CENTRE D'ETUDES EN MACROECONOMIE ET FINANCE INTERNATIONALE
Université de Nice - Sophia Antipolis
http://www.unice.fr/CEMAFI

Is there a clearly identifiable distribution function of individual poverty scores ?

Valérie BERENGER* & Franck CELESTINI**

July 2004

Abstract :

The goal of this paper is to define a multidimensional poverty score for each household belonging to the same society. We will then be able to answer the following question: As for income, is it possible to characterize poverty by analysing the nature of the associated probability density function ?

The components of the poverty score are selected adopting the capability approach of Sen. Based on the "Totally Fuzzy and Relative" (TFR) approach, the method permits to obtain individual poverty scores lying in between 0 and infinity. We apply the method to the data from the 1986 and 1993-1994 French Surveys of Living Conditions. In both cases, the probability density function of the poverty scores follows an exponential distribution characterised by a single parameter. This result makes evident the self organisation of the society relative to this poverty score.

As actual poverty measures continue to be based on income distributions, we propose to examine the relationship between our multidimensional poverty score and income. The intuitive negative correlation is recovered, measured and fully analysed. In peculiar, we can use our poverty score distribution in order to estimate the poverty line that gives the best agreement between income-based and multidimensional definitions of poverty.

JEL Classification : D31, I32.

Keywords : POVERTY, MULTIDIMENSIONAL APPROACH TO POVERTY, INCOME DISTRIBUTION FUNCTIONS, FUZZY SETS, CAPABILITY APPROACH.

E-mail :berenger@unice.fr, franck.celestini@unice.fr

Paper presented at the "4th Conference on the Capability Approach. Enhancing Human Security", University of Pavia, Italy, 5-7 september 2004

Preliminary version. Comments are welcome.
Not to be quoted without the author's permission.

* CEMAFI (Centre d'Etudes en Macroéconomie et Finance Internationale), University of Nice-Sophia Antipolis
** LPMC (Laboratoire de Physique de la Matière Condensée), CNRS UMR 6622, University of Nice-Sophia Antipolis.

1.Introduction

Pareto (1897) was the first to state that income distribution is fitted by a power law whatever the country and time considered. Even if today the real nature of the income distribution is disputed, this work has opened a new research area consisting in identifying and analysing the income probability density function. The measure of poverty is not as direct as the one of income and several poverty indices have been proposed. Indices like the ones of Sen (1976), Foster Greer and Thorbecke (1984), and Watts (1968) use the individual income distribution to affect a poverty score to the overall society under consideration. There is no information about the poverty of each individual or household and the poverty distribution is not accessible. The goal of this paper is to build a poverty score for each household belonging to the same country and to examine the nature of the associated density probability function. In the second section, we justify the interest of analysing personal distribution of multidimensional poverty scores by referring to the studies of personal income distribution and to other emphasizing the definition and measurement of multidimensional poverty. In the third section, we present the method used to obtain multidimensional poverty scores based upon a recently fuzzy approach to the measurement of poverty. The section 4 analyses the results obtained from the application of the proposed method to the French data from the Survey of Living Conditions for the years 1986 and 1993. Finally, section 5 examines the relationship between our multidimensional poverty score and income. Concluding comments are then given in section 6.

2. From Pareto power law distribution of personal income to the need of identifying a functional distribution for multidimensional poverty scores.

Traditionally, poverty has been defined as a lack of income and has been associated with the study of personal income. Pareto was the first to analyse and estimate a model of personal income distribution. By demonstrating that the distribution of personal income was best fitted by a power law whose parameter could be interpreted as an index of inequality, Pareto advocated that only economic growth could achieve a reduction of income inequality.

Although, this finding has been largely disputed, it has opened a fruitful area of research consisting in analysing and identifying theoretical distribution functions associated with income. Certainly due to its normative implications, this topic of research has been left aside by economists for nearly half a century. The failure to obtain a distribution that can best fit the overall incomes explains the interest of defining and constructing inequality indexes like the Gini and Lorenz Curve.

The study of personal income distribution has however stimulated a good deal research among statisticians, mathematicians and probability theorists[1]. The former have proposed alternative models of income distribution. Among the most famous, Gibrat (1931) proposed the "law of proportionate effect" and showed that income was governed by a log-normal distribution. Recently, this topic has attracted the interest of econophysicians. Among the latter, Dragulescu A. and Yakovenko V.M. (2001) established that income would follow an exponential law for the great majority of the population in the UK and the individual states of the USA.

Beside these studies, the poverty concept has considerably evolved during the last three decades and has given rise to new definitions that underlined the multidimensional and vague nature of poverty on theoretical and empirical grounds.

The main critics of the traditional definition based on income and consumption data have been well documented. The most common pointed out the limitations of income as a proxy of welfare or its dual poverty and the arbitrariness inherent in the identification of the poor according to a poverty line defined in reference to the mean or to the median income of the society. According to the poverty income indexes derived from this approach, individuals whose income is near the threshold from below and above are respectively considered as poor and non poor although they do not experiment a significant difference in their standard of living.

New definitions have emerged since the seminal works of Townsend (1979) and Sen (1985). As is now well recognized, poverty is a multidimensional and vague phenomenon. Income is only one of these dimensions. In order to take into account various aspects of poverty, Townsend (1979) selected several indicators supposed to capture the general living conditions of individuals and to include non monetary aspects. The capability approach proposed by Sen provides a theoretical framework applicable to the context of poverty analysis. Following Sen, well-being can not be determined by the possession of goods but is defined as a combination of functionings. The three concepts of capability approach are commodities, functionings and capabilities. Commodities are goods and services. They can be viewed as instruments to reach the goal of increased well-being. Functionings are related to what an individual is able to do and to be with his commodities. They correspond to the

[1] For more details, see Dagum C. (1990/1999) and Kakwani N. (1980) b .

achievements of a person. The various “doings and beings” depend on personal, environmental and social characteristics. Capability, closely linked to functioning, is a set of vectors of functionings and includes the individual’s opportunities and freedom to lead a type of life than another one. Capabilities can be viewed as the abilities to achieve certain functionings. Poverty is then defined as a default of capabilities to convert commodities into functionings. The operationalization of this approach needs to take into account human dimensions of poverty and particularly to classify the indicators selected according to the distinction between the core concepts of the capability approach.

Multidimensional approaches have opened a new field of research in defining and searching adequate methodological tools for measurement using micro data from census or surveys. The fuzzy set theory appears as a useful tool to deal with the multivariate and the vague nature of poverty.

This methodology consists in extending the traditional definition of the membership function to a given set. Precisely, instead of partitioning the population between poor and non poor as it is the case in the traditional money metric measure of poverty, fuzzy set theory has the advantage of taking into account a continuum of situations between these two extremes. Based on the selection of a vector of indicators covering all relevant areas of living conditions of the households of a given society, fuzzy set permits to obtain for each household a poverty score. The latter represents the degree of deprivation or the degree of his membership to the subset of poor. It leads then to the determination of a poverty index for the whole population as generalisation of the head count ratio.

This methodology was first applied by Cerioli and Zani (1990). Their work has been followed by others with some new theoretical aspects. Cheli and Lemmi (1995) offer notable contributions. They propose a “totally fuzzy and relative” procedure to the measurement of deprivation using the distribution functions of the variables accounted in the multidimensional definition of poverty, instead of comparing the situation of individuals to absolute norms as did Cerioli and Zani. The contribution of Chiappero Martinetti (2000) constitutes a reference as she used fuzzy set with an explicit association with Sen’s approach.

Surprisingly, as for income, no attempt has been made in the context of such multivariate analysis of poverty to characterize the distribution of poverty scores by an appropriate theoretical density function. Such an attempt would permit to use graphic devices in order to have a more revealing picture of the degree of poverty relative to each household and its distribution. Such a device would allow us to detect possible laws from data and to obtain some information on the organisation of poverty in a given society. The study of functional distribution of multidimensional poverty scores could have useful applications in poverty comparisons across times, regions and countries.

In the following section, we propose to define a method based on the TFR approach that allows the analysis of the distribution of poverty scores.

3. Derivation of individual multidimensional poverty scores

In order to study the distribution of multidimensional poverty scores, it is necessary to propose a method to define poverty scores lying, as for income, between 0 and infinity.

We consider i ? [1, N] households. For each household, we select j ? [1, v] indicators about ownership of durable goods, the housing basic characteristics, the quality of housing and the ownership of assets. The deprivation according to a given indicator may be interpreted as a symptom of poverty. According to the capability approach, these indicators relate to commodities and their characteristics and may be viewed a priori as reflecting the specific use of the possession of income that a person can make. The full list of the v = 17 questions considered in this study is given in the appendix.

Considering the set of j indicators and $a_j^{(l)}$, with l = 1…m, the possible values of modalities taken by j, we can define the function $a_j(i)$ as the value attributed to the answer of the household i to the question j. For each indicator or question, the values associated with each modality are chosen in order that the higher degree of poverty of the household i regarding the answer at question j is, the higher the value of $a_j(i)$. We choose $a_j(i)$ to be a positive integer ($a_j(i)$? 0), the value $a_j(i) = 0$ hence corresponds to the lowest risk of poverty.

As an example, suppose that question j is: “Do you have a colour television, a black and white television or no television at all?” Depending on the answer, $a_j^{(l)}$ can take the value 0, 1 and 2 (0 for a colour, 1 for a black and white and 2 for no television).

Once recorded the answers of the N households, we then compute for each indicator the normalised probability density functions $p_j(a_j)$ (1 ? j ? v).

In the TFR approach, Cheli and Lemmi proposed that the degree of poverty associated to the indicator (or question) j should directly be proportional to the cumulative distribution function :

$$P_j(a_j) = \int_0^{a_j} p_j(x)\,dx \qquad (1)$$

This realistic assumption is based on the fact that the feeling of poverty of a household is directly related to the number of households owning a good that it does not own by itself. In other words, this approach stresses and takes into account the relative nature of the poverty feeling. Within the TFR approach, the degree of poverty or equivalently the value of the membership function to the subset of poor, $s_j(i)$ of the household i as regarded to indicator j satisfies the following specification :

$$s_j(i) = \begin{cases} 0 & \text{if } a_j(i) = a_j^{(1)} \\ s_j(a_j^{(l-1)}) + \dfrac{P_j(a_j^{(l)}) - P_j(a_j^{(l-1)})}{1 - P_j(a_j^{(1)})} & \text{if } a_j(i) = a_j^{(l)},\ l = 2\ldots m, \end{cases} \quad (2)$$

where $s_j(a_j^{(l-1)})$ represents the degree of deprivation attached to the modality (l-1) for indicator j. With this formulation, the degree of deprivation $s_j(i)$ lies in the interval [0, 1] and increases with risk of poverty. However, as is the case for income or wealth, we need to introduce a poverty score that is not limited but that naturally lies between 0 and infinity. For this purpose, we propose the new following definition of $s_j(i)$:

$$s_j(i) = \ln\left(\frac{1}{1 - P_j(a_j(i))}\right), \quad (3)$$

where $s_j(i)$ is still an increasing function of $P_j(a_j)$ but is no longer restricted to the interval [0,1]. Even if numerous alternative expressions can be proposed, this one seems to be the simplest and introduces a logarithmic function often present in the measure of different human sensitivities[2].

Finally, the degrees of poverty assessed according to each of the v deprivation indicators need to be reduced to one dimension in order to obtain the multidimensional score of poverty s(i) of each household i.

[2] In social sciences, logarithmic functions are common for example in the information theory. Beside they also describe the sound intensity felt by the hears or the different colours by the eyes.

The degree of poverty s(i) is then defined as the weighted average with respect to the v indicators :

$$s(i) = \sum_{j=1}^{v} w_j s_j(i), \tag{4}$$

where w_j is the weight attached to the j-th indicator.

The weight w_j is chosen to be an inverse function of the average degree of poverty : $\bar{s}_j = \frac{1}{N} \sum s_j(i)$.

This means that an important weight is given to a variable j associated to a very widespread good in the society. In other words, the more a good is owned the poorer is the household that does not own this good. The idea is of the same type than the one used above for justifying the proportionality between $s_j(i)$ and the cumulative distribution $P_j(a_j)$. The expression for the weights w_j is proposed to be :

$$w_j = \frac{\ln\left(1 + \frac{1}{\bar{s}_j}\right)}{\sum_j \ln\left(1 + \frac{1}{\bar{s}_j}\right)} \tag{5}$$

This expression satisfies the inverse relation between the weight and the mean score and uses another logarithmic function. The form $\ln\left(1 + \frac{1}{\bar{s}_j}\right)$ is chosen to prevent the occurrence of negative weights. Indeed, in contrast to the classical TFR approach, scores can be greater than 1. The denominator ensures the normalisation of the weights avoiding a trivial dependence of s(i) with v. Starting from the $a_j(i)$ functions, a multidimensional poverty score s(i) is evaluated for each of the N households.

4. The case of France for the years 1986 and 1993

We apply the method to the case of France for the years 1986 and 1993. We used two different databases both coming from an INSEE Survey of Living Conditions and distributed by the LASMAS-CNRS. The size of the household sample is N = 13154 for the year 1986 and N = 13280 for the year 1993. The same indicators are selected for the two samples and the full list is given in appendix. Using the method described above, we compute the poverty scores and the associated probability density function $\rho(s)$. In order to identity the nature of $\rho(s)$, we use the "rank ordering method[3]". The N poverty scores are reordered by decreasing values: s_1 and s_N are respectively the largest and the smallest values. The rank ordering method consists in identifying the relation between the n-th larger value s_n and its rank n.

In figure 1, we represent the s_n values as a function of their rank n for the years 1986 and 1993. We know[4] that for an exponential probability density function of the form :

$$\rho(s) = \frac{1}{\sigma_s} \exp\left(-s/\sigma_s\right) \quad (6)$$

The n-th larger value of s satisfies :

$$s_n = -\sigma_s \ln\left(\frac{n}{N}\right) \quad (7)$$

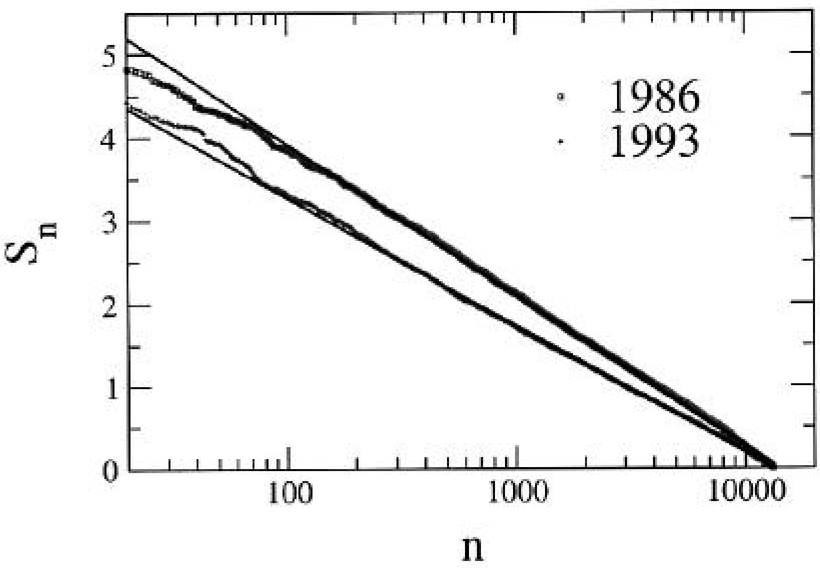
Using a semi-logarithmic representation, the data points are fitted by two different straight lines for the two different years. Except for roughly the 100-th largest scores, the agreement between the computed scores and the fitting functions is good.

---

[3] For more details, see chapter 6 in Sornette D. (2000), *Critical Phenomena in Natural Sciences*, Springer-Verlag, Berlin, p.137-150.

[4] See in Sornette D. (2000), p.141-43.

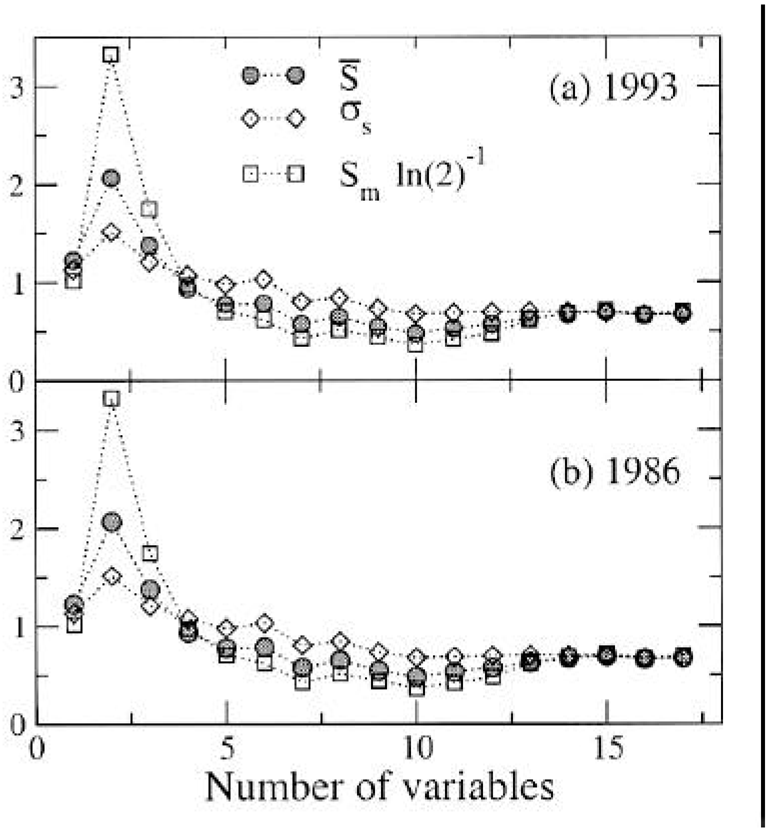
Figure 1: *Rank ordering of the multidimensional poverty scores with logarithmic horizontal scale for the years 1986 and 1993*



Source : Authors' calculations from *French Surveys of Living Conditions,* INSEE (1986, 1993-1994)
Note *: Straight lines are best fits to the exponential distribution.*

More than 99 % of the household scores follow the exponential distribution. This agreement demonstrates the relevance of our computed scores because they are distributed accorded to a well-defined distribution law which outlines the way the society organises itself relatively to this poverty score. The exponential distribution is a function that it is characterised by a single parameter ? $_s$. Its mean is equal to its standard deviation and is also proportional to its median s_m : $\bar{s}$? ? $_s$? $s_m / \ln ?2?$ In figure 2, we plot these three parameters as a function of the number of indicators v considered for each household i.e. as a function of the number of questions v. The distribution is computed in the same way than above, except that we restrict our analysis to the v-th first questions. As it is shown in figure 2, the three parameters are rather different when a few numbers of indicators are considered. As v increases, they take closer and closer values. Finally for a number of variables roughly greater than 13, $\bar{s}$, ? $_s$ and $s_m / \ln ?2?$ reach a constant and almost equal value. The same comment stands for the other year considered.

Figure 2: $\bar{s}$, $?_s$ and $s_m / \ln ?2?$ *as a function of the number of indicators defining multidimensional poverty scores*

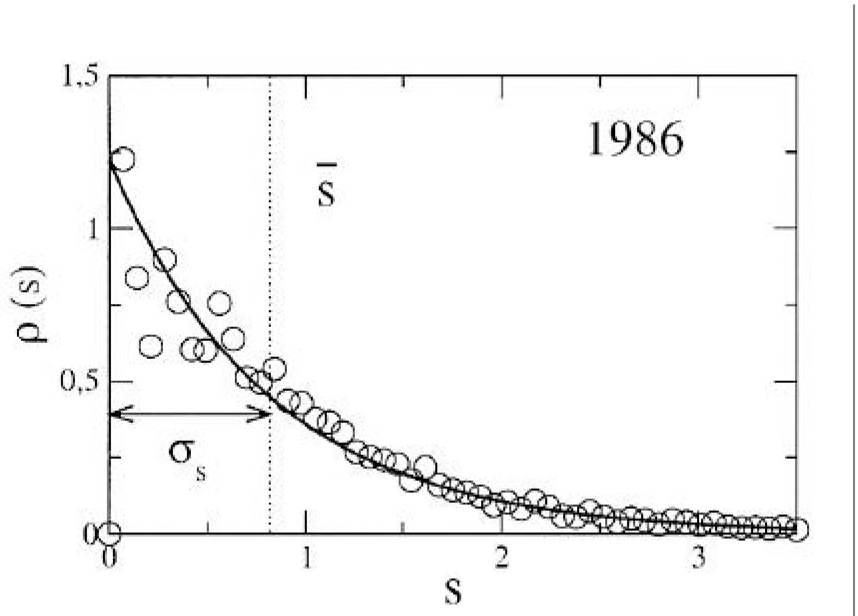


Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

This plot confirms the hypothesis of an exponential distribution of poverty and also demonstrates that a sufficient number of indicators need to be taken into account in order to reach this distribution. For a weak number of attributes, the statistic is too poor to clearly identify the exponential distribution. This result should be recovered in other multidimensional measures of poverty.

The main property of this organisation type is certainly that we fully characterised the poverty of the sample society by a single parameter. Indeed, as stressed above, the exponential distribution is defined by the unique value of $?_s$. This is illustrated in figure 3 where is plotted the exponential distribution function of the poverty scores for the year 1986.

Figure 3: *Probability density function of poverty scores for the year 1986.*



Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986)

Note *: Circles correspond to data from the French Survey and the full line is the best fit to an exponential distribution.*

All other well-known distributions (Gaussian, log-normal, power law, etc) have at least two indicators characteristic of the different moments. Thus, it is important to note how our multidimensional poverty scores lead to one of the simplest probability density function. In this case, the knowledge of ?$_s$ is sufficient to obtain all other possible indicators from the distribution density function. As a consequence, the comparison of poverty across societies or across a chosen period is straightforward. One just has to compare the different values obtained for ?$_s$. This comparison or variation study is usually not so easy because the underlying distribution density is not as simple as the exponential. In these cases, methods like the stochastic dominance are required to test if the same ranking of the distribution is obtained whatever the poverty line and the index of poverty chosen.

For the situation of France, the best fit to the household poverty scores for the two years considered gives the following values for ?$_s$: ?$_s$ = 0.819 for 1986 and ?$_s$ = 0.672 for 1993 which means a relative decrease of the mean poverty score of 17.9 %. According to the properties of exponential distribution function, the same relative variation stands for the standard deviation and higher moments.

This result has important implications for the political choices relative to poverty. It shows that one can not decide for example to reduce the mean poverty while increasing standard deviation of poverty. The knowledge of this exponential organisation of poverty can be used to predict how a global change in the poverty situation is going to be felt by the different households composing the society.

Let us assume the following variation of the mean poverty: at time t the poverty distribution is characterised by an exponential distribution with a mean value $?_{st}$ while at time t + dt the mean value decreases to $?_{st} + d?_{st}$[5]. The relative poverty decrease is then defined as $r = d?_{st}/?_{st}$. According to this finding, when analysing the effect of a global policy without considering any individual increase or decrease in the poverty rank between t and t + dt, we can easily see that the relative individual poverty score variation is the same for each household and is equal to the mean relative variation (r).

This result is directly connected to the exponential nature of the distribution. Conversely, considering for example a power law distribution, a mean variation would not be equally distributed among each individual or household. In that sense, the exponential distribution can be viewed as an egalitarian probability density distribution.

5. Statistical comparison between income-based measure of poverty and multidimensional poverty score

As the multidimensional score of poverty is defined as a function of the main relevant indicators of living conditions of households, it is intended to give a different picture of poverty than the one based on income as the single indicator. Nevertheless, as the indicators selected in this analysis may be considered as indicators of possible effect of symptom of insufficient economic resources, a certain similarity between the two approaches is also possible. A comparative analysis is then required in order to validate one of these two possibilities.

Therefore, we propose to investigate the relationship between our multidimensional measure of poverty and the ones based on the income distribution in order to see if the two different approaches indicate two different subsets of poor households.

From the same sample survey used to establish the poverty scores, we extract the income I(i) of the i-th household. It is the total income coming from the sum of all the individuals composing the household corrected using OECD equivalence scale. For the

[5] Note that a poverty decrease corresponds to a negative value of $d?_{st}$.

largest incomes, the probability density function ?(I) of incomes roughly follows a Pareto law :

$$? ?I? ?\; I^{? ?? ?1?}, \qquad (8)$$

where the exponent ? is often named Pareto index. For the intermediary and lowest incomes, no distinct law can be identified. This is certainly due to the fact that we consider the total income of the household and not the distribution of individual incomes. In a recent study, Dragulescu and Yakovenko (2000) pointed out that an exponential distribution of income can be found in the USA using individual rather than household incomes. We nevertheless decide to take into account the total household income because our poverty scores have been established for households. The correlation between the two approaches should then be more pronounced when considering the household rather the individual income.

At the i-th household is associated the poverty score s(i) defined in the previous section and the income I(i). We calculate the global correlation between these two random variables :

$$C?s, I?? \frac{\mathrm{cov}?s, I?}{?_s ?_I}, \qquad (9)$$

with $?_s$, $?_I$ and cov(s,I) denoting respectively the standard deviations and covariance of s and I. For the two years considered, table 1 gives the mean, the standard deviation and the Pareto exponent for the income distribution together with the correlation coefficient just defined above and the rank coefficient correlation described below.

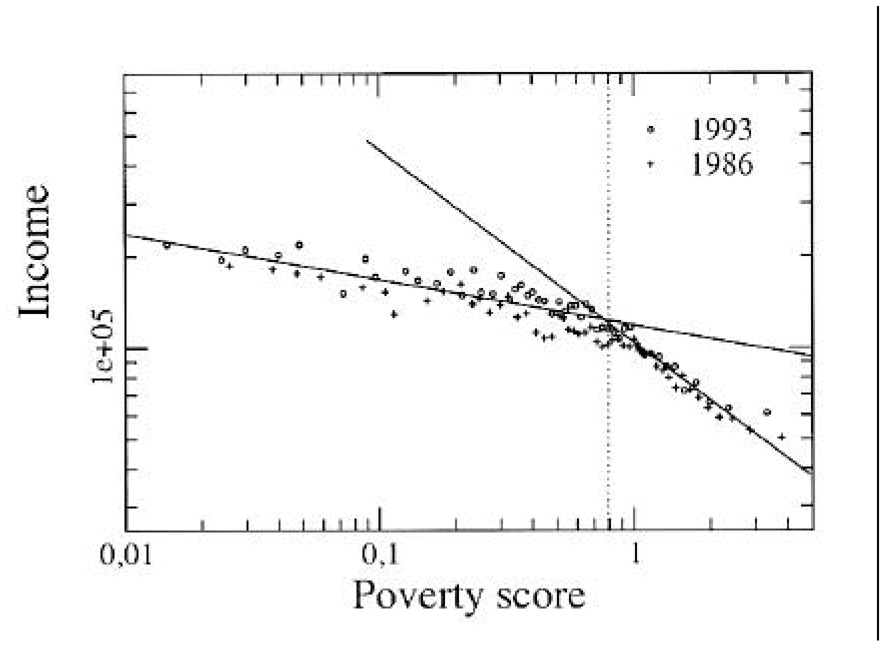
Table 1: *Parameter of the income distribution and correlation coefficients.*

Year	**1986**	**1993**
Mean of I	114267	137373
Mean of s(i)	88047	111006
? ?	3.04	2.82
C(s, I)	-0.34	-0.32
C(R_s, R_I)	0.44	0.41

Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

The values of the correlation coefficients confirm the intuitive negative correlation between income and poverty: the lower the income is the greater the poverty is. To further quantify this relationship, we represent in figure 4 the I(i) versus s(i) plot for the two years[6].

Figure 4: *Relationship between income and poverty scores*



Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

Note *: In the log-log representation, the lines are best fits to power laws for the large and low income regions.*

We may observe two regimes: for large incomes and low poverty scores, the correlation is clearly smaller than in the region of low incomes and large poverty scores. Income and poverty are more correlated when households have low incomes, and above a certain level of income the correlation tends to disappear. The log-log representation of figure 4 permits to estimate a mean power law relation between the income and the poverty scores for the two regions :

$$I ?i?? \; s ?i?^{? ?} , \qquad (10)$$

---

[6] For clarity, we do not represent the overall n points associated to the N households but subset averages.

where ? is the exponent characterising the relationship. The best fit to the data gives ? ? 0.1 in the large income region and ? ? 0.5 in the low income region. The relation between income and poverty score is weaker for larger incomes. The influence of the income on the poverty status is reasonably shown to be more important for the low incomes. For the low incomes, we find that the income roughly scales as the inverse square-root of the poverty score. In the same manner that there is a region where the correlation is almost zero because the income is large, one can wonder about the existence of a third region where the correlation could disappear because of the large amount of poverty. Unfortunately, homeless households are not considered in the survey so that the current data are not appropriate for such a study of extreme poverty.

The comparison is now carried out by measuring the rank correlation of income and poverty variables. We denote by $R_s(i)$ the rank of the poverty score of the household i and $R_I(i)$ the reverse rank of income of the same household[7]. The Bravais-Pearson correlation coefficient $C(R_s, R_I)$ often used in the analysis of rank correlation is defined as :

$$C(R_s, R_I) ? \frac{\mathrm{cov}(R_s, R_I)}{?_{R_S} ?_{R_I}} \qquad (11)$$

The values obtained for the two years are listed in table 1.

The positive Bravais-Pearson correlation coefficients confirm the negative correlation between the two variables s(i) and I(i). Their values are close to the ones obtained in a study of the relationship between unidimensional and multidimensional approaches to the measurement of poverty in Europe (Costa M. (2003)). No relevant difference is found for the variations measured for the two different years.

The principal poverty indexes are based on income distribution analysis. Several indicators can be obtained, the main one being the proportion of poor p or the headcount ratio. The number of poor (p) is identified by determining arbitrarily a poverty line, i.e. an income value I_l. Households having income less than I_l are then considered as poor. This poverty line is often expressed as a ratio of the median of the income distribution: I_l = ? I*. The choice of the poverty line is empirical, for example the value ? = 0.6 is frequently used for European poverty studies.

[7] Poverty scores and incomes are ranked in reverse order because of the negative correlation just discussed above.

We can nevertheless think that there is an optimal value of ? that should give a measure of poverty more closely linked to the multidimensional one.

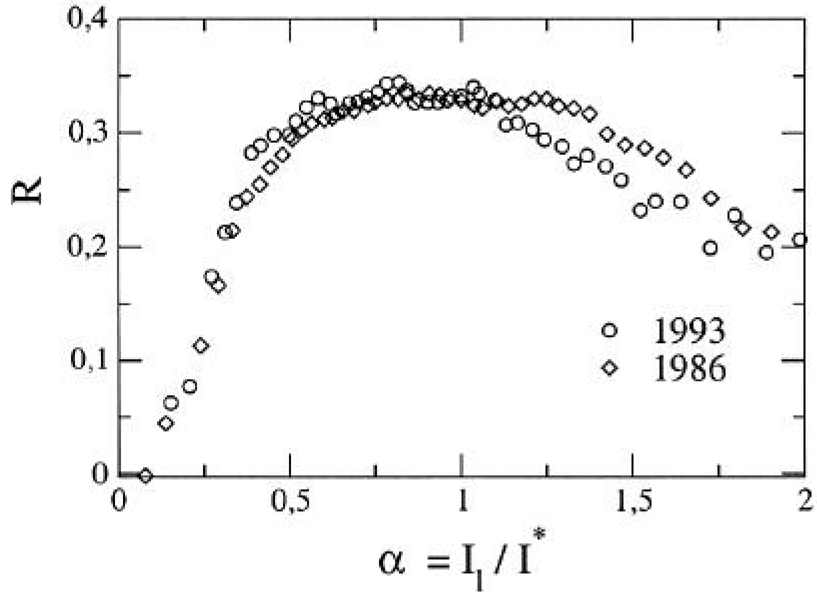
In the following, we use our poverty score distribution in order to study the influence of the poverty line on the measure of the different income based poverty indexes.

We fix a value of ? that selects a first subset of households considered as poor because of their lowest incomes. The number of households composing this subset is N_p = N * p, p being the headcount ratio. We then build a second subset of households of the same size N_p. This time, households are selected as a function of their poverty scores : the households with the N_p-th largest poverty scores belong to this second subset. Our goal is to see if there is an optimal value of ? that gives the best agreement between the two subsets. We can easily count the number of households N_c that belong to the two subsets, but we can not directly look at this quantity to obtain conclusions on the correlation between the two subsets. Indeed, N_c trivially increases as the size of the subsets increases. The relevant quantity R measuring agreement between the two subsets is then defined as :

$$R \,?\, \frac{N_c \,?\, p^2 N}{N_p \,?\, p^2 N} \qquad (13)$$

In this expression, p^2N is the number of common households that one would find if the two subsets were completely uncorrelated. This is the case if households were randomly distributed among the two subsets. N_p is the number of common households if the two subsets are exactly the same, N * p = N_p is the size of the two subsets. Hence R varies between 0 and 1, the larger the value of R is, the better the agreement between the two subsets and the more significant the poverty measurements are. In figure 5, we represent the values obtained for R as a function of ? .

Figure 5: *Value of R as a function of $\alpha = I_l / I^*$ for the two years considered.*



Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

For the two years considered, we can see that the correlation is optimal for values α in between roughly 0.6 and 1.2.

This means that the poverty line has to be chosen in between these two values in order to obtain an optimal measurement of the poverty headcount ratio. However, the correlation coefficients remain quite low indicating that the multidimensional poverty score captures other aspects of poverty.

Other indexes are calculated to measure poverty from the income distribution. We focus on indicators that are the most frequently used in the current poverty studies. The Watts index (1968) is defined as :

$$W = \frac{1}{N} \sum_{i=1}^{q} \ln \left(\frac{I_l}{I_i} \right), \qquad (14)$$

where the sum runs over the q-th households having an income less than the poverty line I_l. Foster, Greer and Thorbecke (1984) have defined a class of poverty indexes given by :

$$P_{\alpha} = \frac{1}{N}\sum_{i=1}^{q}\left(\frac{I_l - I(i)}{I_l}\right)^{\alpha} , \qquad (15)$$

where α is a parameter which puts more weight on the standard of living of the poorer of the poor for values above 1.

For $\alpha = 0$, we obtain the poverty headcount ratio (P_0) while the indices defined for $\alpha = 1$ and $\alpha = 2$ are respectively the poverty gap (P_1) and the poverty severity (P_2) known as the squared poverty gap. These two indexes capture respectively the degree to which mean income of the poor differs from I_l and differences in income levels among the poor. Sen (1976) has introduced an index taken into account the inequality within the poor subset. It is given by :

$$S = \frac{2}{(q+1) N I_l}\sum_{i=1}^{q}\left(I_l - I(i)\right)(q+1-i) \qquad (16)$$

We finally also consider the Gini index among the poor expressed as :

$$G = \frac{1}{2 I_p N_p (N_p - 1)}\sum_{i=1}^{q}\sum_{j=1}^{q}\left| I(i) - I(j)\right|, \qquad (17)$$

where I_p is the mean income of the subset of poor households.

We compute the values of these five indices for the two years considered together with their relative variation ?. The data are presented in table 2 for a poverty line chosen to be $I_l = 0.6\ I^*$.

Table 2: *Values of different poverty indices and their relative variations.*

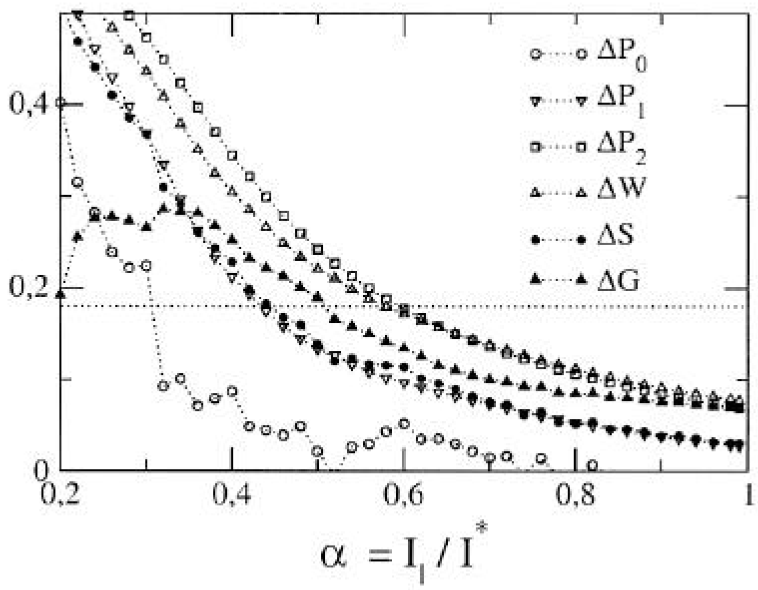
Index	**W**	**P_0**	**P_1**	**P_2**	**S**	**G**
1986	0.136	0.24	0.085	0.045	0.214	0.118
1993	0.113	0.227	0.077	0.037	0.185	0.105
? ?	0.172	0.051	0.096	0.177	0.134	0.113

Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

In the previous section, we found a value of 0.179 for the relative variation of the single parameter ?$_s$ characterising our poverty score distributions. This value is close to the ones found for the Watts and P_2 = FGT(2) indicators. Insofar as the Sen and Gini indexes are concerned, their relative variation is weaker than the one found for?? $_s$ but of the same order. However, as we may observe, the relative variation of the headcount ratio, ? P_0 is weak as compared to the other indicators. For the case of France, for the two years considered, the relative variation of poverty found through our model seems to be reliable when compared to the other poverty indicators. The headcount ratio seems to be a qualitatively different measure of poverty in the sense that its variations are rather different than the variation of other indexes.

We finally study how the variation of these indexes varies with the value of the poverty line I_l = ? I*. We represent in figure 6 these variations as a function of ? .

Figure 6: *Relative variation of the income based poverty indexes as a function of ? .*



Source : Authors' calculations from *French Surveys of Living Conditions*, INSEE (1986, 1993-1994)

Note *:The horizontal dotted line gives the value obtained for the relative variation of ? $_s$ upon the same period.*

The influence of the poverty line on the overall indexes values is weaker for the larger values of ? . As it has been demonstrated above, values of ? greater than roughly 0.6 give the best agreement with our poverty scores and seem to be more appropriate to obtain a reliable measure of poverty.

6. Conclusion

In this paper, we have tried to transpose analysis of personal income distribution to a multivariate measurement of poverty. We have dealt with the possibility of extracting a law from multidimensional scores of poverty analogous to the power law identified by Pareto from income data. We then proposed a method based on fuzzy set approach in order to define a poverty score lying between 0 and infinity. Using data from French Surveys of Living Conditions, we found that the multidimensional poverty score is distributed according to an exponential law for the quasi-whole population for the two years considered. The main property of this organisation type is that we can fully characterise the poverty of a sample society by a single parameter. Indeed, the exponential distribution is defined by a unique value of $?_s$. As a consequence, the evolution of poverty upon a chosen period is straightforward. One just has to compare the different values obtained for $?_s$. For the situation of France, we have found a relative decrease in the mean poverty score of 17.9%.

The relationship between income-based measure of poverty and the multidimensional poverty score has been carried out using rank correlation analysis. The estimated correlation indices are quite low indicating that the multidimensional score brings different information on poverty than income. As the correlation between the two approaches is sensitive to the value of the poverty line, we have estimated the range of values that give the best agreement. A value corresponding to 60% of median income seems to be the most appropriate to obtain a reliable measure of poverty. In this case, except for the headcount ratio, the relative variations of the most common indexes of poverty using income data at the two different points of time considered are of the same order as the ones of the multidimensional score of poverty.

However, the promising results obtained suggest a need for further analysis on two levels.

The first one is conceptual and concerns the selection of indicators defining the multidimensional poverty score. The method proposed has only been performed using information on the possession of durable goods and housing conditions, i.e, commodities. We are aware of their rather limited significance in order to assess individual well-being according to the capability approach. The method could easily be used to operationalize Sen’s concepts. It could be applied on different vectors of functionings ranked by domain as health, education, ability to enjoy leisure and the extent of social relations, etc. The identification of particular distribution functions relative to a set of indicators may provide useful and efficient information for the implementation of socioeconomic policies to deal with the eradication of poverty.

The second level of analysis is empirical. It suggests we examine the validity of the exponential law of the multidimensional poverty score for the whole population in all countries and at all times. As is the case for income, the application of the method proposed using data from different countries may lead to other distributions of poverty score or display possible breakdown in the distribution limiting the validity of the exponential law to a certain range of poverty scores. The heterogeneity in the organisation of poverty in a given society or across countries may reveal different subsets of population according to sociocultural and ethnics factors. A study that we are carrying out at present using data from another country seems to validate the method proposed and the conjectures evoked above.

REFERENCES :

BERENGER V. et SILBER J. (2003) “La mesure du niveau de vie et de la qualité de vie en France », dans *Mondialisation et Régulation Sociale*, Euzeby C et alii, Paris : l’Harmattan, p.191-206.

BOETTCHER H. (1994), « The Use of Fuzzy Set Techniques in the Context of Welfare Decisions », in *Models and Measurement of Welfare and Inequality*, ed. by Eichhorn W., (ed.), Springer-Verlag. Heidelberg, p.891-899.

CERIOLI A and ZANI S. (1990), “A Fuzzy Approach to the Measurement of Poverty“, in *Income and Wealth Distribution, Inequality and Poverty*, ed. by Dagum C. and Zenga M., Springer-Verlag, Heidelberg, p.272-284.

CHELI B. and LEMMI A. (1995), “A totally Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty”, *Economic Notes by Monte dei Paschi di Siena*, vol.24, (1), p.115-134.

CHELI B., GULLINI G., LEMMI A and PANNUZZI N. (1994), “Measuring Poverty in Transition via TFR Method : the Case of Poland in 1990-1991”, *Statistics in Transition*, vo1,(5), p.585-636.

CHIAPPERO MARTINETTI E. (2000), “A multidimensional Assessment of Well-Being Based on Sen’s Functioning Approach”, *Società Italiana di Economia Publica Working Paper*, Rivista Internazionale di Scienze Sociali, n°2, 2000.

COSTA M. (2003), “A comparison between Unidimensional and Multidimensional Approaches to the Measurement of Poverty”, *IRISS Working Papers* Series, 2003-02.

DAGUM C. (1990), “Generation and Properties of Income Distribution Functions”, in *Income and Wealth Distribution, Inequality and Poverty*, Studies in Contemporary Economics, ed. By Dagum C. and Zenga M., Springer-Verlag, p.1-17.

DAGUM C. (1999), “Linking the functional and personal distributions of income, in *Handbook on Income Inequality Measurement*, ed. By Silber J., Kluwer Academic publishers, Boston, p.101-132.

DRAGULESCU A. and YAKOVENKO V.M. (2001), “Evidence for Exponential Distribution of Income in the USA”, *The European Physical Journal* B, 20, p.585-589.

DUBOIS D. and PRADE H. (1980), *Fuzzy Sets and Systems: Theory and Applications*, Academic Press, Boston.

FOSTER J.E., GREER J. and THORBECKE E. (1984), “A Class of Decomposable Poverty Indices”, *Econometrica*, vol.59, p.687-709.

INSEE (1993-1994), *Enquête des Conditions de Vie des Ménages,* source : LASMAS-Idl C.N.R.S.

KAKWANI N. (1980) a, “On a Class of Poverty Measures”, *Econometrica*, vol.48 (2), p.437-446.

KAKWANI N. (1980) b, *Income Inequality and poverty: Methods of Estimation and Applications*. Oxford University Press, New York.

LELLI S. (2001) “Factor Analysis vs. Fuzzy Sets theory: Assessing the influence of Different Techniques on Sen’s Functioning Approach”, présenté à la conférence sur “Justice and poverty: Examining Sen’s Capability Approach”, St. Edmund’s College, Cambridge, juin 2001.

MICELI D. (1998), “Measuring Poverty using Fuzzy Sets”, *NATSEM, Discussion paper* n°38, p.1-41.

PARETO V., (1897), *Cours d’Economie Politique*, Lausane, ed. by Bousquet G. and Busino G., Droz (1965).

QIZILBASH M. (2002), “A Note on the Measurement of Poverty and Vulnerability in the South African Context”, *Journal of International Development*, 14, p.757-772.

SEN A. (1976), “Poverty : an Ordinal Approach to Measurement”, *Econometrica*, vol.44, p.219-231.

SEN A. (1985), *Commodities and Capabilities*, Oxford University Press, Oxford India Paperbacks.

SEN A. (1992), *Inequality Reexamined*, Harvard University Press, New Delhi.

SILBER J. (1999), *Handbook on Income Inequality Measurement*, Kluwer Academic publishers, Boston.

SORNETTE D. (2000), *Critical Phenomena in Natural Sciences*, Springer-Verlag, Berlin.

ZADEH L.A. (1965), “Fuzzy Sets”, *Information and Control*, n°8, p.338-353.

ANNEX : LIST OF DEPRIVATION INDICATORS CONSIDERED FOR EACH HOUSEHOLD USING THE INSEE-FRENCH SURVEYS OF LIVING CONDITIONS.

? Television : a_1

? Number of rooms in the accommodation : a_2 is function of the ratio of number of rooms per individual

? Telephone in the accommodation : a_3

? Water in the accommodation : a_4

? Bathroom in the accommodation: a_5

? Toilets in the accommodation: a_6

? Kitchen in the accommodation: a_7

? Owning a washing machine: a_8

? Bad insulation: a_9

? Humidity in the house: a_{10}

? Existence of bad smells: a_{11}

? Perception of the household situation according to the general aspect of the house: a_{12}

? Noise around the house: a_{13}

? Owning a car: a_{14}

? House owned or rented: a_{15}

? Banking account cheques: a_{16}

? Cut of electricity, water and gas: a_{17}