Measuring and communicating uncertainty of poverty indicators at regional level

TIZIANA LAURETI, ILARIA BENEDETTI

2020 edition

STATISTICAL WORKING PAPERS

eurostat

Measuring and communicating uncertainty of poverty indicators at regional level

LAURETI TIZIANA, BENEDETTI ILARIA

2020 edition

Manuscript completed in November 2020

Luxembourg: Publications Office of the European Union, 2020

Theme: Population and social conditions
Collection: Statistical working papers

PDF: ISBN 978-92-76-28361-4 ISSN 2315-0807 doi: 10.2785/313140 KS-TC-20-010-EN-N

Abstract

During the last years, there has been an increasing interest in providing uncertainty measures for poverty indicators both at national and regional level. After reviewing the literature on the various methodological approaches for measuring uncertainty in poverty indicators, this report presents standard error estimation results for official income-poverty and income-inequality poverty indicators by using 2018 European Union Statistics on Income and Living Conditions (EU-SILC) surveys. Eight European countries are selected on the basis of the sample design structure, which varies from a simple random sampling in one stage to a stratified two-stage random sampling. Both linearization and re-sampling methods, i.e. bootstrap and jackknife techniques, are used with the aim of comparing the different uncertainty measures obtained. Overall, our analysis shows that there is not a unique variance estimation method that performs well with any sampling design and any complex estimator. Yet, the general closeness of the results from entirely different methodologies can be considered quite remarkable. Moreover, the issue of communicating uncertainty of official poverty indicators is addressed by reviewing the current practice adopted by various National Statistical Institutes. Finally, we formulate some recommendations to ameliorate the way in which uncertainty is communicated.

Keywords: Poverty Indicators, measuring uncertainty, sub-national estimations, linearization, jackknife.

Authors: Tiziana Laureti ([1]) , Ilaria Benedetti ([2]) .

Acknowledgement: On behalf of Eurostat - Unit B1 the study was coordinated by Mátyás Mészáros and Dario Buono, with contributions from Maria Serfioti. The authors would like to thank to all contributors from other Eurostat units, as well, for their helpful comments. The study was implemented by GOPA, as contractor of Eurostat for the framework contract on Methodological Support (ref. 2018.0086).

([1]) Department of Economics, Engineering, Society and Business, University of Tuscia, laureti@unitus.it
([2]) Department of Economics, Engineering, Society and Business, University of Tuscia, i.benedetti@unitus.it

Table of contents

1 Introduction

Over the last decades, there has been an increasing interest in identifying comparable indicators for measuring poverty and social exclusion in the EU as well as in providing uncertainty measures both at national and regional level. In the context of the Europe 2020 Strategy, poverty indicators play an essential role in informing and supporting responsible evidence-based policies towards the Union's key goals which relate to inclusive and sustainable growth and reduction of poverty, inequalities and social exclusion. In the framework of the OECD's Better Life Initiative, the economic well-being (material living condition) is identified as one of the pillars for understanding and measuring people's well-being following a multi-dimensional approach (OECD, 2013).

The EU Statistics on Income and Living Conditions (EU-SILC) survey, launched by Eurostat in 2004, constitutes the main data source for constructing indicators of poverty and inequality, such as the at-risk-of-poverty rate (AROP) and the Gini coefficient, in the multi-country comparative context of the EU. Given that all the indicators based on EU-SILC are sample estimates, they should be reported along with estimates of standard errors and confidence intervals. Given the key role played by poverty indicators in designing and monitoring social progress in the EU, it is paramount to produce and communicate to the public measures of the associated inherent and unavoidable uncertainty of point estimates. To this respect, by referring to the content of intermediate and official EU-SILC Quality reports, the new EU regulation on Social Statistics (2019/1700), amending the Commission Regulation No. 28/2004, requires that countries should provide estimates of standard error along with the EU-SILC main target indicators. Therefore, it is essential that the indicators used for measuring poverty have the necessary high quality, especially in terms of their accuracy as well as reliability, timeliness and usability. In this framework, a crucial role is played by estimates of poverty and social policy indicators at regional or sub-national level (NUTS2 and NUTS1) to be used for bench-marking and assessing the efficiency of regional policies (Betti et al., 2012 and Piacentini, 2014). To this respect, Eurostat and a number of stakeholders are exploring various methodological approaches for improving the quality of statistics on income and living conditions based on the EU-SILC survey, especially regarding their accuracy (Atkinson et al., 2017). Measuring uncertainty is a complex and challenging task, which can involve the use of sophisticated statistical methods and econometric techniques as well as the adoption of subjective judgment to quantify indicators' uncertainties.

The present document focuses on estimating standard errors for poverty indicators at national and regional level. Among the various official poverty indicators used at EU level, we selected income-poverty measures which belong to the class of the well-known Foster-Greer-Thorbecke (FGT) measures, that is the at-risk-of-poverty rate and the relative median at-risk-of-poverty gap and income-inequality measures (quintile share ratio and the Gini coefficient). Section 2 illustrates the main characteristics of these indicators. Section 3 reviews the different approaches for estimating the variance of the official poverty indicators. These methods can be classified into two approaches: 'direct' methods, which rely on analytic variance formulas through linearization (Alper and Berger, 2015) and 'resampling' methods, such as the Jackknife Repeated Replication and Boostrap, which consist of resampling a high number of 'replications' from the original sample in order to empirically derive a sampling distribution (Davidson and Flachaire, 2007 and Verma and Betti, 2010). Section 4 describes the EU-SILC survey by considering its main characteristics, including the different sampling designs adopted by countries. Previous projects focused on uncertainty estimation of EU-SILC indicators are illustrated. In Section 5 computational aspects and estimation results of standard error both at national and sub-national level are described with a discussion of the main findings. In this report we considered eight countries according to their EU-SILC sampling designs. More specifically, we selected five countries using the two-stage stratified sampling design, that is Italy, the United Kingdom, Portugal, Belgium and Ireland; two countries using the one-stage sampling design (Germany and Sweden) and the Norway which adopts a simple sampling design without stratification. Focusing on the computing standard errors for regional indicators, we selected those countries for which NUTS1 and NUTS2 are considered as planned domains in the sample design. Nevertheless, in a few cases, standard errors are particularly high due to the reduced number of secondary sampling units (SSUs). This suggests that the number of sampling units in each region should be increased or other methods should be used for improving the accuracy of point estimates. After a brief review of the current practices adopted by NSIs in communicating uncertainty for official estimates, in Section 6 suggestions for communicating uncertainty in the national quality reports are provided. Finally, in Section 7 we draw some conclusions and suggest lines of research and future developments.

2 Estimating uncertainty measures for official poverty indicators

2.1 Choosing poverty indicators for variance estimation

Over the last years, the interest in statistical inference for inequality and poverty measures has been increased. The notion of poverty in the European Union, initially proposed by the Council of the European Communities in 1975 refers to *"individuals or families whose resources are so small as to exclude them from the minimum acceptable way of life of the member state in which they live"*. Regarding the conceptualisation of poverty, the EU Commission follows the social inclusion approach with the aim of monitoring progress towards a set of EU objectives for social protection and social inclusion which have been jointly defined with the EU Member States. After the Lisbon Council of March 2000, social cohesion became the most challenging responsibility for the European Union. In order to monitor the progress towards the reduction of poverty a first set of indicators, called the Laeken Indicators, was set up at the Laeken Council in December 2001 by following the methodological principles defined by the Laeken Council and the Social Protection Committee Indicators Sub-Group (ISG)([3]) . These indicators are key in the Europe 2020 strategy and are considered consistent indicators based on harmonized definition of income, thus allowing international comparisons and reliable measurements of social cohesion. Moreover, they are considered as a reliable source on which policy makers can base their decisions.

In income analyses it is often assumed that precision is not a serious issue since these analyses are based on very large samples. However, various empirical studies have shown that this may not be true in some cases (see for example Maasoumi (1997)). The issue of uncertainty measurement is crucial since we are often interested in knowing whether poverty has increased or decreased over time or in comparing poverty differences among geographical areas or among various socio-economic groups. In these cases, it is essential to have information about the sampling variability of the estimates. The formulae for calculating standard errors for poverty measures depends on the statistic to be computed. Many poverty indicators consist of headcounts with a fixed poverty threshold. Standard errors for all poverty measures of the well-known Foster-Greer-Thorbecke (FGT) class (Foster et al., 1984), may be estimated by using similar formulae to those used for proportion or mean. Kakwani (1993) provided distribution-free asymptotic confidence intervals and statistical inference for poverty measures that are applied to analyse poverty in Cote d'Ivoire. The author found that observed differences in values of poverty measures may lead to misleading conclusions without carrying out statistical tests. An indicator that has been widely used when comparing poverty is the proportion of population falling below a fraction α of the $\beta - th$ quantile of a distribution. However, the estimation of standard errors for such proportions is a challenging task since the quantile must first be estimated before estimating the proportion falling

([3]) The Social Protection Committee (SPC) is an EU advisory policy committee for Employment and Social Affairs Ministers in the Employment and Social Affairs Council (EPSCO). It monitors social conditions in the EU and the development of social protection policies in Member States and promotes discussion and coordination of policy approaches among national governments and the Commission. In 2001 the SPC established the Indicators' Sub-Group (ISG) of the SPC to support its activities, in particular by providing technical and analytical support, especially with regard to indicators. The role of the ISG is to develop and define EU social indicators to monitor member countries' progress towards the commonly agreed EU objectives for Social Protection and Social Inclusion, to carry out analytical work based on agreed indicators and develop analytical frameworks to support policy reviews conducted by the SPC, and to contribute to the improvement of social statistics at EU level, particularly through development of the EU-SILC.

below a share of this estimated quantile (Berger and Skinner, 2003). Therefore, in the case of complex statistics, such as the at-risk-of-poverty rate, where the poverty threshold is estimated on the basis of the survey data, the computation of standard errors is confronted with many difficulties. Indeed, in this case there are two main sources of variability: one is due to the estimated threshold and the other one comes from the estimated proportion given the estimated threshold (Verma and Betti, 2011).

Several authors have derived formulae for standard errors in the case that the poverty line is estimated as a share of average or median income (for example Preston (1995)) while others have introduced alternative considerations such as stochastic dominance over a range of poverty lines (Davidson and Duclos, 2000) and the need to treat household size as a random variable (Thuysbaert, 2008). More specifically, when indices are used in comparing poverty among regions or over time, it should be recognized that each index establishes a particular ordering over income distributions, thus the ordering of a given set of distributions may depend on the choice of the index. The use of poverty dominance criteria allows making poverty comparisons in a more robust way since dominance ensures that all indices of some well-defined class unanimously prefer one distribution to another. For example, the popular FGT class of additive poverty indices are clearly related to the criteria of stochastic dominance, as was noted by Foster and Shorrocks (1988).

However, Thuysbaert (2008) stated that the limited empirical application of the theoretical results of stochastic dominance may be a result of the fact that the application is not always straightforward due to the presence of a stochastic weighting variable. This situation occurs when equivalent household income is obtained by using a weighted measure based on the number of household members. Therefore, by considering a bivariate distribution function defined over income and weight, Thuysbaert (2008) derived the limiting distributions of the decomposable poverty measures and of the ordinates of stochastic dominance curves. In this way, the poverty line may be allowed to depend on the income distribution. The suggested procedure is also illustrated by using Belgian data. Preston (1995) derived exact small and large sample distributions for proportions of a sample falling below given fractions of sample median income (or other sample quantiles). The resulting standard error formulae are then used to assess the statistical significance over time in relative income poverty in the UK, using data close to those underlying the statistics published by the Department of Social Security. Berger and Priam (2016) proposed the use of direct variance estimators. The main assumption underlying such estimators is that sample units have been selected with replacement, which considerably simplifies the estimation of the variances (Berger and Skinner, 2003). The proposed variance estimators are simple and flexible and can accommodate a wide class of sampling designs using standard statistical techniques.

However, in the framework of non-linear statistics, the variance of non-linear estimators cannot be given in closed form in most cases. In addition, unbiased estimates of variances of non-linear estimators may do not exist and therefore exact variance formulae cannot be estimated by conventional methods. In these cases, it is necessary to rely on approximate variance estimation techniques, including the linearization and resampling methods.

The FGT poverty measures P_α are calculated according to equation 1:

$$P_\alpha\left(y,z\right)=\frac{1}{n}\sum_{i=1}^{q}\left(\frac{z-y_i}{z}\right)^\alpha \tag{1}$$

where α is a real positive number, $y = (y_1, y_2, \ldots, y_n)$ is a vector of properly defined income in increasing order, $z > 0$ is a predefined poverty line, n is the total number of individuals under analysis, while $\frac{z-y_i}{z}$ is the normalised income gap of individual i and q is the number of individuals having income not greater than the poverty line z. The parameter α can be seen as a parameter of 'poverty aversion': the higher α, the higher the relevance assigned to the poorest poor. Therefore, the FGT class is based on the normalized gap $g_i = \frac{z-y_i}{z}$ of a poor person i, which is the income shortfall expressed as a share of the poverty line. Viewing g_i^α as the measure of individual poverty for a poor person, and 0 as the respective measure for non-poor persons, P_α is the average poverty in the given population.

The case $\alpha = 0$ yields a distribution of individual poverty levels in which each poor person has poverty level 1; the average across the entire population is simply the headcount ratio, denoted by P_0 or H. The case $\alpha = 1$ uses the normalized gap g_i as a poor person's poverty level, thereby differentiating among the poor; the average becomes the poverty gap measure P_1. The FGT class has certain advantages due to its simple structure—based on powers of normalized shortfalls—which facilitate communication with policymakers. Its axiomatic properties are sound and include the helpful properties of additive decomposability and subgroup consistency, which allow poverty to be evaluated across population subgroups in a coherent way (Foster et al., 2010).

The original FGT paper published in 1984 did not present the associated tools for formulating statistical tests and computing standard errors, but since then the literature has provided a steady stream of inference-based research for poverty estimation. Xu (2007) suggested that the sample counterparts to the FGT and other decomposable measures can be represented as a "*U-statistic*" (or some function of a U-statistic). Since as already mentioned the FGT class of additive poverty indices are related to the criteria of stochastic dominance, Davidson and Duclos (2000) derived the asymptotic sampling distribution of various estimators frequently used to order distributions in terms of poverty, welfare and inequality, including estimators of most of the poverty indicators currently in use. Davidson and Duclos (2000) also derived the sampling distribution of the maximal poverty lines up to which one may confidently assert that poverty is greater in one distribution than another. Kakwani (1993) provided distribution-free asymptotic confidence intervals and statistical inference for FGT poverty measures. Verma and Betti (2011) developed formulae and algorithms for the Taylor linearization and Jackknife Repeated Replication (JRR) methods covering the FGT class of poverty measures. Davidson and Flachaire (2007) used standard bootstrap methods funding that they perform very well with the FGT poverty measures and give accurate inference in finite samples.

Considering the above-mentioned properties of the FGT measures, in this report, among the Laeken indicators, we selected two income-poverty indicators that belongs to the FGT class and two income-inequality measures. More specifically we will focus on:

- **at-risk-of-poverty rate [ilc_li02]**. This indicator shows the share of persons with an equivalised disposable income below the risk-of-poverty threshold. This indicator is frequently disaggregated by age and gender, by household type, by tenure status, and work intensity. The AROP is a complex statistic since it is based on a poverty threshold computed from the median of the income distribution, that is:

$$AROP = P(X < 0.6q_{0.5}) \tag{2}$$

  Where $q_{0.5}$ is the median equivalised disposable income. Therefore, the at-risk-of-poverty threshold (ARPT) [ilc_li01] needs to be estimated first, which is set at 60% of the national median equivalised disposable income([4]) :

$$ARPT = 0.6q_{0.5} \tag{3}$$

  Then the AROP rate is defined as the proportion of persons with an equivalised disposable income below the ARPT:

$$\widehat{AROP} = \frac{\sum_{i \in I < \widehat{ARPT}} w_i}{\sum_i w_i} \tag{4}$$

- **relative median at-risk-of-poverty gap [ilc_li11]**. The indicator is calculated as the distance between the median equivalised total net income of persons below the at-risk-of-poverty threshold and the at-risk-of-poverty threshold itself, expressed as a percentage of the at-risk-of-poverty threshold;

- **quintile share ratio(S80/S20) [ilc_di11]**. This indicator is the ratio of total income received by the 20% of the country's population with the highest income (top quintile) to that received by the 20% of the country's population with the lowest income (lowest quintile);

([4]) For each person, the equivalised disposable income is defined as his/her total household disposable income divided by equivalised household size. The equivalised household size is defined according to the 'modified-OECD scale', which gives a weight of 1.0 to the first adult, 0.5 to other household members aged 14 or over and 0.3 to household members aged under 14. Each person in the same household receives the same equivalised disposable income

- **Gini coefficient [ilc_di12]**. The Gini coefficient measures the extent to which the distribution of income within a country deviates from a perfectly equal distribution. A coefficient of 0 expresses perfect equality, where everyone has the same income, while a coefficient of 100 expresses full inequality where only one person has all the income.

2.2 Issues in estimating sampling errors for regional poverty indicators

Even if there has been an increasing interest in the measurement of poverty and social exclusion across European countries at national level, there is also an urgent need for regional indicators. National level indicators are not necessarily appropriate or sufficient for regional analysis (Betti et al., 2012). Survey estimates are required not only for the whole population but also separately for many subgroups in the population. However, poverty information at detailed territorial levels comes at the cost of higher uncertainty levels. Information on the magnitude of sampling errors is therefore essential in deciding the degree of detail with which the survey data may be meaningfully tabulated and analysed.

In order to estimate regional indicators, it is essential to choose the type of units to serve as 'regions'. To this aim, the Nomenclature of Territorial Units of Statistics (NUTS) classification appears to be the most appropriate choice for EU countries. The NUTS classification covers each country exhaustively, providing a hierarchical set of units (NUTS level 1, 2 and 3) for which data can be linked across different levels. Countries are the highest classification units. Despite the fact that NUTS units are not defined exactly the same way in different countries and may differ in size and homogeneity, the NUTS classification system provides a widely used framework which allows us to improve the comparability of the resulting statistical information.

Moreover, although EU-SILC uses harmonized methods and variable definitions, each country can choose a specific sampling design according to the socio-economic and geographical structure of the country and population. The sampling design most commonly used by EU countries is the two-stage stratified sampling design. Stratifying a population means dividing it into non-overlapping subpopulations, called strata. Independent samples are then selected within each stratum. Stratification serves the purpose of increasing the representativeness of the sample and, at the same time, decreases the standard error. Most of the EU-SILC samples have been stratified by geographical region, which generally improves the accuracy of estimates. In addition, many of them have been clustered by, for instance, the so-called "Census Areas". Yet, as underlined by Osier (2009) although clustering reduces data-collection costs, it also tends to decrease the precision of estimates because the population units in a cluster are likely to be more homogeneous to each other than units of a simple random sample. In several countries, the EU-SILC survey is designed to produce direct accurate estimates at the macro level (NUTS 1-2 levels). Spatial variation is a relatively neglected dimension in poverty analysis both at the national and at the EU level. Since high national poverty rates may be accompanied by concentration of poverty in specific regions or, on the contrary, by widespread poverty across regions, it is paramount to estimate income-poverty and income-inequality indicators at regional level. Yet, standard error and confidence intervals of regional poverty indicators have been analysed only in a limited number of studies probably due to the lack of full documentation of the sample design variables in the EU-SILC dataset (Verma and Betti, 2010, Betti et al., 2012, Piacentini, 2014). In general terms, the relative magnitude of sampling error increases as we move from estimates for the total population to estimates for individual subgroups such as territorial regions. The quantification of these sampling errors is therefore essential in deciding the degree of detail with which the survey data may be meaningfully tabulated and analysed. Various methods for improving the precision of sampling error estimates of statistics based on small but complex samples, have been discussed by Verma et al. (2017). Therefore, for the purpose of estimating reliable income-poverty and income-inequality indicators at sub-national level, we focus on countries for which the geographical areas, represented by NUTS1 and NUTS2 levels, are considered as sample design domains. In this case, the above-mentioned variance estimation techniques, such as linearization, bootstrap, JRR, can be adapted for application at the sub-national level (Verma et al., 2017). Laureti and Rao (2019) and Biggeri et al. (2017) presented an empirical evaluation of uncertainty in measuring well-being at local level by focusing on the accuracy of the AROP and using the 2017 wave of EU-SILC for Italy in which detailed information on the sample design where provided by the Italian National Statistical Institute (ISTAT) thanks to a collaboration with the Dagum Inter-University Center. The authors used a generalized linearization method based on the concept of influence function (Deville, 1999) which allows to deal with non-linear statistics for which the linearization method cannot be used. In this framework, an interesting line for future research on the measurement of uncertainty in regional poverty indicators is the introduction of spatial price statistics into the definition of poverty threshold. An accurate measurement of price level differences across regions within a country is essential for a better assessment on regional disparities, thus enabling policy makers to adequately identify and address areas of intervention. Indeed, regional values of economic indicators such as Gross Domestic Product (GDP), income and poverty levels, should be adjusted for regional price differences measured by spatial consumer price indexes, following the same logic according to which the economic well-being of different countries is compared by taking into account international purchasing power parities (Laureti and Rao, 2019).

3 Methodological approaches: a review

3.1 Linearization approach

3.1.1 Common linearization methods

In order to estimate the variance of complex (non-linear) statistics, a long-established procedure, i.e. the linearization approach, may be used (Deville, 1999, Demnati and Rao, 2004, Osier, 2009, Verma and Betti, 2011). The Taylor linearization method (TLM) for variance estimation leads to proper results for statistics which can be expressed by functions which are continuously differentiable up to order two and are asymptotically normal. Linearization methods approximate the non-linear estimator by a linear function after which standard variance formulas for the given design can be applied to this linear approximation, justified on the basis of asymptotic properties of large populations and samples. For each statistic of interest, TLM seeks for each sample unit a linearized "indicative" variable, such that the variance of the total of the variable in question approximates the variance of the complex statistic of interest. Therefore, this method can handle all the non-linear statistics which can be expressed as a regular function of estimated totals or ratios. More specifically, the idea underlying this method is to approximate a non-linear statistic with a linear function of estimated totals. As a result, a variance estimator of the non-linear statistic is specified by a variance estimator of its linear approximation, that can be easily calculated. Intuitively, the linearization approach rests on the assumption that the sample-to-sample variation of a non-linear statistic around its expected value is small enough to be considered linear. This assumption is particularly correct when samples are large even if there is no definite evidence on how large a sample should be for the linear approximation to be valid (Osier, 2009). However, the TL approach cannot deal with all non-linear statistics and indicators based on EU-SILC due to the complex mathematical expressions which characterized these indicators. As an example, the at-risk-of-poverty rate is calculated on the basis of a poverty line which is estimated from sample observations, thus the indicators become more complex than a mere proportion. Consequently, variance estimation for the at-risk-of-poverty rate should take into account both the randomness which is brought by the at-risk-of-poverty threshold and that of the estimated proportion of "poor" individuals given the poverty threshold. In addition, Osier (2009) stated that there is some degree of covariance between the at-risk-of-poverty threshold and the at-risk-of-poverty rate which should be accounted for. Therefore, for statistics which cannot be expressed as a smooth function of estimated totals, other methods should be used for variance estimation.

3.1.2 The generalized linearization method based on influence function

In the context of the linearization approach, one alternative method for variance estimation is based on the concept of influence functions, introduced by Hampel (1974). An influence function measures the asymptotic bias caused by contamination in the observations on whose basis the statistic is estimated. In other words, it gives a picture of the infinitesimal behaviour of the asymptotic value of a statistic. Deville (1999) proposed that the precision of non-linear statistics in sampling designs be estimated using the generalized linearization method based on the concept of influence function. In addition to encompassing more non-linear statistics than the TL method, the linearization based on influence functions does not involve more calculations since the derivation rules for influence functions are similar to the rules for computing the derivative of a function in standard differential calculus. In order to define the influence function, Deville (1999) used a measure *M* with unit mass for each point of the population *T*. According to Deville's definition, the measure M is positive, discrete, with a total mass N while the total mass is equal to 1 for the influence function proposed by Hampel (1974). A function of interest θ can be presented as a functional $T(M)$ that associates for each measure a real number or a vector. More formally, let *U* denote a population of size *N* and let *M* be the measure which allocates a unit mass to each of the units x_k in *U*. We seek to estimate a population parameter θ which can be expressed as a functional *T* of the measure *M*:

$$\theta = T\left(M\right) \tag{5}$$

The specialization of the general measure *M* into a discrete measure turns the functional *T*, predefined on a continuum, into a discrete functional, in the same way as the total *Y* is defined as the sum of all y_k over the given finite population. The influence function of a functional *T*, is defined as:

$$IT\left(M,k\right) = \lim_{t \to 0} \frac{T\left(M + t\delta_k\right) - T\left(M\right)}{t} \tag{6}$$

where δ_k is the Dirac measure for unit *k*; $\delta_k(i) = 1$ if and only if $k = i$

A natural way to estimate 6 from a sample *S* of the population consists of plugging an estimated measure $\hat{M}$ of *M* into 6:

$$\hat{M} = T(\hat{M}) \tag{7}$$

The estimated measure $\hat{M}$ allocates the sample weight $w_i(s)$ for all units *i* in *S*, and 0 otherwise:

$$\widehat{M}(i) = \widehat{M}_i = \begin{cases} w_i(s) \text{ for } i \in S \\ 0 \text{ for } i \notin S \end{cases} \tag{8}$$

Deville (1999) showed that under broad assumptions, the substitution estimation of a functional $T(M)$ is linearizable. A linearized variable is $v_k = IT(M, x_k)$ where *IM* is the influence function of *T* in *M*.

This influence function $v_k = IT(M, x_k)$ is a linearized variable of $T(M)$ in the sense that it allows for the approximation of the function of interest.

The main result of this generalised linearization theory is that the variance of the so-called plug in estimator 6 can be approximated with a linear statistics:

$$Var[T(\widehat{M})] \approx var\left[\sum_{i \in S} w_k(s) v_k\right] \tag{9}$$

Osier (2009) explained the main features of the linearization approach based on influence functions and derived estimated standard errors, confidence intervals and design effect coefficients for the main target indicators (i.e. Laeken indicators) based on the EU-SILC. However, as underlined by Graf and Tillé (2014) the starting point of the approach suggested by Deville (1999) is the population parameter and not the estimator to be used for the evaluation using the sample. When the estimator being used follows naturally from the population parameter expression, as in the case of the total *Y* approached by the Horvitz-Thompson estimator, the procedure is unambiguous. Yet, inaccuracy may arise if the total Y is estimated using the ratio estimator with an auxiliary variable *x*. In this case, the approach proposed by Deville (1999) will yield a constant influence function equal to 1 since it does not specify the form of the total estimator to use.

Demnati and Rao (2004) suggested an alternative method that may avoid these problems for deriving Taylor linearization variance estimators that is based on representing Taylor linearization in terms of partial derivatives with respect to design weights. This method leads to variance estimators with good conditional properties (Antal et al., 2011). In the context of cross-sectional samples of complex design and of reasonably large size, Verma and Betti (2011) obtained linearized variables for estimating sampling errors of complex non-linear statistics involved in the analysis of poverty and income inequality by extending the use of standard variance estimation formulae (developed for linear statistics such as sample aggregates) to non-linear statistics. For any complex statistic, Verma and Betti (2011) developed a linearized variable λ_i such that the simple expression for its variance approximates the variance of the complex statistic. More specifically, the linearized variables for poverty measures involve reference to the density function at various points in income distribution, such as at the median or at the poverty line. Therefore, the derivative of the distribution function at a certain point $y\alpha, f\alpha = f(y\alpha) = (dF/dy)y\alpha$ can be estimated by means of density estimation techniques, such as the kernel estimator. Verma and Betti (2011) also compared the performance of the linearization technique with that of the Jackknife repeated replication method because of its widespread use and especially because the procedure has been officially adopted by Eurostat for the EU Statistics on Income and Living Conditions survey. Graf and Tillé (2014) used the generalized linearization technique based on the concept of influence function, following Osier (2009), to estimate the variance of complex statistics such as the Laeken indicators. By carrying out simulations, using the R language, Graf and Tillé (2014) showed that the use of Gaussian kernel estimation method for estimating an income density function results in a strongly biased variance estimate. Therefore, the authors proposed two other density estimation methods that significantly reduce the observed bias.

The linearization of the AROP rate based on the influence function is also provided by Münnich and Zins (2011). Many parameters can be expressed in the form of equation 6, for instance:

- the population total *Y* of a variable *y*:

$$Y = \sum_{i \in U} y_i = \sum_{i \in U} y_i \times M(i) = \int y dM = T(M) \tag{10}$$

- the ratio *R* of two population totals *X* and *Y*:

$$R = \frac{Y}{X} = \frac{\int y dM}{\int x dM} = T(M) \tag{11}$$

- the cumulative distribution function. Let inc_i be an income distribution over the population *U*. The cumulative distribution function *F* at *x* is the share of population elements whose income is lower than *x*:

$$F(x) = \frac{\sum_{i \in U} 1(inc_i \leq x)}{N} = \frac{\int 1(inc \leq x)\, dM}{\int dM} = T(M) \tag{12}$$

  where the function $1(inc_\bullet \leq x)$ is equal to 1 for all *i* whose inc_i is lower than *x* and 0 otherwise.

- the ARPT, that is 60% of the median income $MED(\hat{M})$): $ARPT = 0.6 \times MED(M) = T(M)$, where the median income $MED(M)$ splits the income distribution into halves: $F[M, MED, (M)] = 0.5$, where $F(M, .)$ designates the cumulative income distribution function.

If we consider the population total of a variable *y*, the so-called plug-in estimator can be written as:

$$\hat{Y} = T(\hat{M}) = \int y d\hat{M} = \sum_{i \in s} w_i(s) y_i \tag{13}$$

Likewise, if we consider the ratio 11 of two population totals *X* and *Y*, we get:

$$\hat{R} = T(\hat{M}) = \frac{\int y d\hat{M}}{\int x d\hat{M}} = \frac{\sum_{i \in s} w_i(s) y_1}{\sum_{i \in s} w_i(s) x_1} \tag{14}$$

Concerning to the cumulative income distribution, we obtain:

$$\hat{F} = T(\hat{M}) = \frac{\int 1(inc \leq x) d\hat{M}}{\int d\hat{M}} = \frac{\sum_{i \in s} w_i(s) 1(inc_i \leq x)}{\sum_{i \in s} w_i(s)} \tag{15}$$

Finally, regarding the ARPT, we have the following estimator:

$$\widehat{ARPT} = T(\hat{M}) = 0.6 \times MED(\hat{M}) \tag{16}$$

where the estimated median income $MED(\hat{M})$ satisfies $F\left[\hat{M}, MED, \hat{M}\right] = 0.5$ and:

$$F\left(\hat{M}, x\right) = \frac{\sum_{i \in s} w_i(s) \times 1(inc_i \leq x)}{\sum_{i \in s} w_i(s)} \tag{17}$$

By following the rule of derivation stated in Deville (1999), the influence function for the AROP at *k* is derived by Osier (2009):

$$IARPR_k(M) = \frac{1}{N}\left[inc_k \leq ARPT(M) - ARPT(M)\right] -$$

$$\frac{0.6}{N} \times \frac{F'[\widetilde{ARPT(M)}]}{F'[\widetilde{MED(M)}]} \times \left[1(inc_k \leq MED(M)) - 0.5\right] \tag{18}$$

Where $F'[\widetilde{ARPT(M)}]$ and $F'[\widetilde{MED(M)}]$ are the values of the derivative of the cumulative income distribution function $\widetilde{F}$ at the points $ARPT(M)$ and $MED(M)$ (median income), respectively. These two quantities can be interpreted as the income densities at $ARPT(M)$ and $MED(M)$.

The influence function 18 can be regarded as a sum of two terms: the first term is the influence function that would be obtained assuming the ARPT (M) is constant, while the second term is a "correction" which takes into account the fact that the ARPT threshold is estimated from sample observations.

3.1.3 Estimating equations and other approaches

Another approach to linearization is based on the use of estimating equations. Estimating Equations (EE) is a technique which can be applied to derive both point estimates and their corresponding linearized values used for variance estimation (Binder, 1991). A general formulation of the EE approach for large sample complex surveys is given in Binder and Patak (1994). More specifically, Kovacevic and Yung (1997) applied the EE method for estimating the coefficient of variation and the exponential measure of inequality. The advantage of this approach when compared to other resampling alternatives is that it can be used under a wide class of sampling designs and does not require intensive computations. Betti and Gagliardi (2018) present a practical methodology for variance estimation for multidimensional measures of poverty and deprivation of households and individuals, derived from sample surveys with complex designs and large sample sizes. The authors applied this methodology in a multi-domain and comparative context by considering measures based on fuzzy representation of individuals' propensity to deprivation in monetary and diverse non-monetary dimensions. The basic idea of the fuzzy approach is to treat poverty and deprivation as a matter of degree, replacing the conventional poor/non-poor dichotomy. An individual degree of poverty is determined by the person's place in the income distribution (see Betti and Gagliardi (2018) for further details).

3.2 Replication methods

An alternative approach to variance estimation of nonlinear statistics from complex samples is provided by replication methods based on repeated resampling of the parent sample. This class of methods includes the Bootstrap, Jackknife Repeated Replication (JRR), and Balanced Repeated Replication (BRR) and is based on measures of observed variability among replications of the full sample. The basic requirement is that the full sample is composed of several subsamples or replications, each with the same design and reflecting complexity of the full sample, numbered using the same procedures. Although a replication differs from the full sample only in size, its own size should be large enough to reflect the structure of the full sample and for any estimate based on a single replication to be close to the corresponding estimate based on the full sample. Simultaneously, the number of replications available should be large enough so that the comparison among replications gives a stable estimate of the sampling variability in practice. The various resampling procedures available differ in the manner in which replications are generated from the parent sample and the corresponding variance estimation formulae evoked (such as the Balanced Repeated Replication (BRR) and the bootstrap, apart from JRR).

Bootstrap inference for inequality measures have been carried by various researchers (Mills and Zandvakili, 1997, Biewen, 2002). The bootstrap is a method for recovering the distribution of a statistic by employing Monte Carlo simulation methods to approximate the small sample distribution (see for example Efron and Stein (1981), Tibshirani and Efron (1993) for detailed expositions of the bootstrap). This method provides a numerical approximation to the distribution of interest, F, that is similar to a high-order Edgeworth expansion which can represent considerable improvements over Normal approximations. In the context of inequality measurement, the bootstrap was first applied by Mills and Zandvakili (1997) who considered the use of bootstrap for computing interval estimates and performing hypothesis tests for decomposable measures of economic inequality. They provided two applications of the suggested approach, using the Gini coefficient and Theil's entropy measures of inequality based on the data from the US Panel Study of Income Dynamics. Bootstrap intervals are computationally inexpensive and easy to calculate, the same method applies to all the inequality measures used in the literature, and the bootstrap method automatically considers any bounds that apply to a particular measure. Further, since bootstrap intervals computed using the percentile method have a clear Bayesian interpretation, they provide a straightforward solution to the Behrens-Fisher problem of comparing means from two distributions. Given the potential advantages from bootstrapping, it appears worthwhile to consider its use as a tool for statistical inference for inequality measures. Indeed, Biewen (2002) proved the validity of the bootstrap method for various indicators used in the context of inequality, mobility and poverty measurement.

Originally introduced as a technique of bias reduction (Durbin, 1959), the Jackknife method has by now been widely tested and used for variance estimation (Verma and Betti, 2011). For a detailed description of the Jackknife method see Efron and Stein (1981). Like other resampling procedures, the JRR method estimates the sampling error from comparisons among sample replications which are generated through repeated resampling of the same parent sample. Each replication needs to be a representative sample in itself and must reflect the full complexity of the parent sample. Nevertheless, as the replications are generally not independent but are overlapping, special procedures are required for constructing replications in order to avoid bias in the resulting variance estimates. The JRR variance estimates take into account the effect on variance of aspects of the estimation process which are allowed to vary from one replication to another, including complex effects such as those of imputation and weighting.The JRR method proved to give satisfactory results for means, ratios and functions of ratios, which are by far the most commonly encountered statistics in survey analysis. Regarding more complex statistics, the method can be safely used for indicators named as "U statistics". The JRR procedure performs well when the parameter of interest is a smooth function of sample aggregates while it may not provide a consistent variance estimator for non-smooth statistics, such as the median or other quantiles of the income distribution. Shao and Wu (1989) demonstrated that this drawback of the basic Jackknife method can be corrected by using a more general form which involves the construction of replications by deleting a number of observations simultaneously, the appropriate number to be deleted depending on the degree of smoothness of the measure.

3.3 Evaluating the relative performance of different methods

For poverty indicators computed using the EU-SILC surveys, which are generally based on reasonably large samples but with complex designs, Taylor linearization is well-established in the literature. However, the linearization method is not always the most practical procedure for variance estimation for the type of statistics and samples being considered. Moreover, when non-linear statistics are involved, as it is frequently the case in poverty analysis, it is difficult to identify a closed form for estimators of their standard errors. In these cases, the class of methods based on the idea of resampling or replication, such as the Bootstrap and Jackknife Repeated Replication (JRR), may provide an alternative.

The S80/S20 constitutes a very interesting class of inequality indices in their capacity to detect perturbations at different levels of an income distribution. Yet, variance estimation is not straightforward, especially when dealing with complex sampling designs (Langel and Tillé, 2011). Several studies use variance estimators based on the linearization approach suggested by Deville (1999). Inference for the S80/S20 using this approach has already been conducted by Osier (2006) and Osier (2009) while bootstrap methods have been analysed by Antal et al. (2011). In a similar framework, Brzezinski (2014) studied finite-sample performance of asymptotic inference for richness measures, suggesting that the asymptotic inference for the richness headcount ratio and the concave richness indices is satisfactory even for samples of moderate size. In many cases, standard bootstrap inference gives a small improvement over the asymptotic inference, but both approaches can be considered reliable for samples of 1000 or larger. It is worth noting that the performance of the standard bootstrap can be improved in some cases using a semi-parametric bootstrap. The estimation of the variance of the Gini coefficient indicator can be based both on linearization methods and resampling-based methods (Giorgi and Gigliarano, 2017, Langel and Tillé, 2013). Linearization techniques are aimed to obtain the asymptotic distribution of the Gini index which, in the case of simple survey sample, was discussed in Hoeffding (1992) as an application of his general results on U-statistics. Other approaches related to linearization methods for estimating the standard error of the Gini estimator are based on the influence function of the Gini index (Cowell and Victoria-Feser, 2003) and on estimating equations (Kovacevic and Yung, 1997). Since the Gini index is usually estimated through data from complex surveys, as in the case of EU-SILC surveys, several works have focused on deriving an expression of the Gini index by also considering survey design in the estimation process (Bhattacharya (2007); Langel and Tillé (2013)).

A systematic procedure for the derivation of linearized variables for the estimation of sampling errors of income inequality measures, including Gini and S80/S20 was developed by Verma and Betti (2011). As an alternative, estimation of the variance of the Gini index estimator can be based on re-sampling methods. Bootstrap techniques have been implemented for estimating the Gini variance in various studies (see for example Palmitesta et al., 2000). In the context of inequality measurement, the bootstrap was applied by Mills and Zandvakili (1997), Palmitesta et al. (2000), Palmitesta and Provasi (2006), Biewen (2000) and Biewen (2002) who recommended its use rather than asymptotic methods especially in applications where the sample size is not large. The consistency of the bootstrap can be shown by using general results on U-statistics (Xiquan, 1986). Several authors have suggested using the jackknife technique to approximate a standard error for the Gini coefficient. The jackknife approach was firstly used by Manfredi (1974), who proved that applying this method could lead to getting less biased estimates than those obtained with traditional methods (Giorgi and Gigliarano, 2017). Schechtman (1991) and Yitzhaki (1991) suggested the use of jackknife estimators of the variance of the plug-in Gini estimator based on the influence function while Ogwang (2000) proposed a fast algorithm for a jackknife estimation of the Gini coefficient's variance.

Since linearization, jackknife and bootstrap are estimating the same quantity, that is the variance of income-poverty and income-inequality indicators, one may ask if it is possible to identify conditions under which some estimators perform better than others. In this context, several comparisons, both theoretical and empirical (by simulation), have been carried out in the literature. Rao and Wu (1985) and Rao and Wu (1988) demonstrated that different Jackknife and Balanced Repeated Replication estimators are very close to one another and that the jackknife is closest to the linearization estimator, followed by the bootstrap estimator. Overall, the jackknife and linearization methods tend to exhibit similar performance. They are more stable for smooth functions but inconsistent for non-smooth functions. Bootstrap is applicable for all statistics and can provide more-accurate one-sided confidence intervals and better balance of the tail probabilities of two-sided confidence intervals. However, this versatility comes at the price of lesser stability. Davidson and Flachaire (2007) studied finite-sample performance of asymptotic and bootstrap inference for inequality and poverty measures founding that neither asymptotic nor standard bootstrap perform very well with inequality measures, even in very large samples. The reasons for this poor performance are related to the fact that inference for inequality measures is very sensitive to the exact nature of the upper tail of the income distribution. Indeed, an issue frequently encountered with heavy-tailed distributions like Generalized Beta, Dagum and Pareto distributions, which describe real-word income data, is that extreme observations are frequent in data sets, thus causing difficulties in the bootstrap technique. Similarly, Cowell and Flachaire (2007) examined the statistical performance of inequality measures in the presence of extreme values in the data showing that these indices are very sensitive to the properties of the income distribution. In contrast, Davidson (2012) demonstrated that, unless the tails are very heavy, or the Gini index itself is large, the bootstrap can yield acceptably reliable inference. Linearisation is efficient in computation and may perform better in some cases. However, the derivation of the computational forms is much more complex and in some cases may not be available in a form readily usable. Replication method can handle such complexity much more easily. Verma and Betti (2011) stated that their simplified procedures for linearization and jackknife perform well when the parameter of interest is a smooth function of sample aggregates, such as the class of FGT indicators (Ravallion et al., 1994). Contrastingly, as in the case of bootstrap an often well-known drawback of the jackknife method is that it may not provide a consistent variance estimator for non-smooth statistics, such as the median or other quantiles of the income distribution.

The performance of the various estimation methods for the variance of income-poverty and income-inequality measures is also strongly influenced by the type of sample design. Indeed, it has been found that the jackknife method works satisfactorily for variance estimation in the case of Gini but not for all sampling designs. The same consideration holds for the bootstrap approach. Berger (2008) compared numerically the jackknife estimators with two linearization estimators showing, on the basis of a simulation study, that linearization technique proposed by Kovacevic and Yung (1997) and the generalised jackknife are asymptotically equivalent and consistent under mild conditions. As underlined by Graf et al. (2011) and Ollila (2004) there is not a unique method working well with any sampling design and any complex estimator. Berger (2008) use jackknife and linearization to estimate the variance of the Gini coefficient, allowing for the effect of the sampling design. The complexity of the sampling error estimates of these measures is further increased by the fact that the empirical income distribution from which they are derived is itself subject to sampling variability. However, in the case of the replication methods, the main requirement is efficient and accurate code for repeated computation of the involved measures (Goedemé, 2013). This is a subject-matter specific requirement. In a different way, the same applies to the linearisation approach: the computational forms are measure-specific. As we will explain in the next section we had to manipulate the data and write specific programs for estimating standard errors for inequality and poverty measures in order to adapt formulae to the characteristics of sampling design.

4 The EU-SILC survey

4.1 Main characteristics of the EU-SILC

The EU-SILC survey is currently implemented in 37 countries([5]) . Every year in Europe more than 200,000 households and 500,000 individuals are interviewed and the complete microdata are sent to Eurostat by NSIs (Atkinson et al., 2017). The EU-SILC survey is the main source for the compilation of statistics on income, social inclusion and living conditions since it collects comparable multidimensional micro-data on: (a) income, (b) poverty, (c) social exclusion, (d) housing, (e) labour, (f) education, (g) health. In terms of statistical units two types of variables measured and analysed are thus involved in EU-SILC: (a) variables at household level (education, basic labour information and second job) and (b) variables at personal level (health, access to health care, detailed labour information, activity history and calendar of activities). These "target variables" are either compiled from registers (register variables) or collected from the sampled units (observation variables). EU-SILC is organised under a framework regulation and is thus compulsory for all EU Member States. EU-SILC is based on the idea of a "common framework" in contrast with the concept of a "common survey". The common framework is defined by harmonised lists of target primary (annual) and secondary (every four years or less frequently) variables, by a recommended design for implementing EU-SILC, by common requirements (for imputation, weighting, sampling errors calculation), common concepts (household and income) and classifications (ISCO, NACE, ISCED) aiming at maximising comparability of the information produced. The common framework is defined in the legislative background of the project, the Council and European Parliament framework Regulation, and the implementing Commission Regulations. SILC provides two types of annual data:

- Cross-sectional data pertaining to a given time or a certain time period with variables on income, poverty, social exclusion and other living conditions,
- Longitudinal data pertaining to individual-level changes over time, observed periodically over a four year period. They aim at measuring gross (micro-level) change and elucidating the dynamic processes of social exclusion and poverty (Verma and Betti, 2006).

There are two kinds of variables in EU-SILC: the primary and secondary variables. The primary (target) variables are collected every year, whereas secondary variables are collected every five years or less frequently in the so-called ad-hoc modules. Based on these variables, additional variables (derived variables) are calculated for each statistical unit-observation, to support the computation of indicators used for monitoring poverty and social inclusion in the EU within the Europe 2020 strategy.

Many revisions and improvements have been made to the EU-SILC survey over the years. The latest change is the adoption of the Regulation 2019/1700 which establishes a common framework for European statistics relating to persons and households, based on data at individual level collected from samples (IESS Regulation). The new EU-SILC legal acts require:

([5]) The national questionnaires received from countries are available here.

- improved timeliness, with shorter deadlines for EU-SILC data submission;
- reformulated precision requirements both at national and regional level (NUTS2) for the at-risk-of-poverty-or- social-exclusion indicator and the persistent-risk-of-poverty rate;
- additional/ changed EU-SILC variables;
- data collection in three frequencies: nucleus, three-year module and six-year module;
- prolongation of the longitudinal panel.

The aim of the revision process is to obtain reliable statistics both at national as well as at regional level (where better comparability is required from the REGULATION (EU) 2019/1700). In this framework it is important that aggregated data be made available for comparable territorial units such as NUTS2. In order to establish comparable regional statistics, data on the territorial units should be provided in accordance with the NUTS classification.

4.2 EU-SILC sampling designs across European countries

The Framework Regulation requires the selection of nationally representative probabilistic samples. Units used in the sample selection may be addresses, households or individuals; each unit is selected with a known probability and collected according to the design chosen by the country. As already mentioned, variables are collected by register or interview survey. In most of the countries (the so-called "survey" countries), both income and non-income variables are collected by interviewing all household members. On the other hand, a set of countries use population registers with income information. For all household members, registers provide information for income, education and housing. In this case, income variables are fully collected by register and the questionnaire is more focused on qualitative questions. In "register" countries the *selected respondent* is only one person in each household and he/she answers to most non-income questions. Randomised selection procedures must be used to ensure that a representative sample of persons is obtained from the representative sample of households. Table 1 shows the sampling units used for each country.

Sampling unit	Country
Dwelling/Address	CZ, DE, ES, FR, HR, LV, LU, MT, AT , PL, PT, RO, UK
Household	BE, BG, EE, IE, EL, IT, CY, LT, HU, SK, CH
Individuals	DK, NL, SI, FI, SE, ES, NO

Table 1: Sampling units chosen by Country

In order to ensure both longitudinal and cross-sectional representativeness, Eurostat has suggested using an integrated structure that is a rotational panel([6]) . Rotational design refers to the sample selection which is based on a number of sub-samples. The sample for any one-year consists of four sub-samples, which have been in the survey for 1-4 years. From year to year, some sub-sample are maintained, while others are dropped and replaced by new sub-sample. With this methodology, cross-sectional and longitudinal statistics are produced from essentially the same set of sample observations. Although EU-SILC uses harmonized methods and definitions in order to establish reliable comparisons between EU Member States, there are considerable differences in sample designs among EU countries. The specific sampling design is chosen according to the socio-economic characteristics of the population and geographical structure of the country. Every year, each EU country sends to Eurostat information on the sampling design used, on the strata and primary sampling units (PSU) from which each household is drawn. Table 2 shows the sampling design chosen by countries according to the EU comparative quality reports([7]) .

([6]) All countries adopted the four-year rotational design recommended by Eurostat, except for France and Norway where a longer panel duration (eight and nine years, respectively) is used.
([7]) For further details you find here the quality reports.

Table 2: Sampling design chosen by Country

Type of sampling design	Countries
Stratified sampling design:	
One stage sampling	DK, DE, EE, CY, LT, LU, AT, SK, FI, CH
Multi-stage sampling	BE, BG, CZ, IE, EL, ES, FR, IT, LV, HU, NL, PL, PT, RO, SI, UK, HR
Sampling design without stratification	MT, SE, IS, NO

As illustrated above, most of the EU-SILC samples are stratified by geographical region, which generally increases the accuracy of estimated indicators. In addition, many sampling design consider cluster sampling thus reducing data-collection costs but in some cases also decreasing the precision of estimates. The loss of precision may be due to the fact that the population elements in a cluster are likely to be more homogeneous to each other than elements of a simple random sample. To this respect, the Design Effect (Deff), that is, the ratio of the variance of a statistic with a complex sample design to the variance that would be obtained with a simple random sample of same size, can be estimated in order to measure the combined effect of design components, such as stratification, clustering or unequal weights. Stratification serves the purpose of increasing the representativeness of the sample and decreasing the risk that sub-groups in the population remain unrepresented. If the variance between strata is large in terms of the relevant variable, the stratification contributes to decreasing the standard error. Clustering can seriously increase the standard error if the variance within clusters is small compared to the between-cluster variance with respect to the variable of interest. In contrast, if clustering is neglected, standard errors will be under-estimated and relations which are not statistically significant may appear to be significant (Trindade and Goedemé, 2020). The geographical clustering could reduce the costs of the survey (Sturgis, 2004) allowing the collection of the interviews in a limited number of geographical areas. The computation of standard errors usually requires a complete description of the implemented sample design and accurate description of stratification and clustering variables regarding "Primary strata" and "PSUs". If strata are not taken into account, confidence intervals can be overestimated and the researcher might be unduly conservative.

In order to take into account the sample design when estimating poverty and social exclusions indicators, the following variables should be considered: household id (DB030), the stratification variable (DB050) and primary sampling unit variable (DB060). Every household will receive a household number (HB030). This number is the base on which which the Household ID and the Personal ID are constructed. It should be a sequential number and it should not contain other information. This number must be unique for all the years of the survey. The DB050 provides an identification code for the strata in case the target population is stratified at the first stage of the sample design. DB050 refers only to explicit strata([8]) . The DB060 provides an identification code for the PSUs. Every selected PSU should receive a value that is unique across all PSUs that have been selected in EU-SILC and that remains the same for the entire duration of EU-SILC. According to the information required for a complete description of the sampling design, EU countries can be divided into three groups:

- Belgium, Bulgaria, Czech Republic, Ireland, Greece, Spain, France, Croatia, Italy, Latvia, Hungary, the Netherlands, Poland, Portugal, Romania, Slovenia and the United Kingdom, where the variable DB050 (primary strata) can be used for strata specification and DB060 (Primary Sampling Unit) for cluster specification;
- Germany, Estonia, Cyprus, Lithuania, Luxembourg, Austria, Slovakia, Finland, Switzerland, where the variable DB050 can be used for strata specification and DB030 for cluster specification;
- Denmark, Malta, Sweden, Iceland, Norway, variable DB030 can be used for cluster specification and no strata are specified.

([8]) Variable DB050 is included in the D-file of the EU-SILC survey.

4.3 Previous projects for uncertainty estimation of EU-SILC based indicators

As EU-SILC is a sample-based survey, a great variety of errors can seriously affect the accuracy of all estimates. Several frameworks exist to classify sources of errors. Generally non-sampling errors in EU-SILC survey are described (Verma et al., 2010, Goedemé, 2013) while survey data are subject to errors from diverse sources. Information on sampling errors is of crucial importance in proper interpretation of the survey results, and for the purpose of evaluating and improving the sample design, including sample size.

Measuring sampling errors is an important step in assessing the accuracy, as confidence intervals in which the population value lies with a high probability can be easily derived. Over the last years, important progress has been made in the assessment of data accuracy and survey errors in EU-SILC survey by also taking comparability of the results across the national surveys as a basic requirement. This is the result of several research projects developed by National Statistical Institute (NSIs) and Eurostat such as the Network for the Analysis of EU-SILC (Net-SILC), an ambitious 18-partner Network bringing together expertise from both data producers and data users, and the Advanced Methodology for European Laeken Indicators (AMELI), focused on methodological aspects of Laeken indicators especially regarding their impact on policy making.

The Net-SILC was established in 2008 by Eurostat and includes various institutions and researchers. The main results of Net-SILC have been presented at an international conference that took place in Warsaw, on 25-26 March 2010. In line with the previous Network (Net-SILC1), the aim of Net-SILC2 is to elaborate methodology for the analysis of the EU-SILC data, and to develop common tools and approaches regarding various aspects of data production. The main objective of the Net-SILC2 is to develop a practicable set of recommendations both for data producers and data users regarding standard error estimation. Those recommendations include suggestions concerning the procedures for computing standard errors at NSI's level and non-NSI's level. Since, the EU-SILC User Data Base version does not convey enough information which would allow data users outside NSIs to compute reliable standard errors estimates for a given set of indicator, Goedemẽ et al. (2013) includes recommendations concerning how to improve the quality of sampling design variables. In this framewok, in order to estimate standard errors for the main EU-SILC indicators, Osier (2009) implemented linearization techniques by using the package Poulpe([9]) . Although the package Poulpe is powerful enough to take into account the main sample design features, this solution turned out to be difficult to implement from the second wave onwards, especially because of the rotational structure of the sample.

In 2018, the Net-SILC reached the third edition, called Net-SILC3([10]) The aim of the first working package of Net-SILC3 is to provide analyses of non-sampling errors in the EU-SILC survey. It is designed to identify the main sources of non-sampling errors, to describe the nature and impact of each type of error and to produce guidance for reducing them. The preliminary findings from Net-SILC3 work packages have been presented at the Unit non-response and weighting and Item non-response and imputation international workshops at University of Essex, held in the period 20-22 February 2019. Therefore they are particularly suitable to illustrate the results of our analyses at national and sub regional level.

([9]) POULPE is an SAS macro program for sampling variance estimation. It is very exact on applied formulae but quite demanding to use. POULPE takes into account the impact of calibration on variance estimation (Caron, 1998)

([10]) It was built on the basis of the solid research carried out in the context of the Net-SILC1 and Net-SILC2 networks. The final outcome of Net-SILC2 was a book on "Monitoring social inclusion in Europe" (eds. Atkinson et al., 2017), which can be downloaded free of charge.

5 Computational aspects and results

5.1 Data manipulation and selection of countries

The computations were made within Eurostat premises using the EU-SILC PDB for the year 2018 to allow developing performance analyses of various standard error estimation methods with the aim of improving official statistics. This source of data includes full information on primary strata and almost complete information on PSUs for all countries. In our analysis we used the additional variables already computed by Eurostat in order to ease further statistical computations. We referred to eight countries according to their sampling designs: we selected five countries using the two-stage stratified sampling design, that is Italy, the United Kingdom, Portugal, Belgium and Ireland; two countries using the one-stage sampling design (Germany and Sweden) and the Norway which adopts a simple sampling design without stratification. In addition to different sample designs and sampling units, these countries show important differences as regards the auxiliary information and variables considered in the sampling design.

5.1.1 The Bootstrap method and Laeken package

The first method we decided to use for variance estimation is the Bootstrap. This analysis has been performed by using the R package Laeken (Alfons and Templ, 2012). The function `variance()` included in the Laeken package provides a flexible framework for estimating the variance for the selected poverty indicators. One of the most convenient features of the package Laeken is that point and variance estimation can be obtained for different sub-domains using a single command. In our case, we estimated standard errors for poverty indicators for NUTS1 and NUTS2 regions in addition to the overall national values by including the "Breakdown" argument in the function. Moreover, since stratified sampling designs have been considered, in the Laeken package the specification of the design argument allows separate re-sampling within each strata. Let $X = (X_1; ...; X_n)^{'}$ denote a survey sample with n observations and p variables. Then the naive bootstrap algorithm, which is called also the Standard Normal Approach, for estimating the variance of poverty indicator can be summarized as follows:

- Draw S independent samples $X_1^*, ..., X_S^*$ from X;
- Compute the bootstrap replicate estimates $\widehat{\theta}_s^* = \widehat{\theta}(X_s^*)$ for each X_s^* $s = 1, ..., S$ where $\widehat{\theta}$ denotes an estimator for the poverty indicator of interest. The sample weights are always needed for the computation of the bootstrap replicate estimates;
- Estimate the variance $V\left(\widehat{\theta}\right)$ by using the variance of the S bootstrap replicate estimates:

$$\widehat{V}\left(\widetilde{\theta}\right) = \frac{1}{S-1}\sum_{s=1}^{R}\left(\widehat{\theta}_s^* \frac{1}{s}\sum_{r=1}^{S}\widehat{\theta}_s^*\right)^2 \tag{19}$$

5.1.2 The Jackknife Repeated Replication method (JRR)

The second method used in order to obtain variance estimation for poverty indicators is the JRR. This method is based on comparisons among replications generated through repeated re-sampling of the same sample. Verma and Betti (2011) provide a general description of JRR and other practical variance estimation methods in large-scale surveys. In order to estimate variance with JRR some assumption are required: the sample selection is independent between strata, two or more primary selections are drawn (independently and with replacement) from each stratum and the number of primary selections should be large enough.

Moreover, some redefinitions of units and strata in the given sample are necessary to ensure that the basic requirement of at least two independent selections per stratum is met (Verma and Betti, 2011). Primarily, such redefinition involves grouping or collapsing of PSUs (and possibly also of strata) in the sample in order to obtain larger and more uniform computational units for variance estimation. This method implies that at least four PSUs are included in each stratum. We followed the procedure suggested by Verma and Betti (2011) thus requiring that the number of replications should be equal or at least similar to the number of PSUs in the sample. More specifically, in order to follow a unique procedure for all the countries, the number of replications is equal to 2,000. Each JRR replication is formed by eliminating one sample PSU from a particular strata at a time and increasing the weight of the remaining sample PSUs in that stratum appropriately. In this way it is possible to obtain an alternative but equally valid estimate to that obtained from the full sample. Let j indicate a sample PSU and k indicates the stratum; $a_k \geq 2$ is the number of PSUs in stratum k, assumed to be selected independently. Let λ be a full sample estimate of any complexity, and λ_{kj} the estimate produced using the same procedure after eliminating primary unit j in stratum k and increasing the weight of the remaining (a_{k-1}) units in that stratum by $g_{kj} = \frac{w_k}{w_k - w_{kji}}$ with $w_k = \sum_j w_{kj}, w_{kj} = \sum_i w_{kji}$ the sum of the sample weights of ultimate units i in PSU j.

This means that the weights for individual units are redefined in replication kj as follows:

- For a unit i not in stratum k: $w^{'}_{kji} = w_{kji}$;
- For a unit i in stratum k but not in PSU j: $w^{'}_{kji} = g_{kj} w_{kji}$;
- For a unit i in stratum k and PSU j: $w^{'}_{kji} = 0$.

Let λ_k be the simple average of the λ_{kj} over the a_k values of j in k. The variance of λ is then estimated as:

$$Var(\lambda) = \sum_k \left[(1 - f_k) \frac{a_k - 1}{a_k} \sum_j (\lambda_{kj} - \lambda_k)^2 \right] \tag{20}$$

In this case the finite population correction $(1 - f_k)$ is close to 0. The aggregate quantity λ_{kj} for replication (kj) has been computed by taking a weighted sum of values for individual units with modified weights w_{kji}. With the delete one-PSU at a time JRR model, each term in $Var(\lambda)$ is the contribution of a single PSU to the variance of the whole sample. The average contribution per PSU (or per replication) is then summed over all replications to obtain an estimate of total variance.

5.1.3 The Linearization method

In order to estimate the variance of non-linear indicators the linearisation approach, based on the concept of influence function, has been used as proposed by Osier (2009) and Berger et al. (2017). As specified in section 3.3.1 of Deliverable 4.A, this method has been firstly introduced in robust statistics by Hampel (1974). This method was implemented by Eurostat to produce variance estimates for the EU-SILC social indicators: it simplifies the estimation of the variances and can be easily extended to cover multi-stage designs by using the ultimate cluster approximation. With the ultimate cluster approach, the variance between PSUs is used as an approximation of the total sampling variance. The variance of the linear approximation can be used as an approximation of the variance of the non-linear indicator considered. The "vardpoor" package in R has been used in order to obtain linearized versions of the non-linear poverty measures used (the at-risk-of-poverty rate, Gini coefficient, income quintile share ratio and relative median at risk of poverty gap). Suppose θ is a complex non-linear indicator. The variance of an estimator $\widehat{\theta}$ of θ is estimated by:

$$\widehat{V}\left(\widehat{\theta}\right) = \sum_{h=1}^{H} \frac{n_h}{n_h - 1} \sum \left(z_{hi.} - \overline{z}_{h..}\right)^2 \tag{21}$$

Where $z_{hi.} = \sum_{j=1}^{m_{hi}} \omega_{hij} \cdot z_{hij}, \bar{z}_{h..} = n_h^{-1} \left(\sum_{i=1}^{n_h} z_{hi.}\right)$ and z_{hij} is the value of a linearised variable. We introduced a criterion to proper consider the case in which a strata includes only a PSU. Since as already emphasized, we do not have information on the Order of selection of PSU variable, we decided to regroup (collapse) PSUs. In this way, each stratum contains at least four sample PSUs – even if the minimum number required for the computation of variance is two.

5.2 Standard errors estimation results

This section illustrates estimated standard errors of poverty measures for the eight countries selected for our empirical application both at national and regional level. Table 3 provides the following information for each country: achieved sample size, number of NUTS1 and number of NUTS2.

Table 3: Countries sample size and NUTS classification

Country	n	NUTS1	NUTS2
Two stage stratified sampling			
IT	45,767	5	21
UK	38,776	12	42
PT	33,935	3	7
BE	13,767	3	11
IE	11,130	1	3
One stage stratified sampling			
DE	25,259	16	38
Sampling design without stratification			
NO	14,315	1	7
SE	14,403	3	8

As we can observe from Table 3, sample size varies greatly from one country to another. This is an essential factor for explaining national differences as regards to the accuracy of income-poverty and income-inequality indicators. In general, there is a positive correlation trend between effective sample size and relative standard error. As reported in Section 2, regarding the income-inequality measures, we selected the Gini coefficient and the quantile share ratio (S80/S20) indicators; as for the income-inequality indicators, we selected the At-risk-of-poverty rate (AROP) and the Relative median at-risk-of-poverty gap (RMPG).

5.2.1 Results at national level

Tables 4 - 6 compare, when available, our point and standard error estimates with the official figures obtained from Eurostat. In each table, the first four columns illustrate our estimation results by comparing bootstrap (B), JRR and linearization (L) methods, while the last two columns include the official Eurostat estimates.

Regarding our estimates, the general closeness of the results from three entirely different methodologies can be considered quite remarkable. Similar results were found for the jackknife by Verma and Betti (2011) from a national survey based on a complex (stratified, two-stage, weighted) sample. In addition it is essential to note that our standard error results show a satisfactory level of reliability for all the methods used, since the estimated coefficients of variation are lower than 5% as underlined by Ardilly (2006). Indeed, even if precision thresholds are generally survey-specific and depend on the required reliability and resource-related political decision, specifying what degree of precision is an important step when planning a sample survey. Our results also confirm previous finding concerning the poor performance of the boostrap method for estimating standard errors for income-inequality indicators (Gini coefficient and S80/S20) in the case of the UK and Belgium. The estimation results may be influenced by the dimension of strata and the implicit stratification used by Belgium and the United Kingdom, where PSUs are descendingly sorted according to variables strictly related to income. Using the jackknife and linearization methods we are able to better capture both the reduced number of PSUs and, to some extent, the implicit stratification by defining computational strata. However, our findings demonstrate that there is not a unique method working well with any complex estimator and sampling design implemented in the various countries (Ollila, 2004).

Looking at the uncertainty measures for poverty indicators for Italy, the relative standard errors of estimates tend to be lower when bootstrap is used. If we consider the AROP, the relative standard error ranges from 1.63% (for bootstrap) to 2.42% (for jackknife). Standard errors obtained using the jackknife procedure are similar to those obtained using linearization for almost all the poverty indicators with the exception of the Gini coefficient. Yet, bootstrap may under-estimate the true value of standard error because it may be less able to reproduce strata and PSUs thus being similar to a simple random sample.

In the case of the United Kingdom, the linearization method produces lower standard errors than bootstrap and jackknife for the AROP and S80/S20. Contrastingly, bootstrap seems to perform well only in the case of RMPG. For Portugal, standard errors for the Gini coefficient are very similar among the three methods that were used. For the AROP and S80/S20 jackknife produces higher values than bootstrap and the linearization methods. As already mentioned, the jackknife performs very well for Belgium and Ireland for the various measures.

Focusing on countries adopting a one-stage sampling structure (Germany) and a simple random sampling design (Norway and Sweden) we can observe that the three estimation methods produce similar values for the standard errors of the various poverty measures. However, the jackknife seems to poorly perform in the case of Germany not only for the two income-inequality indicators (Gini coefficient and S80/S20) but also for the AROP.

Table 4: AROP: Point and standard errors estimated values vs. official figures

	Own estimated values				Official figures	
	value	B s.e.	JRR s.e.	L s.e.	value	s.e. (1)
Italy	19.9	0.324	0.482	0.407	20.3	0.534
United Kingdom	18.7	0.445	0.409	0.402	18.6	0.452
Portugal	17.5	0.476	0.693	0.484	17.3	0.516
Belgium	14.7	0.699	0.529	0.603	16.4	1.178
Norway	13.2	0.276	0.536	0.373	12.9	0.529
Germany	16.0	0.235	0.415	0.251	16.0	0.366
Sweden	16.2	0.253	0.592	0.371	16.4	0.614
Ireland	16.9	0.980	0.961	0.902	14.9	0.842

Source: Own calculations based on data extracted on 14/06/2020, and Eurostat official figures extracted on 25/09/2020

(1) The standard errors are from the Annex 4 of the EU quality report publicly available on 21/10/2020

Table 5: GINI: Point and standard errors estimated values vs. official figures

	Own estimated values				Official figures	
	value	B s.e.	JRR s.e.	L s.e.	value	s.e. (1)
Italy	33.3	0.322	0.368	0.369	33.4	0.101
United Kingdom	34.0	1.155	0.459	0.460	34.2	0.318
Portugal	33.6	0.415	0.435	0.373	32.1	0.147
Belgium	26.4	0.818	0.218	0.526	25.6	0.504
Norway	25.0	0.205	0.382	0.246	24.8	0.161
Germany	30.9	0.508	1.374	0.608	31.1	0.194
Sweden	26.3	0.214	0.128	0.371	27.0	0.207
Ireland	30.5	1.089	1.117	1.034	28.9	0.365

Source: Own calculations based on data extracted on 14/06/2020, and Eurostat official figures extracted on 25/09/2020

(1) The standard errors provided by Eurostat on 13/07/2020

Table 6: S80/S20: Point and standard errors estimated values vs. official figures

	Own estimated values				Official figures	
	value	B s.e.	JRR s.e.	L s.e.	value	s.e. (1)
Italy	6.1	0.134	0.146	0.142	5.9	0.014
United Kingdom	5.8	0.268	0.132	0.121	5.4	0.017
Portugal	5.6	0.110	0.145	0.100	5.7	0.007
Belgium	3.8	0.155	0.007	0.096	3.8	0.026
Norway	3.8	0.055	0.093	0.056	3.9	0.010
Germany	5.0	0.152	0.428	0.188	4.5	0.020
Sweden	3.9	0.045	0.371	0.074	4.3	0.019
Ireland	4.6	0.202	0.201	0.179	4.6	0.028

Source: Own calculations based on data extracted on 14/06/2020, and Eurostat official figures extracted on 25/09/2020

(1) The standard errors provided by Eurostat on 13/07/2020

Tables 4 - 6 report point estimates and uncertainty measures for AROP, GINI and S80/S20 poverty indicators. It is worth noting that our estimates do not match exactly the official figures calculated by Eurostat and reported in the EU quality report regarding AROP and obtained from Eurostat for the other indicators.

The observed differences between our findings and official figures may be caused by various factors which are related to data and variable definitions as well as to the software used for performing estimates. With regard to the differences in point estimates, the main factor affecting our estimates is related to the different dataset used for our analysis, more specifically we used the harmonized EU-SILC dataset produced by Eurostat that may differs from the dataset used by NSIs. In addition, it is worth noting that Eurostat frequently receives data revisions that are not promptly reported in UDB, which is only published twice a year. Another aspect related to the accuracy of measurement at level of individual units needs to be considered. Although the EU-SILC survey is the most important comparative microdata source on household income in Europe, the impact of conceptual or definitional differences in data collection and data treatment procedures adopted by NSIs may be significant thus causing a loss of accuracy of measurement and comparability of income variables among the various countries. The impact of the above mentioned differences is seen markedly in the case of income regarding the inclusion or exclusion of imputed rent and income from production for own consumption. In addition, different methods may be used for outlier detection and data error correction. Moreover, countries adopt different methods regarding the conversion process which may concern the omission of income components and/or misallocation of income components within income target variables. For example, detailed information concerning the allocation of payments for fostering children is lacking for the United Kingdom and Italy. This income component, which should be treated as employee cash and near-cash income, is not included under PY010 nor in any other variable. Therefore, even if various methodological guidelines and quality reports are available to assess procedural comparability of the EU-SILC income variables at EU level, not all countries adhere to the standard definition suggested in the Eurostat guidelines (Trindade and Goedemé, 2020). In addition complete information on how the national income components are constructed and classified is often lacking. As a result, the quality of collected information is not uniform across countries.

Another reason that can justify the above differences between our estimates and the official figures could derive from the software and packages that are being used. Eurostat use the SAS statistical software, while in our analysis we used the open-source software R. More specifically, to obtain point estimates we used the Laeken package, which provides standard functions for estimating a set of Laeken indicators. It should be highlighted that the developed R codes have been fine-tuned remotely and ran by Eurostat staff on the EU-SILC micro-data for the production of the results. For this reason, we could not check the composition and the distribution of income variables of the EU-SILC data in order to examine the top end of the distribution for identifying top-values set and potential outliers that influence the estimation of income-inequality indicators. More specifically, the presence of large values at the upper end of the distribution does not affect point estimates of poverty indicators, but they can markedly affect the estimated indicators of inequality such as the Gini coefficient and S80/S20 (see Section 3). Therefore, variance of the estimates can also become greatly inflated. These factors affect cross countries comparability. Computational aspects need also to be considered when explaining the differences among our estimates and official figures. More specifically, in our analysis some aspects of sample structure have been redefined in order to make variance computation possible, efficient and stable. Due to the unavailability of detailed information regarding the order of unit selection, we had to regroup units by considering individual ID in order to meet the basic requirement of practical methods of variance estimation for complex samples. Sometimes non-response can result in the disappearance from the sample of whole PSUs. This can disturb the structure of the sample, such as increasing the heterogeneity of the PSUs in some strata. This problem arises more frequently and seriously when computing sampling errors at sub-national level.

5.2.2 Results for NUTS1 regions

In the following sub-section we present the results obtained by estimating the poverty indicators and their standard errors at NUTS1 level([11]) . All income-poverty indicators are based on country poverty lines. Therefore, the income distribution is considered separately at the level of each country, in relation to which a poverty line (ARPT) is defined and the AROP and the RMPG are computed. It is worth noting that we refer to countries for which the NUTS1 geographical entities are considered planned domain in the sampling design. As a matter of fact, since the available sample sizes at NUTS1 level become small, sampling error tends not only to be high, but also estimates of sampling error tend to be more complex and subject to high levels of variability. It is possible that in some strata only 1 or 2 PSUs determine the total variability associated with the point estimates. In order to remove this effect and increase the size of strata, computational units have been used when computing jackknife and linearization. In this section we illustrate the relative performance of the three methods for estimating uncertainty measures at NUTS1 level by providing tables for all the selected countries and by using bar plot for the UK and Germany where the number of NUTS1 regions is particularly high.

It is worth noting that we encountered computational problems when estimating the RMPG indicator for those countries adopting a one stage stratified sampling design and sampling design without stratification. As a result, we do not illustrate relative standard errors estimation results obtained using linearization for Sweden and Germany.

Our findings show that the standard errors obtained using the jackknife method are greater than those obtained using the other two methods. Table 7 illustrates the estimated relative standard errors for the selected income-poverty indicators and income-inequality indicators for Italy at the NUTS1 level.

([11]) All the tables below for NUTS1 and NUTS2 are containing values of own calculations using the data extracted on 14/06/2020.

Table 7: Estimated standard errors EU-SILC 2018 for income poverty – Italy NUTS1

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
North-west	12.8	15.99	5.48	5.31
South	32.8	6.11	4.22	3.28
Insular	35.7	6.49	6.17	5.55
North-east	10.9	20.61	6.82	6.77
Central	17.2	10.98	6.67	5.02
RMPG				
North-west	27.4	6.53	12.76	7.26
South	33.3	4.22	4.15	9.99
Insular	34.5	8.52	7.85	19.70
North-east	20.1	8.54	26.66	5.88
Central	25.3	8.14	12.37	6.07
GINI				
North-west	31.6	2.08	2.80	2.60
South	33.9	1.46	1.90	1.93
Insular	35.9	3.13	1.73	3.47
North-east	29.1	1.61	4.18	1.97
Central	33.2	1.17	1.52	1.91
S80/S20				
North-west	5.3	3.75	4.88	3.90
South	6.6	3.69	5.44	4.66
Insular	7.8	10.02	11.40	13.77
North-east	4.6	2.75	2.93	2.93
Central	5.7	2.27	3.46	3.31

The estimated values of the AROP lies between 10.9 in the North-East NUTS1 region and 35.7 in the Insular regions of Italy. As regards variance estimation, the bootstrap method is less stable among regions and the obtained standard errors are higher than those obtained using the other two methods. On the contrary, the linearization method gives more accurate estimates when compared to the other methods. The estimated value of the RMPG lies between 20.1 in the North-East regions and 34.5 in the Insular regions of Italy. Considering the jackknife method, we obtained higher values of standard error for the North-East regions where it reaches the value of 26.65. The Gini coefficient and the S80/S20 achieved minimum estimated values in the North-East of Italy (29.7 and 4.6 respectively) while the maximum values are reported for the insular regions (35.88 and 7.8 respectively). For these indicators, the relative standard errors estimates for the three approaches are close for the various NUTS1.

Table 8: Estimated standard errors EU-SILC 2018 for monetary poverty indicators – The UK NUTS1

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
North East (England)	19.0	17.39	12.18	7.95
North West (England)	23.6	10.71	4.61	4.50
Yorkshire and the Humber	22.5	11.96	9.51	5.79
East Midlands (England)	18.6	14.65	10.13	7.11
West Midlands (England)	20.5	12.79	6.79	6.64
East of England	18.1	14.96	10.67	6.90
London	16.5	16.70	10.14	7.74
South East (England)	12.0	23.81	6.34	9.22
South West (England)	14.6	19.47	7.61	10.02
Wales	20.3	13.78	8.02	7.98
Scotland	18.1	13.88	6.34	6.59
Northern Ireland	28.9		11.95	3.92
RMPG				
North East (England)	16.2	20.91	32.44	38.23
North West (England)	24.8	11.32	7.10	8.84
Yorkshire and the Humber	22.2	7.24	7.09	9.06
East Midlands (England)	22.4	8.95	7.09	11.71
West Midlands (England)	23.9	6.95	16.19	7.50
East of England	14.6	14.36	13.20	14.90
London	26.6	7.23	17.40	8.94
South East (England)	23.6	9.56	10.51	8.48
South West (England)	19.9	14.82	15.57	12.34
Wales	23.3	11.35	16.21	11.82
Scotland	21.8	9.53	3.86	8.63
Northern Ireland	36.6		21.01	2.56

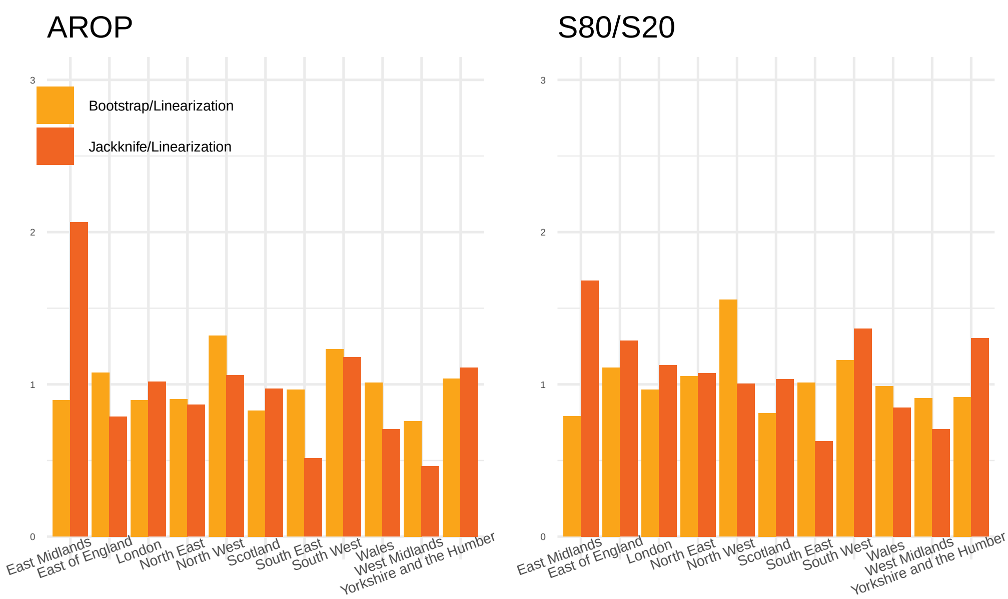

Figure 1: Estimated relative Standard Errors for income-poverty indicators: The UK NUTS1

Table 8 and Figure 1 illustrate the estimated relative standard errors for income-poverty indicators in the United Kingdom at NUTS1 level. The estimated AROP lies between 12 and 28.9: the minimum value is observed in South East England while the highest value is observed in Northern Ireland. Regarding the AROP uncertainty measures, the performance of the bootstrap is less satisfactory than jackknife and linearization methods. More specifically, the relative standard error for South East England (for which the sample size is equal to 4,429 units) obtained using the bootstrap method assumes the highest value compared to the jackknife and linearization. The linearization method produces more accurate and precise variance estimates when compared with jackknife method. The estimated value of the RMPG lies between 14.6 and 36.6 in the UK NUTS1. The minimum value is observed in East England while the highest value is observed in Northern Ireland (in this region the bootstrap variance estimate is not available). In the case of AROP, relative standard error results from the three variance estimation approaches are quite close.

Table 9 and Figure 2 illustrate the estimated relative standard errors for income-inequality indicators in the UK at NUTS1 level. The estimated values of the Gini coefficient and the S80/S20 slightly vary among the UK regions except for London and Northern Ireland. As regards relative standard errors for both indicators it is worth noting that the three methods perform similarly. However, for the East Midland region, the jackknife produces the highest value of standard error.

Table 9: Estimated standard errors EU-SILC 2018 for income inequality indicators – The UK NUTS1

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
GINI				
North East (England)	28.8	4.04	3.87	4.47
North West (England)	32.6	3.13	2.51	2.37
Yorkshire and the Humber	29.7	2.47	2.64	2.37
East Midlands (England)	30.2	2.76	6.36	3.08
West Midlands (England)	31.8	4.91	2.98	6.46
East of England	33.5	2.93	2.14	2.72
London	38.9	3.61	4.09	4.02
South East (England)	32.7	3.16	1.69	3.28
South West (England)	29.2	3.31	3.17	2.69
Wales	29.1	3.39	2.37	3.36
Scotland	30.9	2.91	3.42	3.51
Northern Ireland	60.8		4.57	3.42
S80/S20				
North East (England)	4.2	6.50	6.62	6.16
North West (England)	5.6	6.50	4.19	4.18
Yorkshire and the Humber	4.6	4.11	5.83	4.48
East Midlands (England)	4.7	4.53	9.65	5.74
West Midlands (England)	5.2	8.16	6.35	8.98
East of England	5.6	5.03	5.82	4.53
London	7.6	7.26	8.46	7.50
South East (England)	5.3	5.10	3.16	5.04
South West (England)	4.5	5.18	6.11	4.47
Wales	4.6	5.68	4.88	5.76
Scotland	5.1	4.88	6.21	6.01
Northern Ireland	17.7		8.22	5.04

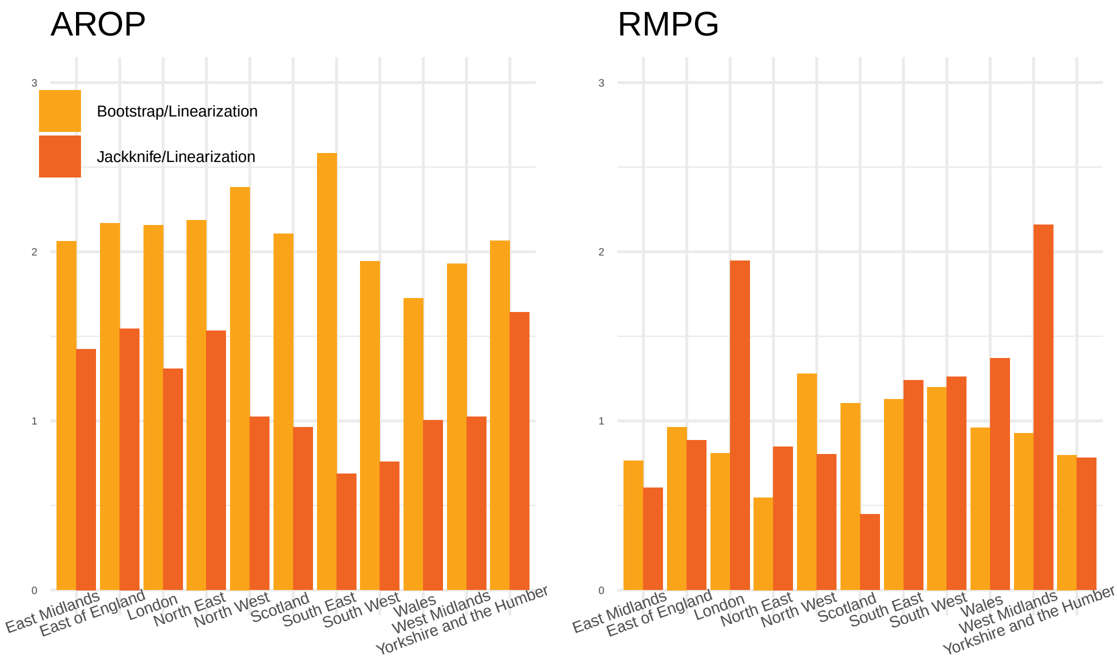

Figure 2: **Estimated relative Standard Errors for income-inequality indicators: The UK NUTS1**

Table 10: **Estimated standard errors EU-SILC 2018 – Portugal NUTS1**

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Continente	15.5	11.90	4.56	4.70
Região Autónoma Dos Açores	23.5	12.54	11.38	9.05
Região Autónoma Da Madeira	24.7	10.63	8.28	7.08
RMPG				
Continente	24.3	3.81	3.45	4.11
Região Autónoma Dos Açores	31.6	10.64	10.67	9.20
Região Autónoma Da Madeira	27.2	9.60	6.13	6.68
GINI				
Continente	32.9	1.42	1.38	1.18
Região Autónoma Dos Açores	38.1	3.54	2.17	2.64
Região Autónoma Da Madeira	33.6	2.40	2.59	1.94
S80/S20				
Continente	5.3	2.13	2.29	1.92
Região Autónoma Dos Açores	7.2	8.35	7.53	5.73
Região Autónoma Da Madeira	5.8	4.52	5.91	4.26

Table 10 reports points and uncertainty estimates for income-poverty and income-inequality indicators for NUTS1 of Portugal. Focusing on point estimates, the AROP value varies between 15.5 observed in Continente region and 24.7, observed in the Regia Autónoma Da Madeira. The Gini coefficient, S80/S20 and RMPG achieve maximum estimated values in the Região Autónoma Dos Açores. As regards the standard errors estimations, the relative precision of estimates obtained from bootstrap tends to be lower than the other two methods for most of the regions.

Table 11: Estimated standard errors EU-SILC 2018 – Belgium NUTS1

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Brussels	26.0	9.70	2.59	4.70
Flanders	9.3	10.58	9.93	9.05
Wallonia	18.3	8.43	9.78	7.08
RMPG				
Brussels	20.3	13.31	7.29	8.07
Flanders	12.2	16.72	12.94	11.17
Wallonia	15.9	10.34	8.67	12.66
GINI				
Brussels	32.6	3.63	2.07	1.58
Flanders	24.5	3.25	3.24	3.17
Wallonia	25.4	2.69	2.68	2.84
S80/S20				
Brussels	5.1	7.37	2.03	2.89
Flanders	3.5	3.66	2.99	3.53
Wallonia	3.6	4.03	3.92	4.29

Table 11 reports point estimates and uncertainty measures for selected poverty indicators for Belgium at NUTS1 level. As regards the point estimates, the selected monetary poverty and income inequality indicators achieve maximum estimated values in the Brussels region, while the lowest values are observed in the Flanders region. Brussels region seems to present higher income inequality indicators than the other two regions in Belgium. Overall, the difference between the various relative standard errors is slight among the methods, except for Bruxelles where bootstrap produces higher values for all the indicators.

Table 12: Estimated standard errors EU-SILC 2018 – Sweden NUTS1

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Ostra Sverige	14.5	2.79	3.68	4.04
Södra Sverige	17.1	2.13	6.06	3.30
Norra Sverige	17.5	3.39	6.15	5.08
RMPG				
Ostra Sverige	19.2	5.29	7.36	
Södra Sverige	19.9	3.57	7.18	
Norra Sverige	17.3	6.79	19.55	
GINI				
Ostra Sverige	27.5	1.40	2.16	2.35
Södra Sverige	26.3	1.20	3.24	2.16
Norra Sverige	23.0	1.17	4.63	1.75
S80/S20				
Ostra Sverige	4.2	1.70	4.69	3.08
Södra Sverige	4.0	1.78	5.34	3.02
Norra Sverige	3.3	1.36	7.04	2.34

Estimation results of uncertainty measures for income-poverty and income-inequality indicators for Sweden are reported in Table 12. For the EU-SILC survey, Sweden adopts a sampling design structure without stratification. For this country the RMPG uncertainty measure calculated with linearization method is not available. Focusing on point estimates, the AROP value varies between 14.5 observed in Ostra Sverige region and 17.5 observed in Norra Sverige region. For the selected income inequality indicators, the highest values are observed in Ostra Sverige (27.5 for the Gini coefficient and 4.2 for the S80/S20 indicator respectively), while the lowest values are observed in Norra Sverige (23.0 for the Gini coefficient and 3.3 for the S80/S20 indicator respectively). From Table 12 we can observe that bootstrap method produces lower standard error estimates when compared with linearization and jackknife methods. Yet, it is important to note that this method may under-estimate the true value of variance. Tables 13, 14, 15 and Figures 3 -4 report point estimates and uncertainty measures of AROP, Gini coefficient and S80/S20 for Germany at NUTS1 level.

Table 13: Estimated standard errors EU-SILC 2018 for AROP – Germany NUTS1

Region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Baden-Württemberg	12.1	5.00	19.05	6.08
Bavaria	14.4	3.80	10.50	4.54
Berlin	18.4	7.63	7.26	8.16
Brandenburg	15.5	5.85	16.22	7.36
Free Hanseatic City of Bremen	28.4	9.70	11.84	8.84
Hamburg	24.0	9.28	13.47	8.81
Hessen	12.1	6.01	10.39	7.23
Mecklenburg-Vorpommern	23.4	6.29	9.12	6.65
Lower Saxony	15.2	4.94	17.52	5.19
North Rhine-Westphalia	15.8	2.66	6.47	3.27
Rhineland-Palatinate	20.8	6.36	16.02	5.75
Saarland	16.2	13.55	13.21	13.95
Saxony	18.4	5.60	8.13	5.51
Saxony-Anhalt	16.3	8.04	22.96	8.94
Schleswig-Holstein	22.5	5.68	7.08	5.76
Thuringia	18.4	6.93	18.02	7.77

As shown in Table 13, the estimated values of the AROP lies between 12.1 in Baden-Wurttemberg and Hessen regions and 28.4 in the Free Hanseatic City of Bremen region of Germany. From the same Table and Figure 3, we can observe that the jackknife is less stable than the other two methods among regions. It reaches the maximum value of 22.95 in the Saxony-Anhalt, which is one of the smallest regions with a sample size equal to 639 units. Contrastingly, linearization performs well compared with the other methods.

Table 14: Estimated standard errors EU-SILC 2018 for Gini indicator – Germany NUTS1

Region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Baden-Württemberg	28.4	1.93	2.94	2.02
Bavaria	33.7	5.12	2.88	5.67
Berlin	26.8	4.11	5.68	4.51
Brandenburg	25.3	2.35	3.02	2.81
Free Hanseatic City of Bremen	28.2	5.26	3.24	4.21
Hamburg	30.6	3.92	4.79	4.34
Hessen	40.8	2.71	2.95	3.06
Mecklenburg-Vorpommern	26.1	3.09	3.55	3.28
Lower Saxony	27.4	1.56	7.93	1.92
North Rhine-Westphalia	29.4	1.25	3.62	1.79
Rhineland-Palatinate	31.6	2.55	3.21	2.66
Saarland	23.7	4.55	3.61	4.71
Saxony	23.7	2.35	5.20	2.80
Saxony-Anhalt	27.7	5.08	5.02	6.17
Schleswig-Holstein	36.5	6.02	9.05	7.27
Thuringia	25.6	2.82	6.89	3.29

The point estimates of the Gini coefficient lie between 23.7 in the Saxony and Saarland regions and 40.8

in the Hessen region of Germany. Focusing on relative standard errors, we can observe that the Jackknife produces higher values than those obtained with the other two methods in most of the NUTS1 regions.

Table 15: Estimated standard errors EU-SILC 2018 for S80/S20 indicator – Germany NUTS1

Region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Baden-Württemberg	4.2	3.25	16.01	3.44
Bavaria	7.7	28.66	3.18	28.65
Berlin	3.7	5.09	32.84	5.31
Brandenburg	3.6	3.50	9.70	3.72
Free Hanseatic City of Bremen	4.2	8.57	5.84	4.30
Hamburg	5.0	7.27	7.63	7.76
Hessen	6.9	12.46	9.92	14.36
Mecklenburg-Vorpommern	3.8	4.65	11.20	5.09
Lower Saxony	4.0	2.35	8.25	2.68
North Rhine-Westphalia	4.5	1.85	7.39	2.56
Rhineland-Palatinate	5.3	6.17	9.28	5.48
Saarland	3.3	9.05	5.61	7.37
Saxony	3.3	3.19	8.83	3.21
Saxony-Anhalt	3.9	5.90	10.19	6.78
Schleswig-Holstein	9.2	12.87	23.41	32.73
Thuringia	3.7	4.89	10.64	5.31

The estimated value of the quintile share ratio lies between 3.3 in the Saxony and Saarland regions and 9.2 in the Schleswig-Holstein region of Germany. The jackknife method provides higher value of the variance for 8 out of 16 NUTS1 compared to bootstrap and linearization. On the contrary, bootstrap and linearization methods tend to exhibit similar performance with some differences at regional level. Indeed, for the Fee Hanseatic city of Bremen and Saarland regions, which are the smallest regions of Germany, the bootstrap method provides less accurate estimates.

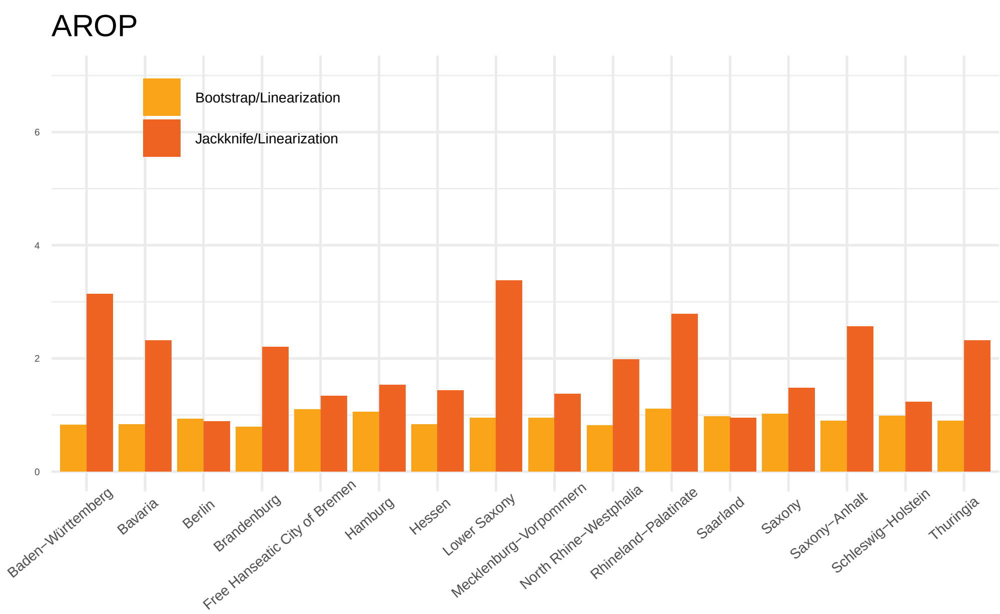

Figure 3: Estimated relative Standard Errors for income-poverty indicators: Germany NUTS1

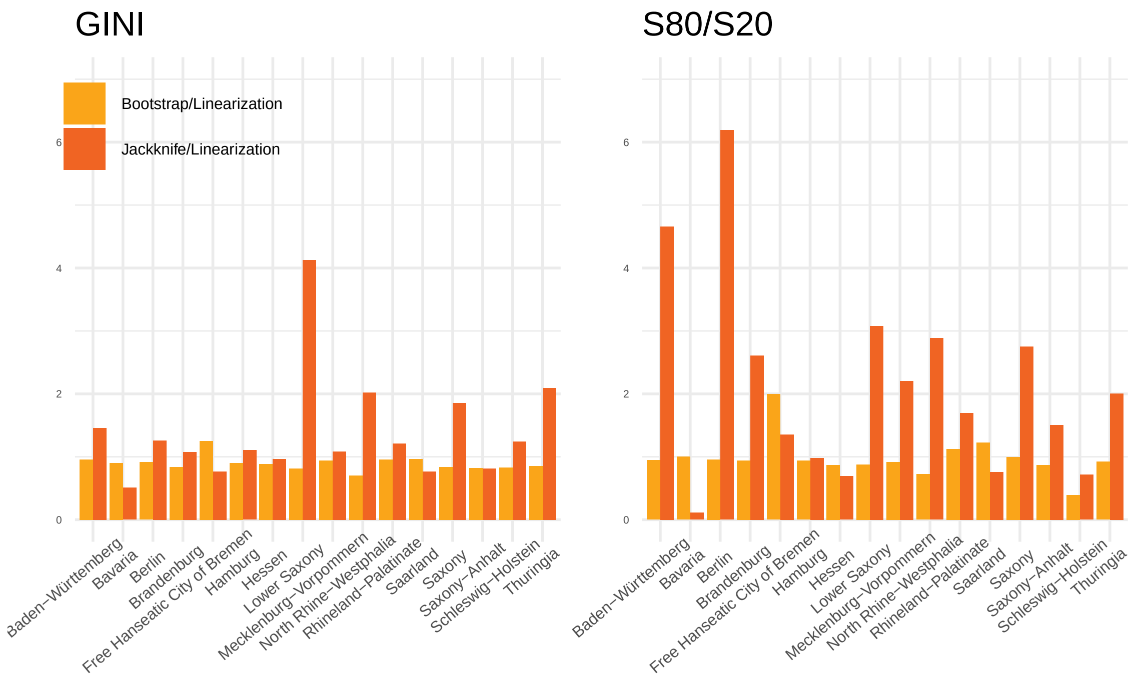

Figure 4: Estimated relative Standard Errors for income-inequality indicators: Germany NUTS1

5.2.3 Results for NUTS2 regions

We considered NUTS2 regions to provide a common framework that enhances comparability of the estimated relative standard errors. As expected, uncertainty measures for income-poverty and income-inequality indicators at NUTS2 level are higher than those obtained for NUTS1 regions especially for the UK and Germany where the number of NUTS2 is particularly high. Estimated standard errors obtained by applying the various methods are higher for those regions with a reduced number of sampling units. Indeed, for the UK the number of units is lower than 1,000 for 26 out of 41 NUTS2 regions while concerning Germany only 7 NUTS2 contains more than 1,000 units. In these cases, the relative standard errors obtained using jackknife could be also affected by the sensitivity of this method to the construction of strata as well as to the presence of outliers in the income distribution, especially regarding the income-inequality indicators([12]) . In order to ameliorate the problem of small sample sizes and produce regional estimates with reduced sampling error, various procedures can be implemented. It is advisable to improve size and unit allocation and/or using auxiliary information for computing small area estimation. Linearization seems to produce lower relative standard errors than re-sampling techniques for all the regions and countries considered. There are few exceptions in Italy (as can be seen in Tables 16 - 19 reported in the Appendix and Figures 5-6), regarding AROP for Calabria, RMPG for Puglia and S80/220 for Molise, in Portugal (Table 27 and Figures 7 - 8), regarding RMPG in Norte region, in Belgium (Table 28 and Figures 9 - 10) regarding RMPG for Liège, Gini and S80/S20 for Hainaut and in Norway (Table 30 and Figures 11 - 12) regarding AROP and Gini for Hedmark og Oppland.

([12]) In order to reduce the upward bias of jackknife we have set the number of replications to 2,000.

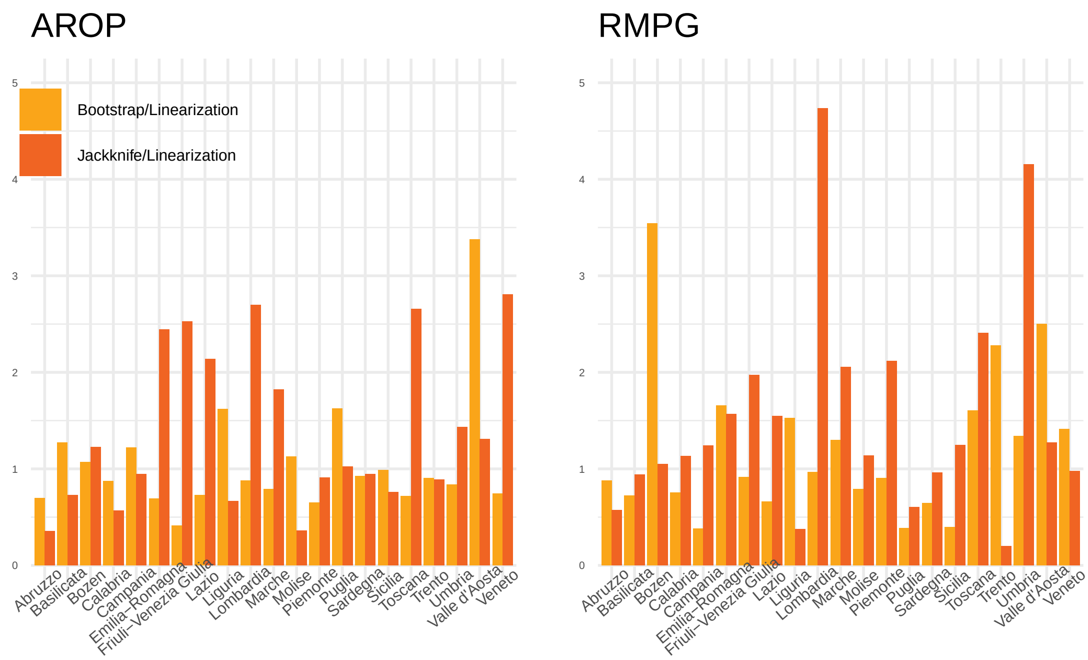

Figure 5: Estimated relative standard errors for income-poverty indicators: Italy NUTS2

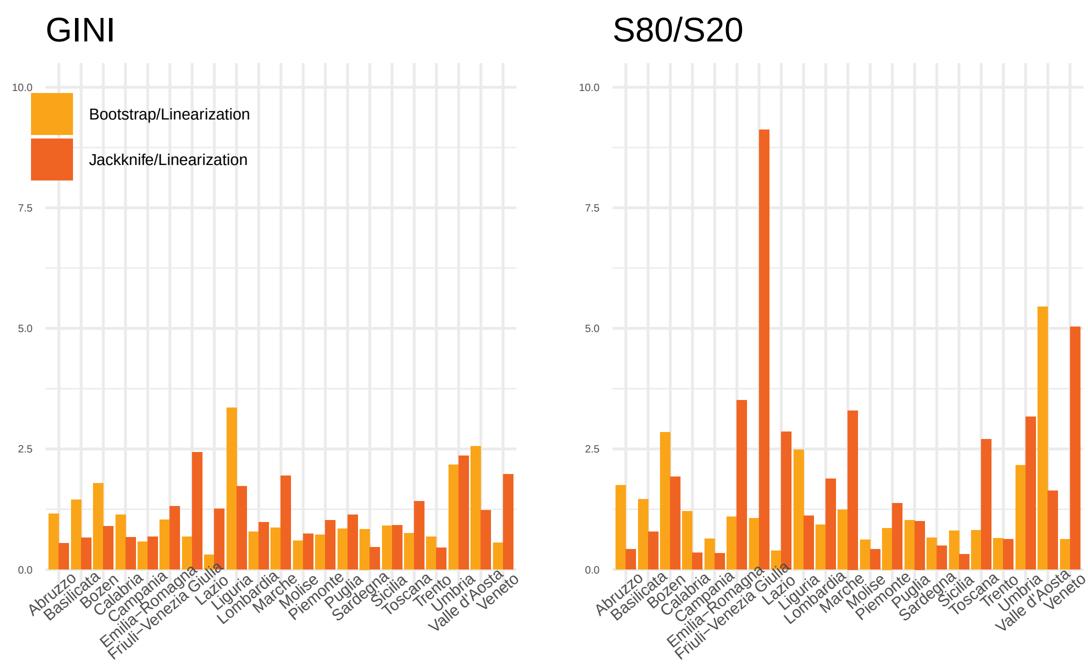

Figure 6: Estimated relative standard errors for income-inequality indicators: Italy NUTS2

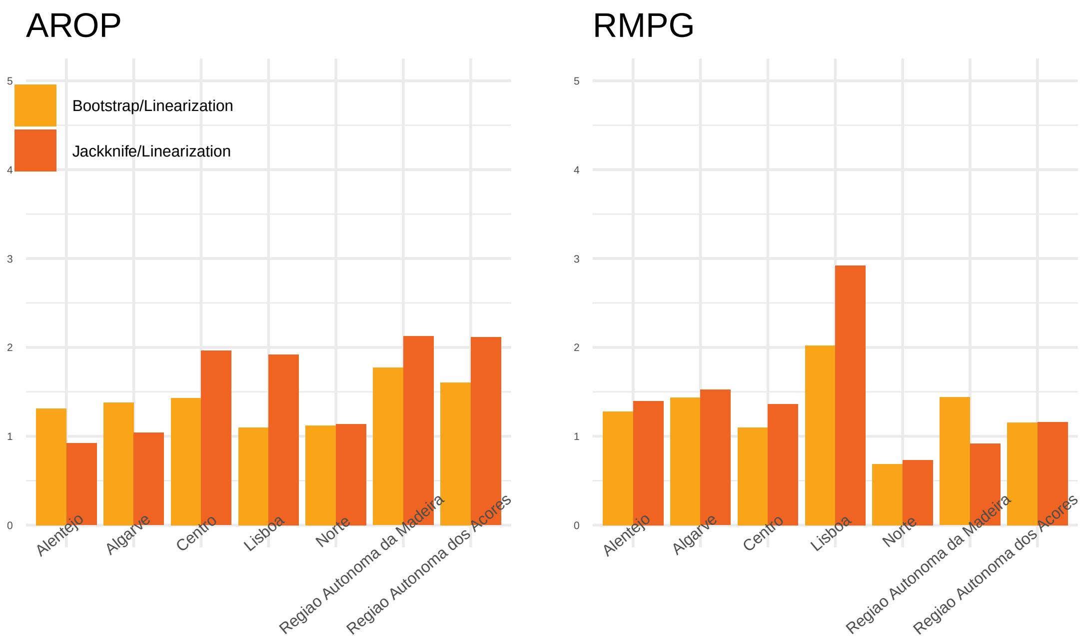

Figure 7: Estimated relative standard errors for income-poverty indicators: Portugal NUTS2

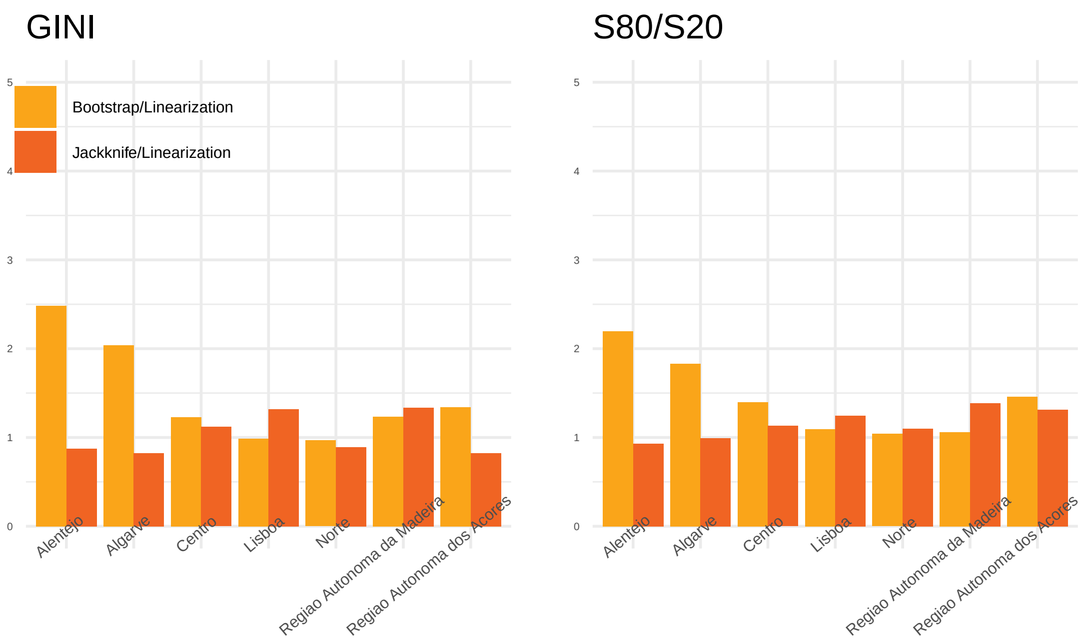

Figure 8: Estimated relative standard errors for income-inequality indicators: Portugal NUTS2

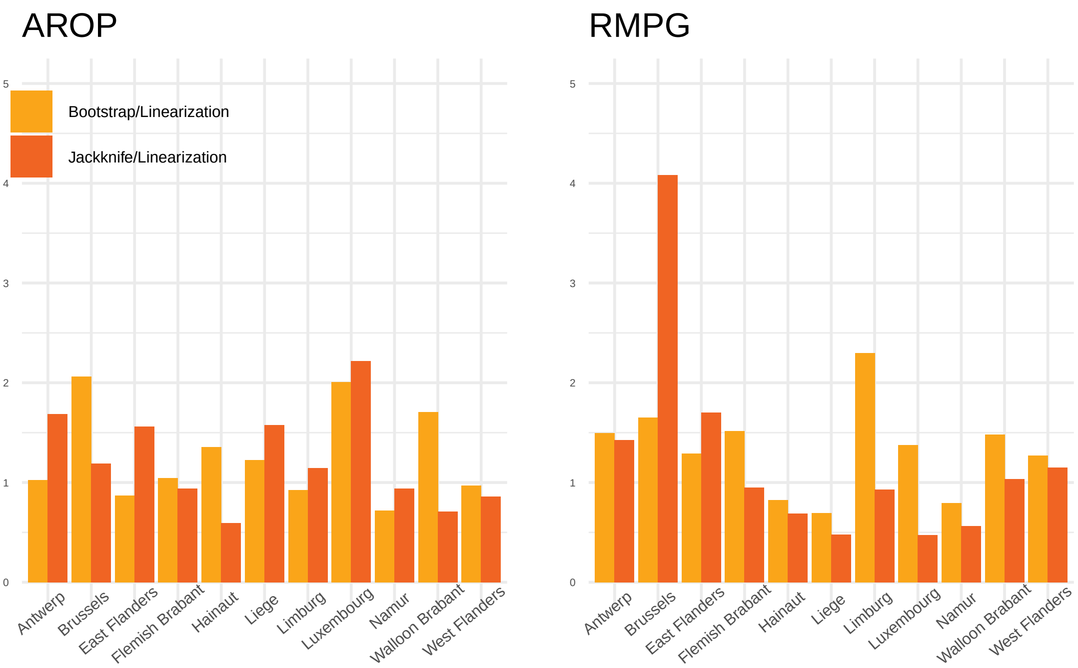

Figure 9: Estimated relative standard errors for income-poverty indicators: Belgium NUTS2

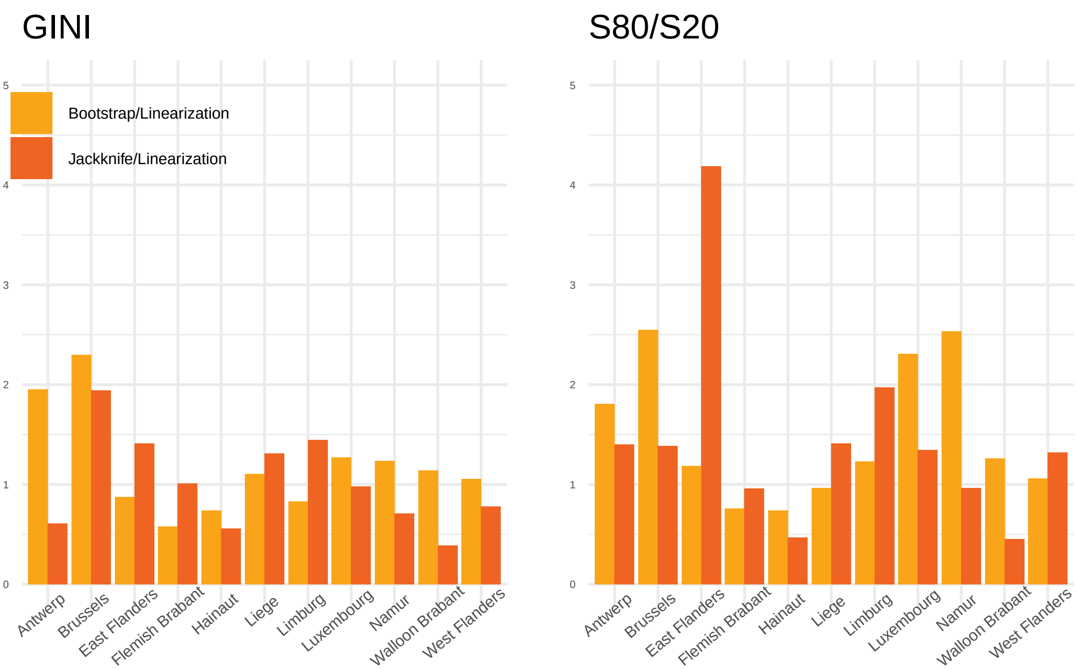

Figure 10: Estimated relative standard errors for income-inequality indicators: Belgium NUTS2

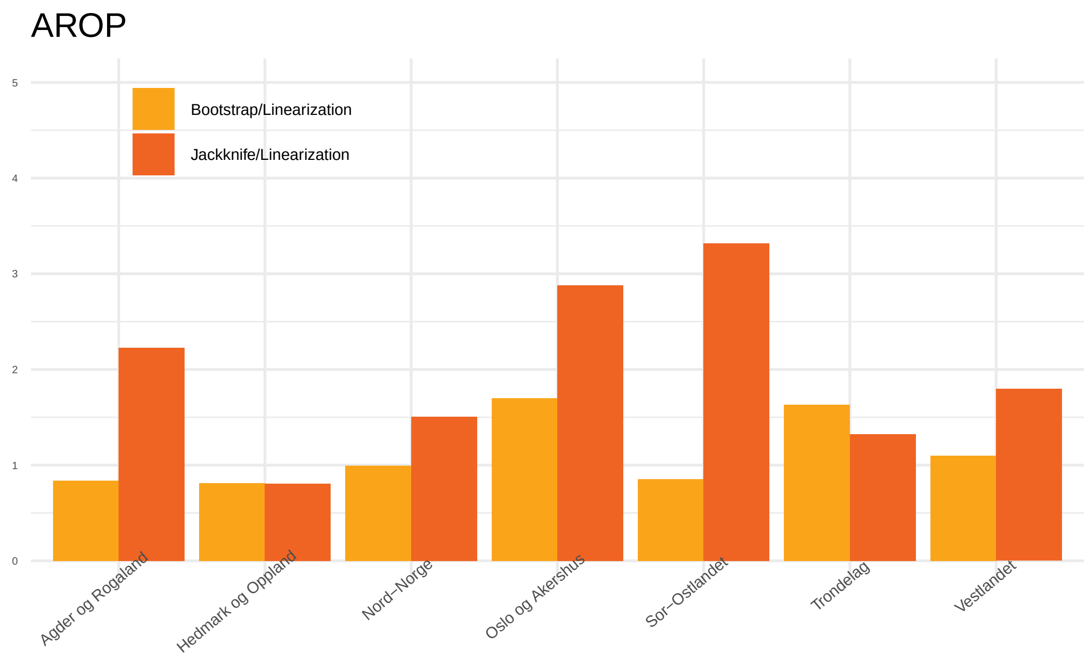

Figure 11: Estimated relative standard errors for income-poverty indicators: Norway NUTS2

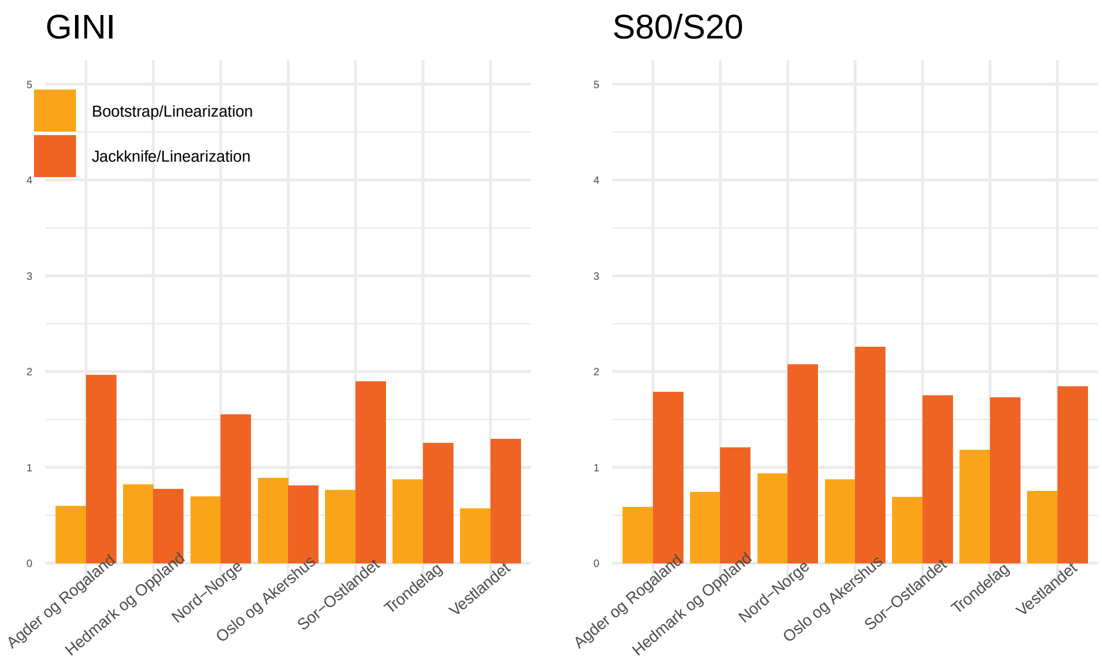

Figure 12: Estimated relative standard errors for income-inequality indicators: Norway NUTS2

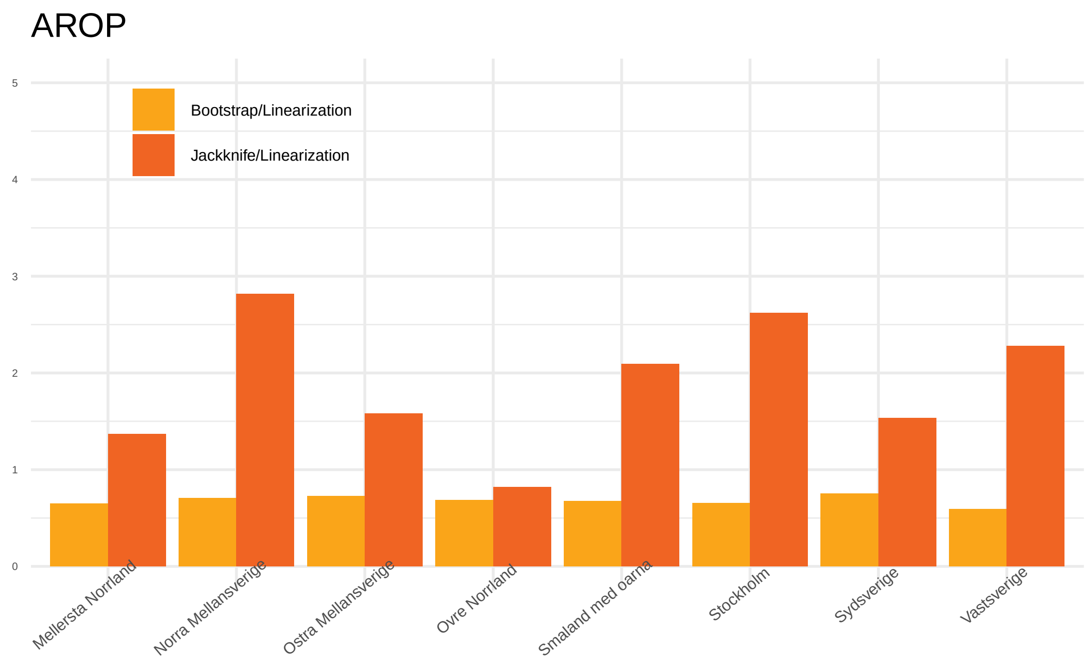

Figure 13: **Estimated relative standard errors for income-poverty indicators: Sweden NUTS2**

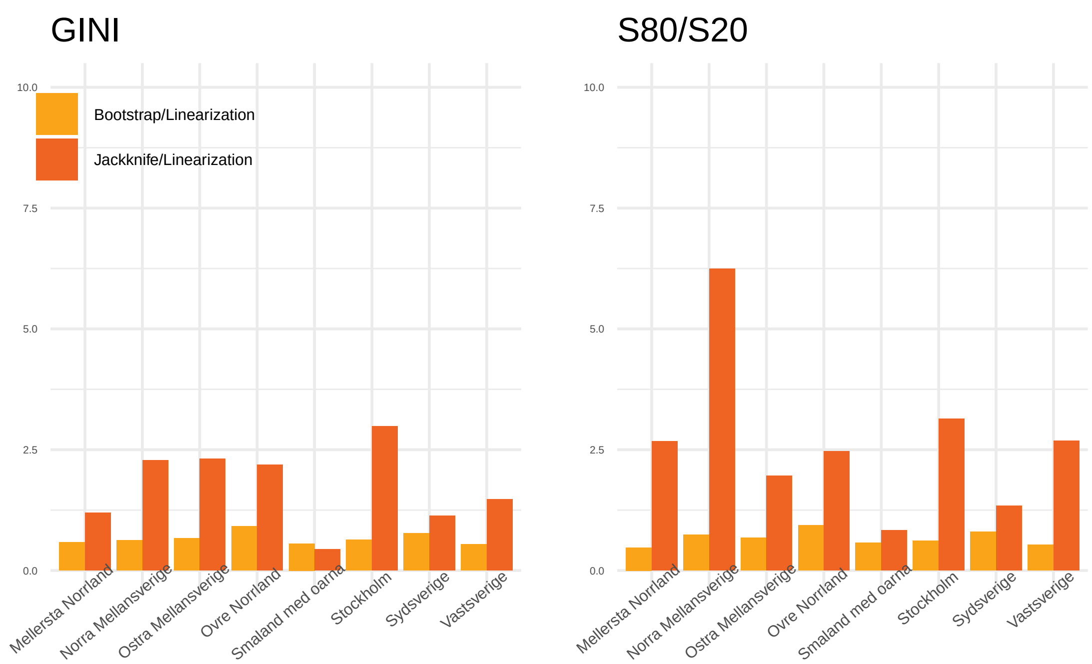

Figure 14: **Estimated relative standard errors for income-inequality indicators: Sweden NUTS2**

6 Communicating uncertainty in poverty indicators

Communicating uncertainty aims at increasing the transparency of the scientific assessment process and provide the risk managers with a more informed evidence based on reporting on the strengths and weaknesses of the evidence (EFSA Scientific Committee et al., 2017). It is clear that a transparent and open communication requires that uncertainty is communicated, yet Frewer et al., 2002 found that uncertainty related to the scientific process is more readily accepted than uncertainty due to lack of action by the government. This finding suggests that uncertainty communication is less likely to cause public alarm if it is accompanied by information on what actions are being taken by the pertinent authorities to address that uncertainty. However, when a measure of uncertainty is known, it is usually not presented or little communicated (Van der Bles et al., 2019). Untill now, specific guidance on how to communicate uncertainty has not yet been developed and no consensus has emerged as a general recommendation for communicating uncertainty among different target audiences. To avoid that audience often incorrectly interpret confidence intervals or other measures of uncertainties (Greenland et al., 2016), communicators of uncertainty should have relationships with the audience they are communicating to. Indeed, the source of uncertainty could have different effects on different audience and different sources of uncertainty can lead to different forms of communication.

Although several statistical organisations have started to invest in identifying ways to measure and communicate data uncertainty, this is only being done randomly. The EU-SILC implementing regulations specify quality criteria concerning the detailed content of intermediate and final quality reports. More specifically, the regulation states that NSIs should provide the following information for equivalised disposable income and for the unadjusted gender pay gap:

- effective sample size for the common cross-sectional EU indicators based on the cross-sectional component of EU-SILC,
- standard errors for the common cross-sectional EU indicators based on the cross-sectional component of EU-SILC.

Recently, the European Regulation 2019/1700 of 10 October 2019 established a common framework for European statistics relating to persons and households, based on data at individual level collected from samples, amending Regulations (EC) No 808/2004 (EC) No 452/2008 and (EC) No 1338/2008 of the European Parliament and of the Council, and repealing Regulation (EC) No 1177/2003 of the European Parliament and of the Council and Council Regulation (EC) No 577/98.

6.1 Current practices adopted by NSIs

NSIs communicate uncertainty of poverty measures by providing a quantification of the magnitude of the variability in their national quality reports. However, to the authors' knowledge, only few EU countries are currently communicating measures of uncertainty when publishing new releases on poverty and living conditions. In 2018, ISTAT adopted official figures of absolute and relative poverty using Household Budget Survey data for various subgroups in the population. In this report, ISTAT adopted a numerical communication of uncertainty for relative poverty incidence by regions and geographical area. As illustrated in Figure 15, standard errors and confidence intervals are provided for the years 2016 and 2017 (Istat, 2018).

	2016				2017			
	Incidence	Error	Confidence interval		Incidence	Error	Confidence interval	
	(%)	(%)	Lower limit	Upper limit	(%)	(%)	Lower limit	Upper limit
Italy	**10.6**	**3.0**	**10.0**	**11.2**	**12.3**	**2.5**	**11.7**	**12.9**
Piemonte	6.0	11.4	4.6	7.3	6.8	12.5	5.2	8.5
Valle d'Aosta/Valleè d'Aoste	4.8	21.8	2.8	6.9	4.4	21.2	2.6	6.3
Liguria	11.1	12.9	8.3	14.0	8.5	11.7	6.5	10.4
Lombardia	5.0	11.7	3.8	6.1	5.5	8.7	4.5	6.4
Trentino Alto Adige/Südtirol	*	*	*	*	4.9	14.2	3.5	6.3
Bolzano-Bozen	*	*	*	*	*	32.9	0.6	2.9
Trento	*	*	*	*	7.8	15.8	5.4	10.2
Veneto	5.5	12.7	4.2	6.9	6.1	12.7	4.6	7.7
Friuli Venezia Giulia	10.4	14.6	7.4	13.3	6.9	13.5	5.1	8.7
Emilia Romagna	4.5	16.0	3.1	5.9	4.6	17.1	3.1	6.2
North	**5.7**	**5.5**	**5.1**	**6.3**	**5.9**	**5.0**	**5.3**	**6.5**
Toscana	3.6	21.9	2.0	5.1	5.9	13.4	4.4	7.5
Umbria	11.8	13.4	8.7	14.9	12.6	14.5	9.0	16.2
Marche	8.9	17.4	5.8	11.9	8.8	13.0	6.5	11.0
Lazio	9.7	12.7	7.3	12.1	8.2	9.4	6.6	9.6
Centre	**7.8**	**8.9**	**6.5**	**9.2**	**7.9**	**6.2**	**6.9**	**8.8**
Abruzzo	9.9	14.8	7.0	12.8	15.6	13.2	11.6	19.6
Molise	18.2	18.2	11.7	24.6	21.0	10.1	16.9	25.2
Campania	19.5	9.7	15.8	23.2	24.4	7.9	20.6	28.1
Puglia	14.5	10.1	11.6	17.4	21.6	7.1	18.6	24.6
Basilicata	21.2	13.4	15.7	26.8	21.8	10.5	17.4	26.3
Calabria	34.9	6.0	30.8	39.1	35.3	6.5	30.7	39.8
Sicilia	22.8	6.9	19.7	25.9	29.0	6.1	25.5	32.5
Sardegna	14.0	18.5	8.9	19.1	17.3	10.8	13.6	21.0
South and Islands	**19.7**	**4.1**	**18.2**	**21.3**	**24.7**	**3.3**	**23.1**	**26.3**

* The value is statistically unreliable given the small sample size

Figure 15: Relative poverty incidence estimates for Italy at sub-national level reporting standard error and confidence interval

Source: ISTAT, La povertà in Italia. Statistiche Report 2018

Since statistical institutes have different users, this communication method based on numerical communication may be confusing especially for layman users since it assumes a certain level of statistical literacy. In order to enhance the users' interpretation of the data in 2018 and 2019, ISTAT briefly described the uncertainty related to the official figures as follows: *Confidence interval and absolute and relative sampling error: Knowing the estimate Y^* of a Y parameter of the population and the estimate of the absolute sampling error associated with it, it is possible to construct a confidence interval which, with confidence level α, includes within it the value of the Y parameter being estimated. The magnitude of this interval is a function of the absolute sampling error of a k value that depends on the shape of the sample distribution of the* *estimator and the value chosen for the confidence level α. For large samples, reference is commonly made to the normal distribution and there is, for example, for $\alpha = 0.05$, that $k = 1.96$. The magnitude of the confidence interval, and therefore the degree of uncertainty on the parameter Y in the population, is equal to $2k$ times the absolute sampling error. The estimation of the absolute sampling error is a statistic to evaluate the sample error and is equal to the mean squared deviation of the parameter's estimator Y^*. The coefficient of variation of the estimator is instead the relative sampling error, generally expressed as a percentage* (Istat, 2018). However, a simpler verbal information method describing probabilities may be preferred by layman user, since as underlined by Druzdzel (1989) the general public tends to prefer to conceptualize the uncertainty in verbal form.

Turning to EU-SILC based indicators, Statistics Austria publishes on a regular basis standard errors and confidence intervals of the main poverty indicators at national level. These estimates are disaggregated also by gender and age. On the contrary, regional indicators on poverty are not regularly published. As shown in Figure 16, the latest release is based on EU-SILC data for the year 2007.

At-risk-of-poverty rate and confidence intervall for Austrian provinces

	Armutsgefährdungsquote		Konfidenzintervall 95%				befragte Haushalte	Personen in befragten Haushalten
			untere Grenze		obere Grenze			
	in %	in 1.000	in %	in 1.000	in %	in 1.000		
Österreich	**12,0**	**989**	**11,2**	**917**	**12,9**	**1.060**	**6.806**	**16.684**
Burgenland	13,7	37	9,0	24.036	18,5	49.290	247	606
Kärnten	10,8	59	6,5	35.557	15,0	81.667	475	1.156
Niederösterreich	10,4	167	8,6	137.843	12,3	196.716	1.287	3.212
Oberösterreich	8,1	114	6,3	88.265	9,9	140.253	1.233	3.207
Salzburg	10,1	55	7,1	38.696	13,1	70.904	440	1.125
Steiermark	13,6	160	11,2	132.031	16,0	188.265	1.014	2.547
Tirol	10,1	71	7,5	52.184	12,8	89.535	584	1.493
Vorarlberg	13,0	50	9,4	35.984	16,6	63.428	307	832
Wien	17,4	276	15,0	238.439	19,8	314.202	1.219	2.506

Q: STATISTIK AUSTRIA, EU-SILC 2007.

Figure 16: Austria AROP estimates at sub-national level reporting confidence interval
Source: Statistiks Austria, EINKOMMEN, ARMUT UND LEBENSBEDINGUNGEN

Statistics Belgium publishes confidence intervals for the common cross-section EU poverty indicators in annual Quality Reports. However, no explanation is provided to help users in understanding uncertainty presentation. Nevertheless, Statistics Belgium planned to use Small Area Estimation methods for the estimation of poverty indicators at NUTS2 level starting from EU-SILC 2018 by using administrative data, in order to produce reliable and stable results at NUTS2 level, as expected by Eurostat. Indeed, the EU-SILC survey was planned to provide results at national level and therefore some NUTS2 sampling sample are very small.

Regarding the UK, the Office for National Statistics (ONS) provides a detailed explanation of uncertainty measures such as standard errors, confidence interval, coefficient of variation and statistical significance and how they affect estimates from surveys used for producing official figures. Figure 17 provides estimates of uncertainty for Gini coefficient .

		Disposable income			
		Lower bound	Published estimate	Upper bound	CV^2
All individuals	*Gini coefficient (%)*	32.5	34.7	37.0	3.3

Figure 17: 95% confidence intervals for statistics on Gini coefficient for individuals, 2018/19
Source: Office for National Statistics of the UK

In addition, as reported in Figure 18 the ONS published provisional estimates of Gini coefficient in the UK for the financial year ending 2020 by comparing provisional and final estimates. Using visualization method, the ONS highlights the accuracy of the provisional estimates of income inequality measures.

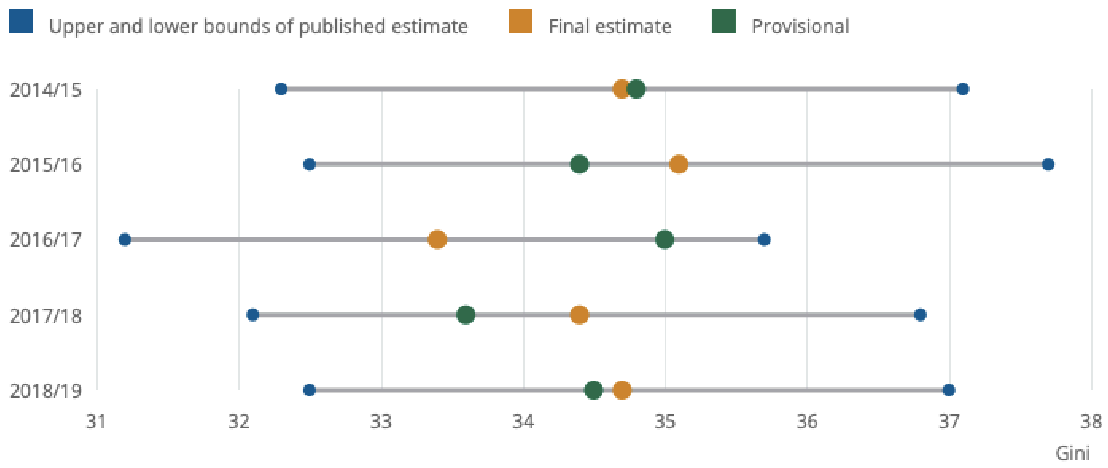

Figure 18: Provisional estimates of income inequality
Source: Office for National Statistics of the UK

Focusing on estimations of poverty indicators at local level, the ONS designed experimental statistics which are calculated using a model-based method to produce two estimates of the percentage of households in poverty: before housing costs and after housing costs at the middle layer super output area (MSOAs) level in England and Wales. Figure 19 shows the estimated percentage of households in poverty (with 95% confidence intervals) for the 27 MSOAs in the Calderdale Local Authority Districts LAD.

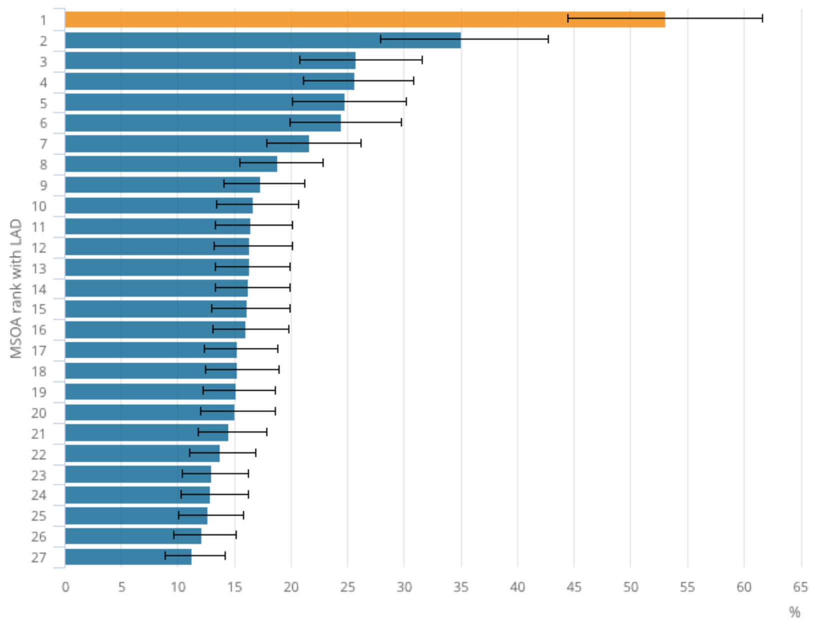

Figure 19: Percentage of households in poverty in UK
Source: Office for National Statistics of the UK

The Swiss Institute of Statistics publishes a report providing point estimates and confidence intervals for regional indicators of poverty but not on a regular basis. The latest report is based on EU-SILC data for the year 2014 and has been published in June 2017. Results for the AROP at regional level are reported in Figure 20.

Region	Sample size	AROP	Lower bound	Upper bound
Region Lémanique	2739	13.2	10.6	15.9
Espance Mitteland	3788	17.4	14.7	20.2
Suisse du Nord-Ouest	2254	14.1	10.2	18.1
Zurich	2734	8.1	6.1	10.1
Suisse orientale	2128	14.3	11.1	17.5
Suisse centrale	1460	10.4	7.2	13.7
Tessin	550	26.3	17.0	35.6

Figure 20: Swisse AROP estimates at sub-national level reporting confidence interval
Source: Swisse Institute of Statistics, Quality report

From the above illustrated practices adopted by NSIs, it is clear that communication of uncertainty is a challenging task and implementation of uncertainty communication activity should be gradually applied by NSIs in their assessment process.

6.2 Practical suggestions for communicating uncertainty in the Quality reports

In this sub-section we suggest methods for presenting uncertainty in the Eurostat national quality reports by considering different types of audience. In the national quality reports, poverty estimates are often shown in tables, with a numerical indication of their precision. However, point estimates and standard errors for income-inequality indicators such as Gini coefficient and S80/S20 are not reported in the national quality reports. It is worth noting that clearly communicating uncertainty measures for these figures may not be trivial, due to the possibility that this information may be mis-interpreted by the general public. Indeed, information regarding standard errors or relative standard errors is often shown without an explanation of the meaning of the uncertainty range (Griethe et al., 2006). The effects of uncertainty communication depend not only on the characteristics of the target audience and on the relationship between the audience and the communicator, but also on the topic or source of the uncertainty. Important differences between individuals, including the level of expertise, prior attitudes, numeracy skills, education level, might mean that the same communication of uncertainty affects people in a different manner (Van der Bles et al., 2019). Therefore improving verbal communication should be a good practice in order to understand the magnitude of the uncertainty among the general public. Some factors relating to the form of the communication should be considered: for non-expert audience (general public) we suggest to improve the format of uncertainty communication, in terms of verbal statements and the medium of the communication, by including not only print, but also online, broadcast or verbal conversation. The approach adopted by the ONS could be followed by other NSIs when compiling their quality reports for official poverty indicators. A verbal explanation of the meaning of the underlying and inherent uncertainty of official poverty indicators could be an efficient communication tool for non-expert users. They may be basically based on two kinds of measures: an interval measure (confidence or credible intervals) and probability distribution. As regards the interval measure, the error bars, where each point estimates are visualised with a bar and the confidence interval is plotted as an interval on top of each bar, are widely used in scientific and other publications. This kind of plot are particularly useful when standard errors are obtained using cross-sectional data. A case study is presented to illustrate the type of information that is usually published by NSIs. Focusing on Italy, Figure 22 shows point estimates, standard error and confidence intervals for various poverty measures. More specifically in the 2018 national quality report for Italy the following measures are reported by gender and age: AROPE, AROP, several material deprivation and very low work intensity.

	AROPE			At Risk of Poverty			Severe Material Deprivation			Very Low Work Intensity		
	Ind. Value	s.e.	Half CI (95%)	Ind. Value	s.e.	Half CI (95%)	Ind. Value	s.e.	Half CI (95%)	Ind. Value	s.e.	Half CI (95%)
Total	27,29	0,53	1,04	20,30	0,46	0,90	8,53	0,44	0,86	11,25	0,44	0,85
Male	26,14	0,58	1,15	19,40	0,44	0,86	8,60	0,48	0,95	10,18	0,48	0,94
Female	28,39	0,59	1,15	21,16	0,43	0,84	8,47	0,47	0,93	12,34	0,50	0,98
Age0-17	30,56	0,87	1,70	26,25	0,81	1,58	8,12	0,66	1,29	7,35	0,51	1,01
Age18-64	29,03	0,68	1,34	20,54	0,47	0,93	9,15	0,56	1,09	12,43	0,49	0,97
Age 65+	20,19	0,62	1,21	15,28	0,49	0,96	7,17	0,40	0,78			

Figure 21: Italian quality report: point and standard error estimates
Source: National Quality Report for Italy, year 2018

In addition to this numerical information, bar chart may be used to visualise uncertainty measures computed for different sub-populations. Error bars may be a good solution due to their widespread use and for the lack of better visualisation methods. Points and confidence intervals for AROP, computed for different sub-population, may be illustrated as in Figure 22.

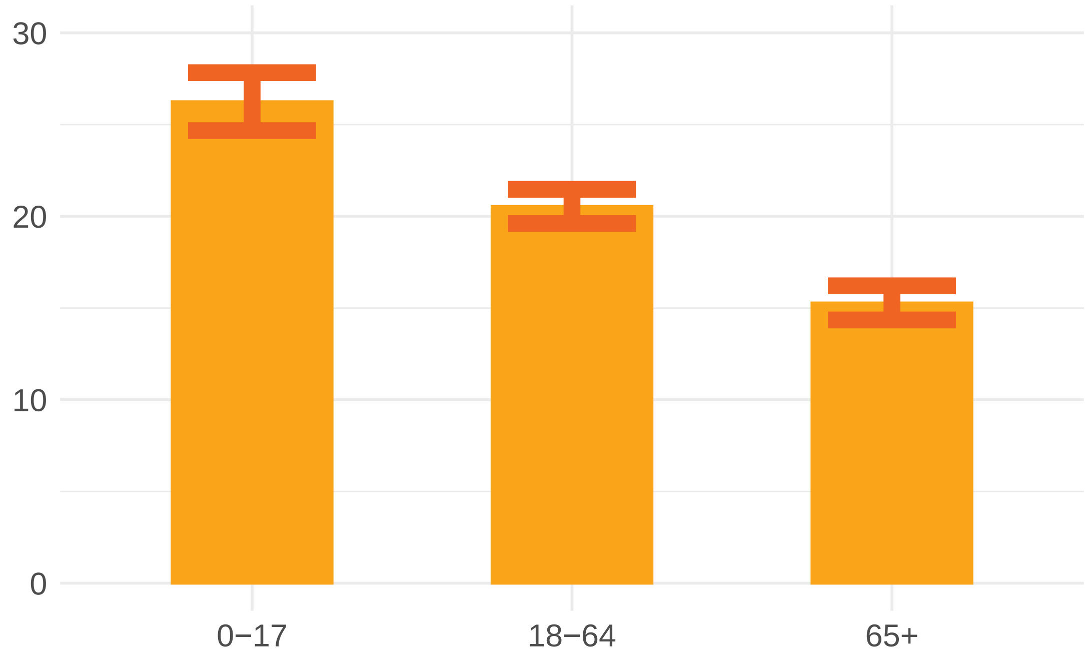

Figure 22: AROP: Error bars for AROP by age
Source: Own Elaboration from National Quality Report Italy, year 2018

However, error bars do not provide indication of the underlying distribution of the number. For this reason we suggest continuous-outcome visualizations approach by referring to a probability distribution which describes a set of possible values for the estimated poverty indicators that are consistent to varying degrees with the data we saw and what our model assumes. For the sake of explaining uncertainty, a statistical model may be created that allows pointing to intervals. This kind of uncertainty can be defined as the difference between the estimated and the true population value. The measurement of poverty is accurate, but not exact, since it is an estimate based on a sample of total population and it is therefore affected by sampling and non-sampling errors. As an example, Figure 23 shows the probability distribution for the AROP in Italy for the year 2018. Therefore, this plot may help users in understanding the uncertainty in what the AROP was in 2018: there was a 95% chance the AROP was between 19.39% and 21.2%. Values that are more consistent are assigned a higher probability.

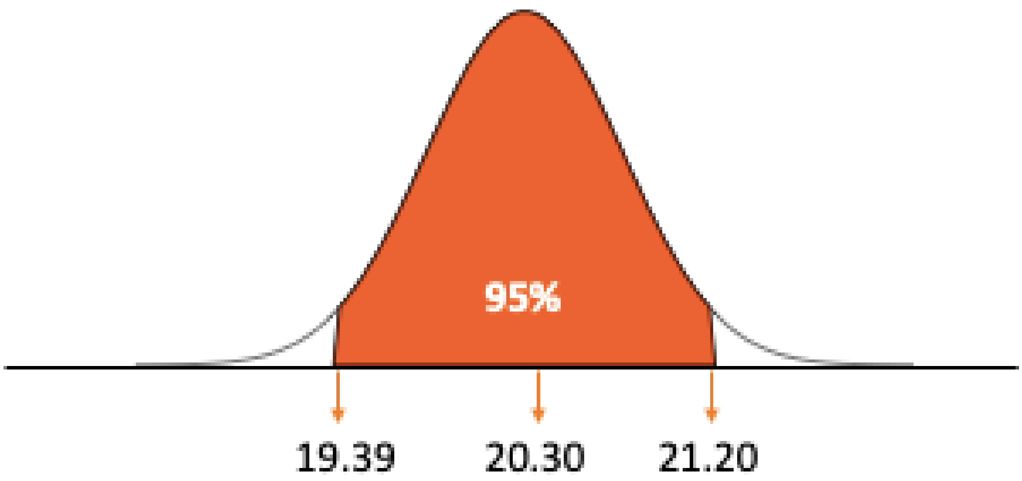

Figure 23: Probability distribution of uncertainty for AROP (Italy 2018)
Source: National Quality Report for Italy, year 2018

It is worth noting that official poverty indicators are inherently uncertain due to the way they are computed and this may change the interpretation of the meaning of intervals or probability distributions. In this context, the effect of uncertainty communicating on decision making is particularly relevant.

7 Conclusion

This report focuses on the issue of measuring and communicating uncertainty in poverty measures on EU countries by using EU-SILC surveys. Information about the sampling variability of point estimates is essential when comparing poverty differences among geographical areas or among various social-economic groups and when understanding if poverty rate has increased or decreased over time. The computation of standard errors for the main poverty measures is a complex task due to the characteristics of these indicators which are often expressed as non-linear statistics. Moreover, standard errors estimation should reflect as much as possible the complexity of EU-SILC surveys, which involve stratification, geographical clustering, unequal probabilities of selection, post-survey weighting adjustments and rotating samples. If these aspects are not considered, standard errors can be under-estimated, thus resulting in wrong interpretations.

First, we briefly reviewed the various studies focused on comparing the linearization approach with re-sampling methods for the AROP, RMPG, Gini coefficient and S80/S20. Since the linearization, jackknife and bootstrap are estimating the same quantity, that is the variance of poverty indicators, one may ask if it is possible to identify conditions under which some estimators perform better than others. In this context, several comparisons, both theoretical and empirical, have been carried out in the literature. In this report we provided an empirical application based on EU-SILC survey for assessing the relative performance of these methods even if we cannot know the true value of the variance. In our analysis, we focused on eight countries chosen according to their sampling designs. More specifically, we selected five countries using the two-stage stratified sampling design, that is Italy, the United Kingdom, Portugal, Belgium and Ireland; two countries using the one-stage sampling design (Germany and Sweden) and the Norway which adopts a simple sampling design without stratification. With the aim of reflecting the main features of the sample design an "analytic" approach for variance estimation could be used. Nevertheless, since the EU-SILC indicators are nonlinear statistics, they had to be linearized to allow variance calculations. The Taylor linearisation approach for approximating variance of non-linear statistics is a long-established procedure. However, for statistics which cannot be expressed as a smooth function of estimated totals, other methods should be used for variance estimation. In this report we referred to the concept of influence functions.

Bootstrap and jackknife are also used to obtain relative standard errors to be compared with those derived from linearization.The jackknife approach is a valid alternative to the linearization approach since it is simpler technically. However, this method requires the specification of the sample structure and the appropriate definition of computational strata. It involves repeated computation of the estimates (for which sampling errors are required) over different (often numerous) sample replications. Then variance of any statistic is estimated simply from variability in the estimate itself across the replications. The final variance estimation formula does not depend on the particular statistic involved. The presence of large values at the upper end of the distribution does not affect point estimates of poverty, but it can markedly affect the estimated indicators of inequality such as the Gini coefficient and S80/S20. Therefore, variance of the estimates can also become greatly inflated.

Moreover in order to obtain strata containing at least four PSUs, we have manipulated the data by constructing computational strata. However, information regarding the order of PSUs is not available. Therefore in order to obtain computational strata we ordered them according to their individual ID. We considered this new structure when computing jackknife and linearization but not when implementing bootstrap. Therefore, bootstrap may under estimate relative standard error of the various poverty indicators since it may replicate strata with a single PSU, thus reducing the variability of estimates.

Therefore, several factors including the different dataset used by NSIs when producing national quality report may affect our estimation results and their differences with respect to the official figures. Our findings demonstrate that there is not a unique method working well with any complex estimator and sampling design implemented in the various countries (Ollila, 2004). Yet, the general closeness of the results from these three entirely different methodologies can be considered quite remarkable. Bootstrap method seems to perform poorly when estimating standard errors for income-inequality indicators (Gini coefficient and S80/S20) especially in the case of the UK and Belgium. The estimation results may be influenced by the implicit stratification used by Belgium and the United Kingdom, where PSUs are sorted in descending order according to variables strictly related to income. Using the jackknife and linearization methods we are able to better capture both the reduced number of PSUs and, to some extent, the implicit stratification by defining computational strata.
Contrastingly, the relative standard errors obtained using jackknife could be influenced by the sensitivity of this method to the construction of strata as well as by the exact nature of the upper tail of the income distribution, especially regarding the income-inequality indicators.

As expected, uncertainty measures for income-poverty and income-inequality indicators at NUTS2 level are higher than those obtained for NUTS1 regions especially for the UK and Germany. This effect is due to the reduced number of sampling units in each NUTS2 regions which characterize these two countries. In addition, when sample sizes become small, sampling error tends not only to be high, but also estimates of sampling error tend to be more complex and subject to high levels of variability. In order to ameliorate the problem of small sample sizes and produce regional estimates with reduced sampling error, various procedures can be implemented. It is advisable to improve size and unit allocation and/or using auxiliary information for computing small area estimation. Overall, linearization proved to be the best performing method for estimating relative standard errors for poverty indicators. However, all linearization approaches rest on the asymptotic assumption that the sample size is large enough for the linear approximation to be valid. Moreover, linearization method is not always the most practical procedure for variance estimation for the type of statistics and samples being considered thus requiring high level of expertise for implementing computational formulae.

Another critical aspect in measuring uncertainties of poverty indicators is how to communicate them in a “comprehensive” manner, in terms of fully capturing the uncertainties, but also in a “understandable” way so that different users and readers of these data correctly infer and interpret the uncertainties communicated to them. Increasing attention has been paid to this aspect in literature (Spiegelhalter et al., 2011, Van der Bles et al., 2019). NSIs communicate uncertainty of poverty measures by providing a quantification of the magnitude of variability in their national quality reports. However, to the authors’ knowledge, only few EU countries currently communicate measures of uncertainty when publishing new release on poverty and living conditions. To this respect, useful suggestions for communicating uncertainty in the national quality reports are provided.

8 References

References

Alfons, A. and Templ, M., *Estimation of social exclusion indicators from complex surveys: The R Package Laeken*, KU Leuven, Faculty of Business and Economics Working Paper, 2012.

Alper, M. O. and Berger, Y. G., *Variance estimation of change in poverty rates: an application to the Turkish EU-SILC survey*, Journal of Official Statistics, 2015, 31(2):155–175.

Antal, E., Langel, M., and Tillé, Y., Variance estimation of inequality indices in complex sampling designs, In *Proceedings of the th World Statistical Congress*, 2011.

Ardilly, P., *Les techniques de sondage*, Paris: Technip, 2006.

Atkinson, A. B., Guio, A.-C., and Marlier, E., *Monitoring social inclusion in Europe*, Publications Office of the European Union Luxembourg, 2017.

Berger, Y., Osier, G., and Goedemé, T., *26standard error estimation and related sampling issues*, Monitoring social inclusion in Europe, 2017, page 465.

Berger, Y. G., *A note on the asymptotic equivalence of Jackknife and linearization variance estimation for the Gini coefficient*, , 2008.

Berger, Y. G. and Priam, R., *A simple variance estimator of change for rotating repeated surveys: an application to the EU-SILC household surveys*, Journal of the Royal Statistical Society: Series A (Statistics in Society), 2016, 179(1):251–272.

Berger, Y. G. and Skinner, C. J., *Variance estimation for a low-income proportion*, Journal of the Royal Statistical Society: series C (applied statistics), 2003, 52(4):457–468.

Betti, G. and Gagliardi, F., *Extension of JRR method for variance estimation of net changes in inequality measures*, Social Indicators Research, 2018, 137(1):45–60.

Betti, G., Gagliardi, F., Lemmi, A., and Verma, V., *Subnational indicators of poverty and deprivation in Europe: methodology and applications*, Cambridge Journal of Regions, Economy and Society, 2012, 5(1):129–147.

Bhattacharya, D., *Inference on inequality from household survey data*, Journal of Econometrics, 2007, 137(2):674–707.

Biewen, M., *Income inequality in Germany during the 1980s and 1990s*, Review of Income and Wealth, 2000, 46(1):1–19.

Biewen, M., *Bootstrap inference for inequality, mobility and poverty measurement*, Journal of Econometrics, 2002, 108(2):317–342.

Biggeri, L., Laureti, T., and Polidoro, F., *Computing sub-national PPPS with CPI data: an empirical analysis on Italian data using country product dummy models*, Social Indicators Research, 2017, 131(1):93–121.

Binder, D. A., Use of estimating functions for interval estimation from complex surveys, In *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 1991. pages 34–42.

Binder, D. A. and Patak, Z., *Use of estimating functions for estimation from complex surveys*, Journal of the American Statistical Association, 1994, 89(427):1035–1043.

Brzezinski, M., *Statistical inference for richness measures*, Applied Economics, 2014, 46(14):1599–1608.

Caron, N., *Le logiciel poulpe: aspects méthodologiques*, INSEE: Actes des Journées de Méthodologie. Available at: http://jms. insee. fr/files/documents/1998/513_1-JMS1998_S3-1_CARON_P173-200. PDF (accessed May 2015), 1998.

Cowell, F. A. and Flachaire, E., *Income distribution and inequality measurement: The problem of extreme values*, Journal of Econometrics, 2007, 141(2):1044–1072.

Cowell, F. A. and Victoria-Feser, M.-P., *Distribution-free inference for welfare indices under complete and incomplete information*, The Journal of Economic Inequality, 2003, 1(3):191–219.

Davidson, R., *Statistical inference in the presence of heavy tails*, The Econometrics Journal, 2012, 15(1):C31–C53.

Davidson, R. and Duclos, J.-Y., *Statistical inference for stochastic dominance and for the measurement of poverty and inequality*, Econometrica, 2000, 68(6):1435–1464.

Davidson, R. and Flachaire, E., *Asymptotic and bootstrap inference for inequality and poverty measures*, Journal of Econometrics, 2007, 141(1):141–166.

Demnati, A. and Rao, J., *Linearization variance estimators for survey data*, Survey Methodology, 2004, 30(1):17–26.

Deville, J. C., *Variance estimation for complex statistics and estimators: linearization and residual techniques*, Survey methodology, 1999, 25(2):193–204.

Druzdzel, M. J., *Verbal uncertainty expressions: Literature review*, Pittsburgh, PA: Carnegie Mellon University, Department of Engineering and Public Policy, 1989.

Durbin, J., *A note on the application of Quenouille's method of bias reduction to the estimation of ratios*, Biometrika, 1959, 46(3/4):477–480.

Efron, B. and Stein, C., *The Jackknife estimate of variance*, The Annals of Statistics, 1981, pages 586–596.

Foster, J., Greer, J., and Thorbecke, E., *A class of decomposable poverty measures*, Econometrica: journal of the econometric society, 1984, pages 761–766.

Foster, J., Greer, J., and Thorbecke, E., *The Foster-Greer-Thorbecke (FGT) poverty measures: 25 years later*, The Journal of Economic Inequality, 2010, 8(4):491–524.

Foster, J. E. and Shorrocks, A. F., *Poverty orderings*, Econometrica: Journal of the Econometric Society, 1988, pages 173–177.

Frewer, L. J., Miles, S., Brennan, M., Kuznesof, S., Ness, M., and Ritson, C., *Public preferences for informed choice under conditions of risk uncertainty*, Public understanding of science, 2002, 11(4):363–372.

Giorgi, G. M. and Gigliarano, C., *The Gini concentration index: a review of the inference literature*, Journal of Economic Surveys, 2017, 31(4):1130–1148.

Goedemé, T., *How much confidence can we have in EU-SILC? Complex sample designs and the standard error of the Europe 2020 poverty indicators*, Social Indicators Research, 2013, 110(1):89–110.

Goedemē, T. et al., The EU-SILC sample design variables: critical review and recommendations, Technical report, 2013.

Graf, E. and Tillé, Y., *Variance estimation using linearization for poverty and social exclusion indicators*, Survey Methodology, 2014, 40(1):61–79.

Graf, M., Wenger, A., and Nedyalkova, D., *Quality of EU-SILC data*, AMELI deliverable, 2011, 5.

Greenland, S., Senn, S. J., Rothman, K. J., Carlin, J. B., Poole, C., Goodman, S. N., and Altman, D. G., *Statistical tests, P-values, confidence intervals, and power: A guide to misinterpretations*, European journal of epidemiology, 2016, 31(4):337–350.

Griethe, H., Schumann, H., et al., The visualization of uncertain data: Methods and problems., In *SimVis*, 2006. pages 143–156.

Hampel, F. R., *The influence curve and its role in robust estimation*, Journal of the American Statistical Association, 1974, 69(346):383–393.

Hoeffding, W., A class of statistics with asymptotically normal distribution, In *Breakthroughs in statistics*, pages 308–334. Springer, 1992.

Istat, Income, living conditions and fiscal burden of households, pages 91–110, 2018.

Kakwani, N., *Statistical inference in the measurement of poverty*, The Review of Economics and Statistics, 1993, pages 632–639.

Kovacevic, M. S. and Yung, W., *Variance estimation for measures of income inequality and polarization-an empirical study*, Survey Methodology, 1997, 23:41–52.

Langel, M. and Tillé, Y., *Statistical inference for the quintile share ratio*, Journal of Statistical Planning and Inference, 2011, 141(8):2976–2985.

Langel, M. and Tillé, Y., *Variance estimation of the Gini index: revisiting a result several times published*, Journal of the Royal Statistical Society: Series A (Statistics in Society), 2013, 176(2):521–540.

Laureti, T. and Rao, D. P., *Measuring spatial price level differences within a country: Current status and future developments*, Studies of Applied Economics, 2019, 36(1):119–148.

Maasoumi, E., *Empirical analyses of inequality and welfare*, Handbook of applied econometrics, 1997, 2:202–245.

Manfredi, G., *L'applicazione del Jackknife nella stima del rapporto di concentrazione r di Gini*, Annali dell'Istituto di Statistica, Universita di Bari, 1974, 38:15–36.

Mills, J. A. and Zandvakili, S., *Statistical inference via bootstrapping for measures of inequality*, Journal of Applied econometrics, 1997, 12(2):133–150.

Münnich, R. and Zins, S., *Variance estimation for indicators of poverty and social exclusion*, Work package of the European project on Advanced Methodology for European Laeken Indicators (AMELI). Available at: http://www. uni-trier. de/index. php, 2011.

OECD, OECD framework for statistics on the distribution of household income, consumption and wealth, 2013.

Ogwang, T., *A convenient method of computing the Gini index and its standard error*, Oxford Bulletin of Economics and Statistics, 2000, 62(1):123–129.

Ollila, P., *A theoretical overview for variance estimation in sampling theory with some new techniques for complex estimators*, Tilastokeskus, 2004.

Osier, G., Variance estimation: the linearization approach applied by Eurostat to the 2004 SILC operation, In *Eurostat and Statistics Finland Methodological Workshop on EU-SILC, Helsinki*, 2006.

Osier, G., Variance estimation for complex indicators of poverty and inequality using linearization techniques, In *Survey Research Methods*, 2009. volume 3, pages 167–195.

Palmitesta, G. and Provasi, C., *Asymptotic and bootstrap inference for the generalized Gini indices*, Metron, 2006, 64:107–124.

Palmitesta, P., Provasi, C., and Spera, C., *Confidence interval estimation for inequality indices of the Gini family*, Computational Economics, 2000, 16(1-2):137–147.

Piacentini, M., *Measuring income inequality and poverty at the regional level in OECD countries*, , 2014.

Preston, I., *Sampling distributions of relative poverty statistics*, Journal of the Royal Statistical Society: Series C (Applied Statistics), 1995, 44(1):91–99.

Rao, J. and Wu, C. J., *Inference from stratified samples: Second-order analysis of three methods for nonlinear statistics*, Journal of the American Statistical Association, 1985, 80(391):620–630.

Rao, J. N. and Wu, C., *Resampling inference with complex survey data*, Journal of the american statistical association, 1988, 83(401):231–241.

Ravallion, M. et al., *Poverty comparisons*, volume 56. Taylor & Francis, 1994.

Schechtman, E., *On estimating the asymptotic variance of a function of U statistics*, The American Statistician, 1991, 45(2):103–106.

Shao, J. and Wu, C. J., *A general theory for Jackknife variance estimation*, The Annals of Statistics, 1989, pages 1176–1197.

Spiegelhalter, D., Pearson, M., and Short, I., *Visualizing uncertainty about the future*, science, 2011, 333(6048):1393–1400.

Sturgis, P., *Analysing complex survey data: Clustering, stratification and weights*, Social research UPDATE, 2004, (43).

Thuysbaert, B., *Inference for the measurement of poverty in the presence of a stochastic weighting variable*, The Journal of Economic Inequality, 2008, 6(1):33–55.

Tibshirani, R. J. and Efron, B., *An introduction to the bootstrap*, Monographs on statistics and applied probability, 1993, 57:1–436.

Trindade, L. Z. and Goedemé, T., *The comparability of the EU-SILC income variables: review and recommendations*, Luxembourg: Publications Office of the European Union, 2020, 2020.

EFSA Scientific Committee, Hardy, A., Benford, D., Halldorsson, T., Jeger, M. J., Knutsen, H. K., More, S., Naegeli, H., Noteborn, H., Ockleford, C., et al., *Guidance on the use of the weight of evidence approach in scientific assessments*, EFSA Journal, 2017, 15(8):e04971.

Van der Bles, A. M., van der Linden, S., Freeman, A. L., Mitchell, J., Galvao, A. B., Zaval, L., and Spiegelhalter, D. J., *Communicating uncertainty about facts, numbers and science*, Royal Society open science, 2019, 6(5):181870.

Verma, V. and Betti, G., *EU Statistics on Income and Living Conditions (EU-SILC): Choosing the survey structure and sample design*, Statistics in Transition, 2006, 7(5):935–970.

Verma, V. and Betti, G., *Data accuracy in EU-SILC*, Income and living conditions in Europe, 2010, page 57.

Verma, V. and Betti, G., *Taylor linearization sampling errors and design effects for poverty measures and other complex statistics*, Journal of Applied Statistics, 2011, 38(8):1549–1576.

Verma, V., Betti, G., and Gagliardi, F., *An assessment of survey errors in EU-SILC*, Luxembourg: Eurostat, 2010.

Verma, V., Lemmi, A., Betti, G., Gagliardi, F., and Piacentini, M., *How precise are poverty measures estimated at the regional level?*, Regional Science and Urban Economics, 2017, 66:175–184.

Xiquan, S., *A note on bootstrapping the U-STATISTICS*, Chinese Journal of Applied Probability and Statistics, 1986, 2.

Xu, K., *U-statistics and their asymptotic results for some inequality and poverty measures*, Econometric Reviews, 2007, 26(5):567–577.

Yitzhaki, S., *Calculating Jackknife variance estimators for parameters of the Gini method*, Journal of Business & Economic Statistics, 1991, 9(2):235–239.

Annex

Table 16: Estimated standard errors EU-SILC 2018 for AROP indicator – Italy NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Piemonte	14.7	6.04	8.48	9.32
Valle d'Aosta	19.9	39.72	15.41	11.77
Liguria	14.3	16.18	6.66	9.98
Lombardia	11.2	7.03	21.52	7.98
Abruzzo	22.5	8.61	4.36	12.35
Molise	23.4	11.22	3.58	9.96
Campania	40.5	4.78	3.70	3.92
Puglia	26.7	8.11	5.10	4.99
Basilicata	28.5	10.56	6.04	8.28
Calabria	34.5	8.33	5.42	9.52
Sicilia	39.8	5.66	4.36	5.73
Sardegna	24.4	8.28	8.45	8.95
Bolzano	11.6	27.92	31.94	26.08
Trento	11.4	15.26	14.92	16.83
Veneto	12.0	7.79	29.47	10.50
Friuli-Venezia Giulia	8.0	7.62	46.65	18.46
Emilia-Romagna	10.5	7.28	25.68	10.50
Toscana	14.5	6.08	22.42	8.44
Umbria	13.6	12.79	21.83	15.25
Marche	15.9	6.85	15.77	8.65
Lazio	20.1	4.85	14.21	6.65

Table 17: Estimated standard errors EU-SILC 2018 for relative mean Poverty Gap indicator (RMPG) – Italy NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Piemonte	31.1	10.05	23.49	11.09
Valle d'Aosta	46.1	21.72	11.07	8.69
Liguria	30.0	14.74	3.63	9.64
Lombardia	22.5	10.20	50.12	10.58
Abruzzo	38.8	15.11	9.83	17.21
Molise	22.7	29.57	42.55	37.43
Campania	34.2	5.78	18.91	15.20
Puglia	30.3	7.04	11.06	18.28
Basilicata	34.1	12.02	15.68	16.67
Calabria	34.1	14.98	22.56	19.92
Sicilia	34.5	9.35	29.51	23.68
Sardegna	33.0	12.04	17.86	18.59
Bolzano	18.6	102.85	30.44	29.04
Trento	29.2	42.47	3.68	18.65
Veneto	20.1	15.63	10.81	11.08
Friuli-Venezia Giulia	18.3	13.32	28.69	14.53
Emilia-Romagna	21.5	16.46	15.59	9.94
Toscana	21.7	16.68	25.00	10.39
Umbria	19.6	29.28	90.86	21.86
Marche	23.4	26.89	42.51	20.69
Lazio	28.0	5.39	12.67	8.17

Table 18: **Estimated standard errors EU-SILC 2018 for Gini indicator – Italy NUTS2**

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Piemonte	30.5	2.04	2.91	2.85
Valle d'Aosta	33.0	13.86	6.67	5.43
Liguria	33.7	7.42	3.82	2.21
Lombardia	31.4	2.94	3.71	3.77
Abruzzo	30.3	5.24	2.46	4.54
Molise	29.2	2.31	2.84	3.86
Campania	35.6	1.98	2.33	3.44
Puglia	32.2	2.77	3.72	3.27
Basilicata	33.2	6.15	2.79	4.26
Calabria	35.2	5.36	3.17	4.72
Sicilia	36.0	3.57	3.62	3.92
Sardegna	33.8	5.56	3.03	6.61
Bolzano	30.9	9.68	4.88	5.42
Trento	31.1	4.93	3.32	7.30
Veneto	29.3	2.00	7.16	3.62
Friuli-Venezia Giulia	25.7	2.38	8.47	3.48
Emilia-Romagna	29.2	2.96	3.75	2.86
Toscana	31.2	2.70	5.06	3.56
Umbria	32.0	7.89	8.54	3.63
Marche	30.4	3.23	7.25	3.73
Lazio	35.6	0.92	3.83	3.04

Table 19: **Estimated standard errors EU-SILC 2018 for Quantile Share Ratio (S80/S20) indicator – Italy NUTS2**

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Piemonte	5.2	3.88	6.20	4.51
Valle d'Aosta	6.7	28.05	8.42	5.15
Liguria	5.8	13.69	6.17	5.52
Lombardia	5.1	5.31	10.77	5.73
Abruzzo	5.7	14.22	3.44	8.16
Molise	4.6	6.72	4.61	10.94
Campania	7.4	5.48	2.93	8.61
Puglia	5.8	6.53	6.41	6.37
Basilicata	6.2	13.16	7.03	9.03
Calabria	7.4	14.43	4.16	11.95
Sicilia	8.0	13.70	5.38	17.02
Sardegna	6.3	8.94	6.66	13.57
Bolzano	4.7	25.45	17.21	8.95
Trento	5.1	9.06	8.77	13.97
Veneto	4.7	3.19	25.44	5.06
Friuli-Venezia Giulia	3.7	4.83	41.34	4.54
Emilia-Romagna	4.6	5.03	16.17	4.61
Toscana	5.0	4.11	13.68	5.06
Umbria	4.9	12.32	18.07	5.70
Marche	5.1	6.86	18.26	5.54
Lazio	6.7	2.42	17.64	6.17

Table 20: Estimated standard errors EU-SILC 2018 for AROP indicator – The UK NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Tees Valley and Durham	23.0	15.26	10.05	9.56
Northumberland and Tyne and Wear	15.5	15.95	20.07	16.41
Cumbria	28.6	15.81	7.03	10.52
Greater Manchester	25.6	9.81	8.48	6.97
Lancashire	21.6	11.57	14.55	10.92
Cheshire	15.5	16.92	13.38	17.03
Merseyside	24.3	13.12	8.43	8.49
East Yorkshire	17.4	21.65	21.79	15.23
North Yorkshire	21.4	15.88	17.55	14.32
South Yorkshire	27.1	13.20	10.30	10.04
West Yorkshire	22.4	11.44	27.47	10.03
Derbyshire and Nottinghamshire	17.7	11.77	15.55	10.54
Leicestershire	20.1	13.25	19.42	10.32
Lincolnshire	17.5	18.81	6.75	23.26
Herefordshire	12.4	18.87	16.96	17.96
Shropshire and Staffordshire	20.5	12.24	17.54	12.59
West Midlands	24.5	8.85	10.54	8.69
East Anglia	21.2	10.72	15.14	9.62
Bedfordshire and Hertfordshire	14.0	15.58	20.96	15.96
Essex	17.4	12.54	17.74	11.92
Inner London - West	17.5	16.73	12.78	15.11
Inner London - East	18.2	19.00	21.31	18.21
Outer London - East and North East	20.2	11.53	8.41	13.29
Outer London - South	12.9	16.69	11.53	17.71
Outer London - West and North West	13.3	18.31	19.82	19.06
Berkshire Oxfordshire	9.6	18.38	13.00	25.05
Surrey, East and West Sussex	12.8	16.37	18.47	15.31
Hampshire and Isle of Wight	12.1	14.35	17.82	19.87
Kent	14.6	15.88	8.03	16.60
Gloucestershire and Bristol area	13.1	12.42	23.54	18.87
Dorset and Somerset	17.0	14.63	19.96	12.66
Cornwall and Isles of Scilly	19.5	25.11	7.63	22.38
Devon	12.7	18.37	22.62	24.42
West Wales and The Valleys	20.1	10.64	11.93	9.66
East Wales	20.8	11.05	9.11	14.53
North Eastern Scotland	19.0	22.68	16.62	20.27
Highlands and Islands	23.2	20.17	9.94	12.74
Eastern Scotland	15.8	10.45	17.59	11.93
West Central Scotland	20.0	12.87	13.10	11.27
Southern Scotland	17.2	16.99	8.82	15.20
Northern Ireland	28.9		11.95	3.92

Table 21: **Estimated standard errors EU-SILC 2018 for relative mean poverty gap (RMPG) indicator – The UK NUTS2**

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Tees Valley and Durham	20.8	30.36	26.03	25.87
Northumberland and Tyne and Wear	11.1	34.20	27.48	33.36
Cumbria	20.9	25.22	34.61	36.15
Greater Manchester	24.8	15.38	9.99	14.91
Lancashire	29.1	17.02	10.51	14.01
Cheshire	32.3	32.01	13.72	13.02
Merseyside	32.8	16.60	4.23	14.02
East Yorkshire	24.9	12.20	21.27	18.52
North Yorkshire	27.6	23.82	14.09	24.30
South Yorkshire	19.3	24.19	13.00	28.88
West Yorkshire	22.2	9.74	24.80	11.23
Derbyshire and Nottinghamshire	25.2	11.98	15.76	12.73
Leicestershire	19.5	15.62	39.03	29.10
Lincolnshire	24.0	17.44	6.44	18.72
Herefordshire	20.7	9.40	8.85	18.86
Shropshire and Staffordshire	24.8	14.43	15.28	14.04
West Midlands	23.9	9.15	12.94	11.57
East Anglia	14.9	25.01	33.04	25.61
Bedfordshire and Hertfordshire	15.7	28.11	37.34	25.62
Essex	11.3	33.66	13.54	28.52
Inner London - West	27.3	29.74	12.01	18.25
Inner London - East	36.5	12.21	16.66	15.59
Outer London - East and North East	22.2	16.36	30.44	19.60
Outer London - South	22.3	24.81	33.27	22.24
Outer London - West and North West	21.7	22.04	24.55	17.02
Berkshire Oxfordshire	27.1	14.35	33.93	20.23
Surrey, East and West Sussex	20.7	29.79	31.68	18.62
Hampshire and Isle of Wight	24.2	17.36	34.08	18.89
Kent	13.6	52.92	95.55	27.65
Gloucestershire and Bristol area	16.3	18.82	19.10	19.94
Dorset and Somerset	25.2	26.06	26.81	16.66
Cornwall and Isles of Scilly	19.7	45.43	4.33	31.08
Devon	25.6	17.54	20.41	13.48
West Wales and The Valleys	25.7	12.95	21.13	14.03
East Wales	21.5	17.09	13.75	18.26
North Eastern Scotland	32.6	32.93	48.19	22.68
Highlands and Islands	24.2	23.78	8.03	29.96
Eastern Scotland	20.6	15.12	60.86	13.11
West Central Scotland	20.1	17.99	14.69	15.28
Southern Scotland	21.8	19.41	4.56	17.95
Northern Ireland	36.6		21.01	2.56

Table 22: Estimated standard errors EU-SILC 2018 for Gini coefficient indicator – The UK NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Tees Valley and Durham	29.39	6.05	5.81	4.49
Northumberland and Tyne and Wear	28.13	5.72	3.06	6.99
Cumbria	31.72	5.85	3.82	8.70
Greater Manchester	33.51	5.34	4.75	3.29
Lancashire	30.89	6.31	4.70	5.59
Cheshire	31.95	5.65	4.68	5.65
Merseyside	31.70	5.56	4.02	5.63
East Yorkshire	27.08	4.21	6.19	5.01
North Yorkshire	32.82	5.91	3.15	5.55
South Yorkshire	27.93	4.84	6.38	5.46
West Yorkshire	30.00	3.52	6.19	3.33
Derbyshire and Nottinghamshire	30.02	4.02	6.57	4.39
Leicestershire	30.40	4.68	4.30	5.04
Lincolnshire	30.10	7.37	6.13	7.93
Herefordshire	26.38	5.68	14.49	6.60
Shropshire and Staffordshire	33.05	4.31	3.56	4.52
West Midlands	33.12	8.90	5.52	11.93
East Anglia	33.15	3.87	2.02	4.09
Bedfordshire and Hertfordshire	32.27	4.85	5.07	4.32
Essex	34.43	5.46	4.83	4.51
Inner London - West	46.69	7.47	4.07	8.76
Inner London - East	43.01	7.21	4.89	8.10
Outer London - East and North East	33.93	4.73	6.94	5.04
Outer London - South	33.47	5.56	3.33	5.59
Outer London - West and North West	34.33	3.94	9.09	3.85
Berkshire Oxfordshire	32.71	4.84	6.69	5.07
Surrey, East and West Sussex	33.68	6.49	4.07	7.68
Hampshire and Isle of Wight	30.10	4.41	3.24	4.55
Kent	31.59	4.92	7.24	6.32
Gloucestershire and Bristol area	31.34	5.04	3.07	3.89
Dorset and Somerset	29.20	6.43	7.38	5.24
Cornwall and Isles of Scilly	27.61	6.56	2.25	5.06
Devon	24.38	5.57	7.91	5.65
West Wales and The Valleys	28.21	4.28	4.40	4.67
East Wales	30.42	4.34	3.02	4.48
North Eastern Scotland	35.69	10.82	14.15	14.74
Highlands and Islands	29.13	11.87	4.97	7.84
Eastern Scotland	30.87	4.33	5.63	5.52
West Central Scotland	30.81	4.84	5.02	4.19
Southern Scotland	28.34	4.18	5.24	6.05
Northern Ireland	60.80	0.00	4.57	3.42

Table 23: Estimated standard errors EU-SILC 2018 for S80/S20 indicator – The UK NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Tees Valley and Durham	4.6	10.55	6.67	7.93
Northumberland and Tyne and Wear	3.6	8.32	9.64	10.37
Cumbria	4.3	13.84	6.99	15.46
Greater Manchester	5.9	10.78	7.33	5.57
Lancashire	5.3	14.32	8.98	7.91
Cheshire	5.1	17.38	10.72	13.64
Merseyside	5.3	10.89	12.14	11.45
East Yorkshire	4.0	10.67	15.18	7.72
North Yorkshire	5.6	9.92	6.43	9.61
South Yorkshire	4.0	10.54	20.52	12.30
West Yorkshire	4.7	6.42	12.95	5.57
Derbyshire and Nottinghamshire	4.6	7.83	11.72	7.27
Leicestershire	4.9	8.02	11.93	9.67
Lincolnshire	4.3	14.86	16.76	14.11
Herefordshire	3.6	10.93	25.69	11.34
Shropshire and Staffordshire	5.2	11.41	10.63	8.18
West Midlands	5.4	13.59	9.88	17.06
East Anglia	5.3	5.70	4.80	6.24
Bedfordshire and Hertfordshire	5.4	8.57	28.21	9.71
Essex	5.6	10.99	10.96	7.51
Inner London - West	12.0	18.94	5.68	24.08
Inner London - East	9.2	17.46	7.88	19.47
Outer London - East and North East	5.7	9.28	10.62	9.03
Outer London - South	5.8	10.94	5.49	10.58
Outer London - West and North West	5.9	9.10	22.51	6.84
Berkshire Oxfordshire	5.3	7.97	9.97	9.62
Surrey, East and West Sussex	5.6	11.05	8.58	11.09
Hampshire and Isle of Wight	4.6	9.56	9.40	7.31
Kent	4.9	8.36	8.54	9.95
Gloucestershire and Bristol area	4.7	8.56	7.92	7.14
Dorset and Somerset	4.8	15.94	12.66	8.29
Cornwall and Isles of Scilly	4.2	12.25	6.99	9.01
Devon	3.4	8.99	18.51	9.20
West Wales and The Valleys	4.3	7.14	7.44	7.10
East Wales	4.7	9.80	8.19	9.84
North Eastern Scotland	6.8	23.75	23.45	38.74
Highlands and Islands	4.4	26.96	7.80	10.14
Eastern Scotland	4.9	7.39	12.24	9.21
West Central Scotland	4.9	8.39	14.85	7.17
Southern Scotland	4.4	9.49	7.19	10.14
Northern Ireland	17.7	0.00	8.22	5.04

Table 24: Estimated standard errors EU-SILC 2018 for AROP indicator – Germany NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Stuttgart	12.0	7.91	10.29	9.45
Karlsruhe	12.5	10.15	21.41	12.12
Freiburg	11.6	14.10	13.20	15.92
Tübingen	12.1	11.03	22.70	14.18
Oberbayern	12.6	8.35	8.42	9.43
Niederbayern	22.6	8.21	20.15	9.65
Oberpfalz	16.4	15.12	19.22	14.98
Oberfranken	15.6	10.56	15.17	11.54
Mittelfranken	15.0	9.75	28.31	10.44
Unterfranken	10.4	16.19	43.64	20.52
Schwaben	13.4	9.79	24.80	12.24
Berlin	18.4	7.63	13.10	8.16
Brandenburg	15.5	5.85	6.57	7.36
Bremen	28.4	9.70	7.71	8.84
Hamburg	24.0	9.28	12.33	8.81
Darmstadt	10.2	8.85	26.28	11.36
Gießen	14.4	10.93	14.41	13.32
Kassel	15.1	11.30	16.67	12.45
Mecklenburg-Vorpommern	23.4	6.29	7.68	6.65
Braunschweig	18.6	9.80	12.03	8.95
Hannover	15.1	8.23	18.15	9.83
Lüneburg	12.2	10.85	21.74	13.58
Weser-Ems	15.0	9.40	10.07	10.11
Düsseldorf	15.5	5.08	10.53	6.15
Köln	12.8	6.20	9.83	8.27
Münster	18.0	7.23	12.89	7.62
Detmold	14.6	10.84	10.76	11.64
Arnsberg	18.4	4.55	12.48	5.79
Koblenz	21.7	7.78	7.17	8.50
Trier	21.6	14.18	16.33	14.00
Rheinhessen-Pfalz	19.8	11.33	13.25	9.20
Saarland	16.2	13.55	12.92	13.95
Dresden	20.3	6.74	10.54	7.94
Chemnitz	15.0	13.11	25.99	10.87
Leipzig	19.3	10.57	8.99	10.76
Sachsen-Anhalt	16.3	8.04	22.96	8.94
Schleswig-Holstein	22.5	5.68	14.70	5.76
Thüringen	18.4	6.93	18.02	7.77

Table 25: Estimated standard errors EU-SILC 2018 for Gini Coefficient – Germany NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Stuttgart	30.1	3.03	7.81	3.38
Karlsruhe	25.9	3.78	8.24	3.88
Freiburg	26.6	3.68	6.03	4.31
Tübingen	28.9	3.54	6.57	4.27
Oberbayern	32.8	4.13	2.79	4.04
Niederbayern	31.0	3.76	3.34	4.63
Oberpfalz	28.0	5.94	5.46	6.68
Oberfranken	28.1	6.43	4.80	6.89
Mittelfranken	51.9	2.38	5.02	2.94
Unterfranken	26.9	4.37	4.66	4.96
Schwaben	28.7	5.05	3.54	4.98
Berlin	26.8	4.11	7.50	4.51
Brandenburg	25.3	2.35	5.79	2.81
Bremen	28.2	5.26	8.43	4.21
Hamburg	30.6	3.92	3.22	4.34
Darmstadt	46.2	3.25	4.16	3.41
Gießen	31.1	6.06	3.50	4.18
Kassel	25.8	3.06	8.04	3.69
Mecklenburg-Vorpommern	26.1	3.09	2.89	3.28
Braunschweig	29.1	3.51	6.62	4.52
Hannover	27.1	2.62	7.75	3.14
Lüneburg	25.7	3.29	8.46	3.78
Weser-Ems	27.5	3.41	8.45	3.86
Düsseldorf	28.7	1.55	5.43	1.96
Köln	30.7	1.97	3.90	2.87
Münster	31.3	5.99	4.08	7.74
Detmold	25.5	3.80	5.65	5.81
Arnsberg	28.1	2.01	9.38	3.21
Koblenz	32.2	4.47	3.06	4.50
Trier	32.8	6.08	6.14	6.69
Rheinhessen-Pfalz	30.6	3.23	5.45	3.66
Saarland	23.7	4.55	3.26	4.71
Dresden	23.9	3.85	3.87	4.84
Chemnitz	20.2	3.54	4.58	3.94
Leipzig	27.0	3.81	6.30	5.05
Sachsen-Anhalt	27.7	5.08	5.02	6.17
Schleswig-Holstein	36.5	6.02	9.48	7.27
Thüringen	25.6	2.82	6.89	3.29

Table 26: Estimated standard errors EU-SILC 2018 for S80/S20 indicator – Germany NUTS2

NUTS2 region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
Stuttgart	4.6	6.31	17.19	6.72
Karlsruhe	3.6	5.19	16.17	5.52
Freiburg	3.9	6.03	13.51	6.18
Tübingen	4.3	5.60	29.78	6.34
Oberbayern	5.7	7.05	4.34	7.38
Niederbayern	4.4	8.86	25.56	7.23
Oberpfalz	4.9	17.51	7.12	15.04
Oberfranken	4.3	8.75	10.58	8.24
Mittelfranken	4.5	12.00	24.31	69.49
Unterfranken	3.7	8.87	20.23	9.10
Schwaben	4.4	18.32	11.17	8.84
Berlin	3.7	5.09	15.62	5.31
Brandenburg	3.6	3.50	10.58	3.72
Bremen	4.2	8.57	17.74	4.30
Hamburg	5.0	7.27	14.63	7.76
Darmstadt	8.3	16.82	8.35	19.49
Gießen	4.7	8.23	6.81	5.28
Kassel	3.7	7.35	17.01	7.01
Mecklenburg-Vorpommern	3.8	4.65	14.19	5.09
Braunschweig	4.4	4.68	18.65	5.58
Hannover	4.1	4.54	13.87	4.70
Lüneburg	3.6	5.01	9.19	4.77
Weser-Ems	4.0	5.50	20.17	5.40
Düsseldorf	4.3	2.83	20.13	3.31
Köln	4.8	3.19	14.29	4.82
Münster	4.9	8.65	9.96	11.48
Detmold	3.9	6.43	91.87	8.97
Arnsberg	4.2	3.20	16.23	4.12
Koblenz	5.6	7.95	7.07	8.20
Trier	5.2	13.17	10.44	11.39
Rheinhessen-Pfalz	5.3	7.87	8.19	7.93
Saarland	3.3	9.05	13.70	7.37
Dresden	3.4	5.32	12.64	5.74
Chemnitz	2.6	4.60	49.22	3.47
Leipzig	3.8	6.74	12.50	6.62
Sachsen-Anhalt	3.9	5.90	10.19	6.78
Schleswig-Holstein	9.2	7.75	8.68	32.73
Thüringen	3.7	4.89	10.64	5.31

Table 27: Estimated standard errors EU-SILC 2018 – Portugal NUTS2

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Norte	17.8	5.91	6.01	5.27
Algarve	16.8	8.68	6.54	6.29
Centro (PT)	17.0	7.62	10.45	5.33
Lisboa	11.1	9.66	16.92	8.81
Alentejo	15.3	8.53	6.00	6.51
Região Autónoma dos Açores	23.5	8.64	11.38	5.38
Região Autónoma da Madeira	24.7	6.90	8.28	3.89
RMPG				
Norte	23.3	4.93	5.27	7.18
Algarve	26.2	13.69	14.54	9.54
Centro (PT)	26.4	6.43	7.98	5.85
Lisboa	25.3	12.65	18.28	6.26
Alentejo	20.4	14.67	16.04	11.48
Região Autónoma dos Açores	31.6	10.64	10.67	9.20
Região Autónoma da Madeira	27.2	9.60	6.13	6.68
GINI				
Norte	30.9	2.39	2.20	2.47
Algarve	33.9	6.11	2.46	3.00
Centro (PT)	32.2	2.73	2.49	2.22
Lisboa	33.6	2.01	2.69	2.04
Alentejo	31.8	5.38	1.89	2.17
Região Autónoma dos Açores	38.1	3.54	2.17	2.64
Região Autónoma da Madeira	33.6	2.40	2.59	1.94
S80/S20				
Norte	4.9	3.83	4.06	3.68
Algarve	5.6	8.61	4.66	4.70
Centro (PT)	5.1	4.96	4.02	3.55
Lisboa	5.6	3.90	4.44	3.56
Alentejo	4.8	7.65	3.24	3.49
Região Autónoma dos Açores	7.2	8.35	7.53	5.73
Região Autónoma da Madeira	5.8	4.52	5.91	4.26

Table 28: Estimated standard errors EU-SILC 2018 – Belgium NUTS2

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Brussels	26.0	9.70	5.59	4.70
Antwerp	9.5	18.57	30.59	18.14
Limburg	8.8	25.58	31.73	27.68
East Flanders	9.8	13.39	23.96	15.37
Flemish Brabant	10.8	23.59	21.16	22.57
West Flanders	7.0	25.47	22.62	26.25
Walloon Brabant	20.9	27.15	11.31	15.90
Hainaut	18.4	15.23	6.68	11.23
Liège	16.6	15.41	19.87	12.60
Luxembourg	17.2	19.89	21.97	9.91
Namur	20.5	16.13	21.12	22.43
RMPG				
Brussels	20.3	13.31	32.94	8.07
Antwerp	15.9	33.06	31.43	22.08
Limburg	11.3	60.59	24.54	26.34
East Flanders	11.5	21.89	28.88	16.99
Flemish Brabant	11.4	39.02	24.49	25.72
West Flanders	17.4	30.01	27.20	23.63
Walloon Brabant	17.3	57.76	40.41	38.98
Hainaut	18.0	14.04	11.73	17.08
Liège	13.4	12.99	8.97	18.65
Luxembourg	28.1	27.32	9.42	19.88
Namur	12.5	29.70	21.15	37.41
GINI				
Brussels	32.6	3.63	3.06	1.58
Antwerp	25.4	9.98	3.12	5.11
Limburg	21.3	3.72	6.48	4.48
East Flanders	23.2	3.02	4.87	3.45
Flemish Brabant	28.2	5.64	9.81	9.72
West Flanders	21.4	5.70	4.23	5.42
Walloon Brabant	28.8	6.84	2.34	5.99
Hainaut	24.2	3.97	3.00	5.36
Liège	25.0	4.66	5.55	4.23
Luxembourg	26.4	11.64	8.97	9.16
Namur	24.6	6.44	3.69	5.21
S80/S20				
Brussels	5.1	7.37	4.00	2.89
Antwerp	3.5	12.12	9.39	6.70
Limburg	2.9	5.01	8.01	4.07
East Flanders	3.3	5.34	18.85	4.50
Flemish Brabant	4.2	7.92	10.04	10.44
West Flanders	3.0	6.75	8.40	6.36
Walloon Brabant	4.3	17.03	6.16	13.52
Hainaut	3.4	6.19	3.92	8.35
Liège	3.3	5.46	7.96	5.64
Luxembourg	3.9	23.46	13.70	10.16
Namur	3.5	15.92	6.08	6.29

Table 29: **Estimated standard errors EU-SILC 2018 – Sweden NUTS2**

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Stockholm	12.4	4.11	16.47	6.28
Östra Mellansverige	17.1	3.74	8.17	5.17
Småland med öarna	17.7	4.79	14.82	7.09
Sydsverige	21.1	3.58	7.30	4.75
Västsverige	14.3	3.38	12.97	5.70
Norra Mellansverige	17.9	5.10	20.34	7.22
Mellersta Norrland	15.7	7.70	16.29	11.90
Övre Norrland	18.2	6.09	7.28	8.85
GINI				
Stockholm	28.9	2.13	9.95	3.33
Östra Mellansverige	24.5	1.25	4.34	1.87
Småland med öarna	27.9	3.84	3.00	6.86
Sydsverige	27.3	1.84	2.68	2.37
Västsverige	24.9	1.60	4.36	2.95
Norra Mellansverige	22.2	1.57	5.69	2.49
Mellersta Norrland	21.9	2.49	5.12	4.27
Övre Norrland	24.8	2.76	6.57	3.00
S80/S20				
Stockholm	4.4	2.77	14.13	4.49
Östra Mellansverige	3.6	2.15	6.23	3.17
Småland med öarna	4.2	4.69	6.79	8.17
Sydsverige	4.2	4.01	6.68	4.98
Västsverige	3.7	2.11	10.64	3.96
Norra Mellansverige	3.1	2.30	19.46	3.11
Mellersta Norrland	3.1	3.34	18.78	7.01
Övre Norrland	3.6	3.75	9.85	3.99

Table 30: **Estimated standard errors EU-SILC 2018 – Norway NUTS2**

Indicator/region	Stat	Boots. Rel. SE (%)	Jackk. Rel. SE (%)	Linear. Rel. SE (%)
AROP				
Oslo og Akershus	4.4	2.87	4.88	1.69
Hedmark og Oppland	3.1	3.67	3.63	4.51
Sør-Østlandet	3.3	2.14	8.31	2.51
Agder og Rogaland	3.6	2.32	6.14	2.76
Vestlandet	3.4	2.82	4.61	2.57
Trøndelag	3.8	5.50	4.46	3.37
Nord-Norge	3.8	3.53	5.34	3.54
GINI				
Oslo og Akershus	27.59	1.67	1.52	1.88
Hedmark og Oppland	22.93	3.65	3.44	4.43
Sør-Østlandet	23.02	1.72	20.19	2.24
Agder og Rogaland	23.96	1.40	4.58	2.33
Vestlandet	23.50	1.63	3.70	2.85
Trøndelag	24.54	2.65	3.82	3.03
Nord-Norge	24.89	1.98	4.39	2.83
S80/S20				
Oslo og Akershus	4.4	2.87	7.42	3.29
Hedmark og Oppland	3.1	3.67	5.96	4.92
Sør-Østlandet	3.3	2.14	5.41	3.09
Agder og Rogaland	3.6	2.32	7.02	3.93
Vestlandet	3.4	2.82	6.87	3.73
Trøndelag	3.8	5.50	8.04	4.65
Nord-Norge	3.8	3.53	7.81	3.76

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:
– by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),
– at the following standard number: +32 22999696 or
– by email via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications at: https://op.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

EU law and related documents

For access to legal information from the EU, including all EU law since 1952 in all the official language versions, go to EUR-Lex at: http://eur-lex.europa.eu

Open data from the EU

The EU Open Data Portal (http://data.europa.eu/euodp/en) provides access to datasets from the EU. Data can be downloaded and reused for free, for both commercial and non-commercial purposes.

PDF: KS-TC-20-010-EN-N

Measuring and communicating uncertainty of poverty indicators at regional level

During the last years, there has been an increasing interest in providing uncertainty measures for poverty indicators both at national and regional level. After reviewing the literature this report presents standard error estimation results at regional level for official income-poverty and income-inequality poverty indicators for eight selected European countries which vary in the sample design structure from a simple random sampling in one stage to a stratified two-stage random sampling.

For more information
https://ec.europa.eu/eurostat/

Publications Office
of the European Union

PDF: ISBN 978-92-28361-4