence of each country income vector operates only through its value measured by the simple indicators, thus allowing their dynamic analysis, because the basic framework is the same, providing a stable weighting scheme over the initial set of indicators. Also, an analysis of the differential facts involved in the individual measuring characteristics could be possible, taking into account the second principal component.

On the other hand, let us suppose now that null hypothesis of stable correlation structure has been rejected and, therefore, at least one variance-covariance matrix is different. In such a case, it may still be possible to find out another way of solving the problem of comparison, using an adaptation of an algebraic method to locate the closest vector to the common space generated by principal components, proposed in Krzanowski (1979, 1982), named the *Common Space Analysis* procedure[13].

The aforementioned adaptation of Krzanowski's method is the following. If all the first eigenvectors associated to $\{\mathbf{R(t)},\ t \in T\}$ were close to each other, it would be possible to find out a vector located in their neighborhood. Using only the first principal components, Theorem 3 included in Krzanowski (1979, pg. 705) permits to assure that the vector we are looking for is the first eigenvector ($v$) of the matrix:

$$H = \sum_{t \in T} u_1'(t) \cdot u_1(t),$$

which maximizes

$$B = \sum_{t \in T} \cos^2 \delta_t ,$$

[13] An equivalent technique in a more descriptive framework, can be found in Keramidas, Devlin and Gnanadesikan (1987).

where $\delta_t$ is the angle between $u_1(t)$ and *v*. This solution is valid only if the first eigenvectors associated to {**R(t)**, $t \in T$} are close, in such a manner that the angles between *v* and each of them are small enough. At this point, it seems reasonable to expect such behavior when we are dealing with indicators trying to measure the same concept. Finally, the alternative synthetic inequality indicator would be named the *common space-based synthetic indicator*, defined as:

$$Z_{CS}(x(t)) = \sum_{j=1}^{p} b_j^* \cdot I_j(x(t)) = \sum_{j=1}^{p} \left( \frac{v_j / s_j}{\sum_{h=1}^{p} (v_h / s_h)} \right) \cdot I_j(x(t)), t \in T . \quad (5)$$

Now, it comes evident how if the first proposed synthetic indicator (equation 4) is adequate, the second (equation 5) must be very close to it. Nevertheless, in contexts where high correlations among the indicators should be expected, the second approximation provides an interesting alternative when the first one fails (in cases where sample oscillations are important).

## 3. DATA DESCRIPTION

The computation of inequality indexes will be accomplished using data from the European Community Household Panel (ECHP). ECHP is a longitudinal survey of households and individuals, centrally designed and coordinated by the Statistical Office of the European Communities (EUROSTAT) and covering all countries of the European Union. An attractive feature of ECHP is its comparability across countries and over time, as the questionnaire is similar and the elaboration process of the survey is carried out by EUROSTAT (Álvarez-García, Prieto-Rodríguez and Salas, 2002).

As household economic position we have chosen, as a shake of convenience, is total net household income, which is one of the variables included in ECHP. In order to include household size in the analysis, we use the potential equivalence scale proposed in Buhmann *et al.* (1988), using s=0.5 as its elasticity value. It is well known that levels in measured income inequality can vary depending on the choice of equivalence scale, although none of them has been proved to be superior. It is not the purpose of this paper to analyze the influence of equivalence scales on income inequality, but to see the way in which a set of indicators can be aggregated (for further discussion on equivalence scales, see, for example, Coulter, Cowell and Jenkins, 1992, Buhmann, Rainwater, Schmaus, and Smeeding, 1988, or Casas, Domínguez and Núñez, 2003, in the Spanish case).

In order to face a comparative study of inequality in the European countries, in a cross-sectional as well as in a longitudinal sense, net household income has been transformed into US dollars, using exchange rates obtained from EUROSTAT.

A full description of the ECHP dataset in terms of sampling, response rates, weighting procedures, etc., can be found in specialized literature (EUROSTAT publications and web page, Nicoletti and Peracchi, 2002, Ayala and Sastre, 2002, etc.), but it is necessary to point out that we had to exclude some households from the dataset in our analysis because they presented missing values for total net household income. Table 1 shows the initial number of cases in each country and the number of households that were finally selected. It is interesting to notice the large amount of households from Sweden for which this variable is not available. Despite Layte, Maître, Nolan and Whelan (2000) indicate that they had excluded Luxembourg because it must be frequently treated as an exceptional case, we haven't found empirical evidence to discard this case, or any other. Although Austria, Finland and Sweden were not included in the first waves of the ECHP, we have decided to include them in those waves where their data are available, in order to enrich the comparative results.

Finally, in this paper, we have taken into account the information from waves 1 to 8, which correspond to years 1994 to 2001. As it is well known, income data of each wave is always referred to the previous year, thus they give us information about years 1993 to 2000.

**Table 1**

*Total sample sizes and sample sizes for households with total net income, in brackets. ECHP Countries, Waves 1 to 8.*

| Country | Code | Wave 1 1993 | Wave 2 1994 | Wave 3 1995 | Wave 4 1996 | Wave 5 1997 | Wave 6 1998 | Wave 7 1999 | Wave 8 2000 |
|---|---|---|---|---|---|---|---|---|---|
| **Denmark** | DK | 3482 (3478) | 3223 (3218) | 2955 (2951) | 2745 (2740) | 2512 (2505) | 2387 (2381) | 2281 (2273) | 2283 (2279) |
| **Netherlands** | NL | 5187 (5139) | 5110 (5035) | 5179 (5097) | 5049 (5019) | 4963 (4922) | 5023 (4981) | 5008 (4976) | 4851 (4824) |
| **Belgium** | BE | 3490 (3454) | 3366 (3343) | 3210 (3191) | 3039 (3013) | 2876 (2863) | 2712 (2691) | 2571 (2555) | 2362 (2342) |
| **France** | FR | 7344 (7108) | 6722 (6679) | 6600 (6555) | 6176 (6142) | 5866 (5849) | 5610 (5594) | 5345 (5331) | 5345 (5268) |
| **Ireland** | IE | 4048 (4038) | 3584 (3569) | 3173 (3164) | 2945 (2935) | 2729 (2723) | 2378 (2372) | 1951 (1944) | 1760 (1757) |
| **Italy** | IT | 7115 (6915) | 7128 (7004) | 7132 (7026) | 6713 (6627) | 6571 (6478) | 6370 (6273) | 6052 (5989) | 5606 (5525) |
| **Greece** | GR | 5523 (5480) | 5220 (5173) | 4907 (4851) | 4604 (4543) | 4211 (4171) | 3986 (3952) | 3918 (3893) | 3916 (3895) |
| **Spain** | ES | 7206 (7142) | 6522 (6449) | 6267 (6133) | 5794 (5714) | 5485 (5439) | 5418 (5301) | 5132 (5048) | 4966 (4950) |
| **Portugal** | PT | 4881 (4787) | 4916 (4870) | 4849 (4807) | 4802 (4167) | 4716 (4666) | 4683 (4645) | 4633 (4606) | 4614 (4588) |
| **Austria** | AT | - (-) | 3380 (3367) | 3292 (3281) | 3142 (3130) | 2960 (2952) | 2815 (2809) | 2644 (2637) | 2544 (2535) |
| **Finland** | FI | - (-) | - (-) | 4139 (4138) | 4106 (4103) | 3920 (3917) | 3822 (3818) | 3104 (3101) | 3115 (3106) |
| **Sweden** | SE | - (-) | - (-) | - (-) | 5891 (5286) | 5807 (5208) | 5732 (5165) | 5734 (5116) | 5680 (5085) |
| **Germany** | DE | 6207 (6196) | 6336 (6329) | 6259 (6252) | 6163 (6156) | 5962 (5955) | 5847 (5845) | 5693 (5687) | 5563 (5559) |
| **Luxembourg** | LU | 1011 (1010) | 2978 (2976) | 2472 (2471) | 2654 (2651) | 2523 (2521) | 2552 (2551) | 2373 (2373) | 2428 (2428) |
| **United Kingdom** | UK | 5126 (5041) | 5032 (4999) | 5011 (4991 | 4965 (4958) | 4996 (4975) | 4951 (4935) | 4890 (4866) | 4819 (4779) |

## 4. ANALYSIS OF THE RESULTS

### 4.1. Inequality trends comparison among European countries.

The corresponding weighting schemes to compute the inequality synthetic indexes based on ACP are presented in Table 2 for each cross-sectional wave. They have been obtained from the aforementioned equation 3. We can appreciate that the weighting scheme is quite stable. That gives us the hint that it may be possible to consider that correlation structures are the same all over the period analyzed.

**Table 2**
*Weighting schemes for the computation of the cross-sectional synthetic inequality indexes based on the first Principal Component.*

| Inequality Index | Wave 1 1993 | Wave 2 1994 | Wave 3 1995 | Wave 4 1996 | Wave 5 1997 | Wave 6 1998 | Wave 7 1999 | Wave 8 2000 |
|---|---|---|---|---|---|---|---|---|
| **ATKIN05** | 0.348 | 0.316 | 0.317 | 0.311 | 0.395 | 0.313 | 0.327 | 0.314 |
| **ATKIN1** | 0.172 | 0.157 | 0.163 | 0.163 | 0.219 | 0.170 | 0.173 | 0.166 |
| **ATKIN2** | 0.034 | 0.021 | 0.033 | 0.020 | 0.032 | 0.028 | 0.023 | 0.026 |
| **CV2NORM** | 0.059 | 0.046 | 0.050 | 0.056 | 0.045 | 0.047 | 0.033 | 0.050 |
| **GINI** | 0.090 | 0.082 | 0.065 | 0.079 | 0.108 | 0.079 | 0.085 | 0.081 |
| **PIETRA** | 0.251 | 0.207 | 0.207 | 0.209 | 0.177 | 0.208 | 0.215 | 0.203 |
| **TH1NORM** | 0.046 | 0.171 | 0.166 | 0.162 | 0.025 | 0.155 | 0.146 | 0.160 |

In order to prove the validity of our intuition, we shall test the equality of the correlation matrices obtained from data matrix in each wave. Nevertheless, applying M-Box Test on standardized data, we can reject null hypothesis about correlation matrices equality (see Tables 3a and 3b). This fact leads us to take the second alternative presented in methodology section, and so we are going to compute the Common Space-based synthetic indicator.

**Table 3a**
*Box's M Test on equality of correlation matrices.*

| Wave | Rank | Log of determinant |
|---|---|---|
| **1993 correlation matrix** | 7 | -51.057 |
| **1994 correlation matrix** | 7 | -63.521 |
| **1995 correlation matrix** | 7 | -62.571 |
| **1996 correlation matrix** | 7 | -63.204 |
| **1997 correlation matrix** | 7 | -50.170 |
| **1998 correlation matrix** | 7 | -63.159 |
| **1999 correlation matrix** | 7 | -62.456 |
| **2000 correlation matrix** | 7 | -63.770 |
| **Pooled correlation matrix** | 7 | -50.761 |

**Table 3b**
*Results of M-Box Test.*

| | | |
|---|---|---|
| Box's M | | 994.925 |
| F | Approx. | 3.910 |
| | df1 | 196.000 |
| | df2 | 11998.570 |
| | Sig. | 0.000 |

Once Common Space Analysis procedure has been used, the following common space-based synthetic inequality indicator is obtained using the corresponding eigenvector. As it was described, this synthetic indicator can be expressed as a convex linear combination of the simple ones included in the selected initial set, and so it can be used to develop dynamic inequality analysis.

$$Z^*(r,t) = 0{,}3511\,\mathrm{ATKIN05}(r,t) + 0{,}1824\,\mathrm{ATKIN1}(r,t) + 0{,}0291\,\mathrm{ATKIN2}(r,t) + 0{,}0499\,\mathrm{CV2NORM}(r,t) + 0{,}0905\,\mathrm{GINI}(r,t) + 0{,}2195\,\mathrm{PIETRA}(r,t) + 0{,}0776\,\mathrm{TH1NORM}(r,t).$$

Table 4 shows the angles between the obtained eigenvector using the Common Space Analysis procedure and each one of the eigenvectors associated to the first component of correlation matrices.

**Table 4:** *Angles between common space eigenvector and each cross-sectional eigenvectors*

| YEAR | RADIANS | DEGREES |
|---|---|---|
| 1993 | 0.11 | 6.06 |
| 1994 | 0.06 | 3.46 |
| 1995 | 0.06 | 3.31 |
| 1996 | 0.08 | 4.53 |
| 1997 | 0.17 | 9.71 |
| 1998 | 0.04 | 2.55 |
| 1999 | 0.04 | 2.48 |
| 2000 | 0.04 | 2.19 |

It can be observed that all of these angles are quite small, with the greater one value around 0.17 radians (9.71º). So, we can admit the common-based indicator to be close enough to all the cross-sectional first component analysis indicators, as a result of the proposed method.

Keeping in mind the construction of this global synthetic indicator (equation 3), it is easy to note how its weighting scheme depends on the standard deviations associated to the simple indexes compounding the initial set. So, Table 5 shows the sample standard deviations of these simple indexes, using all the cases involved, with no temporal consideration.

It can be noticed how the smaller the standard deviation of the simple index, the greater its weight into the global synthetic indicator. In this sense, the smaller standard deviation is associated to the Atkinson index with 0,5 as its inequality aversion coefficient (ATKIN0.5) and its weight into the global synthetic indicator is found to be the greater. The second index, in decreasing order, is found out to be the Pietra (PIETRA) index and the third one is the Atkinson inequality index, with 1 as its aversion parameter (ATKIN1). However the greatest standard deviation corresponds to the Atkinson index with 2 as aversion inequality parameter (ATKIN2) and, consequently, it shows the smaller participation (2,9%) on the global synthetic indicator's system of coefficients.

**Table 5:** *Inequality indexes standard deviations* .

| **Inequality Indexes** | **Standard Deviation ($s_j$)** |
|---|---|
| ATKIN05 | 0.021614 |
| ATKIN1 | 0.041242 |
| ATKIN2 | 0.163351 |
| CV2NORM | 0.111622 |
| GINI | 0.072120 |
| PIETRA | 0.033036 |
| TH1NORM | 0.085359 |

Álvarez-García, Prieto-Rodríguez and Salas (2002) present a general overview of the income inequality results in European Union countries, during the convergence process to Monetary Union (from 1993 to 1996, they use data of the four first waves of the ECHP). These authors classify the thirteen countries present in at least three out of the four ECHP waves considered (excluding Finland and Sweden, included in ECHP from 1996 and 1997 waves, respectively). In their work, the following classification of countries into five different groups according to the income inequality is proposed. First of all, Denmark is the country where the lowest inequality rate was found during the first four waves. The second group was composed of The Netherlands, Germany, Austria and Luxembourg. United Kingdom, Ireland, Belgium, France, Italy and Spain constituted the third group, meanwhile Greece and Portugal were the fourth and fifth groups, remaining as the most unequal countries. We have extended this analysis to the last five available waves.

Using the Table 6 below, we can compare the inequality levels observed from the fourth wave on, when we account 15 countries in the EU. So, it can be observed how inequality has been reduced more in Ireland (10,3%), which is followed by France (9,1%), Greece (6,9%), Austria (3,4%), Germany (1,6%), The Netherlands (1,3%), Italy

(1%), Portugal (0,8%) and Spain (0,5%). Spain is the country where inequality have decreased in a lesser proportion through the last four waves.

Inequality has been increased in the remaining countries. Finland stands out as the country where inequality has increased in greater figures (27,6%). The other countries, in decreasing order, are Sweden (13%), United Kingdom (11,4%), Belgium (9,4%), Denmark (9,2%) and Luxembourg that stands as the country where inequality has increased in a smaller account, fixed at 3,1%.

**Table 6:** *Common Space Inequality Indicator values for each Country in the ECHP*

| Country | Wave 1 | Wave 2 | Wave 3 | Wave 4 | Wave 5 | Wave 6 | Wave 7 | Wave 8 |
|---|---|---|---|---|---|---|---|---|
| DK | 0.132901 | 0.128313 | 0.125970 | 0.114077 | 0.125760 | 0.126869 | 0.127744 | 0.124596 |
| NL | 0.150738 | 0.174021 | 0.181375 | 0.146635 | 0.201368 | 0.153902 | 0.127755 | 0.144622 |
| BE | 0.190569 | 0.196945 | 0.192756 | 0.165811 | 0.162294 | 0.198411 | 0.207977 | 0.181347 |
| FR | 0.262793 | 0.182273 | 0.171622 | 0.176947 | 0.166247 | 0.178956 | 0.164818 | 0.160830 |
| IE | 0.191325 | 0.213331 | 0.204622 | 0.206517 | 0.230148 | 0.211082 | 0.183910 | 0.185224 |
| IT | 0.166082 | 0.168599 | 0.161996 | 0.153029 | 0.157786 | 0.156741 | 0.150524 | 0.151630 |
| GR | 0.236627 | 0.217969 | 0.213825 | 0.223148 | 0.225085 | 0.221176 | 0.211200 | 0.207690 |
| ES | 0.213997 | 0.205234 | 0.205250 | 0.218104 | 0.211648 | 0.207297 | 0.213866 | 0.216983 |
| PT | 0.267736 | 0.250266 | 0.241610 | 0.238936 | 0.242250 | 0.232487 | 0.228884 | 0.236966 |
| AT | - | 0.162616 | 0.150716 | 0.145460 | 0.139465 | 0.162422 | 0.146218 | 0.140522 |
| FI | - | - | 0.111399 | 0.111382 | 0.123815 | 0.133562 | 0.132383 | 0.142129 |
| SE | - | - | - | 0.129389 | 0.143104 | 0.138616 | 0.146966 | 0.146170 |
| DE | 0.191823 | 0.187816 | 0.168395 | 0.153898 | 0.153176 | 0.148772 | 0.148161 | 0.151458 |
| LU | - | 0.133542 | 0.125557 | 0.129133 | 0.130527 | 0.142037 | 0.132762 | 0.133207 |
| UK | 0.197635 | 0.211177 | 0.190389 | 0.177415 | 0.250009 | 0.202524 | 0.205449 | 0.197614 |

Figure 1 displays the observed trends in inequality, measured through the global synthetic inequality index, over the 15 countries in the EU. An increasing inequality trend in Greece, Ireland and Italy is remarkable from wave 5 to wave 8. However, the opposite effect is observed in Finland and Luxembourg. Nevertheless, Denmark appears to have quite stable inequality figures, from wave 5 to the last of the period.

Furthermore, according to the temporal evolution of the common space-based inequality indicator, a classification method was used to analyse the group structure in data[14] from wave 4 to wave 8 (omission of the three first waves is necessary because

[14] The centroid agglomeration method of hierarchical clustering has been used over the squared euclidean distance dissimilarity matrix.

Austria, Finland an Sweden did not appear, thus not being comparable). The resulting dendrogram is shown in Figure 2.

**Figure 1**

*Common Space Inequality Indicator values for each Country in the ECHP.*

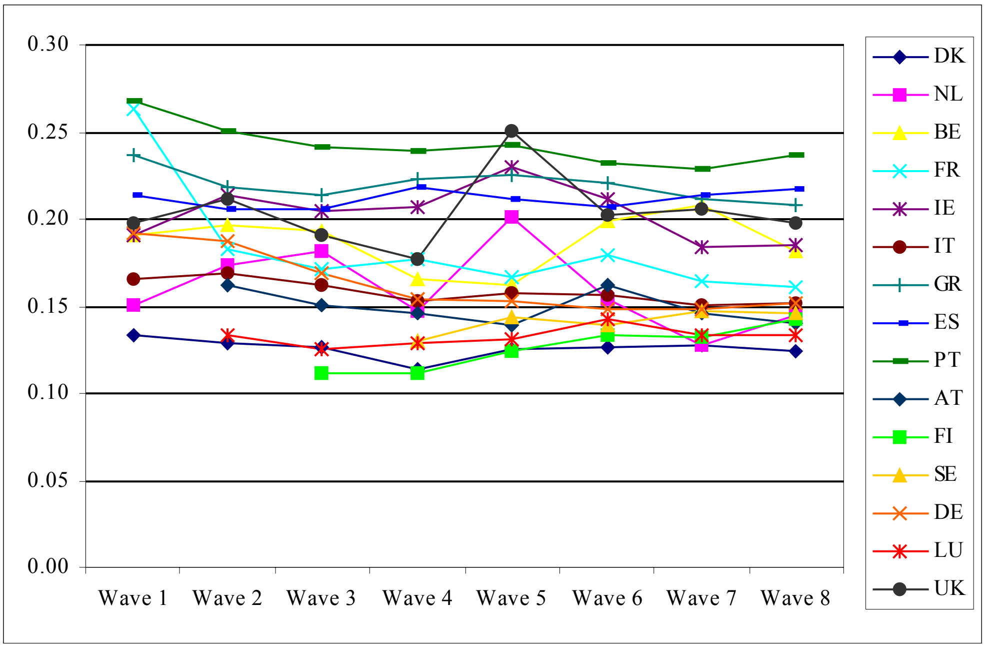
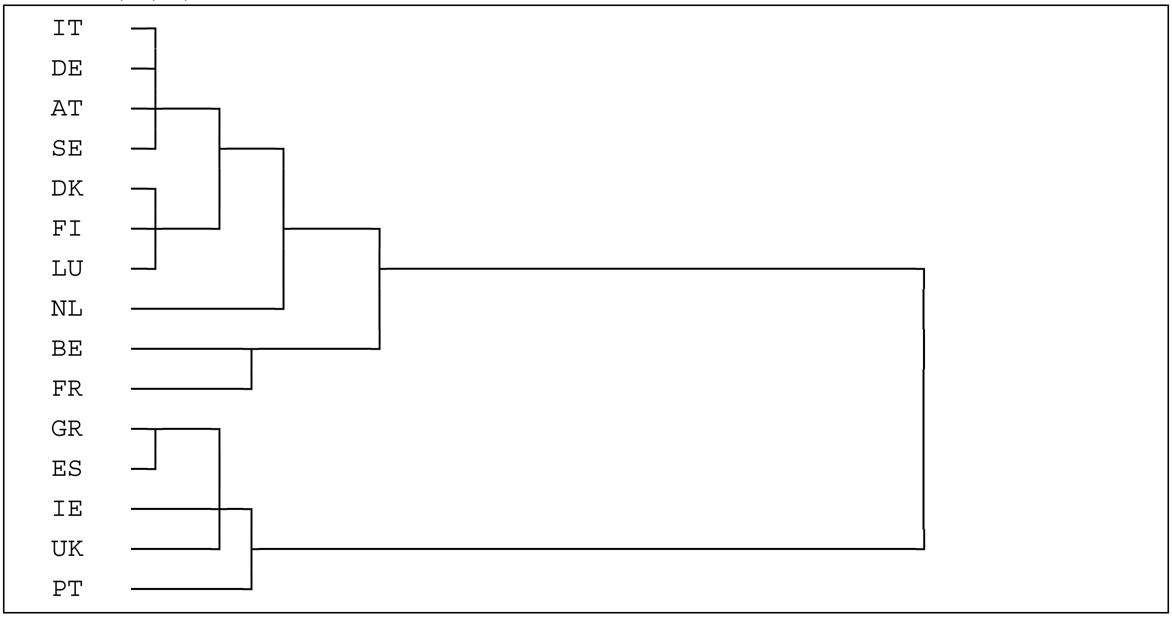



**Figure 2**

*Dendrogram of the countries' common space based inequality index referred to waves 4, 5, 6, 7 and 8.*



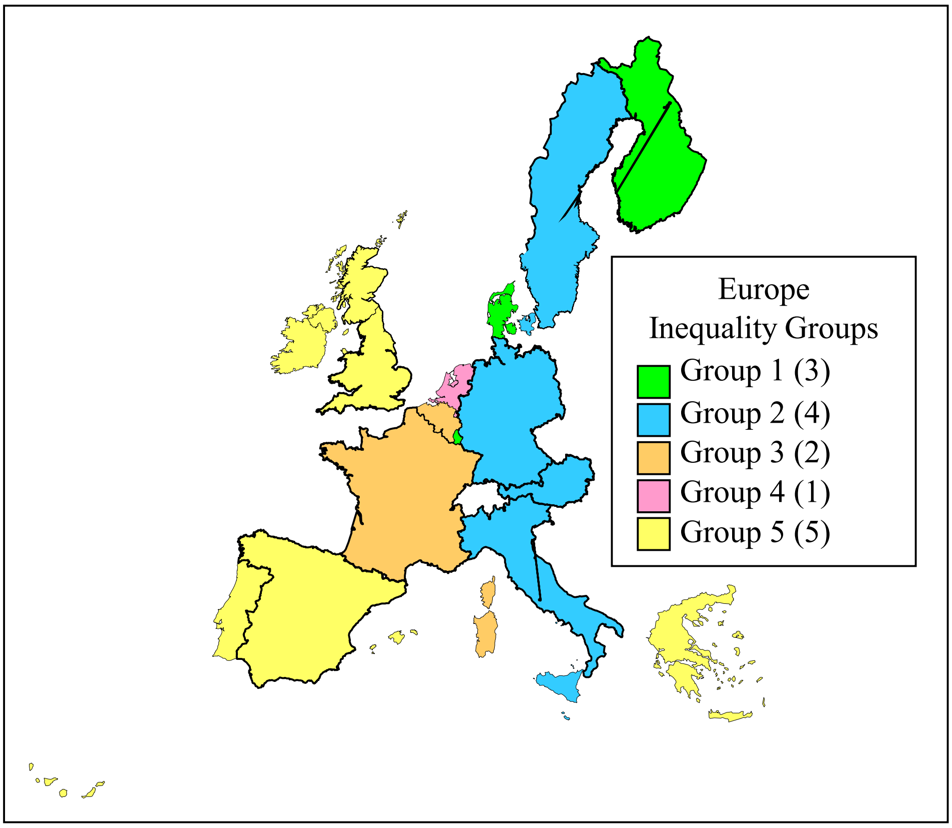

From Figure 2, we can find out the following groups:

- The first group includes Denmark, Finland, Luxembourg. These countries present the lowest inequality rates in the EU.
- The second group comprises Italy, Germany, Austria and Sweden.
- The third group is formed by, Belgium, France which are countries whose inequality is stable and located in the middle of the set of countries.
- The fourth group composed by The Netherlands.
- The fifth group, includes United Kingdom, Ireland, Greece, Spain and Portugal, and presents the greatest inequality indicator levels, thus their income distributions are the most unequal across the EU.

These results are similar to those observed in Figure 1. In Figure 3, the geographical situation of these groups is represented.

**Figure 3**

*Groups of countries derived from the classification according to their inequality level.*



## 5. CONCLUSIONS

In this paper, we have proposed a method to construct a synthetic inequality measure, based on a battery of seven one-dimensional inequality indicators, verifying a set of good properties. The main advantage of the exposed methodology is that we can evaluate inequality among countries, not only in the same period of time, but also in a longitudinal sense.

Using household income data provided by ECHP, from 1994 to 2001 waves (years 1993 to 2000), we have computed all one-dimensional inequality indicators used to elaborate the synthetic index proposed in the methodology. In this case, correlation matrices computed with the indicators in each wave have not turned out to be identical and so a Common Space Analysis-based synthetic indicator has been developed.

Attending to the empirical results, we have accomplished the classification of European countries into five groups according to their inequality degree and trend, measured using the synthetic indicator. The first group, and more equal, is composed by Finland, Luxembourg and Denmark. The second group is composed by Germany, The Netherlands, Austria and Italy. Belgium and France comprise the third group. The Netherlands make up the fourth group. The last group (Spain, Portugal, Greece, United Kingdom and Ireland) exhibits higher levels of inequality.

To sum up, we can extract some general ideas about our empirical findings. First of all, general trends in inequality show a slight convergence of all countries in EU. Second, Nordic countries exhibit a moderate increasing of their measured levels, but they keep themselves into the lower band, while Southern countries (Portugal, Spain and Greece) remain at the upper band. The rest of the countries are placed in the middle, but they present different trend patterns.

## REFERENCES:


**ACHDUT, L. (1996)** "Income inequality, income composition and macroeconomic trends: Israel, 1979-93". *Economica,* 63, 250, S1-S27.

**ÁLVAREZ-GARCÍA, S.; PRIETO-RODRÍGUEZ, J.; SALAS, R. (2002)** "The Evolution of Income Inequality in the European Union". *Papeles de Trabajo.*, 10/02, Instituto de Estudios Fiscales, Serie Economía. Madrid. [http://www.ief.es/Publicaciones/papelest/02/pt10_02.pdf]

**ATKINSON, A.B. (1970)** "On the measurement of inequality". *Journal of Economic Theory,* 2, 244-263.

**AYALA, L.; SASTRE, M. (2002)** "What determines Income Mobility differences across the European Union?", *Working Papers of the Institute for Social and Economic Research*, paper 2002-27. Colchester: University of Essex.

**BISHOP, J.A.; FORMBY, J.P.; SAKANO, R. (1995)** "Lorenz and stochastic-dominance comparisons of European income distributions". *Research on Economic Inequality,* 6, 77-92.

**BISHOP, J.A.; FORMBY, J.P.; SMITH, W.J. (1991)** "Lorenz dominance and welfare: Changes in the U.S. distribution of income, 1967-1986". *Review of Economics and Statistics,* 73, 134-139.

**BOURGUIGNON, F. (1979)** "Decomposable income inequality measures". *Econometrica,* 47, 901-920.

**BUHMANN, B.; RAINWATER, L.; SCHMAUS, G.; SMEEDING, T.M. (1988)** "Equivalence Scales, Well-Being, Inequality and Poverty: Sensitivity estimates across ten countries using the Luxembourg Income Study (LIS) Database". *Review of Income and Wealth*, 42, 381-399.

**CASAS, J.M.; DOMÍNGUEZ, J.; NÚÑEZ, J.J. (2003)** "La pobreza en España: Estudio a partir de curvas I.I.D. y su sensibilidad frente a escalas de equivalencia". In Casas, J.M. and Pulido, A. (Co-ords.) *Información Económica y Técnicas de Análisis en el Siglo XXI*. Ed. I.N.E., Madrid. Pages 161-173.

**COULTER, F.; COWELL, F.; JENKINS, S.P. (1992)** "Equivalence Scale relativities and the extent of Inequality and Poverty", *Economic Journal*, 102, 1067-1082.

**DAGUM, C. (1990)** "Relationship between income inequality measures and social welfare functions". *Journal of Econometrics*, 43, 1-2, 91-102.

**DAVIES, J.B.; HOY, M. (1994)** "The normative significance of using third-degree stochastic dominance in comparing income distributions". *Journal of Economic Theory,* 64(2), 520-530.

**DAVIES, J.B.; HOY, M. (1995)** "Making inequality comparisons when Lorenz curves intersect". *American Economic Review,* 85, 4, 980-986.

**DOMÍNGUEZ, J.; NÚÑEZ, J.J; RIVERA, L.F. (2004)** "Evolución de la desigualdad y la pobreza en España, mediante indicadores sintéticos, a partir de la estructura de covarianza de indicadores simples. El caso de las Comunidades Autónomas de Galicia y Castilla-La Mancha (1997-2000)" *VII Encuentro de Economía Aplicada.* Vigo, Spain.
[Paper available at http://www.revecap.com/viieea/autores/D/123.doc]

**EUROSTAT (2003)** "ECHP UDB Description of variables. Data Dictionary, Codebook and differences between Countries and Waves". DOC. PAN 166 / 2003-12. Eurostat, Luxembourg.
[http://forum.europa.eu.int/irc/DownLoad/k1eMAjJZmSGxH5NtD4GRcD2r6Rgc1SRI/scNqLj4I9If-YFXFO2jv/PAN166-200312.pdf]

**FOSTER, J.E. (1985)** "Inequality Measurement". In *Fair Allocation* (H.P. YOUNG, ed.), Proceedings of Symposia in Applied Mathematics, Vol. 33, Providence, American Mathematical Society, 31-68.

**FOSTER, J.E.; SEN, A. (1997)** *On Economics Inequality. Expanded Edition: Annexe.* Clarendon Paperbacks. Oxford.

**FRICK, J.R.; GRABKA, M.M. (2003)** "Imputed rent and income inequality: A decomposition analysis for Great Britain, West Germany and the U.S.". *Review of Income and Wealth,* 49, 4, 513-537.

**GARCÍA, C.; NÚÑEZ, J.J.; RIVERA, L.F.; ZAMORA, A.I. (2002)** "Análisis comparativo de la desigualdad a partir de una batería de indicadores. El caso de las Comunidades Autónomas españolas en el período 1973-1991". *Estudios de Economía Aplicada*, 20(I), 137-154.

**JENKINS, S.P. (1995)** "Accounting for inequality trends: Decomposition analyses for the UK, 1971-86". *Economica,* 62, 245, 29-63.

**KENDALL, M.; STUART, A. (1977)** *The Advanced Theory of Statistics.* Vol. 1, 4ª ed. C. Griffin. London.

**KERAMIDAS, E.M.; DEVLIN, S.S.; GNANADESIKAN, R. (1987)** "A graphical procedure for comparing the principal components of several covariance matrices". *Communications in Statistics, B. Computation and Simulation,* 16, 161-191.

**KRZANOWSKI, W.J. (1979)** "Between-groups comparison of principal components". *Journal of the American Statistical Association,* 74, 703-707. Correction note (1981), 76, 1022.

**KRZANOWSKI, W.J. (1982)** "Between-groups comparison of principal components – some sampling results". *Journal of Statistical Computation and Simulation,* 15, 141-154.

**LAYTE, R.; MAÎTRE, B.; NOLAN, B; WHELAN, C.T. (2000)** "Persistent and Consistent Poverty in the 1994 and 1995 Waves of the European Community Household Panel Study". *26th General Conference of the International Association for Research in Income and Wealth*. Cracow, 27th August to 2nd September 2000. Paper found on www.econ.nyu/dept/iariw

**LORENZ, M.O. (1905)** "Methods of measuring the concentration of wealth". *Journal of the American Statistical Association,* 9, 209-219.

**LOVELL, M.C. (1998)** "Inequality within and among nations". *Journal of Income Distribution,* 8(1), 5-44.

**MULIERE, P.; SCARSINI, M. (1989)** "A note on stochastic dominance and inequality measures". *Journal of Economic Theory,* 49(2), 314-323.

**NICOLETTI, C.; PERACCHI, F. (2002)** "A Cross-Country comparison of Survey Participation in the ECPH", *Working Papers of the Institute for Social and Economic Research*, paper 2002-32. Colchester: University of Essex.

**NÚÑEZ, J. J. (2002)** "Mayoración de rentas, curvas de Lorenz y medidas de desigualdad económica". *Documento de Trabajo en Estadística Económica,* DTEE 02/07. Dpto. Estadística, Estructura Económica y O.E.I. de la Universidad de Alcalá.

**NYGARD, F.; SANDSTROM, A. (1981)** *Measuring income inequality.* Amqvist & Wiksell International. Stockholm.

**PENA, J.B.(dir.); CALLEALTA, F.J.; CASAS, J.M.; MEREDIZ, A.; NÚÑEZ, J.J. (1996)** *Distribución Personal de la Renta en España*. Ed. Pirámide. Madrid.

**PEÑA, D. (2002)** *Análisis de datos multivariantes.* McGraw-Hill. Madrid.

**RENCHER, A.C. (1995)** *Methods of multivariate analysis.* Wiley Series in Probability and Mathematical Statistics. Wiley Interscience. New York.

**RUÍZ-CASTILLO, J. (1987)** *La medición de la pobreza y la desigualdad en España 1980-81*. Banco de España. Estudios Económicos, 42, Madrid.

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

**SHORROCKS, A. (1983)** "Ranking income distributions". *Economica,* 50, 3-18.

**SHORROCKS, A.; FOSTER, J.E. (1987)** "Transfer sensitive inequality measures". *Review of Economic Studies,* 54, 3, 485-497.