Poverty Analysis in the European Union: A Fuzzy Multidimensional Approach

Giulio Tarditi*
Università degli Studi di Siena

This paper discusses a Fuzzy Multidimensional approach which is developed from a completely relative notion of deprivation, and reports the results obtained from an analysis on the European Community Household Panel. Within this unified framework the classic monetary analysis can be coherently integrated by auxiliary variables which proxy other impoverishment traits such as the possession of durable goods, housing conditions and financial status. The all-comprehensive information-set allows to measure various aspects of what is commonly understood as poverty by identifying different refinements within the notion of poverty itself, however the interconnection between these different aspects still needs to be untangled. [JEL Classification: C42, D31, D63, I32, O15, O52]

1. - Introduction

The analysis of poverty in relation to inequality and Sen's capabilities approach has recently been at the center of the economic literature (Atkinson, 1996). Contrary to physical measurements, the measurement of a subjective feeling such as poverty, although connected to measurable entities such as commodities, presents many difficulties from a practical as well as a theoreti-

* <*giuliotarditi@yahoo.it*>. The Author thanks his supervisor Prof. Gianni Betti and co-supervisor Prof. Achille Lemmi, for the help and encouragement provided. He also thanks the anonymous referees of the *Rivista di Politica Economica* for the useful suggestions they have provided. The author takes full responsibility for any opinion or mistake in the article.

cal point of view. Traditional poverty analysis is totally dependent on the assumption that monetary deprivation is a sufficient proxy variable for all other dimensions of deprivation and that an arbitrary poverty line can divide the poor from the non-poor. The ground-breaking capabilities approach to poverty analysis (Sen, 1985) has fully introduced the multidimensional nature of deprivation and has set relative poverty at the center of attention within first world poverty analysis through the rationale according to which a relative measure in commodities will correspond to an absolute measure in capabilities. The crisp division of the population between poor and non-poor has long been criticized as unrealistic (Cerioli and Zani, 1990) and a *continuum* of living standards from poor to non-poor implies that any predefined cutoff point will be arbitrary (Mack and Lansley, 1985).

In this paper I will use the notion of Fuzzy Set Theory (Zadeh, 1965) and give a measure of how restrictive the crisp assumption is by computing fuzzy measurements at individual level following a complementary methodology (Betti and Verma, 1999)[1] to the Totally Fuzzy and Relative (TFR, Cheli and Lemmi, 1995). Recently the above two methodologies have been unified in the Integrated Fuzzy and Relative method (Betti *et* al., 2006; Betti and Verma, 2008). The European Community Household Panel (ECHP) dataset allows the use of extensive information to define the non-monetary dimensions of our analysis through the use of variables concerning the financial situation, the dwelling conditions and the possession of durable goods.

In the first section traditional poverty (*i.e.* exclusively a monetary phenomenon) is explained and the major shortcomings of such a method are presented. The fuzzy multidimensional alternative follows, discussing two different approaches and focusing on the membership function and the weighting specification. In the third section some first results are presented by a simple

[1] In this method the relative nature of privation is extended by considering the relative position of the statistical units weighted by their income. The analysis of fuzzy variables is then extended by cross-section and longitudinal comparisons, with the introduction of Manifest, Latent, Continuous and Anytime poverty measures.

classification of the data to provide a synoptic general understanding of the dataset over which the analysis is conducted. Finally, preceding the conclusion, the traditional and multidimensional methodologies compare the different results and a synthesis is presented to account for the visible discrepancies and provide a yardstick to quantify the partial nature of the monetary variable.

2. - The Traditional Approach

The traditional approach is centered on the identification, comparison and measurement of a specific subgroup of the population defined as the poor. It necessitates three strategically crucial decisions: the identifying criterion, the equivalence scale and the index formulation.

2.1 *Identification*

In the traditional dichotomous one-dimension poverty analysis, the identification rule must be formulated in order to define a numerical threshold on some monetary dimension. This value, often referred to as poverty line, partitions the sample or population under analysis in two subgroups, the poor and the non-poor. The conclusions that we can derive from such a division are as robust as the tacit assumption that the subgroups are homogeneous, as results can be extended as long as they apply uniformly across the individuals in question.

The criterion of "basic needs" is the most intuitive identification principle, and has been used from the birth of poverty analysis. A family (taken as the unit of analysis) is considered poor if it lacks sufficient resources to consume a predetermined "basket of goods" which is assumed sufficient to satisfy its basic needs. Rowntree (1901) had further proposed a "primary" poverty line that reflected the minimum necessary to maintain physical survival. However, deprivation refers to so much more then physical survival, that the

"correct" composition of the basket of goods had to be largely arbitrary and consequently not easily defendable.

The "food *ratio*" criterion seemed a viable alternative and defined a family as poor whenever the expenditure proportion in food exceeded a prefixed value. Engel (1895) observes that the "food *ratio*" is inversely related to the well-being of a household. The successive step of household comparison may be avoided, but again the threshold value is arbitrary.

From a predominantly absolute poverty measure such as that defined by the basket of goods, the poverty line has evolved into a relative measure, and today most studies within first world poverty rely on relative measures such as predetermined percentages of the median income or specific percentiles of the income distribution concerning the underlying population.

Finally it is worth mentioning that non-objective measures have also been proposed — through open questions (Kapteyn and Van Praag, 1976) or through multiple answers (Deleeck, 1977) — based on the idea that each individual is the best judge for his own poverty level. This statement has been challenged, and some common ground can be found in realizing that the difference between individual perception and objective deprivation can significantly contribute to the understanding of what is colloquially referred to as poverty.

2.2 *Comparison*

Household comparison is crucial to the traditional approach and often is not explicitly stated.

The different needs of different households have to be taken into account, and the household composition is believed to be an admissible way of weighting household incomes (Deaton and Muellbauer, 1980, 1986; Atkinson and Bourguignon, 1987; Ebert, 1996). The monetary variable, whether it be income or consumption, must be collected at the household level under the assumption that resources are more or less equally shared within the family where the economies of scale take place. The

least restrictive assumption one can make is that all components share the same living standard and must be weighted accordingly to their needs. A generally agreed-upon solution does not exist, mainly because the optimal criteria for such scale are interconnected with the objectives of the research, the underlying variability and the scope of the comparison.

Technical measures have been computed by following a normative procedure such as that of basics needs and more specifically such as biological nutrition needs. Thus given a fixed level of well-being, children will have less needs than adults, women less than men, the elderly less than the youth. This approach is absolutely normative in nature, but once we extend beyond nutrition the arbitrariness of what is to be accounted for becomes problematic and one can obtain unjustly discriminating results.

Empirical measures have been proposed following a positive reasoning, statistically inferring the weights from population samples and consequently being "value-free". However this approach poses other difficulties, as estimated well-being functions need not be socially acceptable and can be misinterpreted. An extreme example would be when more numerous households are correlated with lower incomes, hence under extreme circumstances one could infer that family size is inversely proportionate to family needs.

However one decides to compare the needs of different households, the precision of the estimates is only part of the issue. Once the sample information has been obtained, one will often want to compare the results with other samples and using a very detailed equivalence scale will diminish the comparability between different situations. Results obtained through tailored scales will not be comparable to other results without causing a significant bias, therefore one must be aware of the trade-off between the estimation efficiency and the domain over which the results may be used. In an analysis treating diverse populations and centered on the comparison such as the following, the choice has been made for a relatively simple scale in line with many other analyses on the same dataset.

2.3 *Measurement*

Once a poverty line has been defined, and a vector of equivalent incomes computed, we must construct an index that allows to measure the phenomenon from the most relevant perspective. The properties that a poverty index should have to successfully address its purpose can be listed through the following axioms as proposed by Sen (1976):

The monotonicity axiom: a reduction in the income of an individual below the poverty line must increase the measure of poverty;

The transfer axiom: a transfer of income between an individual below the poverty line and a richer individual must increase the measure of poverty;

The relative equity axiom: an increase in the income of an individual below the poverty line must have a greater impact than an equal increase of a richer individual.

Although seldom used, a more restrictive version of the relative equity axiom is:

The ordinal weight axiom: the difference between income and poverty line for the poor must be weighted with respect to the rank, hence proportionally to the degree of poverty.

The traditional poverty analysis approach is commonly based on three indices, each one providing a different interpretation of the income distribution and measuring the Incidence, the Intensity and the Inequality of poverty. They are referred to respectively as Head Count Ratio (HCR), Poverty Gap Ratio (PGR) and Gini coefficient (GIN) (Gini, 1921).

The HCR is the number of poor over the population (*i.e.* the proportion of poor). This measure is very intuitive and gives a correct estimate of the poverty under the tacit assumption that the two partitions defined are completely homogeneous sets with respect to poverty. Its use is widespread although it violates all the above axioms.

The PGR measures the *ratio* between the poverty gap, defined as the difference between the poverty line and the mean income of the poor subpopulation, and the poverty line. This proportion

reflects the minimum income sufficient on average to lift all poor from their situation to the poverty line. The PGR measures the intensity of poverty, and does not violate the monotonicity axiom. This measure confronts us with the question of how to consider the marginal utility of the monetary variable under analysis, that here is set to a constant. Marginal utility is commonly accepted to be increasing in the level of poverty, as the poor obtain more satisfaction from a fixed increase in income than the rich.

The Inequality is taken into account by the GIN, a measure derived from the Lorenz curve (Lorenz, 1905). The GIN is obtained by the difference between the Lorenz curve under complete equality of income and the estimated Lorenz curve, divided by the integral of the Lorenz curve under complete equality of income. By restricting this measure on the subpopulation defined as poor we obtain an index that accounts for an increasing marginal utility of the poor for any monotone transformation.

Amartya Sen proposes a linear formulation of the three indices having noticed their joint importance. This measure is known as the Sen Index (SEN) and is defined as follows:

$$\text{SEN} = HCR\ (PGR+(1-PGR)\ GIN) \tag{1}$$

The SEN proposal is born from Sen's critique (Sen, 1976) to the HCR as a prevailing index to use in poverty analysis, centered on the violation of the monotonicity and the transfer axiom. As successively pointed out by later authors (Shorrocks, 1995), under certain circumstances the SEN would itself violate the transfer axiom[2] unless we restrict the analysis to a weaker version of the transfer axiom in which the number of poor stays constant. This because the SEN in its original formulation establishes a trade-off between the incidence and the intensity of poverty and consequently a transfer from a richer to a poorer household can cause an incidence effect that would outweigh the intensity effect. Sen counter argued (Sen, 1983) that the technicalities of the transfer axiom as initially stated were moderately important, as

[2] Successively referred to as the Pigou-Dalton condition.

— although a relative poverty measure must be adopted — it must always be related to the *"breaking point"* proposed by Townsend (1962). It is this level of discontinuity which we try to approximate when introducing a certain poverty line.

Sen (1983) concludes his article with the following words:

"It is important to know whether the poor, relatively speaking, are in some deeper sense absolutely deprived. It makes a difference."

This concluding remark not only stresses the importance of addressing correctly these issues, but recalls the importance of relative poverty and the capabilities approach, whereby a relative measurement in commodities will give an absolute measure in capabilities.

2.4 *Difficulties*

Difficulties in a traditional poverty analysis arise in both the identification and the aggregation steps. To identify the poor subgroup one must first explicitly formulate a threshold (*i.e.* define some poverty line). The poverty line derives from an absolute poverty measure and grounds its intuitive idea on a specific basket that satisfies some pre-determined "basic needs". However as it has been noticed above, deprivation has a relative nature and is dependent on the context. This means that either by fixing some predetermined threshold, or by fixing a relative statistic of the underlying distribution, the poverty we want to define necessarily is linked to some arbitrary choice of threshold. To this, one must add the risk that the Purchasing Power Parity (PPP) approximation may not correctly account for equal opportunities in different countries.

Equivalence scales are widely used and allow us to weight household income with respect to its composition, as larger households will have lower living standards then smaller households given a specific income level. The more specific the

tool the higher will the cost be in terms of meaningful comparisons between different living standards that would require different specific values.

The monetary variable presents different options, as it is not clear if income is preferable to consumption as a living standards variable. If we are interested in income analysis, generally referred to as disposable income, then it is likely that the answers provided may be biased. Rich households are believed to often understate their incomes, and very poor individuals may overstate it given the intimate nature of the question.

Compared to what one would expect from a living standards variable, the income variable presents greater variability, however it is generally known with precision hence it is easy to estimate and intuitive to interpret.

A consumption analysis on the other hand, might more closely represent what is commonly intended by living standards, but depurates income of the investment component, and involves higher costs for the interviewee.

Given the known difficulty to correctly make inference on monetary variables, additional variables can be extremely valuable, not only for the computation of a multidimensional analysis but also to estimate and validate the truthfulness of the information provided by those household that seem to provide unlikely or incompatible answers.

This further reinforcement of the complexity of poverty analysis continues to undermine the exclusive use of a monetary variable, and it seems unlikely that such a variable is sufficient for a complete analysis.

2.5 *Criticisms*

The HCR as a reference index has often been under attack for its bluntness. This index, widespread to this day given its intuitive appeal, violates the monotonicity axiom (not accounting for the depth of poverty) and the transfer axiom (not distinguishing different poverty situations in which there has

been a transfer of resources from one poorer individual to another). As seen already the Sen's HCR critic stemmed a whole area of research that studied the importance of Sen's Axioms and the disputed "Pigou-Dalton condition". Possibly the most interesting alternative proposed was a convoluted compromise named the Sen-Shorrocks-Thon Index (SST) however it never gained advantage over other measures, and mainly remains a *curiosum*.

The traditional approach is centered on the threshold, whereby the sample is partitioned in two clearly distinguishable subsets. This partition is meaningful as long as the two subgroups are clearly different (polarization), but as those just above the threshold generally do not substantially differ from those just under, it is only an artificial abstraction[3].

The Fuzzy Set approach intends to investigate the information lost by tackling poverty through a dichotomous approach, and search for further insights in what is generally considered a "robustness" problem.

The Multidimensional approach investigates the distortions induced by the simplification of limiting the analysis to the monetary aspect of the individuals' well-being. If the monetary variable perfectly reflects other dimensions then it can be rightfully used as a proxy for living standards. However this tacit hypothesis made by the traditional analysis must be proven, as otherwise a significant bias in some specific dimension of poverty could result, and the picture obtained would be distorted. In this case a supplementary multidimensional analysis would provide informative insights. Difficulties related to the monetary measurement and verification cast further doubts regarding its exclusive effectiveness. Further data such as housing, financial situation or durable goods must be evaluated to verify and possibly open a debate on either its redundancy or its additional value, both of which would be significant findings.

[3] "Poverty is certainly not an attribute that characterizes an individual with its presence or absence, but is rather a vague assessment that manifests itself with different degrees and gradients" (CHELI B. - LEMMI A., 1995) own translation.

3. - The Fuzzy Multidimensional Approach

3.1 *The Fuzzy Aspect: Totally Fuzzy and Relative (TFR)*

Fuzzy Set theory, promoted by Zadeh (1965) and successively developed by Dubois and Prade (1980), provides the mathematical tools to overcome the critiques of the dichotomous aspect of the traditional approach. Given a set X of elements $x \in X$ any fuzzy subset A of X is defined as:

$$A = \{x, f_A(x)\}$$

Where f_A is a membership function (m.f.) of the fuzzy subset A and its value determines the degree of x's membership to A as follows:

$f_A(x) = 0$	x does not belong to A
$0 < f_A(x) < 1$	x partially belongs to A
$f_A(x) = 1$	x completely belongs to A

In poverty analysis we can consider X to be the sample of n individuals and A to be the subset of the poor within that sample. Once we have chosen an adequate set of indicators to fully represent the condition of poverty of the reference population, in order to compute the degree of membership at individual level (Cerioli and Zani, 1990) we define the m.f. as follows:

$$f(x_i) = \frac{\sum_{j=1}^{k} g(x_{ij}) w_j}{\sum_{j}^{k} w_j} \qquad \forall i = 1,2,...n$$

In this example, k[4] variables have been chosen as indicators of k different dimensions of poverty, respectively X_1, X_2,..., X_k. A generic weighting is represented by w_1, w_2,..., w_k and $g(x_{ij})$ is the

[4] The m.f. identifies a complete poverty variable or a partial poverty variable according to whether k is the total number of variables (e.g. Fuzzy Supplementary, Fuzzy Monetary (FM)) or a subset.

measure of deprivation of the *i*-th unit with respect to the *j*-th indicator. The function $f(x_i)$ is the weighted average of the single $g(x_{ij})$ and represents a complete poverty variable calculated for each individual on all indicators. Following the Totally Fuzzy and Relative (TFR) approach proposed by Cheli and Lemmi (1995), each component of the m.f. referring to each single dimension is defined as follows:

$$g(x_{ij}) = \begin{cases} H(x_{ij}) & \text{risk of poverty is increasing in } X_j \\ 1 - H(x_{ij}) & \text{risk of poverty is decreasing in } X_j \end{cases}$$

Where $H(x_{ij})$ is the cumulative distribution function for the variable X_j.

We will divide the case of categorical variables from the one of continuous variables.

Specifying the m.f. for *categorical variables* can be problematic when we have on some dimension high frequencies for high standards. In this case the lowest degree of poverty will be valued by a relatively high estimate and the previous formulation will give all high values in a relatively short range. The problem can be avoided by appropriately rescaling the parameters to standardize the underlying variability in the levels of deprivation within a [0,1] interval:

$$g(x_{ij}) = \begin{cases} 0 & \text{when } x_{ij} = x_j^{(1)} \\ g(x_j^{(\lambda-1)}) + \dfrac{H(x_j^{(\lambda)}) - H(x_j^{(\lambda-1)})}{1 - H(x_j^{(1)})} & \text{when } x_{ij} = x_j^{(\lambda)} \end{cases}$$

Where λ is the index across possible answers available in the dimension X_j and $x_j^{(\lambda)}$ is the generic degree of poverty with respective upper and lower limit $[x_j^{(1)}, x_j^{(\max)}]$.

This new formulation will standardize the membership values in the unit interval proportionally to the deprivation, although in dichotomous variables it would be redundant. Specifying $g(x_{ij})$ in terms of the distribution of X_j each statistical unit is given a "score" on their deprivation that exclusively depends on the relative

proportion of better off statistical units on such a measure of deprivation. This procedure gives this method its categorical ranking system that is completely relative in nature, and relies on the correct specification of the different dimensions that must be taken into account to fully reflect the prevalent lifestyle under analysis.

In the case of *continuous variables*, such as the monetary variable, the previous formulation is not appropriate hence we use the cumulative distribution $H(.)$. To clarify the example we will adopt the notation of y_i to identify the equivalent income of the i-th statistical unit as it is the most common continuous variable. The m.f. is defined by:

(2) $$g\ (y_1) = 1 - H\ (y_1)$$

The $H(.)$ function may either be empirically derived or exogenously introduced in the analysis.

In accordance with the categorical case we assign each individual a degree of poverty within the [0,1] interval that completely depends on the fraction of units that have a higher income. It is important to notice that this method does not make use of any poverty line as high and low values will respectively correspond to poor and non-poor.

Global indices are used to give a synthetic measure of a complete poverty variable by taking the mean value across individuals. We define the global monetary poverty index as:

(3) $$Y = \left(\sum_{i=1}^{n} \left[1 - H(y_i) \right] \right) \Big/ n$$

The calculation of the index on the whole sample is not very informative as the $H(y_i)$ function is by construction within the [0,1] interval and the expected value is 0.5.

$$E[1 - H(y_1)] = 1 - E[H(y_1)] = 0.5$$

This unsatisfying result has been criticized, and considered the limit to the TFR method, as whenever we calculate a global

poverty index on different samples we always obtain the same value. It is not a traditionally cardinal measure but this is not a central issue to regarding its use which is more closely related with the comparison of different subpopulations for which the computed values do differ indeed. It can be argued that the cardinality in traditional measures is however artificial, as the reliance on an arbitrary distinction between poor and non-poor implies arbitrariness in the parameters, hence the cardinal measure varies with the selected line of poverty. This pseudo-cardinal feature can be adapted to the TFR measure through the following change in the formula:

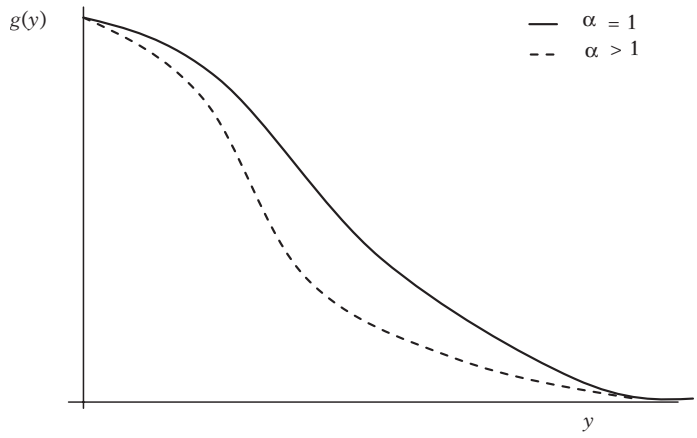
$$g(y_1) = (1 - H(y_1))^{\alpha} \tag{4}$$

An increase in the alpha value will polarize the values assigned to poorer and richer households, hence weighting the poor more as shown in the graph, and by defining an arbitrary value it will model the global poverty index to the HCR through an iterative process. This gives the analysis the pseudo-cardinality feature and allows a direct comparison between the two approaches.

The aggregation of intrinsically different measures, which necessarily determines a possible trade-off, is clearly one of the

GRAPH 1

EFFECT OF VARIATION IN THE ALPHA LEVEL



most fragile points in multidimensional analysis in general, and represents a possible weakness of this method. Methods that stem from TFR generally require the HCR index to establish a relationship between the monetary and the supplementary variable. The researcher can then justify the comparison between different subgroups of the sample, and can evaluate the contribution to poverty of different profiles, such as the working class or rural areas which are particularly at risk. An alternative where a different relationship is established and HCR index need not be calculated is presented in Betti and Verma (2008).

Similarly the global supplementary index can be described by the formula:

$$P = \left(\sum_{i=1}^{n} f(x_i) \right) \Big/ n \tag{5}$$

Where $f(x_i)$ is referred to the relative supplementary m.f. In the TFR approach, the P index represents a generalization of the HCR and can be interpreted as the proportion of poor families deriving from the relative cardinality of the fuzzy subgroup and the cardinality n of the population (Dubois and Prade, 1980). However one must remember that the fuzzy approach does not reason in terms of poor and non-poor, as the P index does not refer to the percentage of poor in the population, but the average level of poverty as measured by the m.f.

3.2 *The Fuzzy Aspect: A Complementary Methodology*

The methodology I have used is centered on the need to extend the conventional classification of poor/non-poor as is presented in (Betti and Verma, 1999)[5]. As in the TFR, one must

[5] The undergraduate thesis from which this article has been drawn was written just after the publication of the complete ECHP dataset. It contributes with an empirical analysis to the valuation of poverty conditions in the EU. Based on the methodologies proposed in CHELI B. - LEMMI A. (1995) e BETTI G. - VERMA V. (1999) the author specifies the necessary for a complete analysis, confronts the results and presents his own conclusions.

replace the dichotomous classification with a degree of membership to the Fuzzy Set of the poor. This degree is confined within the [0,1] interval and proportional to the degree of deprivation. The advantages of such a classification can be summarized in the following points:

i) The methodology allows us to incorporate some aspects of the GIN and SEN index.

ii) A valuation of the relative position of different subgroups within the population can be obtained in a much more intuitive way.

iii) When considering a non-monetary deprivation such as possession of durable goods one may have some and not others, hence a crisp measure is clearly inappropriate.

iv) The approach is particularly helpful in determining entry and exit levels of statistical units within a longitudinal framework.

To determine the poverty index we must link the *i*-th individual's inclination to poverty with his rank and the proportion of income that the individual has within the equivalent income distribution. We define the m.f. as:

$$V_i = \sum_{j=i+1}^{n} \left(y_j \Big/ \sum_{k=2}^{n} y_k \right) \qquad \forall i = 1,2,...,n-1 \quad V_n = 0 \tag{6}$$

This represents the proportion of the sum of equivalent incomes with respect to the individual that has the *i*-th rank in the ordered distribution of incomes. As the previously discussed m.f. also in this case we assign values within the [0,1] interval proportionally to the degree of poverty.

We define a new Fuzzy Monetary (FM) index as:

$$I_i = V_i^{\alpha} \tag{7}$$

This is analogous to the previously illustrated TFR method. As supplementary variables are either dichotomous or discrete in nature we must determine a functional form for identifying the relative score in the discrete case. Dichotomous variables do not need such a specification as we can directly give a score of 0 or

1 to the only two possible outcomes, but in case where we allow for multiple degrees of poverty one must specify how the degrees relate to each other. For M being the number of multiple answers and m being the rank of the degree increasing in the level of poverty we can adopt an equidistant measure on the range as follows:

(8) $$D_{j(m)} = \frac{M - m}{M - 1}$$

Hence Betti and Verma (1998) presents similarities with the TFR method due to the common objectives, but also substantial modifications. In this article the contribution to this branch of research comes not only from the clarification and application of these improvements but also from the comparison and evaluation of the results with respect to the traditional measures. While the above mentioned article conducted an analysis only on the first two years, here the analysis is on the complete dataset, and is comprehensive of Austria, Finland and Sweden which did not participate at the ECHP during the first years.

Having panel data is particularly important for the longitudinal measures which will be introduced at the end of the chapter, but more generally for all those empirical fuzzy and traditional results for which the complete evolution is now available. This gives each result a measure of robustness and can be informative on the relationships between efficiency and equity like in the Irish case. However an organic analysis also poses new questions and introduces new difficulties, such as the accumulation of biases due to attrition or the comparison of temporary poverty with respect to persistent poverty, either monetary or supplementary.

3.3 *The Multidimensional Aspect*

Once a m.f. for a certain dimension has been defined we must identify how to coherently aggregate the different indices corresponding to the different dimensions. The various indicators

that attribute a level of poverty must be defined, in order to calculate the level of deprivation over the whole set of supplementary dimensions.

$$f(x_i) = \frac{\sum_{j=1}^{k} g(x_{ij}) w_j}{\sum_{j}^{k} w_j} = D_i \qquad \forall i = 1,2,...n \tag{9}$$

The weights can be defined in several alternative ways, but they must be a decreasing function of the deprivation level such that poverty in terms of a relatively widespread item has a greater impact on the index than poverty in terms of a scarce item. The proposal I have adopted in the following analysis is given by the logarithm of the inverse of the fuzzy mean of households relatively to the variable X_j (Cerioli and Zani, 1990):

$$w_j = \log\left[n / \sum_{i} g(x_{ij}) \right] \tag{10}$$

The fuzzy proportion of poor statistical units with respect to X_j is given by the mean value with respect to the whole population, and represents the population average membership to the subset of the poor with respect to X_j. The inverse guarantees that as the average level of deprivation grows, the weight given to such dimension is inversely related. The logarithm transformation avoids giving too much weight to rare items, as it would happen in dimensions for which the population under analysis is relatively well endowed. The importance of choosing measures that correctly specify the well-being of the underlying population hence critically influences not only the m.f. but also the weighting factors. Although the monetary and the supplementary fuzzy variables present some common traits, it is as if they were distinct partial indices and independent of one another. This means that to obtain comparable measures we must determine a relationship, which can be done by equating their aggregate value.

$$(11) \qquad D_i' = \left(\frac{\bar{I}}{\bar{D}} \right) D_i \quad \text{where} \quad \bar{D} = \sum_{i=1}^{n} D_i \Big/ n \quad \text{and} \quad \bar{I} = \sum_{i=1}^{n} I_i \Big/ n$$

After this transformation, the average value of the new measure of supplementary poverty D_i' (FS) will be equal to the average value of the measure of monetary I_i (FM) and may be combined to construct a measure of how much these two different aspects overlap.

3.4 *The Comparison among Fuzzy Measures*

Once we have obtained various fuzzy measures with respect to different dimensions — e.g. monetary poverty Fuzzy Monetary (FM) and non-monetary poverty Fuzzy Supplementary (FS) — we can compute individual measures that correspond to particular fuzzy theory operations:

i) Manifest Poverty (MP): the minimum value between FM and FS per unit

ii) Latent Poverty (LP): the maximum value between FM and FS per unit

We can also evaluate longitudinal properties by comparison of the above across time periods identifying the evolution of variability and other statistical properties. Studying the individual maximum and minimum across time periods we obtain the following measures:

i) Continuous Poverty (CP): the minimum value of the measure of poverty across time

ii) Anytime Poverty (AP): the maximum value of the measure of poverty across time

These additional measures give us a new perspective on the information within the dataset and provide valuable information on the whole sample or on specific subgroups. In particular we can evaluate the information overlap for cross-section measurements and some specific dynamic properties such as trends or heteroskedasticity for longitudinal measurements.

4. - Empirical Results

4.1 *European Community Household Panel (ECHP)*

This analysis uses the ECHP[6] dataset that has been monitoring the European Union (EU-15) between 1994 and 2001. The data was collected through the use of questionnaires that all originated from a unique blueprint which was only later adapted to best gather information from the different realities of the participating countries. Although there are some compatibility problems, the ECHP is surely one of the most detailed and self-sufficient datasets for transnational comparisons of living conditions realized up to date. The ECHP has been very successful and has realized its purpose of comprehensively representing the EU population through micro-data collected at both the household and individual level. It is a "panel" database as the interviewed households answer once a year (Lazarsfeld, 1940) and this has many advantages one of which is the possibility to conduct a longitudinal analysis that can shed light on the population dynamics not traceable in a cross-section analysis. It also allows to have more control over the heterogeneity of the individuals, and to obtain greater estimation efficiency by adjusting to various dynamics.

4.2 *The Celtic Tiger*

Many interesting observations can be derived from this multipurpose dataset, however the Irish data is particularly significant as the country underwent an entire economic boom cycle that was later referred to as the Celtic Tiger[7]. As documented by the ECHP data, Ireland was transformed from one of the poorest countries in EU to one of the wealthiest, and this offers the opportunity to analyze the whole process and discuss the link

[6] Coordinated by EUROSTAT (Statistical Office of the European Community).
[7] Term coined in a 1994 Morgan Stanley report by Kevin Gardiner in analogy to the "East Asian Tigers".

between efficiency and equality. Different reasons have been given to the causes of such unexpected growth. The low corporate tax rate is often suggested[8] as one of the major reasons and the substantial net transfer payments provided by the European Community seem to have been wisely used. However growth in a relative poverty framework does not have the same implications as absolute poverty. An increase in the median income given by economic growth corresponds to a proportional increase in the poverty line, which means that if growth leaves behind poorer households, it will cause a worsening in the poverty incidence, intensity and inequality measurements.

4.3 *Traditional Results*

The traditional analysis is centered on the poverty line of the monetary variable and, as we value the relative notion of poverty, we have fixed such value at 60% of the median[9] income. This poverty line is generally referred to as "at-risk-of-poverty" threshold and is often used on the ECHP dataset. PPP[10] disposable income is equivalized through the use of the modified OECD equivalence scale[11].

Graph 2 shows the equivalized median income through different countries and years (waves) for available countries. As the data is in terms of current prices, the comparison should be done across countries instead of across years. This is because the values are in terms of nominal prices and although the PPP counterbalances inflation in different countries, the mean inflation across years is still present as shown by the approximately 3%

---

[8] «The luck of the Irish», *The Economist*, October 14, 2004.

[9] The median is preferred to the mean as it is less sensitive to the pronounced asymmetry of the income distribution.

[10] Purchasing Power Parity at current prices is derived from a fixed basket of goods and allows a meaningful comparison between different currencies.

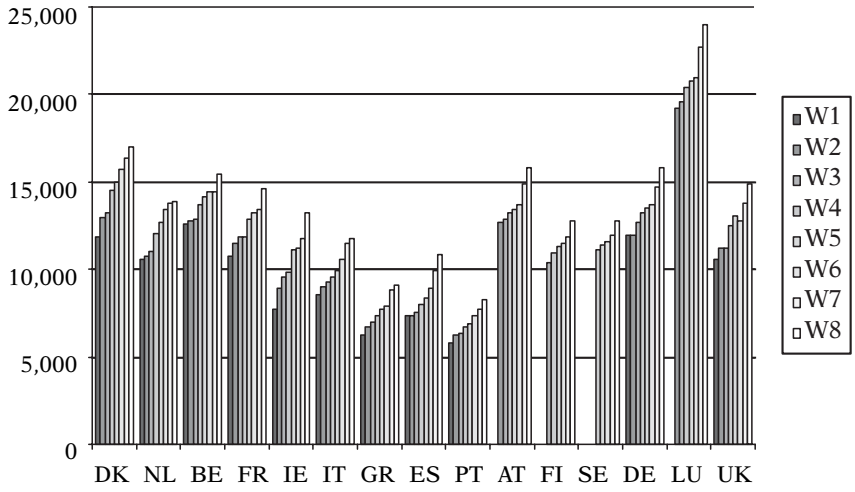
[11] The (Modified) OECD scale assigns a unitary (1) weight to the first household member, 0.5 (0.7) to each additional adult member and 0.3 (0.5) to each child. An alternative scale would be the McClements scale (MCCLEMENTS L.D., 1977).

sustained increase in the median income. The poverty line values are given by a vertical 60% scaling of Graph 2. Greece, Spain and Portugal show values up to one third of those of Luxemburg. Much information is lost in computing the median hence — for illustrative purpose only — Graph 3 reports the cumulative[12] European distribution divided by country for 1994. On the horizontal axis are 100 equidistanced classes between € 0 and € 50,000 a year, and on the vertical axis we have the number of "weighted" interviewees that match such a class. While the countries presented are the same as in Graph 2 the color identification distinguishes only the poorest five.

The HCR, *i.e.* the proportion of the population at risk of poverty, varies between 10% and 20%. We find respectively low values in Finland, Sweden, Denmark, Netherlands and Luxemburg,

GRAPH 2

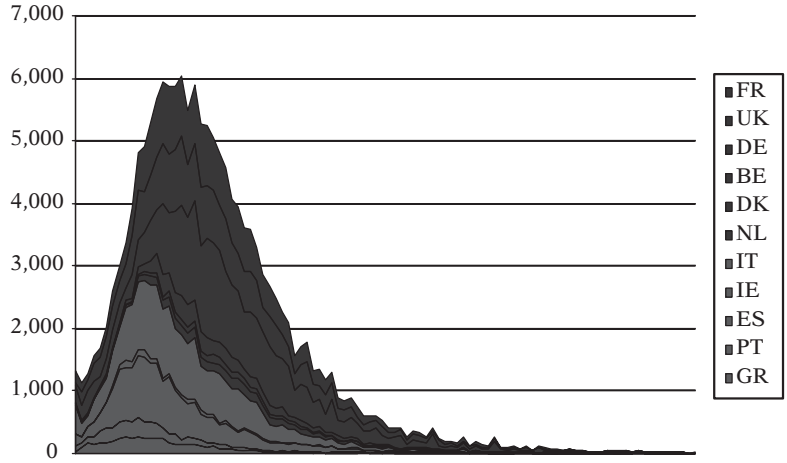
EQUIVALIZED MEDIAN INCOME AT CURRENT PRICES



[12] The country ordering has been chosen to ease the proposed interpretation. The cumulative graph is less traditional, and does have some evident drawbacks, however it is here presented as a powerful illustrative example.

GRAPH 3

COMPARED INCOME DISTRIBUTION (1994)

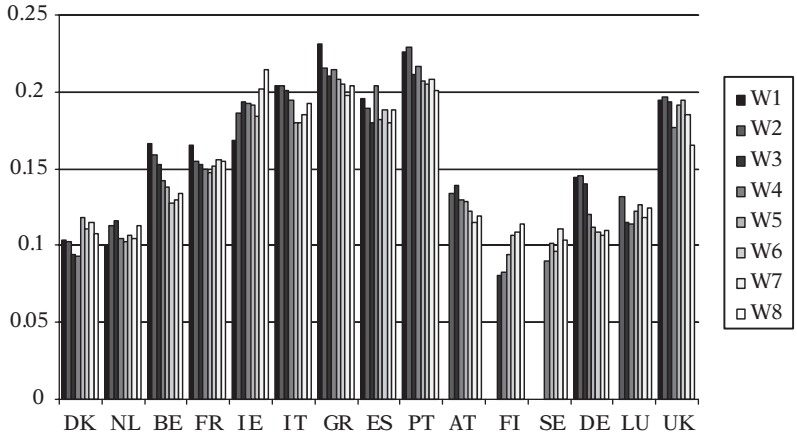


intermediate values in Germany, Austria, Belgium and France, and high values in United Kingdom, Ireland, Italy, Greece, Spain and Portugal. The correlation with equivalized income is evident, as low incomes generally correspond to high HCR even allowing for relative poverty thresholds. This is congruent with our previous intuition that poorer countries have more asymmetrically skewed distributions as seen in Graph 3. The division between north and south is maintained although the Anglo-Saxon model does not perform up to expectations, possibly due to its historically more liberal approach to welfare. If we compare different years the HCR seems to have a general tendency to diminish, except for Finland and Ireland that show a marked increase.

The PGR, *i.e.* the proportion of poverty line that if redistributed to the poor would allow them to reach the poverty line, varies between 15% and 30%. Ireland and Portugal present the most extreme changes (10%), furthermore in opposite direction. Ireland starts with the lowest values in the EU and by the end of the period overtakes Portugal which had started with PGR just second to Greece. The evidence is significant and

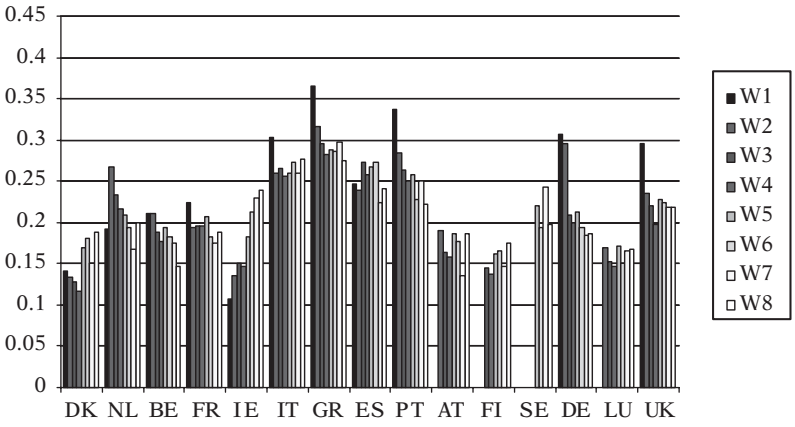
GRAPH 4

HEAD COUNT *RATIO* (HCR)



GRAPH 5

POVERTY GAP *RATIO* (PGR)



although some distortion may lie in different *attrition* patterns, this is not visible from a simple attrition analysis based on the comparison of the original sample and the balanced panel subgroup (see appendix). We can surely attribute some of this

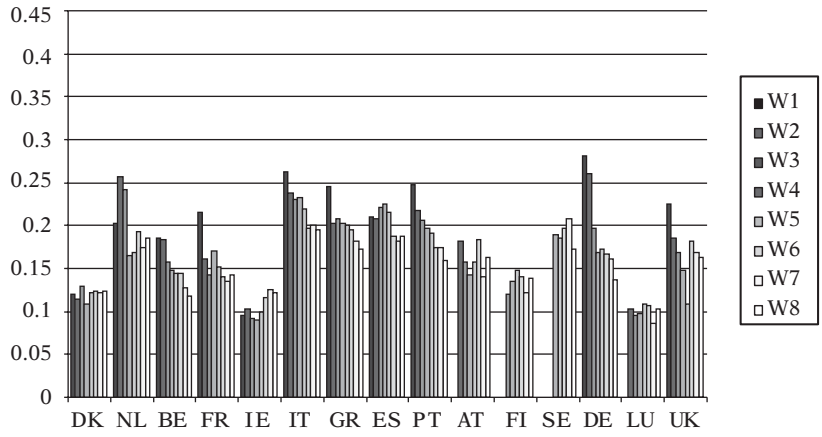
effect to the relative nature of the analysis, as the poverty line in Ireland increases at a higher pace then in Portugal.

It is interesting to notice the common drop in this measure within the first two years, which is probably due to an early exit of the extremely poor households in the sample. In Germany we notice that the PGR drop of 10% is consistent with the HCR measure both of which are fairly stable in the following years. Instead in the UK there is a more marked difference between the two measures especially during the first and the last year.

The GIN index presented in Graph 6 is that referred exclusively to the poor population. Limiting the analysis to the poor subgroup we can picture more clearly the dynamics of the very poor and we are not distracted by distribution changes in the non-poor subgroup for which one could conduct a separate analysis. Ireland shows very low and stable GIN values throughout the eight years, which means that the poor predominantly fall within a specific income bracket that successively shifts further ways from the poverty line as shown by the PGR. This strengthens the conviction that the increase in poverty measures is mainly due to the growth effect.

GRAPH 6

GINI INDEX FOR THE POOR SUBGROUP (GIN)



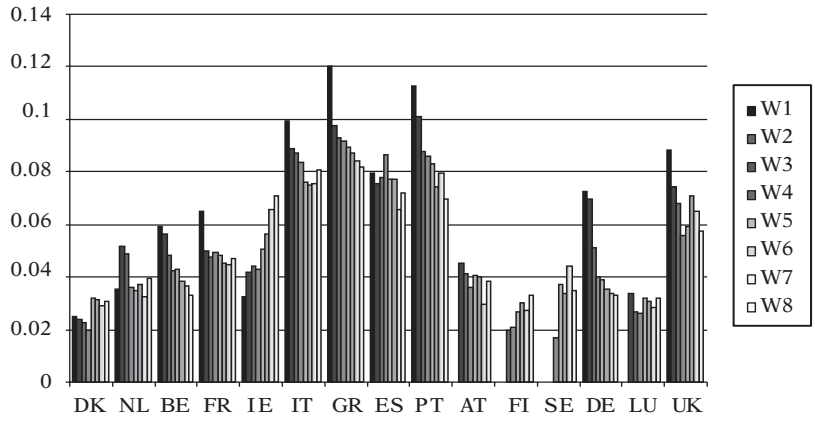
Conversely Germany shows a marked correlation between all three indices which suggests that the sudden change following the first two years is due to an early exit of some very poor individuals. We can propose this interpretation because if poverty where to diminish uniformly from the poverty subgroup, the GIN would not have been affected.

The SEN index in Graph 7 does not add any new information but gives a general overlook on all aspects of the phenomenon. Portugal, Italy, Spain and Greece show very high values as predictable, and Ireland's race to reach southern European levels is worrying.

The bluntness of these indices is self evident, and the results can be ambiguous. Clearly a deeper analysis would necessarily have to take into account the whole income distribution.

GRAPH 7

SEN INDEX (SEN)



4.4 *Fuzzy Multidimensional Results*[13]

The fuzzy multidimensional analysis raises new questions and by either supporting or contradicting the traditional approach

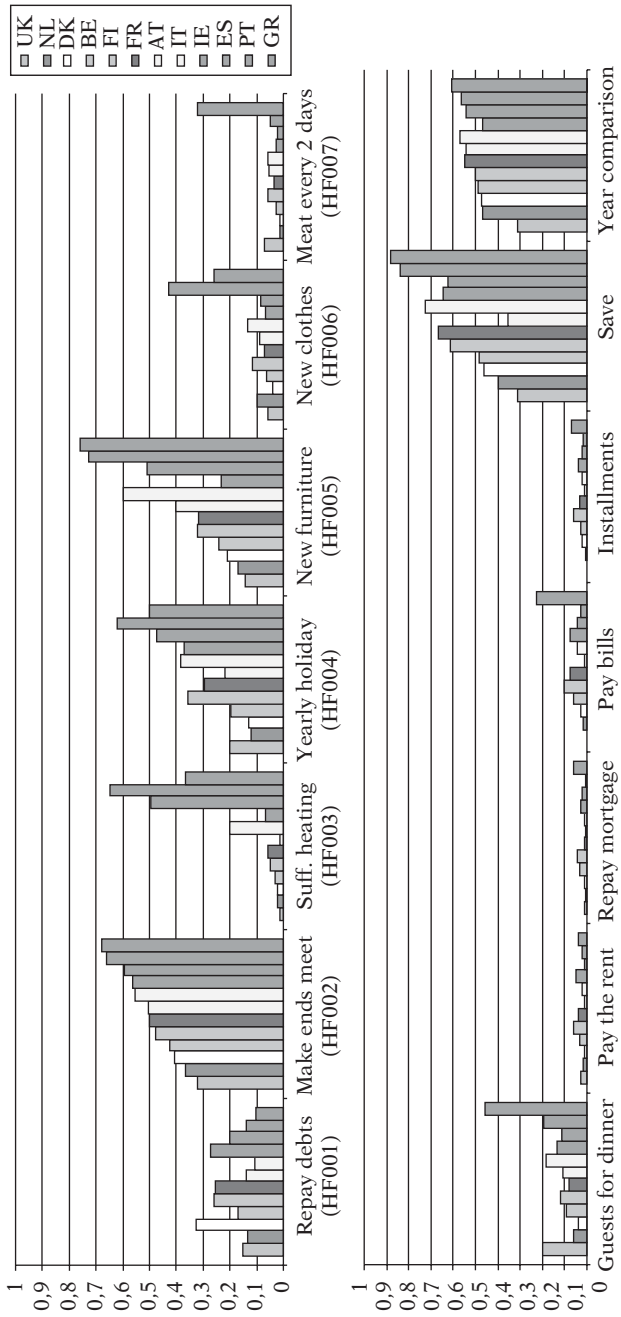
[13] For further the variable specifics consult EUROSTAT DOC. PAN 166 /01 - ECHP UDB Description of variables.

clarifies previously ambiguous answers. The implemented methodology has the advantage of following the relative approach applied in the monetary analysis, and creating a meaningful supplementary poverty measure at a disaggregate level. This plays an important role enabling the researcher to compare different subgroups of the population. The choice of variables to be included, that is intrinsically subjective, has been determined simply by the data available within the areas of interest — financial situation (Graph 8), dwelling conditions (Graph 9), durable goods (Graph 10) — as the primary objective is to present the methodology. Most variables give strong insights in the different living conditions between different countries, and open questions that can have different answers depending on the financial, but also cultural, geographical and climatic background: Insufficient heating in Denmark implies miserable living conditions, but in Spain a comparable deprivation results in many of the answered questionnaires, which suggests a different relative importance.

The first fourteen variables (Graph 8) identify aspects concerned with the financial situation of the interviewee's household. "Making ends meet", saving some income and a valuation of the minimum salary necessary are the indices that show highest values. This can be partly explained by the fact that such questions are subjective, and subject to wishful thinking, therefore they can adapt to many different poverty profiles. However if we were to order the countries by median income, an inversely related pattern would emerge. The inability to repay the rent, the mortgage, the bills, and other installments show very low values, as they can be easily anticipated months before, and when not met imply serious discomforts. We can notice how the relative approach is in line with what is commonly considered poverty as it weights more those causing major discomfort. Questions that ask whether the heating is insufficient show that this is the case predominantly where there are climatic favorable conditions, such as Portugal, Spain and Greece; one could investigate if the pattern is still identifiable at regional level. Our relative approach consents us to accordingly weigh a faulty heating whether it be in Greece

GRAPH 8

FINANCIAL SITUATION (1994)



or in Finland. Holidays, furniture renewal, new clothes and having guests over for dinner have similar rankings, but the size reflects the nature of the expense. While the first two are large, exceptional and fixed the following are moderate, frequent and flexible.

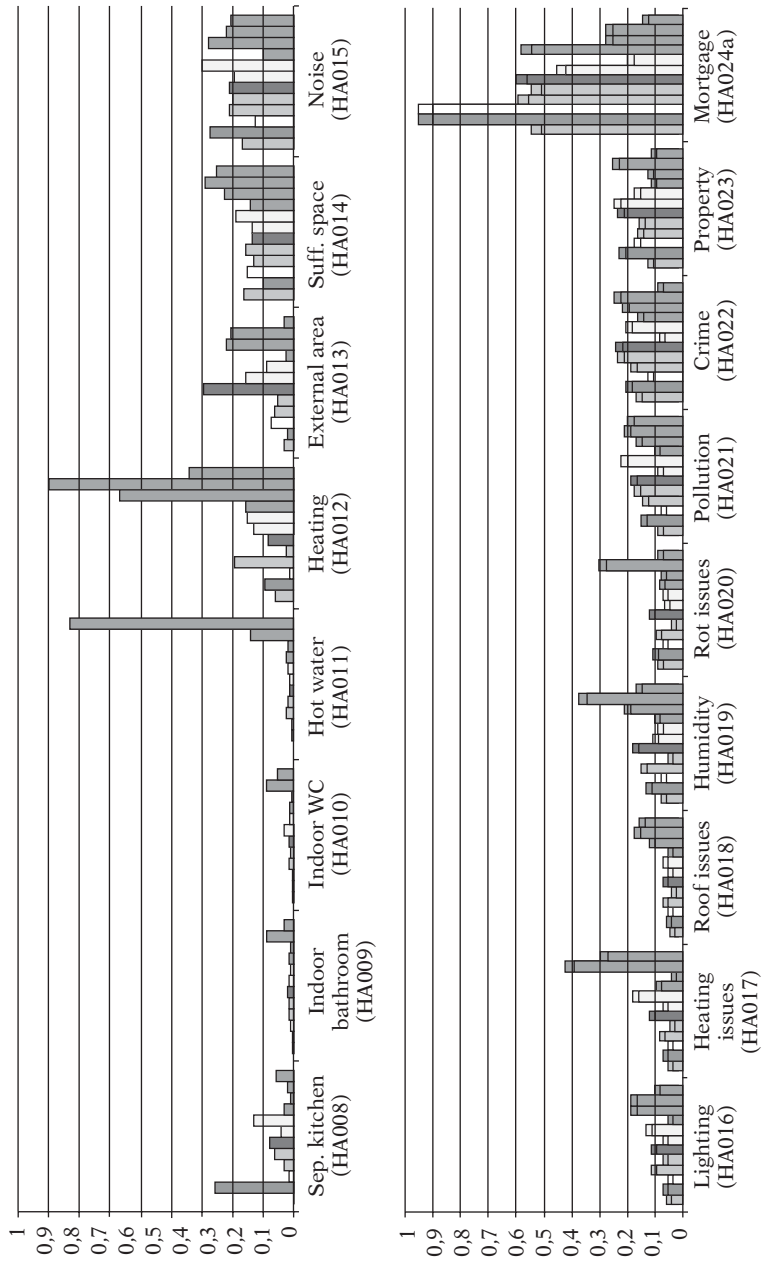
The following seventeen variables (Graph 9) monitor the dwelling conditions. A separate kitchen and an open area attached to the dwelling are largely concerned with the predominant architectural style, and are more difficult to connect to what is commonly meant by poverty. Presence of a heating system presents analogies to the variable in the financial situation which measured the capability to sustain the costs of an adequate heating. Missing bathrooms, WC, hot water, sufficient lighting, functional roofing, humidity and rot are comparatively rare, and present low values (under 10%)[14] apart from Portugal that distinguishes itself for the appalling dwelling conditions. Hot water seems to be an issue in Greece, but the data for the year under analysis is often not considered reliable. Insufficient space, pollution, noise and crime do not present clear patterns. This can be because they are independent of income, they are too complex to be separated at national level, they concern difficult realities to analyze and compare or more likely a combination of the three. The variable concerned with the presence of a house mortgage can be interpreted in different ways. The fact that one does have a debt to repay can be seen as a measure of poverty or wealth, as it generally implies some form of collateral. One could also advocate for neither, but only indicating a choice of inter-temporal consumption facilitated by an efficient financial market that guarantees an efficient allocation of resources between different times and people. We can notice that it is proportional to the income variable, but this could be reflecting the cultural aspect of taking out a loan, more common in northern countries which often have more efficient financial markets.

The last eight variables (Graph 10) are concerned with durable goods and are dichotomous in nature. The durable good

[14] This implies that 10% of the individuals have said to live in such conditions only if the variable is dichotomous.

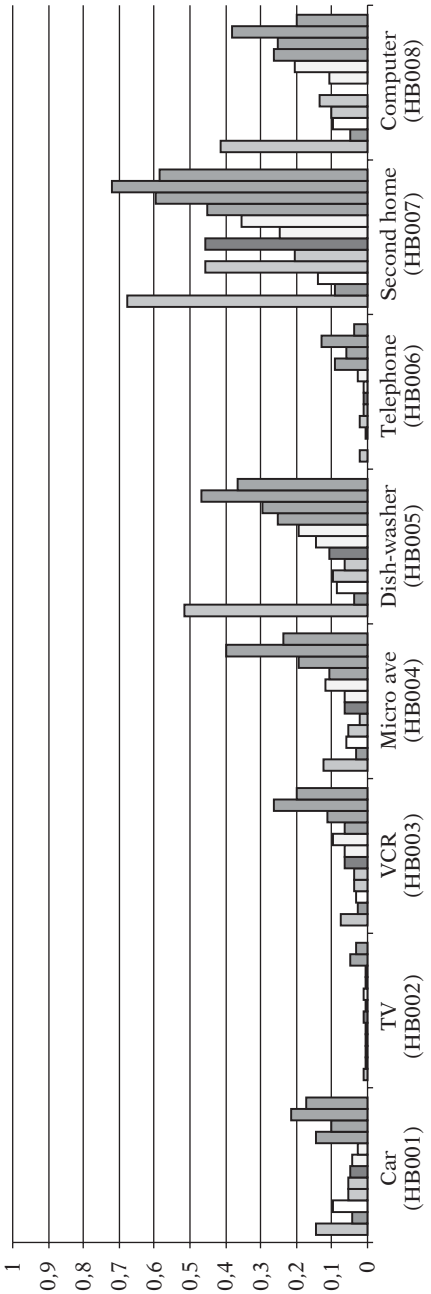
GRAPH 9

DWELLING CONDITIONS (1994)



Graph 10

DURABLE GOODS (1994)



framework allows us to see once more the importance of the relative approach that characterizes this analysis. Owning a good is clearly a sign of wealth, as it allows us to satisfy needs and can generally be sold, thus it could be valued by some value parameter such as the average cost of such item. The intrinsic value however is dependent on the context: an example can be given by the common car. In a social structure as the Italian one, where structural investments have promoted it as cornerstone of mobility a car has a very different value then in the UK where public transportation is a viable alternative, or in Holland where distances are very different. Therefore not having a car means different things in different places and taking the cost value is not necessarily the optimal procedure. Cultural differences can be seen in the relative importance of owning a VCR, a microwave, a dishwasher, or a computer (for which the diffusion proportional to income can be seen as either a cause or an effect). The absence of owning a second house presents high values as could have been expected, this often being a luxurious and expensive item. House prices and costs will be positively correlated with this variable as in the UK.

4.5 *Fuzzy Monetary (FM) and Fuzzy Supplementary (FS)*

The fuzzy multidimensional methodology opens many important questions to which it provides answers that are most valuable when evaluated jointly with the traditional analysis and idiosyncratic information concerning the distinctive issues. In this chapter we try to give a flavor of the main results at EU-15 level, and illustrate the potentialities of further analysis developed by identifying different specific subgroups within the sample.

To fully understand the importance of the fuzzy multidimensional method one must clearly have in mind how the m.f. (Graph 1) relates to the distribution of the fuzzy measures (Graph 11). To be consistent with the HCR measure the FM has a highly polarized m.f. which separates between individuals with high and low deprivation levels within a relatively narrow income

interval. The graph below however must not be confused with the m.f. as it actually represents the distribution of the individual membership values given by the sample distribution, as the increase in FM levels within the [0.9,1] interval clarifies. The FS counterpart on the contrary approaches zero for high levels of deprivation due in part to its highly multidimensional nature. It must be said that the graph is partially incomplete as it shows only a portion of sample with FM values between [0,0.1]. This however is needed to visually compare FM and FS for the OP subgroup.

Having calculated FM and FS measures at an individual level one can follow the traditional approach and identify the "officially poor" (OP) and track them in this new framework. The FM measure gives no significant additional evidence to the dichotomous analysis, but the FS presents a positively skewed distribution, with few individuals having FS above 40%. The difference between those with high levels of FS and the OP population is surprising, proving that the poor in most official statistics will often not reflect such deprivation on a non-monetary scale. The proportion of OP is increasing in the FS measure, but many which are FS-poor are not OP, therefore further analysis is needed to define different poverty profiles. This intuition can be verified at national level. The following graph (Graph 11) presents the sample divided in OP and OP^c by two different viewpoints (FS and FM) therefore the total sample weight size is doubled.[15]

This result, although significant, remains ambiguous as it has two possible explanations:

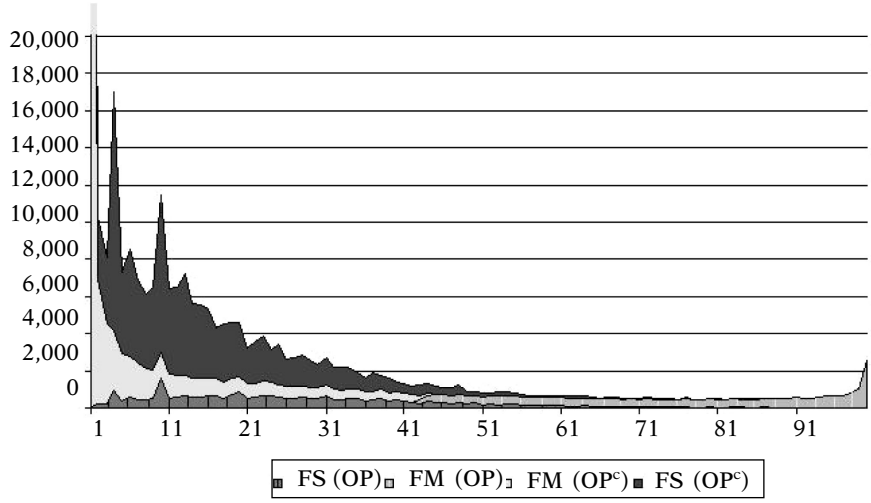
i) Individuals with low incomes (OP) do not entirely correspond to those with low living standards. This further information must be used to guarantee the protection of those who are most in need from a living standards prospective and promote social-cohesion.

ii) The variables used in the analysis are not representative of the needy population, and therefore must be revised and modified

[15] "c" stands for complementary set, therefore as OP are the "officially poor", (OP^c) stands for the "officially non-poor".

GRAPH 11

FUZZY MONETARY (FM) AND FUZZY SUPPLEMENTARY (FS)



in order to describe more adequately the real needs that people suffer.

These two interpretations are not mutually exclusive, and start a breach in the income versus living-standards equality. Deeper questions on the analysis and its purpose are necessarily raised:

What's the use of knowing these further elements? If lower wages do not match lower living standards, how can we develop more focused poverty relief policies? Is the objective of poverty relief policies to guarantee minimum living standards or to provide a safety net for those that have low incomes? On what basis should we fight temporary poverty with respect to permanent poverty?

4.6 *Manifest and Latent Poverty Measures*

If we are comparing two different fuzzy measures, MP and LP measures are by definition complementary. Their objective is to underline a "stronger" and respectively a "weaker" measure of

the membership to the poverty set at individual level. When we compute the LP measure the maximum deprivation between FM and FS is taken into account, therefore it provides an upper bound between the two poverty measures. Being the upper bound, it is expressed by higher figures, but necessarily characterizes a less stringent measurement of poverty as the complete measure is also inclusive of the MP counterpart. This may seem counter intuitive, but taking the maximum value implies that the individual's overall poverty must be lower, hence it specifies a weaker deprivation condition. The complementary analogy indicates that the MP is a lower bound, therefore the underlying privation must be more distressful.

By computing a similar analysis to the FM and FS measures we notice that no $LP(OP)<MP(OP^c)$ but in some cases $MP(OP)<MP(OP^c)$ in line with our previous results. When the global index of the MP/LP[16] is approximately unitary the monetary variable is a good estimate of the supplementary dimensions, as high FM values correspond to high FS values for most individuals. If alternatively the value is low in the [0,1] interval, two distinct analyses must be adopted. The *ratio* compared at national level is in most cases proportional to the traditional poverty measures, and there being a relationship implies that by limiting the analysis to the monetary variable we are distorting the resulting poverty figures inversely to the monetary poverty level.

4.7 *Continuous and Anytime Poverty Measures*

CP and AP measures are computed exclusively on the balanced dataset to provide minimum and maximum values on equal time intervals[17]. As we have four poverty measures this allows us to create a whole set of comparisons that through appropriate underlying assumptions can help us make inference on the inter-temporal dynamics.

---

[16] Computed at individual level, as at aggregate level they are equal by construction.

[17] As the data is on eight years these measures are not complementary.

The AP-CP difference for FM and FS, computed by country, is positive by construction and provides non-trivial information regarding the variability of the respective measures within the time interval. A relatively low monetary variability may be caused by a stable labor market and a relatively low supplementary variability can be sign of established living standards. The data confirm the more active, "catching up" role of poorer countries that present higher variability in both measures and which show higher estimates. This is true for the supplementary but especially for the monetary measure whose higher variability presents analogies with the variability of income with respect to consumption. In Ireland this aspect is even more accentuated, as could be expected.

The AP-CP or the AP/CP measures can be computed for the MP and the LP or the *ratio* MP/LP. This suggests impressive potentialities, however inferring conclusions presents various difficulties, and this line of research has not been investigated beyond the confirmation of previous findings. Substantial research need still be done to explain effects linked to age, gender, territory or any variable present within the ECHP dataset.

Deriving a new variable that incorporates various dimensions at the individual level is difficult to criticize in a multidimensional analysis. The choice of variables presents no general rule of thumb, while the numerical aggregation can be defended in various ways. In this analysis a totally relative approach is used, so that individual deprivation is measured with respect to the whole population of reference. This new variable can be compared with the traditional analysis or with other fuzzy measures, and a study of cross-section and longitudinal properties significantly contributes to identify the underlying phenomenon.

5. - Conclusions

The first result obtained through the use of non-monetary variables is the measurement of the significant difference between monetary income and living conditions. Temporary and permanent

poverty are presented as two different aspects of the same phenomenon and an identification through the use of the different methodologies is proposed. The relative importance between these two aspects is not discussed, but the acknowledgement of these two facets spurs an exploration of the different poverty profiles.

The second result obtained shows how a partial analysis, such as an exclusively monetary one, can significantly distort the results, as the approximation given to the degree of poverty is proportional to the measure itself, and therefore is not independent of income. Verified at national level, this implies that the approximation valid for poor in poor countries is not as valid for poor in rich countries, whom, given the relative nature examined, are also on average richer in absolute terms.

Finally a measure of variability of the monetary and supplementary variable is derived from the balanced sample. The conclusions derived from these measures must be taken cautiously as attrition related issues have a very high impact in this context. An analysis between the balanced and the complete sample has not shown systematic effects, however further analyses are needed.

This research has presented a fuzzy multidimensional methodology that follows relative poverty as the unifying principle through the whole analysis. This line of reasoning is less intuitive then the absolute alternative but, giving importance to the shape of the distribution over the absolute position of its elements, it better describes what is commonly felt to be poverty and is a fundamental requirement for a solid comprehension of poverty analysis. The analysis prefers to work on measurements at individual level and successively aggregate them to compute global measures, as individual values allow for infinite variations on different reference subgroups that are fundamental to identify different poverty profiles in an intuitive fashion.

A picture of the privation in all its aspects must be constructed and monitored, from which we can extract the information to fight poverty in its most extensive meaning, *i.e.* the common sentiment of feeling "incapacitated" to fully participate within a society.

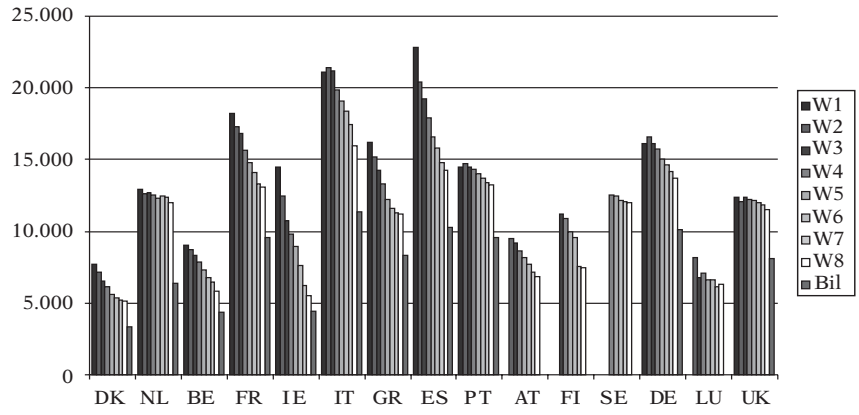
APPENDIX

1.A - Sample Information

The ECHP dataset has been collected by NDUs (National Data Collection Units), generally national statistical institutes, which have autonomously decided the details regarding the sampling procedure in order to best address the reality that each one had to confront. *Ceteris paribus* a nation with very diverse characteristics and complicated patterns to identify will need a larger sample to obtain similar efficiency in the estimates; however as the size of the sample results from separate national decisions rather than from a centralized European one it is difficult to believe that this reason uniquely justifies the sample size. The national sample sizes do not reflect the national proportions of European population and this one of the predominant reasons for the use of ECHP weights[18].

GRAPH 12

YEARLY SAMPLE VARIATION (1994-2001)
COMPARED WITH BALANCED PANEL



[18] W1-W8 correspond to the eight years during which the ECHP has collected information, while Bil shows the number of participants surveyed that have stayed in the sample throughout the eight years.

The difference between the Bil value and the W8 is a measure of the population that has exited and re-entered the sample. The Netherlands seem to have a low attrition, and high re-entry rates as opposed to Ireland which shows a very marked attrition and very low re-entry rates; this could suggest a particularly efficient NDU in following the early exits or a national attitude towards questionnaires. Austria (AT), Finland (FI), Sweden (SE) and Luxemburg (LU) do not have balanced figures as they miss some waves, due to a late entry. The graph also highlights how the initial sample is not proportional to the population of each nation, as Germany has a W1 sample similar to Ireland, when the actual population is twenty-three times that of Ireland.

1.B - Missing Information and Attrition

Statistical analysis is based on the hypothesis that the collected information is complete and correctly reflects reality. Missing data represents a substantial problem that can make the theoretical model quickly drift apart from reality, introducing bias in the calculations. Missing data can be:

i) Complete, if there is no information related to the selected statistical unit.

ii) Partial, if the statistical unit refuses or is incapable of answering a certain questions

iii) Attrition related, if the statistical unit decides to exit the sample

iv) New-entry related, if we have not interviewed the statistical unit the first year

To fix such problems, ECHP imputes the missing data through the use of an elaborate set of weights obtained from those who have answered. The valuation of the impact of partial missing data is very difficult, as ECHP is a complex questionnaire, and the fact that it allows to jump certain questions makes the task even harder. Partial answering is relatively common in questions of economic nature that try to infer the income of a

statistical unit and that is why net or gross income is asked alternatively.

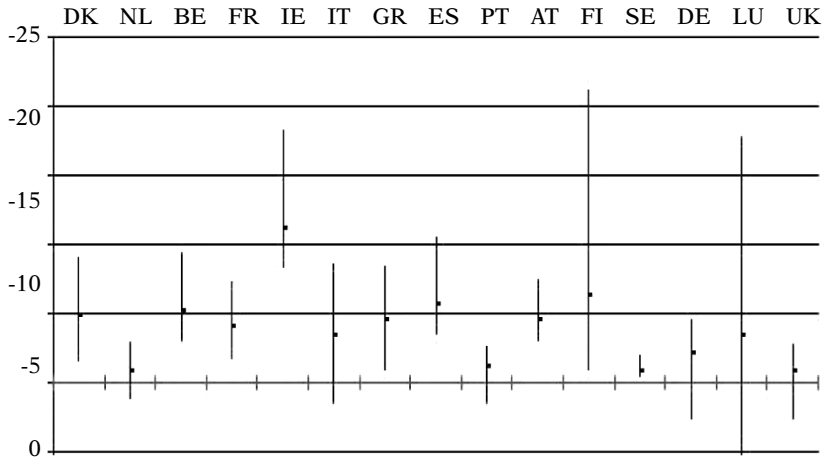
The partial missing information does not cause attrition problems, as imputation is a largely accepted practice, but the imputation relies on the non-verifiable hypothesis that the missing data is random with respect to the covariates that define it, hence the risk of a systematic bias cannot be avoided through imputation (Dempster, Rubin, 1983).

Attrition often can be reduced through the use of split-panel models, where a proportion of the sample is periodically changed. If we look at the missing rates and notice how strongly they differ in different countries we can have a good idea of how quickly a distortion can accumulate its effects. The following graph points out the absolute percentage changes in the sample size and how the maximum, mean an minimum value relate throughout the eight waves. This kind of analysis is brought forward in the complete thesis, but does not present any significantly worrying results.

Maximum values far from the mean indicate outliers which can have different causes and effects then those of sustained attrition. Ireland presents generally high percentage changes, but

GRAPH 13

MAX, MIN AND MEAN YEARLY PERCENTAGE VARIATION



also Finland and Luxemburg must be monitored as a fall in over 15% of the sample could hide some significant distortion to the results.

1.C - The ECHP Weights

The need to weight differently statistical units that compose the sample comes from the different probability of selection, and it can relate to households or individuals. Here are some comments on the automatic use of weights with no contextualization of the objectives:

i) Under certain circumstances, assigning weights to statistical units avoids biased estimates, however this practice is not missing problems of its own (Peracchi, 2002).

ii) To use estimates derived from weighted data can be controversial especially in problems related to particular regression techniques (DuMoushel, Duncan, 1983).

iii) Weighting individuals can be superfluous in situations that have for objective human behavior, instead of a calculated description of the population (Hoem, 1989).

iv) Systematic weights contrary to personalized weights may result in probability values that are outside the logically possible interval (Horowitz, Mansi, 1998).

However weights often will allow us to make meaningful comparisons in a certain temporal instant and their use is widespread. We can further subdivide the weights in three different classes:

i) Design weights: inversely propositional to the probability of selection.

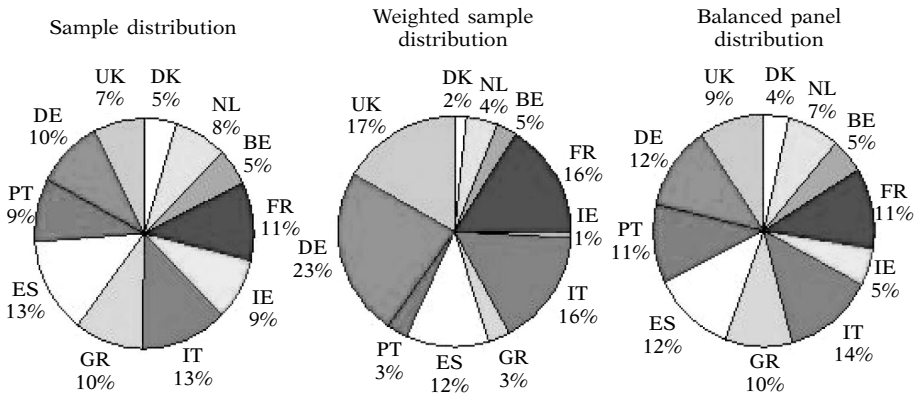
ii) Non-response weights: inversely proportional to the obtained answers by question.

iii) Correction weights: correcting the distribution given some external data.

The weights adopted in this analysis are computed through a step-by-step procedure, and further adjustments have been made to extreme values through EUROSTAT normalization.

GRAPH 14

SAMPLE, WEIGHTED AND BALANCED DISTRIBUTION (1994)



The pie charts in Graph 14 compare the initial sample, with the corresponding weighted sample.

We can notice how the initial sample that does not reflect the population of each nation is successively adjusted. The third pie chart shows the proportion relative to the balanced dataset, and as it seems to be similar to the initial sample, we justify the use of the same weights when discussing results relative to the balanced dataset.

BIBLIOGRAPHY

ATKINSON A.B., «Bringing Income Distribution in from the Cold», presidential address to the Royal Economic Society, *Economic Journal*, March, no. 107, 1996, pages 297-321.

ATKINSON A.B. - BOURGUIGNON F., «Income Distribution and Difference in Needs», in FEIWEL G.R. (ed.), *Arrow and the Foundations of the Theory of Economic Policy*, London, Macmillan, 1987.

BETTI G. - VERMA V., «Measuring the Degree of Poverty in a Dynamic and Comparative Context: A Multi-Dimensional Approach Using Fuzzy Set Theory», *Proceedings of ICCS-VI*, Lahore, Pakistan, no. 11, 1999, pages 289-301.

BETTI G. - CHELI B. - LEMMI A. - VERMA V., «On the Construction of Fuzzy Measures for the Analysis of Poverty and Social Exclusion», *Statistica & Applicazioni*, no. 4, 2006, pages 77-97.

BETTI G. - VERMA V., «Fuzzy Measures of the Incidence of Relative Poverty and Deprivation: A Multi-Dimensional Perspective», *Statistical Methods and Applications*, DOI 10.1007/s10260-007-0062-8, 2008.

CERIOLI A. - ZANI S., «A Fuzzy Approach to the Measurement of Poverty», in DAGUM C. - ZENGA M. (eds.), *Income and Wealth Distribution, Inequality and Poverty*, Studies in Contemporary Economics, Springer, Berlin, 1990.

CHELI B. - LEMMI A., «Totally Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty», *Economic Notes*, no. 24, 1995, pages 115-134.

DEATON A. - MUELLBAUER J., *Economics and Consumer Behaviour*, Cambridge, Cambridge University Press, 1980.

—— - ——, «On Measuring Child Costs: With Applications to Poor Countries», *Journal of Political Economy*, no. 94, 1986, pages 720-722.

DEMPSTER A.P. - RUBIN D.B., «Incomplete Data in Sample Surveys», in MADOW M.G. - OLKIN I. - RUBIN D.B., *Sample Survey*, vol. II, *Theory and Annotated Bibliography*, New York, *Academic Press*, 1983, pages 3-10.

DELEECK H., «Ongelijkheden in de Welvaartstaat - Inequalities in the Welfare State», *De Nederlandsche Boekhandel*, Antwerpen, 1977.

DUMOUCHEL W.H. - DUNCAN G.J., «Using Sample Survey Weights in Multiple Regression Analysis of Stratified Samples», *Journal of Business and Economic Statistics*, no. 7, 1983, pages 441-452.

DUBOIS D. - PRADE H., *Fuzzy Sets and System: Theory and Applications*, Academic Press, New York, 1980.

EBERT U., «Social Welfare when Needs Differ: An Axiomatic Approach», *Economica*, no. 63, 1996, pages 233-244.

ENGEL E., «Die Lebenskosten Belgisher Arbeiter-familien Fruher und Jetzt», *International Statistical Bulletin*, no. 9, 1895.

GINI C., «Measurement of Inequality and Incomes», *The Economic Journal*, no. 31, 1921, pages 124-126.

——, *Variabilità e mutabilità*, 1912, reprinted in PIZZETTI E. - SALVEMINI T. (eds.), *Memorie di metodologica statistica*, Rome, Libreria Eredi Virgilio Veschi, 1955.

HOEM J.M., «The Issue of Weights in Panel Surveys of Individual Behaviour», in KASPRZYK D. - DUNCAN G.J. - KALTON G. - SINGH M.P. (eds.), *Panel Surveys*, Wiley, New York, 1989.

HOROWITZ J. - MANSI C.F., «Censoring of Outcomes and Regressors Due to Survey

non Response: Identification and Estimation using Weights and Imputation», *Journal of Econometrics*, no. 84, 1998, pages 37-58.

KAPTEYN A. - VAN PRAAG B.M.S., «A New Approach to the Construction of Family Equivalence Scales», *European Economic Review*, no. 7, 1976, pages 313-335.

MACK J. - LANSLEY S., *Poor Britain*, Allen & Unwin, London, 1985.

MCCLEMENTS L.D., «Equivalence Scales for Children», *Journal of Public Economics*, no. 8, 1977, pages 191-210.

LAZARSFELD P.F., «Panel Studies», *Public Opinion Quarterly*, no. 4, 1940, pages 122-128.

LORENZ M.O., «Methods of Measuring the Concentration of Wealth», *Publications of the American Statistical Association*, no. 9, 1905, pages 209-219.

PERACCHI F., «The European Community Household Panel: A Review», *Empirical Economics*, Springer, Berlin, no. 27, 2002, pages 63-90.

ROWNTREE B.S., *Poverty: A Study of Town Life*, MacMillan, London, 1901.

SEN A.K., «Poverty: An ordinal Approach to Measurement», *Econometrica*, no. 44, 1976.

— —, «Poor, Relatively Speaking», *Oxford Economic Papers*, no. 35, 1983.

— —, *Commodities and Capabilities*, North-Holland, Oxford, 1985.

SHORROCKS A.F., «Revisiting the Sen Poverty Index», *Econometrica*, no. 63, 1995.

TOWNSEND P., *Poverty in the United Kingdom*, Penguin Books, Middlesex, 1979.

ZADEH L.A., «Fuzzy Sets», *Information and Control*, no. 8, 1965, pages 338-353.