

Alma Mater Studiorum Università di Bologna
Archivio istituzionale della ricerca

On the estimation of fuzzy poverty indices using official survey data. Variance estimation, robustness and computational considerations

This is the final peer-reviewed author's accepted manuscript (postprint) of the following publication:

Published Version:

Crescenzi, F., Mori, L. (2025). On the estimation of fuzzy poverty indices using official survey data. Variance estimation, robustness and computational considerations. JOURNAL OF STATISTICAL COMPUTATION AND SIMULATION, 95(8), 1611-1638 [10.1080/00949655.2025.2465794].

Availability:

This version is available at: <https://hdl.handle.net/11585/1006502> since: 2025-03-07

Published:

DOI: <http://doi.org/10.1080/00949655.2025.2465794>

Terms of use:

Some rights reserved. The terms and conditions for the reuse of this version of the manuscript are specified in the publishing policy. For all terms of use and more information see the publisher's website.

This item was downloaded from IRIS Università di Bologna (<https://cris.unibo.it/>).
When citing, please refer to the published version.

(Article begins on next page)

On the estimation of fuzzy poverty indices using official survey data. Variance estimation, robustness, and computational considerations

Federico Crescenzi (Corresponding author)
Department of Economics, Engineering Society and Business Administration
University of Tuscia, Viterbo, Italy
Via Santa Maria in Gradi, 4, Viterbo, Italy
`federico.crescenzi@unitus.it`

Lorenzo Mori
Department of Statistical Sciences
University of Bologna, Bologna, Italy
Via Belle Arti 41, Bologna, Italy
`lorenzo.mori7@unibo.it`

Abstract

We review the fuzzy approach to poverty measurement by comparing poverty indices having different membership functions proposed in the literature. We simulate two sample surveys from EU-SILC data to examine the statistical properties, sampling errors and robustness to parameter specification of the fuzzy indices considered. Two traditional crisp-set poverty indices (Head Count Ratio and Poverty Gap) are also compared to highlight their close link to the fuzzy approach.

keywords: Sample surveys; fuzzy methods; poverty measurement; jackknife; bootstrap

1 Introduction

The estimation of poverty measures for populations of individuals using sample survey data depends on the choice between two paradigms: the traditional approach, based on crisp-sets, and the fuzzy approach. With the traditional approach, an individual is regarded as poor with respect to a poverty predicate (Y) if he/she belongs to the poor set, usually defined by a rule on Y . The condition of being poor or not-poor with respect to Y is therefore binary. This approach is used for poverty measurement in official statistics. On the other hand, choice of a fuzzy paradigm means not admitting such a binary classification, but rather considering individuals to be more (or less) deprived than others with respect to Y . The use of fuzzy indices to study economic inequalities is discussed in many papers, presented below, and more generally in two reference manuals on poverty: Alkire et al. (2015) and Silber (2023).

It is not the purpose of this paper to discuss the theory or philosophy of fuzzy and crisp paradigms, since the differences between these points of view have already been widely discussed (*e.g.* in Lemmi and Betti 2006; Betti and Lemmi 2008; Garrido 2012; Betti and Lemmi 2021). Rather, we show that some close links exist between certain fuzzy poverty indices and (official) crisp-set indices if we assume a step membership function and consider its expected value. With appropriate parameter specification, some fuzzy indices can be point-wise indistinguishable from official statistical measures, such as the Head Count Ratio (HCR)¹ and the Poverty Gap (PG), while preserving the fuzzy paradigm.

¹In official statistics, the HCR is also referred to as the At Risk of Poverty Rate (ARPR) and it is included on the Laeken indices. However, for consistency with the fuzzy literature, we prefer to call it HCR.

The aim of this paper is to focus on the estimation of fuzzy poverty indices (or measures) from a statistical estimation perspective, building on and going beyond the analysis conducted by Betti et al. (2018) and Kim (2015). In fact, to the best of our knowledge, the literature on fuzzy methods for poverty estimation – and for fuzzy indices of socio-economic deprivation in general – have not yet focused on the statistical properties of these indices, either in terms of their bias or in terms of their sampling errors, and ultimately in terms of their robustness to parameter specification. As previously mentioned, the only two studies that focus on the statistical properties of a fuzzy index are those proposed by Betti et al. (2018) and Kim (2015). However, they only consider two fuzzy indices and focus specifically on the relationship between unidimensional and multidimensional indices. On the other hand, in the traditional approach to poverty and inequality measurement, there is much literature on the issue of variance estimation of poverty and inequality indices (see *e.g.* Kakwani 1993; Duclos et al. 2006; Barabesi et al. 2016; Walter et al. 2022; Barabesi et al. 2024).

Comparison of fuzzy poverty indices is not straightforward because each depends on a so-called membership function that ranks individuals in terms of deprivation with respect to the poverty predicate. The membership function represents the researcher’s view on the poverty phenomenon, so different functions are intrinsically difficult to compare. Moreover, authors often suggest different synthetic measures over the function, making comparison even harder. Our analysis tackles these issues by addressing the Mean Square Error (MSE) of these measures using sample data. In fact, while an approach may be appealing in theory, when it ultimately comes to producing estimates from samples, it is necessary to have estimators with good properties. This however requires a common sample statistic. We consider the expected value of the membership function since it is the statistic reported by many (though not all) fuzzy methods (*e.g.* Betti et al. 2006), and it is applicable to all membership functions.

Here we use a real sample of individuals as synthetic population: the 2019 European Union Statistics on Income and Living Conditions (EU-SILC) microdata from Italy, a primary source for the calculation of the Laeken indices. We then simulate two different sample surveys, the first uses simple random sampling of individuals, the second a two-stage sampling design that resembles that of most household surveys in Europe. We evaluate the bias, the bias in the estimator of the MSE, and the Coefficient of Variation (CV) of our fuzzy indices. To estimate the MSE, we consider a bootstrap approach and an ad-hoc jackknife repeated replications method, as suggested by Betti et al. (2018). For those that depend on parameters, we assess their robustness with respect to mis-specification.

The remainder of the paper is structured as follows. Section 2 offers a brief sketched history of poverty measurement for new readers, not intended as a comprehensive literature review of fuzzy poverty estimation. Section 3 discusses and reviews the analytical details of the fuzzy approach to poverty measurement. Section 4 introduces the simulation set-up and introduces the bootstrap and ad-hoc jackknife approach we use in Section 5. Section 6 discusses the robustness of the fuzzy indices. Section 7 concludes the paper and offers insights for further research.

2 From the poverty line to the fuzzy indices: a brief history of poverty measurement

It is customary to trace the origin of the poverty line back to the work published by Mollie Orshansky in the 1960s, but the concept originates well before, namely in 1887, when Charles Booth presented the results of a survey conducted in London to a meeting of the Royal Statistical Society. This was the very first use of the “line of poverty” concept (Fisher, 1992; Gillie, 1996; Fisher, 1997; Allen, 2013). Introduction of the poverty line made it possible to dichotomize the population into two non-overlapping sets: the set of the poor, those whose incomes do not exceed the threshold, and the set of the not-poor.

Well-known indices for the measurement of poverty referring to income (or consumption) and to the definition of a poverty line have since been suggested in the literature (*e.g.* Sen 1976; Foster 1983; Atkinson 1987; Hagenaars 1987; Pyatt 1987). However, as already argued by Sen (1973), the notion

of poverty – and also inequality – based on this dichotomization may be inherently imprecise. To allocate a person to one of the two sub-sets according to whether his income falls below the poverty line, eliminates differences in the degree of poverty that exist naturally between people. Rather, a gradual transition from severe poverty to wealth seems more realistic. Moreover, the evaluation of individual income is often imprecise, mostly due to the unwillingness of respondents to provide exact information.

In the early 1990s, at the core of the debate on the poor/not-poor dichotomy in Italy, Cerioli and Zani (1990) published what is considered the seminal work on fuzzy theory in poverty measurement based on the fuzzy sets introduced by Zadeh (1965). For the first time a subset in the domain of income is defined where individuals belong to the poor set only to a certain degree. The fuzzy approach had great resonance in Italy, and soon spread beyond its borders. Shortly after, Cheli and Lemmi (1995), proposed the first fuzzy approach based on a distributional membership function. According to this approach, which is introduced formally in Section 3, only the poorest (richest) population unit has a degree of belonging to set of the poor (not-poor) equal to 1 (0), while all other units belong to degrees between 0 and 1. Aiming to modify the index proposed by Cheli and Lemmi (1995), Betti and Verma (1999), noted that poverty analysis based on a poverty line tends to overestimate the extent of mobility near the poverty line itself, and to underestimate mobility far from it.

Qizilbash (2003) again highlighted that unequivocal classification of individuals into the two distinct categories of poor and not-poor was challenging, as it involves an intricate interplay of various socio-economic factors that make the demarcation line between the two groups uncertain, and is further complicated by the Sorites Paradox. This paradox states that a seemingly insignificant incremental change in a particular attribute should not, in theory, suffice to alter one’s poverty status. However, when considered iteratively, a series of such minute alterations may cumulatively lead to the transition of an individual from poverty to non-poverty. The study of multidimensional poverty under the fuzzy paradigm originated in those years (Betti et al., 2006; Belhadj and Limam, 2012; Betti and Lemmi, 2021; Gangopadhyay et al., 2021).

Although changing from a poverty line cutting a population into two sets to a membership function assigning each person a poverty degree may seem intriguing, it has posed theoretical problems that researchers are still trying to answer. Chakravarty (2019) notes the importance of basing an index on axioms, and consequently proposed an index based on five properties that are shown below. Harking back to their working paper dated 1994, Shorrocks and Subramanian (2024) highlighted that like their crisp counterpart, fuzzy indices have crucial properties such as monotonicity, transfer preference, decomposability and scale invariance.

Approximately 30 years later, the major problem of the fuzzy paradigm seems the lack of an appropriate definition of the membership function (Cerioli and Zani, 1990). In this sketched timeline, we try something that has never been attempted before. We do not propose a new membership function, but analyse prominent indices in terms of their statistical estimation properties, robustness and specification of parameters.

3 The fuzzy approach to poverty measurement

The concept of membership function lies at the core of the fuzzy approach to poverty measurement. Given a poverty predicate Y (e.g. equivalized disposable income or expenditure), its support $\mathcal{Y}$, and the (fuzzy) set $\mathcal{A} = [0, 1] \subset \mathbb{R}$, a membership function $\mu : \mathcal{Y} \rightarrow \mathcal{A}$ is a mapping of Y to $\mathcal{A}$. In particular, for an individual predicate y_i , we denote its membership function as $\mu(y_i)$. The belonging to fuzzy sets is then quantified by the membership grade $a \in [0, 1] \subset \mathbb{R}$. The higher the value of a , the greater the membership of the unit to the (poor) set (Zadeh, 1965; Klir and Yuan, 1995). In other words, the greater the membership, the more the unit is considered poor with respect to the predicate.

The traditional approach to poverty measurement is instead based on crisp-sets theory, where the population is divided into the subset of the poor and the subset of the not-poor according to a rule on $\mathcal{Y}$. For example, in official statistics, households or individuals are regarded as poor (the remaining

as not-poor) whenever $y_i < z$, where $y_i \in \mathcal{Y}$ denotes the i -th household or individual level of Y and $z \in \mathcal{Y}$ is the poverty line.

The different types of membership function can be broadly divided into distance-based and distribution-based membership functions. In what follows, to maintain consistency in the notation, we assume without loss of generality that the values of the poverty predicate are arranged in ascending order.

3.1 Distance-based membership functions

The distance-based membership function is a piecewise function of the following kind

$$\mu(y_i) = \begin{cases} 1 & 0 < y_i < z_1 \\ f(y_i) & z_1 \leq y_i < z_2 \\ 0 & y_i \geq z_2 \end{cases},$$

where $f(y_i)$ is a positive decreasing function defined for each i in the sample with $i = 1, \dots, n$, and $z_1, z_2 \in \mathcal{Y}$. For each z_1, z_2 there exists a value a of the membership function, the a -cut, that transforms the fuzzy set into a crisp set (Klir and Yuan, 1995). Given a fuzzy set $\mathcal{A}$ and any number $a \in [0, 1] \subset \mathbb{R}$, the a -cut ($^a\mathcal{A}$) is a crisp set defined as: $^a\mathcal{A} = \{y \mid \mathcal{A}(y) \geq a\}$. That is, the a -cut of a fuzzy set $\mathcal{A}$ is the crisp set containing all the elements of the universal set $\mathcal{Y}$ whose membership grades are greater than or equal to the specified value of a . All the indices presented below admit at least two values of $a \in [0, 1] \subset \mathbb{R}$, ensuring the creation of at least two crisp sets: one representing the absolute poor and one representing the absolute non-poor.

The literature suggests several membership functions depending on the definition of the function $f(\cdot)$. According to Chiappero-Martinetti (2006), distance-based membership functions can be further classified into triangular, trapezoidal and Gaussian. In triangular functions, $f(\cdot)$ is expressed as the combinations of two different functions $f_1(\cdot)$ and $f_2(\cdot)$. In trapezoidal functions, $f(\cdot)$ takes the form $\frac{z_2 - y_i}{z_2 - z_1}$, where z_1 and z_2 are two threshold values. Linear functions can be viewed as a case of the trapezoidal function where z_1 and z_2 are the maximum and minimum values of Y . For Gaussian functions $f(\cdot)$ is expressed in the form of a Gaussian distribution. To the best of our knowledge, there are no Gaussian functions for fuzzy poverty indices.

In chronological order, the most widely used distance-based functions are those proposed in Cerioli and Zani (1990), Belhadj (2011), Zedini and Belhadj (2015), Belhadj (2015), and Chakravarty (2019).

3.1.1 Cerioli and Zani (1990)

The first proposal of a membership function was by Cerioli and Zani (1990) who suggested a trapezoidal function of the kind

$$\mu(y_i) = \begin{cases} 1 & 0 < y_i < z_1 \\ \frac{z_2 - y_i}{z_2 - z_1} & z_1 < y_i < z_2 \\ 0 & y_i \geq z_2 \end{cases}, \quad (1)$$

where the values z_1, z_2 need to be specified by the researcher, and a linear function is used for all intermediate values. An individual therefore belongs to the poor set completely when $\mu(y_i) = 1$, is certainly not poor if $\mu(y_i) = 0$, and shows partial membership to the set for all intermediate values.

3.1.2 Belhadj (2011)

Belhadj (2011) suggested the following membership function

$$\mu(y_i) = \begin{cases} 1 & 0 < y_i < z_{\min} \\ \frac{-y_i}{z_{\max} - z_{\min}} + \frac{z_{\max}}{z_{\max} - z_{\min}} & z_{\min} \leq y_i < z_{\max} \\ 0 & y_i \geq z_{\max} \end{cases}, \quad (2)$$

where $z_{\min}$ and $z_{\max}$ are defined as the upper and lower poverty lines. The upper poverty line corresponds to the level of total expenditure per-capita necessary for households to meet their basic food needs without sacrifice, while the lower poverty line includes the minimum expenditure necessary to meet basic food and non-food needs (see Belhadj (2011) for more details). Remarkably, this membership function is a special case of Equation (1) for given values of the threshold parameters. Henceforth we therefore make no further distinction between Equations (1) and (2).

3.1.3 Zedini and Belhadj (2015)

Zedini and Belhadj (2015) proposed the following trapezoidal membership function based on the previous work of Belhadj and Matoussi (2010). Their aim was to add a novel approach that gives the application of fuzzy approaches a non-parametric perspective. The function is

$$\mu(y_i) = \begin{cases} 1 & b \leq y_i < b' \\ \frac{-y_i}{b'' - b'} + \frac{b''}{b'' - b'} & b' \leq y_i < b'' \\ 0 & y_i < b \cup y_i \geq b'' \end{cases}, \quad (3)$$

where b, b', b'' are percentiles estimated by the bootstrap over a series of fuzzy latent poverty states. The authors² create a set of 100 membership functions corresponding to 100 fuzzy sets. The first membership function has $b = 0$, $b' = \hat{p}_1$ and $b'' = \hat{p}_2$; where $\hat{p}_1$ and $\hat{p}_2$ are the bootstrap estimates of the first and second percentiles of the poverty predicate variable, respectively. For the membership function from 2 to 99 they suggest $b = \hat{p}_{j-1}$, $b' = \hat{p}_j$ and $b'' = \hat{p}_{j+1}$ with $j = 2, \dots, 99$. The last membership function for $j = 100$ is defined with $b = \hat{p}_{99}$, $b' = 0.5(\hat{p}_{99} + \hat{p}_{100})$ and $b'' = \hat{p}_{100}$. The final membership function is then obtained by the cardinality approach and a divisive hierarchical algorithm.

Here it is important to note that in the same paper, the authors, proposed a specific index, which differs from the mean, to estimated with this membership function.

3.1.4 Belhadj (2015)

Belhadj (2015) suggested the following trapezoidal membership function

$$\mu(y_i) = \begin{cases} 1 & y_i < z_1 \\ \mu_1 = 1 - \frac{1}{2} \left(\frac{y_i - z_1}{z_1} \right)^\beta & z_1 \leq y_i < z^* \\ \mu_2 = 1 - \frac{1}{2} \left(\frac{z_2 - y_i}{z_2} \right)^\beta & z^* \leq y_i < z_2 \\ 0 & y_i \geq z_2 \end{cases}, \quad (4)$$

where z^* is the flex point of μ and β is a shape parameter ruling the degree of convexity of the function. The trend is linear when $\beta = 1$.

3.1.5 Chakravarty (2019)

Chakravarty (2019) introduced an axiomatic approach based on five important axioms that a membership function must fulfil, namely, Homogeneity of Degree Zero (A1), Linear Decreasingness (A2),

²For the sake of conciseness, we only report and outline the procedure here, referring readers to Zedini and Belhadj (2015) for full details.

Continuity (A3), Maximality (A4) and Independence of Non-meagre Attribute Quantities (A5). On the basis of such axioms, the author proved that the following triangular membership function:

$$\mu(y_i) = \begin{cases} 1 & y_i = 0 \\ \frac{z_2 - y_i}{z_2} & 0 < y_i < z_2, \\ 0 & y_i \geq z_2 \end{cases} \quad (5)$$

is the only membership function that satisfies all the axioms. In particular, A1 ensures that the function remains unaltered under equi-proportionate variations in quantities of attribute Y . A2 makes a specific assumption about the decreasingness of the membership function. A3 means that μ should vary in a continuous manner with respect to variations in attribute quantities. A4 specifies that μ should achieve its maximum value 1 when the level of the attribute is zero. Finally, (A5) shows insensitivity of μ to the attribute quantities of persons who are certainly non-poor in the attribute through the assumption that the value of the membership function on $[z_2, +\infty]$ is constant.

3.2 Distribution-based membership functions

3.2.1 Cheli and Lemmi (1995)

In the category of distribution-based membership functions, Cheli and Lemmi (1995) proposed the “Totally Fuzzy and Relative” (TFR) approach to defining the membership function for a monetary poverty predicate as:

$$\mu(y_i) = (1 - F_Y(y_i))^\alpha = \left(\frac{\sum_{j>i}^n w_j |y_j > y_i|}{\sum_{j>1}^n w_j |y_j > y_1|} \right)^\alpha, \quad i \in \{1, \dots, n-1\}, \quad \mu(y_n) = 0 \quad (6)$$

where F_Y is the cumulative distribution function of Y calculated for the i -th individual and $\alpha \geq 1$. The definition states that the degree of unit i to belong to the fuzzy set of poor people depends on the proportion of the remaining units that have a value of Y greater than y_i . If Y is income, $\mu(y_i)$ is the proportion of units having an income greater than the unit considered. The authors introduced the parameter α to facilitate comparison with conventional HCR, making it possible to benchmark the fuzzy to the traditional poverty index. Indeed, the α parameter is set so that $E[\mu]$ equals the value of an official statistic (*e.g.* HCR), usually estimated on the same sample.

3.2.2 Betti and Verma (1999)

Starting from the TFR approach, Betti and Verma (1999) suggested replacing the cumulative distribution function of (Y) with the Lorenz curve (L):

$$\mu(y_i) = (1 - L_Y(y_i))^\alpha = \left(\frac{\sum_{j>i}^n w_j y_j |y_j > y_i|}{\sum_{j>1}^n w_j y_j |y_j > y_1|} \right)^\alpha, \quad i \in \{1, \dots, n-1\}, \quad \mu(y_n) = 0 \quad (7)$$

where w_j is the sampling weight of statistical unit j . The parameter α , $\alpha \geq 1$, is usually chosen so that the average fuzzy measure equals an estimate of the HCR of the population. The updated specification offers several benefits. Specifically, $1 - L_Y(y_i)$ is expected to provide a more sensitive measure of income disparities than $1 - F_Y(y_i)$.

3.2.3 Betti, Cheli, Lemmi and Verma (2006)

Betti et al. (2006) considered both the proportion of individuals less poor than the person concerned, $1 - F_Y(y_i)$, and the share of the total equivalized income received by all individuals less poor than the person concerned in proposing the following membership function

$$\begin{aligned}\mu(y_i) &= (1 - F_Y(y_i))^{\alpha-1} (1 - L_Y(y_i)) = \\ &= \left(\frac{\sum_{j>i}^n w_j |y_j > y_i|}{\sum_{j>1}^n w_j |y_j > y_1|} \right)^{\alpha-1} \left(\frac{\sum_{j>i}^n w_j y_j |y_j > y_i|}{\sum_{j>1}^n w_j y_j |y_j > y_1|} \right), \quad i \in \{1, \dots, n-1\}, \quad \mu(y_n) = 0, \quad (8)\end{aligned}$$

where w_j is the sampling weight of statistical unit j . The parameter α , $\alpha \geq 1$, is computed as before. Equation (8) integrates the TFR approach of Equation (6) with information added by use of the Lorenz curve, $1 - L(y_i)$, which is linked to the Gini index. Because the membership function is based on elements of the Lorenz curve, the mean of the membership functions can be expressed in terms of the generalized Gini measure G_α : $HCR = (\alpha + G_\alpha)/(\alpha(\alpha + 1))$.

3.3 A further distinction between membership functions: the economic perspective

Sen's classical economic classification divides economic indices into positive and normative indices (Sen, 1973). Positive indices have a completely objective definition (*e.g.* Theil index, Gini index). Indices are defined as normative if their definition is somewhat subjective (*e.g.* Atkinson index). This distinction may also hold for fuzzy membership functions. As shown below, membership functions can be viewed as positive or normative, depending on how their parameters are defined, *i.e.* whether or not they are objective.

Among positive membership functions we have: Equation (1) in the special case where $z_1 = \min(y_i)$ and $z_2 = \max(y_i)$. Equation (6), Equation (7) and Equation (8) in all cases where parameter α is chosen so that the expected value of the membership function equals the HCR. Equation (3), like the membership function of Zedini and Belhadj (2015), when it has three different parameters b, b' and b'' to be estimated with the bootstrap.

Among normative membership functions we have Equation (1). Equation (4) and Equation (5) as the choice of the parameters is somehow subjective; Equation (2) as the values are estimated using a regression model, which involves subjectiveness in the choice of variables.

3.4 A note on the relationship with crisp-set indices

It is interesting to compare the fuzzy indices discussed in the previous sections with the standard crisp-set indices that are used to produce official statistics. We select two widely used indices that belong to the class of the Foster, Greer and Thorbecke index (FGT) (Foster et al., 1984). In a population set-up, with N units, the indices in this class have the following formulation

$$FGT(\zeta) = \frac{1}{N} \sum_{i=1}^N \left(\frac{z - y_i}{z} \right)^\zeta$$

where z is the poverty line, and $\zeta \geq 0$ is a parameter.

The first index that we consider is the widely used HCR which is obtained as $FGT(0)$, that is: $HCR = \frac{P}{N}$. P is the number of people whose income is below the poverty line z . The second index that we consider is the PG or $FGT(1)$. $PG = \frac{1}{N} \sum_i \frac{(z - y_i)I(y_i \leq z)}{z}$, where $I(\cdot)$ is the indicator function. In terms of their interpretation, HCR denotes the proportion of people in the population that live under conditions of poverty, while PG is the ratio by which the mean income of the poor falls below the poverty line.

The sample counterparts of these two indices are

$$\widehat{HCR} = \frac{\sum_{i=1}^n w_i I(y_i \leq z)}{\sum_{i=1}^n w_i} \quad \text{and} \quad \widehat{PG} = \sum_{i=1}^n \frac{w_i (z - y_i) I(y_i \leq z)}{z}, \quad (9)$$

where w_i is the sample weight of unit i , and n is the sample size. Depending on the sampling design, estimation of the MSE of the indices given in Equation (9) may involve re-sampling procedures like

bootstrap, jackknife repeated replications (Fabrizi et al., 2008), or linearization techniques (Verma and Betti, 2011; Oguz Alper and Berger, 2015).

It is interesting to note that the crisp-set indices given in Equation (9) can be obtained by assuming a membership function of the kind $\mu(y_i) = 1$ if $y_i < z$ and 0 otherwise. Also, note that (9) is equivalent to the average of the index given in Equation (5) (applying Equation (10)) if we let $z_2 = z$.

4 Simulation set-up: data, measures of performance and MSE estimators

4.1 Data and measures of performance

In order to analyse all the different approaches introduced in Section 3, we explore two different scenarios by simulating two different sample surveys from the same population $\mathcal{P}$, i.e. that repeated by the data of the 2019 Italian EU-SILC survey, consisting of $N = 22,186$ households.

The first scenario (hereafter “SRS”) is that of a survey where each sample is selected by simple random sampling without replacement scheme. Each household therefore has an equal inclusion that results in a sampling fraction (n/N) of approximately 2%.

The second scenario is that of a slightly more complex survey design (hereafter “Complex”) that defines strata and primary sampling units (PSUs). The official EU-SILC sampling design is a two-stage random sampling with stratification and distinction between self-representing strata (SR) and non-self-representing strata (NAR). We mimic it as follows. In SR strata, sampling is conducted in one stage, sample households being selected randomly. For each municipality, we selected 5% of the households, with a minimum of 2 households per municipality. In NSR strata, in the first stage, we selected two municipalities from each region, where the probabilities of inclusion are proportional to the size of the municipality. In the second stage, a sample of 4 households was randomly selected from each PSU.

Table 1 – Membership functions and parameter settings in population (Y_p denotes the p -th quantile of the empirical distribution of Y).

Function	Parameters	Values
Belhadj (2015)	z_1	$Y_{0.05}$
	z_2	$Y_{0.95}$
	β	2
Cerioli and Zani (1990)	z_1	$Y_{0.05}$
	z_2	$Y_{0.95}$
Chakravarty (2019)	z_2	$Y_{0.50}$
Cheli and Lemmi (1995)	α	$\sum_{i \in \mathcal{P}} 1(y_i \leq z)/N$
Betti and Verma (1999)	α	$\sum_{i \in \mathcal{P}} 1(y_i \leq z)/N$
Betti et. al., (2006)	α	$\sum_{i \in \mathcal{P}} 1(y_i \leq z)/N$

A key fact to note here is that the comparison over all indices is done by taking the average of the membership functions over the population as the true value. The corresponding sample statistic that we consider is therefore

$$\hat{H} = \frac{1}{\sum_{i=1}^n w_i} \sum_{i=1}^n \mu(y_i) w_i, \quad (10)$$

where w_i is the sample weight of unit i . However, it often happens that rather than estimating a value for the whole population, the interest is in producing estimates for given sub-domains³. In this

³Sub-domains do not necessarily need to be geographical areas, though such are what we explore in this paper.

case we take the 20 Italian Regions divided by the 15 metropolitan areas. For the SRS simulation, the minimum sample size per area is 2, the maximum 54, and the median 14. For the Complex simulation 4 30, and 14, respectively.

These sample sizes are intentionally small, in line with those used in small area estimation simulations (*e.g.* Gonzalez-Manteiga et al., 2008, Chandra et al., 2012 and Liu et al., 2022), since the issue of estimating fuzzy measures also occurs in this context (see for example Betti et al. (2024)).

We evaluate each fuzzy approach in terms of bias, MSE and coefficient of variation of the statistic given in Equation (10). Let $\bar{H}_{j^*}$ be the true area parameter in the j^* -th area, with $j^* = 1, \dots, J^*$, and $\hat{H}_{j^*t}$ its estimate in replicate ($t = 1, \dots, T$, with $T = 500$), the bias of the estimates for each area is

$$Bias_{j^*} = \frac{1}{T} \sum_{t=1}^T (\hat{H}_{j^*t} - \bar{H}_{j^*}).$$

To evaluate the reliability of the estimates we use the Coefficient of Variation and the Coefficient of Variation of the Second order (CV2), a more robust version of CV (Kvålseth, 2017) which ranges in the interval $[0, 1]$. Since, Eurostat (2013) discourages the use of relative measures for percentages or proportions, if possible we prefer to avoid them, as the denominator could be very low, resulting in a high ratio. For a generic area j^* , the CV and the CV2 are

$$CV_{j^*} = \frac{1}{T} \sum_{t=1}^T \frac{\sqrt{\widehat{MSE}(\hat{H}_{j^*t})}}{\hat{H}_{j^*t}} \quad \text{and} \quad CV2_{j^*} = \sqrt{\left(\frac{CV_{j^*}^2}{1 + CV_{j^*}^2} \right)}$$

where $\widehat{MSE}(\hat{H}_{j^*t})$ is the estimated MSE at replicate t . To evaluate the MSE estimation procedure (see sections below), we report the bias of the MSE (BMSE)

$$BMSE = \frac{1}{T} \sum_{t=1}^T (\widehat{MSE}(\hat{H}_{j^*t}) - MSE(\bar{H}_{j^*t})),$$

where $MSE(\bar{H}_{j^*t})$ is the true MSE at replicate t for area j^* . Note that in the following, the values of all performance measures are reported at national and at the area level. As usual, see *e.g.* Tzavidis et al. (2018), at area level, the values of the performance indices are computed as the mean of the indices over the areas. To aid interpretation we also provide a figure with specific area values for each performance measure. The indices are computed with the parameters setting in Table 1.

4.2 Mean Squared Error estimation procedures

In order to estimate the MSE of a sample statistic the jackknife repeated replications and the bootstrap are two widely used methodology in sample surveys. The main difference between the two is that the first one is definable only if all the sample information are available - primary selection units and strata - while the second one is a simplified version which does not necessarily need the sampling information (Field and Welsh, 2007; Alfons and Templ, 2012).

4.3 Non-parametric Bootstrap

The bootstrap is probably the best known re-sampling method for estimating the MSE of a sample statistic. There are many extensions and variants of the bootstrap, which may be parametric or non-parametric. We focus on the non-parametric version of the bootstrap since it is the one provided with the R-package Laeken (Alfons and Templ, 2012) for estimating sampling errors of crisp indices. The procedure consists in sampling with replacement R samples of size equal to the original sample size, calculating a given statistic S_r for each bootstrap sample. The MSE of the bootstrap replicates is then used as an estimate of the unknown MSE of the statistic of interest. The bootstrap MSE estimator is based on the idea that inference about a population from sample data can be modelled by

re-sampling the sample data and performing inference about a sample on the re-sampled data. One of the main problems with the non-parametric bootstrap is determining the number of replicates to use (Chwila and Żądło, 2020). In the following analysis, we use $R = 200$ replicates.

4.4 Jackknife Repeated Replications

Jackknife repeated replication provides a versatile and straightforward computational technique for MSE estimation. Sometimes the survey may be designed so that there exists information on strata, PSUs and computational units. In this situation, the jackknife approach may be an invaluable tool to estimate the MSE of a given sample statistic.

Betti et al. (2018) extended this method to fuzzy poverty measures. Briefly, standard jackknifing involves the following. Let S be a full-sample estimate of any complexity. We use the subscript i to indicate a sample PSU and h to indicate its stratum. Let $S_{(hi)}$ be the estimate produced using the same procedure after eliminating primary unit i in stratum h and increasing the weight of the remaining units in the stratum by an appropriate factor g_h . Let $S_{(h)}$ be the simple average of the $S_{(hi)}$ over the sample units in h . The VAR of S , when the estimator is unbiased, is then estimated as: $VAR(S) = \sum_h \left[(1 - f_h) \times \sum_i g_{(hi)} (S_{(hi)} - S_{(h)})^2 \right]$, where $(1 - f_h)$ is the finite population correction, which is approximately equal to 1 for samples in typical social surveys. Verma and Betti (2011) propose to setting

$$g_{hi} = \frac{w_h}{w_h - w_{hi}} \quad (11)$$

where $w_h = \sum_i w_{hi}$, and $w_{hi} = \sum_j w_{hij}$. This means that in each replication h_i , the weights of individual units are redefined and re-scaled as follows:

$$w_{hij}^1 = \begin{cases} w_{hij} & j \notin h \\ w_{hij}^1 = g_{(hi)} \times w_{hij} & j \in h, j \neq i \\ w_{hij}^1 = 0 & j \in h, j = i \end{cases}$$

Note that $g_{(hi)}$ in Equation (11) is in line with the jackknife MSE estimation for unequal probability sampling proposed by Berger and Skinner (2005).

5 Results

This section reports the results of the two surveys: SRS and Complex. For the first case, we only report estimates of MSE that use the bootstrap, as the jackknife is better suited to more Complex-designs.

5.1 Bias and coefficients of variation

5.1.1 Simple random sampling design

Table 2 reports the mean values of bias, CV and CV2 computed over the areas and at national level for all indices. At national level, we notice that the amount of bias is negligible for almost all indices. This is not surprising as we are considering an estimator of the kind seen in Equation (10). Being unbiased estimators, we may proceed with the variance rather than using the more comprehensive definition of MSE, but to maintain consistency throughout the paper, we continue to refer to it as the MSE.

Figure 1A shows the specific area bias for all the indices considered. Once again, the bias is very low in all areas for all indices and shows similar values for similar indices. For example, those of Cerioli and Zani (1990) and Chakravarty (2019) have similar bias, as do the indices of Cheli and Lemmi (1995), Betti and Verma (1999) and Betti et al. (2006). As well, Figure 1A shows that even for domains with few sampled units, the estimators are approximately unbiased. Comparing the fuzzy

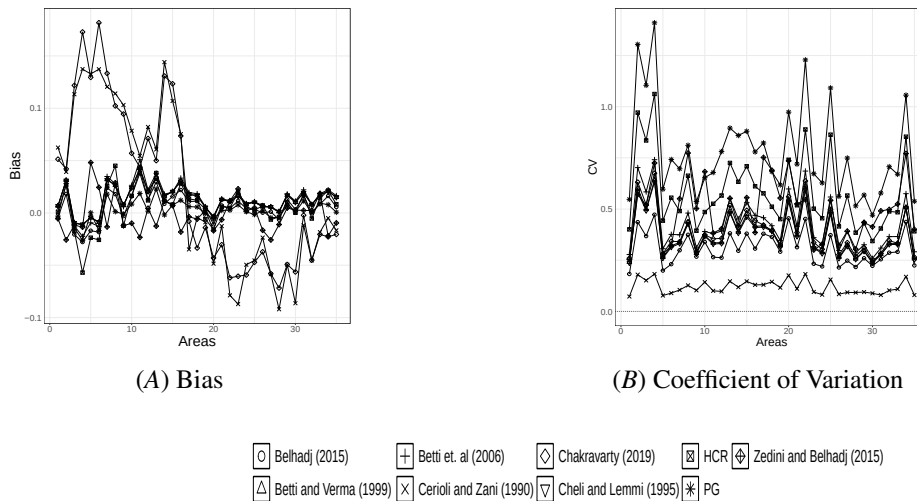
indices with the crisp ones, we see that, the estimation of crisp indices is on the whole less subject to bias than most of the fuzzy indices. At national level, the estimated CV and CV2 seem rather accurate, whereas at area level, Table 2 suggests that the fuzzy indices perform better than the crisp ones (see also Betti et al. (2018)). Figure 1B shows that the distribution-based indices perform almost identically in terms of CV. In the end, it seems that the crisp-set indices attain higher CVs than the fuzzy indices, and that in the fuzzy category, the index of Cerioli and Zani (1990) is preferable in terms of accuracy of the estimates.

With regard to the CVs, it is important to note two additional points. First, the values of these measures are influenced by the accuracies of the variance (numerator) and index (denominator) estimators. Indeed we observe that the CVs do not have a clear descending trend when sample size increases. This is due to the high positive bias of the estimator for almost every index when sample size is low.

Table 2 – SRS simulation: Bias, Coefficient of Variation and, Coefficient of Variation of second order. The "area" measures are computed as the mean over all area values.

	Belhadj (2015)	Cerioli and Zani (1990)	Chakravarty (2019)	Cheli and Lemmi (1995)	Betti and Verma (1999)	Betti et. al (2006)	Zedini and Bel- hadj (2015)	HCR	PG
Bias %									
Areas	0.624	2.243	2.390	1.343	1.350	1.351	-0.188	0.632	0.294
national	0.159	-0.040	-0.064	0.673	0.673	0.673	-0.083	0.673	0.237
CV %									
Areas	30.852	11.884	39.344	37.335	37.941	42.727	49.736	58.123	77.516
National	4.724	1.865	6.062	2.523	2.551	2.467	20.299	2.803	4.933
CV2 %									
Areas	29.239	11.782	36.074	34.508	34.978	38.615	43.616	48.936	59.554
national	4.719	1.864	6.051	2.551	2.550	2.466	19.893	2.802	4.927

Figure 1 – SRS simulation: Bias and Coefficient of Variation by area. Areas are sorted by increasing sample size.



5.1.2 Complex-design

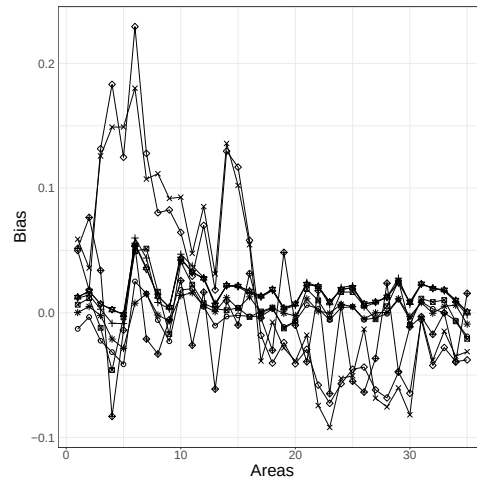
We repeated the same analysis for surveys of Complex-design. Table 3 reports bias, CV and CV2, this time computed by the bootstrap and jackknife estimators of MSE. Once again, the bias is small and slightly positive for every index, at both national and area level, except for the indices of Zedini and Belhadj (2015) and Belhadj (2015) that show negative bias.

Figure 2A reports the bias of the indices by area. The bias clearly tends to zero if the sample size increases. The trends are similar to those observed in the previous simulation. At national level the CVs (CV2s) of the indices are low with both MSE estimators. The Zedini and Belhadj (2015) index is the only one that shows a consistent difference with the CV obtained with the bootstrap and the jackknife. Moving at area level the CVs increase consistently for each index. Figure 3 shows similar performances of the bootstrap and jackknife methods, the latter giving estimates of CV/CV2 slightly lower than those obtained with the bootstrap. The results suggest that fuzzy indices perform better than crisp indices in terms of CV/CV2. The Cerioli and Zani (1990) index is the one with the lowest CV, as seen for SRS case. The Chakravarty (2019) index performs like the distribution-based indices and worse than the Belhadj (2015) index. The figure shows that the CVs for the fuzzy indices tend to be lower than those of crisp-set indices. This result can be seen as a generalization of the findings obtained by Betti et al. (2018), which only considered the HCR and the index proposed by Betti et al. (2006).

Table 3 – Complex simulation: Bias, Coefficient of Variation and Coefficient of Variation of second order by bootstrap and jackknife estimates. The "area" measures are computed as the mean over all area values.

	Belhadj (2015)	Cerioli and Zani (1990)	Chakravarty (2019)	Cheli and Lemmi (1995)	Betti and Verma (1999)	Betti et. al (2006)	Zedini and Bel- hadj (2015)	HCR	PG
Bias %									
Areas	-0.284	2.104	1.973	1.670	1.669	1.616	-0.758	0.574	0.204
National	-0.057	0.364	0.368	1.722	1.722	1.722	-0.557	1.722	0.719
CV-Bootstrap %									
Areas	31.248	11.948	41.194	37.464	38.047	43.191	50.142	60.208	78.991
National	5.788	2.248	7.340	4.094	4.139	4.162	21.284	4.524	5.996
CV-Jackknife %									
Areas	29.709	11.857	37.693	34.794	35.252	39.191	44.373	50.560	60.755
National	4.845	1.361	4.312	4.034	4.030	4.025	7.736	4.386	4.018
CV2-Bootstrap %									
Areas	29.709	11.857	37.693	34.794	35.252	39.191	44.373	50.560	60.755
National	5.778	2.247	7.321	4.090	4.135	4.158	20.818	4.519	5.985
CV2-Jackknife %									
Areas	27.064	10.941	34.160	31.349	31.778	35.520	47.025	51.020	55.766
National	4.839	1.361	4.308	4.030	4.027	4.022	7.713	4.382	4.015

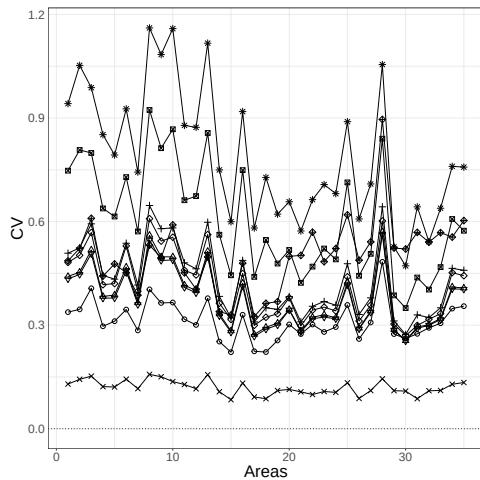
Figure 2 – Complex simulation: values of Bias for each area. Areas are sorted by increasing sample size.



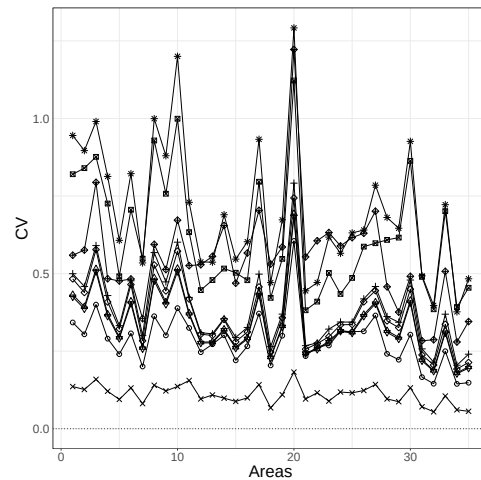
(A) Bias



Figure 3 – Complex simulation: values of Coefficient of Variation by area. Areas are sorted by increasing sample size.



(A) CV (bootstrap)



(B) CV (jackknife)



5.2 MSE estimation

In this section, we analyse the performance of the non-parametric bootstrap and jackknife.

5.2.1 Simple random sampling design

Table 4 reports the values of the BMSE. At area level, the BMSE is quite low for all the indices except that of Zedini and Belhadj (2015) (33.409). Regarding the others, crisp and fuzzy indices perform almost identically with only a few differences worthy of note. The HCR is the index with the highest BMSE (in absolute terms), whereas the index of Cerioli and Zani (1990) show the best non-parametric bootstrap performance (-0.145).

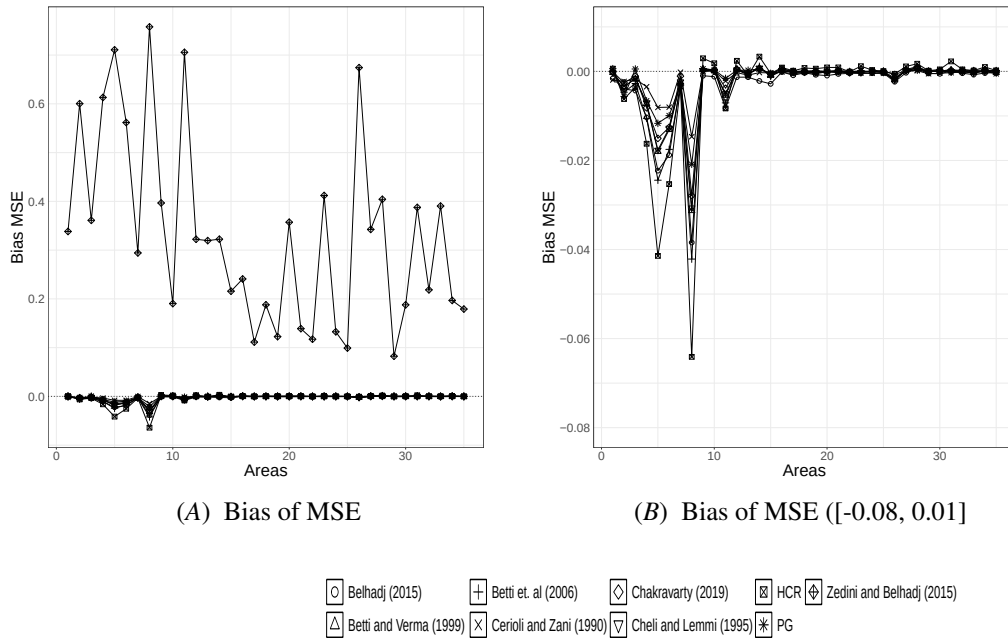
Individual values of the BMSE by area are reported in Figures 4A and 4B, where the index of Zedini and Belhadj (2015) has been removed for easier visualization of the results. Not surprisingly, when the non-parametric bootstrap is used, the BMSE approaches zero as sample size increases. Again, at specific area level, we see that the MSE of all indices is underestimated for all the areas, with almost no differences between fuzzy and crisp indices.

At national level the non-parametric bootstrap performs very well. The BMSEs of both fuzzy and crisp indices are all low. Interestingly, at this level, *i.e.* for very high sample sizes, the non-parametric bootstrap even performs quite well for the index of Zedini and Belhadj (2015).

Table 4 – SRS simulation: BMSE. The measures for the "area" are computed as the mean over all area values.

	Belhadj (2015)	Cerioli and Zani (1990)	Chakravarty (2019)	Cheli and Lemmi (1995)	Betti and Verma (1999)	Betti et. al (2006)	Zedini and Belhadj (2015)	HCR	PG
BMSE%									
Areas	-0.373	-0.145	-0.200	-0.238	-0.242	-0.325	33.409	-0.418	-0.157
National	-0.075	-0.031	0.006	-0.022	-0.022	-0.022	5.582	0.002	-0.001

Figure 4 – SRS simulation: bias of MSE estimator by area. Areas are sorted by increasing sample size.



5.2.2 Complex-design

We investigate empirically whether one of the two variance estimators considered is to be preferred over the other. Table 5 reports the BMSEs of the indices with the two variance estimators. The non-parametric bootstrap clearly outperforms the jackknife estimator. The BMSE by bootstrap is lower than that of the jackknife for all indices. As in the previous simulation, the BMSE is again negative for fuzzy and crisp-sets except for the index of Zedini and Belhadj (2015) and the HCR.

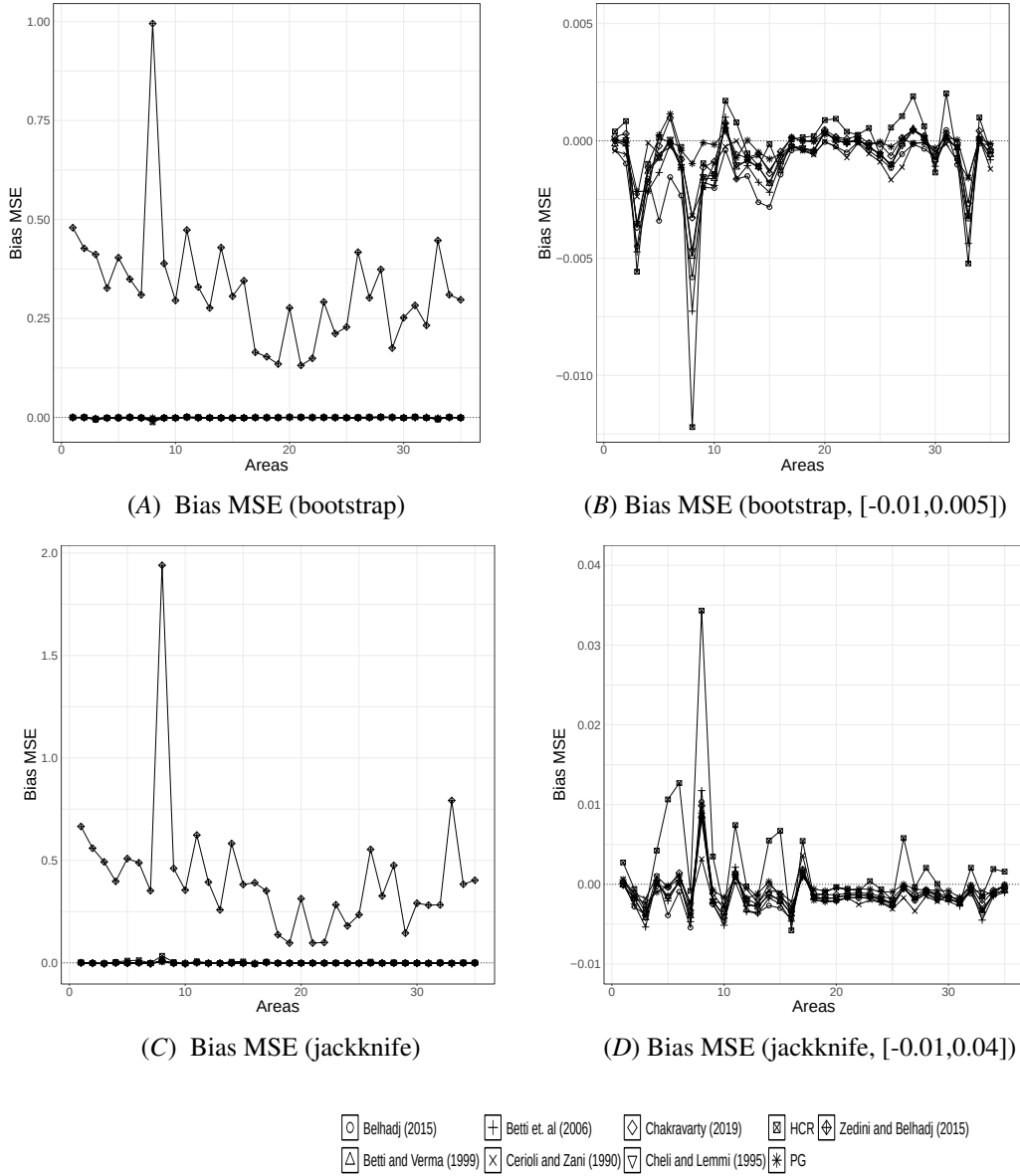
Figure 5 reports the bias of the MSE estimators. For better visualization, the sub-figures on the right again show the MSE bias without the index of Zedini and Belhadj (2015). The area level values show that for large sample sizes, the two variance estimators perform almost identically for every index. Differences between the two estimators are, therefore, found for areas with a very low sample size. In this situation, the bootstrap performs better than the jackknife, but both tend to underestimate the true value. Sub-domains with small sample size should therefore always be considered in this type of analysis. Indeed, as mentioned above and recalling the results of Figure 3, we conclude that the most realistic CVs are those obtained by the bootstrap estimator. To conclude, Figures 5D and 5B clearly show a coherent trend, except in the case of the HCR. In fact, by the jackknife estimator, the BMSE, is positive, as confirmed by Table 5, whereas by the bootstrap it is negative.

At national level, the BMSE is very low for both estimators of the MSE, but the non-parametric bootstrap still performs better than the jackknife. The BMSE of crisp indices is very low when the bootstrap estimator is used. These results are not surprising since the jackknife used here is an ad-hoc estimator for fuzzy measures. It is defined for distribution-based indices, for which it performs almost identically to the bootstrap.

Table 5 – Complex simulation: BMSE by bootstrap and jackknife estimates.

	Belhadj (2015)	Cerioli and Zani (1990)	Chakravarty (2019)	Cheli and Lemmi (1995)	Betti and Verma (1999)	Betti et. al (2006)	Zedini and Bel- hadj (2015)	HCR	PG
BMSE-Bootstrap %									
Areas	-0.130	-0.073	-0.044	-0.073	-0.076	-0.102	32.511	-0.052	-0.025
National	-0.051	-0.033	0.011	-0.026	-0.023	-0.026	6.066	0.006	0.000
BMSE-Jackknife %									
Areas	-0.165	-0.117	-0.081	-0.130	-0.133	-0.160	41.666	0.224	-0.050
National	-0.027	-0.032	0.011	-0.027	-0.026	-0.026	2.356	0.067	0.005

Figure 5 – Complex simulation: bias of the MSE estimator by area. Areas are sorted by increasing sample size.



5.3 Computation time

Table 6 shows the average time needed for the simulation cycles to estimate MSE using the package developed by Crescenzi and Mori (2024). The mean time required to compute the national MSE, excluding sub-areas, is also reported, i.e. without requiring the algorithm to split the population. Both estimators are quite fast but the bootstrap is always faster than the jackknife. The distribution-based estimators and that of Zedini and Belhadj (2015) are slower than the others because at each loop of the MSE estimator it is necessary to rearrange the sample in order to compute the probability density function or to perform a bootstrap in order to obtain the values of b, b' and b'' . If sub-areas are not considered, the computational times are considerably lower. The computational burden of the MSE for the crisp indices is in line with that of the fuzzy indices that do not need a recursive algorithm for their estimation. Given the previous considerations, the computational times and the indication of Osier et al. (2013), we believe that the bootstrap is preferable to the jackknife.

Table 6 – Computation time (seconds)

	Belhadj (2015)	Cerioli and Zani (1990)	Chakravarty (2019)	Cheli and Lemmi (1995)	Betti and Verma (1999)	Betti et. al (2006)	Zedini and Bel- hadj (2015)	HCR	PG
Simple random sampling - Bootstrap									
Areas	$\simeq 0.67$	$\simeq 0.65$	$\simeq 0.68$	$\simeq 4.94$	$\simeq 7.69$	$\simeq 7.97$	$\simeq 134.63$	$\simeq 0.11$	$\simeq 0.14$
National	$\simeq 0.13$	$\simeq 0.16$	$\simeq 0.15$	$\simeq 2.59$	$\simeq 2.19$	$\simeq 5.58$	$\simeq 92.17$	$\simeq 0.06$	$\simeq 0.07$
Complex random sampling - Bootstrap									
Areas	$\simeq 2.26$	$\simeq 2.25$	$\simeq 1.45$	$\simeq 11.39$	$\simeq 11.37$	$\simeq 9.37$	$\simeq 22.15$	$\simeq 0.14$	$\simeq 0.35$
National	$\simeq 0.34$	$\simeq 0.17$	$\simeq 0.15$	$\simeq 7.16$	$\simeq 10.15$	$\simeq 6.37$	$\simeq 13.26$	$\simeq 0.17$	$\simeq 0.17$
Complex random sampling - Jackknife									
Areas	$\simeq 1.14$	$\simeq 0.14$	$\simeq 0.15$	$\simeq 101.51$	$\simeq 154.15$	$\simeq 115.32$	$\simeq 125.35$	$\simeq 0.15$	$\simeq 0.17$
National	$\simeq 0.16$	$\simeq 0.17$	$\simeq 0.16$	$\simeq 71.16$	$\simeq 116.16$	$\simeq 116.51$	$\simeq 116.36$	$\simeq 0.12$	$\simeq 0.24$

6 Robustness analysis

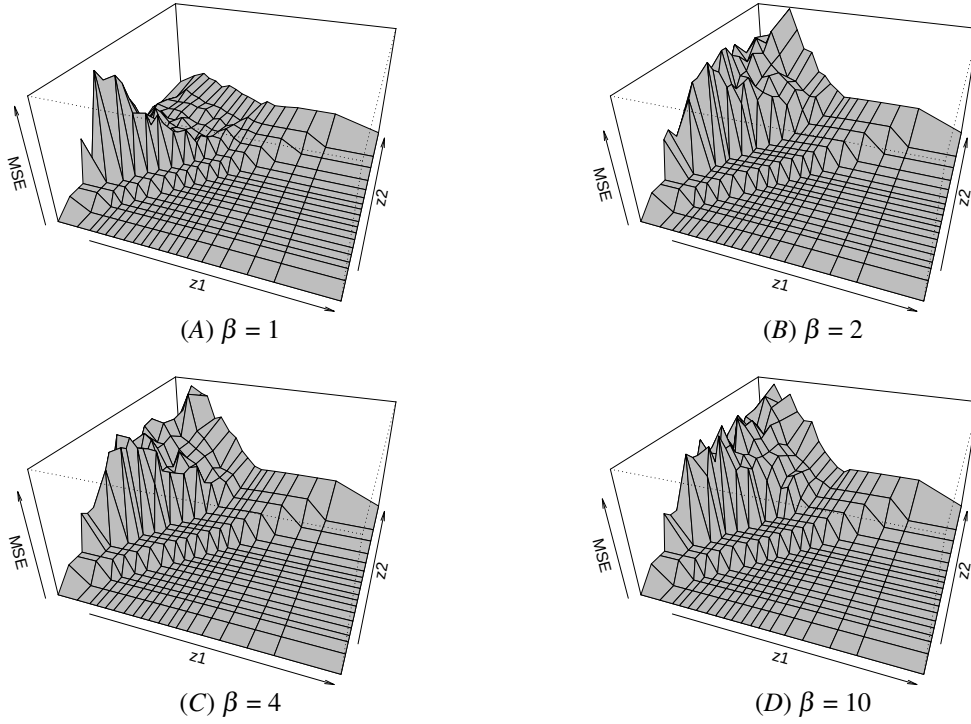
In this section, we analyse the robustness of the fuzzy normative indices. The reader may have noticed that some fuzzy indices depend on parameters that need to be specified and are not estimated from data. Deciding which fuzzy index to use for an analysis is therefore not trivial. One should choose the most robust, i.e. the one least sensitive, to changes the sources of uncertainty and to parameter specifications.

Robustness analysis involves checking the ordering of two or more areas under alternative choices over the parameters specification of the index. At area level, the order of the estimates are checked by “correlation coefficients for ranks” (Alkire and Jahan, 2018). These methods test whether an ordering of more than two areas remains the same when the value of some parameter is altered. The robustness of a ranking is evaluated by calculating a rank correlation coefficient between the initial and the alternative rankings. In this section, we only report the results of the Complex simulation (those of SRS simulation are almost identical and available on request). The MSE of the indices is analysed at national level as a function of the parameters of the membership function in order to understand how the choice of parameters affects its value.

6.1 Impact of the choice of parameters on the MSE

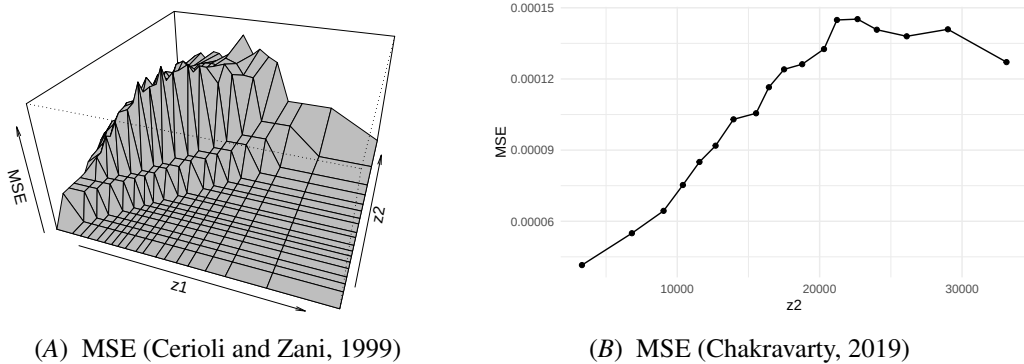
Figure 6 clearly shows that the MSE of the sample average of the membership function of the index of Belhadj (2015) is sensitive to the choice of thresholds. Note that from Equation (4), the portion of the plane (z_1, z_2) to be considered is that of points $\{(z_1, z_2) : z_2 \geq z_1\}$. In particular, setting high values of z_1 will usually decrease the MSE of the estimator while the opposite happens for small values. Parameter β seems to play the role of inflating the MSE in the upper region of the plane.

Figure 6 – MSE of the index of Belhadj (2015).



These pictures also suggest another important aspect to consider when estimating fuzzy poverty indices, that is, a mis-specified selection of thresholds may result in severe biases of the MSE. Suppose that z_1^* and z_2^* are the true values ($b = 1$ without loss of generality) and that $MSE(z_1^*, z_2^*)$ is the corresponding MSE. Suppose that z_1' and z_2' are the values specified by the researcher (again, $\beta = 1$ without loss of generality) and that $MSE(z_1', z_2')$ is the corresponding MSE. The difference between the two MSEs may be non-negligible in some portions of the z_1', z_2' plane. The same considerations apply to the index of Cerioli and Zani (1990), see Figure 7A. Regarding the index of Chakravarty (2019), the MSE does not seem to depend on the choice of the threshold, as long as the threshold is sufficiently large (see Figure 7B).

Figure 7 – MSE of the indices of Cerioli and Zani (1990) and Chakravarty (2019).



6.2 Correlation coefficients for ranks

Suppose there are j^* areas and two set of ranks, r and r^* , for two different specification sets of parameters regarded as the initial and alternative specifications. Also consider r_{j^*} and $r_{j^*}^*$ as the ranks of the area j^* for the two different specification sets of parameters in the same membership function.

The alternative specification may entail a change in the set of α -cut parameters or in β parameters. In line with Noha and El Hage Sleiman (2021), we use the Kendall correlation coefficient - for situations with small sample size and many tied ranks, and the otherwise the Spearman correlation coefficient. Table 7 shows their mean values for the indices of Belhadj (2015), Cerioli and Zani (1990) and Chakravarty (2019). For the former we set $\beta = 2$, whereas for the other two indices, as benchmark we use the indices computed with the parameters set as in Table 1. The correlation coefficients tend to decrease quickly except for Cerioli and Zani (1990), suggesting that this index is more robust than the other two - see also Table 8. More specifically, with the exception of the Cerioli and Zani (1990) index, the others tend to have a correlation ≤ 0.90 , indicating significant modifications in the rankings of the areas when the parameters change. The Belhadj (2015) index, which we studied by modifying the two types of parameter separately, is more stable when the parameter that changes is β . Indeed, in this case, the correlation coefficients drop below 0.9 for every high value of β . On the other hand, when β is fixed and the other parameters change, the correlation coefficients are frequently ≤ 0.8 .

Table 7 – Correlation coefficients for ranks changing z parameters

Index	Correlation	Parameters									
		$Y_{0.02}$	$Y_{0.03}$	$Y_{0.04}$	$Y_{0.05}$	$Y_{0.06}$	$Y_{0.07}$	$Y_{0.08}$	$Y_{0.09}$	$Y_{0.10}$	
		$Y_{0.98}$	$Y_{0.97}$	$Y_{0.96}$	$Y_{0.95}$	$Y_{0.94}$	$Y_{0.93}$	$Y_{0.92}$	$Y_{0.91}$	$Y_{0.90}$	
Belhadj (2015), $\beta = 2$	Kendal	0.74	0.79	0.87	1.00	0.89	0.83	0.79	0.76	0.73	
	Spearman	0.89	0.93	0.97	1.00	0.97	0.95	0.93	0.91	0.89	
Cerioli and Zani (1990)	Kendal	0.94	0.96	0.98	1.00	0.98	0.97	0.96	0.95	0.93	
	Spearman	0.99	0.99	1.00	1.00	1.00	1.00	0.99	0.99	0.99	
Index	Correlation	Parameters									
		$Y_{0.30}$	$Y_{0.35}$	$Y_{0.40}$	$Y_{0.45}$	$Y_{0.55}$	$Y_{0.60}$	$Y_{0.65}$	$Y_{0.70}$	$Y_{0.75}$	
Chakravarty (2019)	Kendal	0.72	0.82	0.91	1.00	0.93	0.87	0.82	0.79	0.76	
	Spearman	0.87	0.94	0.98	1.00	0.99	0.97	0.95	0.92	0.90	

Table 8 – Correlation coefficients for ranks for Belhadj index changing β

Index	Correlation	Parameters: β									
		1	2	3	4	5	6	7	8	9	10
Belhadj (2015)	Kendal	0.97	1.00	0.94	0.89	0.85	0.83	0.81	0.80	0.79	0.78
	Spearman	0.99	1.00	0.99	0.97	0.96	0.94	0.94	0.93	0.92	0.92

7 Conclusions

In this paper we addressed the issue of estimating fuzzy poverty indices from a statistical estimation perspective. We focused on unidimensional indices using a monetary poverty predicate (*e.g.* income), however the analysis can easily be extended to any continuous poverty predicate as well as to multi-dimensional fuzzy indices. Although many books and papers have focused on the formulation of fuzzy indices and on their properties and axioms, the literature lacks a thorough simulation study on their statistical estimation properties with sample data. The Italian national Institute of Statistics (ISTAT) has suggested the use of fuzzy poverty indices several times (*e.g.* Ghellini and Mulas, 2006; De Gregorio and Giordano, 2016). Interestingly, De Gregorio and Giordano (2016) recommended fuzzy methods to quantify irregular work in Italy using survey and administrative data.

Assessment of the properties of the fuzzy indices is therefore of practical importance, and is not limited to poverty data. Using the official 2019 EU-SILC Italian survey data, we simulated two different sample surveys to investigate their statistical properties. We first examined the bias of the sample statistic (*i.e.* the sample mean of the membership function) and the bias of its MSE estimator under simple random sampling from the population. To remain closer to official survey methods, that typically involve stratification and more Complex-designs, we repeated the analysis for this more realistic scenario.

For methods that require parameter specification we drew the surface of the MSE for different values of the parameters. This showed that mis-specification of the parameters can lead to severe distortions when estimating the MSEs. It also showed that inaccurate estimates of fuzzy indices can just be the result of a poor specification of parameters.

At least three reflections emerge from these considerations. First, from simulations demonstrated that there is no unbiased estimator of the MSE when sample size is very small (< 50). Since developing and acquiring reliable sub-population indicator estimates, is increasingly important, in our opinion this problem needs to be addressed in an ad-hoc study. For example, estimation of unbiased MSE could be the first step in obtaining fuzzy small area estimates using area-level models. Furthermore, a closed-form MSE estimator for fuzzy indices could be a research topic alternative to re-sampling methods.

The second reflection, concerning a theoretical background, is that we treated the parameters of the various membership functions as fixed values in the population of individuals unknown to the researcher. Although this is a necessary working assumption for discussing about bias in estimation procedures, the fact of unknown parameters of a population is something that has not been discussed extensively in the literature. The fact that these parameters are set arbitrarily without any statistical procedure makes it difficult to draw a comparison without making assumptions of the kind just mentioned and about which sampling statistic to estimate.

Third, our study assumed the sample mean of the estimated membership function as sample statistic. Although the sample mean is a well-known statistical estimator, not all the methods discussed originally suggested this statistic. The lack of consistency between the various approaches described in the literature and reported here required us to choose the sample mean for two reasons, one statistical and the other because it was the most common statistic chosen by the methods considered. Our results therefore cannot be regarded only as an assessment of the original papers but must also be considered in light of the fact that we required a common ground of comparison.

References

- Alfons, A. and Templ, M. (2012). Estimation of social exclusion indicators from complex surveys: The R package Laeken. *KU Leuven, Faculty of Business and Economics Working Paper, Leuven, Belgium*.
- Alkire, S. and Jahan, S. (2018). *The new global MPI 2018: Aligning with the sustainable development goals*. OPHI, Oxford, UK.
- Alkire, S., Roche, J. M., Ballon, P., Foster, J., Santos, M. E., and Seth, S. (2015). *Multidimensional poverty measurement and analysis*. Oxford University Press, USA.
- Allen, R. (2013). *Poverty lines in history, theory, and current international practice*. University of Oxford, Oxford, UK.
- Atkinson, A. B. (1987). On the measurement of poverty. *Econometrica: Journal of the Econometric Society*, 55:749–764.
- Barabesi, L., Crescenzi, F., and Mori, L. (2024). Theil index estimation by using the influence function with application to income surveys. *Quaderni del Dipartimento di Economia Politica e Statistica, Università di Siena*, 915:1–24.
- Barabesi, L., Diana, G., and Perri, P. F. (2016). Linearization of inequality indices in the design-based framework. *Statistics*, 50:1161–1172.
- Belhadj, B. (2011). A new fuzzy unidimensional poverty index from an information theory perspective. *Empirical Economics*, 40:687–704.

- Belhadj, B. (2015). Employment measure in developing countries via minimum wage and poverty new fuzzy approach. *Opsearch*, 52:329–339.
- Belhadj, B. and Limam, M. (2012). Unidimensional and multidimensional fuzzy poverty measures: New approach. *Economic Modelling*, 29:995–1002.
- Belhadj, B. and Matoussi, M. S. (2010). Poverty in Tunisia: A fuzzy measurement approach. *Swiss Journal of Economics and Statistics*, 146:431–450.
- Berger, Y. G. and Skinner, C. J. (2005). A jackknife variance estimator for unequal probability sampling. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67:79–89.
- Betti, G., Cheli, B., Lemmi, A., and Verma, V. (2006). Multidimensional and longitudinal poverty: an integrated fuzzy approach. In Betti, G. and Lemmi, A., editors, *Fuzzy set approach to multidimensional poverty measurement*, pages 115–137. Springer, Boston, USA.
- Betti, G., Crescenzi, F., Molini, V., and Mori, L. (2024). Estimation of multidimensional poverty in Morocco: A small area estimation approach using meteorological and socio-economic covariates. *Social Indicators Research*, pages 1–31.
- Betti, G., Gagliardi, F., and Verma, V. (2018). Simplified jackknife variance estimates for fuzzy measures of multidimensional poverty. *International Statistical Review*, 86:68–86.
- Betti, G. and Lemmi, A. (2008). *Advances on income inequality and concentration measures*. Routledge, London, UK.
- Betti, G. and Lemmi, A. (2021). *Analysis of Socio-Economic Conditions*. Routledge, London, UK.
- Betti, G. and Verma, V. (1999). Measuring the degree of poverty in a dynamic and comparative context: a multi-dimensional approach using fuzzy set theory. In *Proceedings, iccs-vi*, volume 11, pages 289–300.
- Ceroli, A. and Zani, S. (1990). A fuzzy approach to the measurement of poverty. In *Income and Wealth Distribution, Inequality and Poverty: Proceedings of the Second International Conference on Income Distribution by Size: Generation, Distribution, Measurement and Applications.*, pages 272–284. Springer, Boston, USA.
- Chakravarty, S. R. (2019). An axiomatic approach to multidimensional poverty measurement via fuzzy sets. In Chakravarty, S. R., editor, *Poverty, social exclusion and stochastic dominance*, pages 123–141. Springer, Boston, USA.
- Chandra, H., Chambers, R., and Salvati, N. (2012). Small area estimation of proportions in business surveys. *Journal of Statistical Computation and Simulation*, 82:783–795.
- Cheli, B. and Lemmi, A. (1995). A 'totally' fuzzy and relative approach to the multidimensional analysis of poverty. *Economic Notes*, 24:115–134.
- Chiappero-Martinetti, E. (2006). Capability approach and fuzzy set theory: description, aggregation and inference issues. In Betti, G. and Lemmi, A., editors, *Fuzzy Set Approach to Multidimensional Poverty Measurement*, pages 93–113. Springer, Boston, USA.
- Chwila, A. and Żądło, T. (2020). On the choice of the number of monte carlo iterations and bootstrap replicates in empirical best prediction. *Statistics in Transition new series*, 21:35–60.
- Crescenzi, F. and Mori, L. (2024). FuzzyPovertyR: Estimation of fuzzy poverty measures. *R package version 3.0.2*, CRAN.

- De Gregorio, C. and Giordano, A. (2016). The heterogeneity of undeclared work in Italy: some results from the statistical integration of survey and administrative sources. *Rivista di Statistica Ufficiale*, 2:99–129.
- Duclos, J.-Y., Sahn, D. E., and Younger, S. D. (2006). Robust multidimensional poverty comparisons. *The economic journal*, 116:943–968.
- Eurostat (2013). *Handbook on precision requirements and variance estimation for ESS households surveys - 2013 edition*. Eurostat, Bruxelles, Belgium.
- Fabrizi, E., Ferrante, M. R., and Pacei, S. (2008). Measuring sub-national income poverty by using a small area multivariate approach. *Review of Income and Wealth*, 54:597–615.
- Field, C. A. and Welsh, A. H. (2007). Bootstrapping clustered data. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 69:369–390.
- Fisher, G. M. (1992). The development and history of the poverty thresholds. *Social Security Bulletin*, 55:3–14.
- Fisher, G. M. (1997). The development of the Orshansky poverty thresholds and their subsequent history as the official U.S. poverty measure. Technical report, United States Census Bureau.
- Foster, J., Greer, J., and Thorbecke, E. (1984). A class of decomposable poverty measures. *Econometrica: journal of the econometric society*, 52:761–766.
- Foster, J. E. (1983). *On economic poverty: a survey of aggregate measures*. Institute for Research in the Behavioral, Economic, and Management Sciences, Purdue University, USA.
- Gangopadhyay, D., Nielsen, R. B., and Zahirovic-Herbert, V. (2021). Methodology and axiomatic characterization of a multidimensional and fuzzy measure of deprivation. *Social Indicators Research*, 153:1–37.
- Garrido, A. (2012). A brief history of fuzzy logic. *BRAIN. Broad Research in Artificial Intelligence and Neuroscience*, 3:71–77.
- Ghellini, G. and Mulas, A. (2006). Misure di deprivazione sulla base di analisi fattoriali. In *Rivista di statistica ufficiale*, volume 2, pages 67–88. ISTAT.
- Gillie, A. (1996). The origin of the poverty line. *Economic History Review*, 49:715–730.
- Gonzalez-Manteiga, W., Lombardia, M. J., Molina, I., Morales, D., and Santamaria, L. (2008). Bootstrap mean squared error of a small-area EBLUP. *Journal of Statistical Computation and Simulation*, 78:443–462.
- Hagenaars, A. (1987). A class of poverty indices. *International economic review*, 28:583–607.
- Kakwani, N. (1993). Statistical inference in the measurement of poverty. *The Review of Economics and Statistics*, 75:632–639.
- Kim, S.-G. (2015). Fuzzy multidimensional poverty measurement: An analysis of statistical behaviors. *Social Indicators Research*, 120:635–667.
- Klir, G. and Yuan, B. (1995). *Fuzzy sets and fuzzy logic*. Prentice hall New Jersey, USA.
- Kvålseth, T. O. (2017). Coefficient of variation: the second-order alternative. *Journal of Applied Statistics*, 44:402–415.
- Lemmi, A. A. and Betti, G. (2006). *Fuzzy set approach to multidimensional poverty measurement*. Springer, Boston, USA.

- Liu, Y., Liu, X., Pan, Y., Jiang, J., and Xiao, P. (2022). An empirical comparison of various MSPE estimators and associated prediction intervals for small area means. *Journal of Statistical Computation and Simulation*, pages 1–27.
- Noha, O. and El Hage Sleiman, S. (2021). Testing the robustness of the revised multidimensional poverty index for Arab countries. Technical report, United Nation Bureau Economic and Social Commission for Western Asia.
- Oguz Alper, M. and Berger, Y. G. (2015). Variance estimation of change in poverty rates: an application to the Turkish EU-SILC survey. *Journal of Official Statistics*, 31:155–175.
- Osier, G., D’Alò, M., Aru, J., Boonstra, H. J., Mihályffy, L., Moussallam, K., and Smith, P. (2013). Handbook on precision requirements and variance estimation for ess households surveys. Technical report, ISSN 1977-0375, Eurostat Methodologies and Working papers.
- Pyatt, G. (1987). Measuring welfare, poverty and inequality. *The Economic Journal*, 97:459–467.
- Qizilbash, M. (2003). Vague language and precise measurement: the case of poverty. *Journal of economic methodology*, 10:41–58.
- Sen, A. (1973). *On economic inequality*. Oxford university press, Oxford, UK.
- Sen, A. (1976). Poverty: an ordinal approach to measurement. *Econometrica: Journal of the Econometric Society*, 44:219–231.
- Shorrocks, A. and Subramanian, S. (2024). Fuzzy poverty indices. *Journal of Income Distribution*, 32:1–20.
- Silber, J. (2023). *Research handbook on measuring poverty and deprivation*. Edward Elgar Publishing, UK.
- Tzavidis, N., Zhang, L.-C., Luna, A., Schmid, T., and Rojas-Perilla, N. (2018). From start to finish: a framework for the production of small area official statistics. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 181:927–979.
- Verma, V. and Betti, G. (2011). Taylor linearization sampling errors and design effects for poverty measures and other complex statistics. *Journal of Applied Statistics*, 38:1549–1576.
- Walter, P., Groß, M., Schmid, T., and Weimer, K. (2022). Iterative kernel density estimation applied to grouped data: estimating poverty and inequality indicators from the German microcensus. *Journal of Official Statistics*, 38:599–635.
- Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8:338–353.
- Zedini, A. and Belhadj, B. (2015). A new approach to unidimensional poverty analysis: Application to the tunisian case. *Review of Income and Wealth*, 61:465–476.

Acknowledgment: The work of Lorenzo Mori was funded by the European Union - NextGenerationEU, in the framework of the “GRINS -Growing Resilient, INclusive and Sustainable project” (PNRR - M4C2 - I1.3 - PE00000018 – CUP J33C22002910001). The views and opinions expressed are solely those of the author and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.