How precise are poverty measures estimated at the regional level?

Author's Accepted Manuscript

www.elsevier.com

How precise are poverty measures estimated at the regional level?

Vijay Verma, Achille Lemmi, Gianni Betti, Francesca Gagliardi, Mario Piacentini

PII: S0166-0462(16)30045-X
DOI: http://dx.doi.org/10.1016/j.regsciurbeco.2017.06.007
Reference: REGEC3275

To appear in: *Regional Science and Urban Economics*

Received date: 8 June 2016
Revised date: 30 March 2017
Accepted date: 29 June 2017

Cite this article as: Vijay Verma, Achille Lemmi, Gianni Betti, Francesca Gagliardi and Mario Piacentini, How precise are poverty measures estimated a the regional level?, *Regional Science and Urban Economics* http://dx.doi.org/10.1016/j.regsciurbeco.2017.06.007

How precise are poverty measures estimated at the regional level?

Vijay Verma[1], Achille Lemmi[1], Gianni Betti[1*], Francesca Gagliardi[1], Mario Piacentini[2]

[1]Department of Economics and Statistics, University of Siena, Siena, Italy

[2]Statistics Directorate, OECD, Paris, France

[*]Corresponding author: gianni.betti@unisi.it

Abstract

There is an urgent policy need for regional (subnational) estimates for assessing regional policies and programmes. Often regional indicators, in particular those concerning poverty and social exclusion, have to be derived from surveys with sample size and design determined primarily to serve estimation at the national level. In the specific context of EU-SILC surveys and the Headline Indicator at-risk-of-poverty or social exclusion (AROPE) and its components defined by European Commission, this paper aims to contribute to the methodology for constructing such indicators at the regional level. The main difficulty arises from the smallness of regional samples in national surveys. The paper focuses on two related issues: identifying procedures potentially useful for improving sampling precision of regional estimates; and improving the precision of sampling error estimates of regional statistics based on small but complex samples. In addition to some results presented for a large number of OECD countries, more detailed numerical illustration is provided for two countries (Austria and Spain) based on EU-SILC data.

Keywords: poverty, small area estimation, region, SILC

1. Introduction

In the framework of Europe 2020 strategy of the European Commission, the EU-SILC Headline Indicator at-risk-of-poverty or social exclusion (AROPE) and its components[1] will be included in the budgeting of structural funds which are one of the main instruments for attaining EU and national policy targets. In this context, DG Regional Policy of the European Commission uses regional, i.e. sub-national, level data (NUTS 2, and exceptionally NUTS

[1] EU Statistics on Income and Living Conditions (EU-SILC) is an EU survey aiming at collecting timely (every year) and comparable cross-sectional and longitudinal multidimensional micro data on income, poverty, social exclusion and living conditions. The headline indicator 'people at risk of poverty or social exclusion', consists of the three sub-indicators: monetary poverty, severe material deprivation, and very low work intensity.

1 for a couple of big countries[2]) for the social headline indicators which complement GDP per capita, as recommended by the GDP and Beyond report. For the funding period 2014-2020 these indicators are being used for benchmarking and assessing the efficiency of regional policies and programmes. Therefore, there is an urgent policy need for regional estimates of social policy indicators. The focus has to be on reliably identifying regions with the highest proportion of poor or socially excluded people so as to be able to target policy measures accordingly.

The paper aims to contribute to the methodology for constructing indicators of poverty and social exclusion at the regional (sub-national) level in EU countries.

EU-SILC constitutes the main, practically the unique, data source for constructing AROPE and other indicators of poverty and social exclusion in the multi-country comparative context of the EU. However, EU-SILC, like most other complex population-based surveys, is primarily designed to be represented at the country (rather than at the subnational or regional) level. For one thing, in most countries sample sizes of the EU-SILC survey are too small for directly constructing reliable estimates at the regional level. Special methodologies are involved in producing reliable indicators at the regional level and in estimating the degree of reliability (sampling precision) of the indicators obtained.

This methodological paper focuses on two related issues.

The first objective is to identify procedures potentially useful for improving sampling precision of regional estimates, specifically the headline indicator AROPE and its components using EU-SILC data (Section 2). We identify the following three types of procedures, each with a couple of specific techniques which can help ameliorate the problem of small sample sizes of individual (annual EU-SILC) surveys and yield regional estimates with reduced sampling error:

(1) Improved size, allocation and design of the sample

- Increased total sample size and its disproportional allocation in favour of smaller regions
- Taking regions as not only the reporting but also the *design* domains for the national sample

(2) Improved estimates with techniques using auxiliary information

- Small area (or domain) estimation (SAE)
- Calibration

(3) Adjusting the reporting requirements to be less demanding in certain respects. This involves averaging over time and/or space, such as:

- Cumulation of data or estimates over time, such as over annual waves of a EU-SILC survey

[2] NUTS (*Nomenclature des unités territoriales statistiques*) is a hierarchical system for dividing up the economic territory of the EU.

- Constructing fewer, larger reporting domains by grouping small regions.

The second issue we focus on is the problem that, because the available sample sizes are small, sampling error tends not only to be high, but also estimates of sampling error tend to be complex and subject to high levels of variability. In Section 3 we discuss techniques which can help to improve the precision of sampling error estimates of statistics based on small but complex samples. We identify the following three problem areas, with a couple of specific aspects of each:

(1) Practical variance estimation techniques suited for complex statistics from a complex sample of reasonably large size

- Relative merits and limitations of different variance estimation procedures, including a comparison between linearization and repeated-replication techniques, in particular jackknife repeated replication (JRR)
- Special issues involved in variance estimation of statistics based on cumulation of correlated samples, such as over annual waves of EU-SILC

(2) Special variance estimation problems arising from limitations of the information provided in EU-SILC public-use micro data files

- Lack of information on sample structure in EU-SILC data sets available in the public domain
- On the specification of sample structure variables for computation of sampling errors
- Absence on information for the linking of units, i.e. of same households or persons, across survey waves in the publically available cross-sectional data sets (as distinct from EU-SILC longitudinal data sets where such linkage is possible)

(3) Additional problems arising in applying the country-level variance estimation procedures to the regional level where generally sample sizes are much smaller

- Decomposition of total variance into components: firstly, this involves decomposition into variance under (hypothetical) simple random sampling, and design effect accounting for the effect on variance of complexity (departures from simple random sampling) of the sample design; next, design effect itself can be decomposed according to the aspect of the design giving rise to it
- Separation of components into two types: components which can be easily and directly estimated at the regional level despite the small sample sizes involved; and components which cannot be reliably estimated directly at the regional level but have to be derived ('ported') or even borrowed (copied) from the corresponding estimates from the bigger national sample.

This decomposition of the variance and design effect into components, as it will be described in Section 3.3, is necessary in order to obtain reasonable estimates of variance at regional level. Some component can be computed at the regional level, but others have to be inferred from aggregated or averaged results.

Section 4 illustrates application of the procedures discussed. Numerical results are presented using EU-SILC for Spain and Austria, using pooled data over four annual waves of the rotating panel in each case. The illustration is confined to only two among all the EU-SILC surveys because the objective is mainly to illustrate the methodology developed in this paper. The practical factor limiting the choice to these two particular countries is the availability in these cases of the necessary information on the sample structure, which is generally not available in the EU-SILC public-use data files.

2. Techniques to ameliorate the problem of small sample sizes

This section identifies three types of procedure, each with a couple specific techniques which can help ameliorate the problem of small sample sizes of individual (annual EU-SILC) surveys and yield regional estimates with reduced sampling error.

2.1 Improved size, allocation and design of the sample

(1) Increased total sample size and its disproportional allocation in favour of smaller regions

A general increase of the overall sample size, retaining existing allocation, could be considered as a solution. The main limitation of this procedure is the increased cost and burden of the survey at the national level, and possible negative effect on overall data quality. The real difficulty is that in order to obtain adequate sample sizes for the smallest regions, the increase in overall sample size, while retaining proportionate allocation, is often too large to be a practical option. Of course, it may be a feasible option when the regional population sizes are not too diverse.

An alternative is to re-allocate the sample among the regions. This could be done at regional, say NUTS2, level in two manners: redistribution the existing sample by reducing the sample size of larger regions and increasing that of the smaller regions; or just giving extra sample to smaller regions, without reducing that of larger regions. The first scheme does not increase overall data-collection cost, though there is a (usually) modest reduction in precision of national level estimates as a result of disproportionate allocation. The second scheme can marginally increase the precision of national level estimates, but the real gain is at the level of small regions. Of course, the overall data collection cost is increased, but the increase is normally much less than that with proportional increase in sample size throughout.

(2) Taking regions as not only the sample selection but also the design domains for the national sample

Estimation at the regional level is facilitated by selecting the sample independently within regions, i.e. by taking regions as sample selection domains for the national sample. This ensures that strata and primary sampling units (PSUs) do not cut across regional boundaries. Stratification is performed within regions. In this way, strata do not cut across regions. This requirement is often not difficult to meet, because geographical stratification such as by regions is one of the most common stratification criteria used in national household surveys. The same would apply to PSUs, which are normally defined to lie within strata, and hence within regions as the sample selection domains.

Sampling independently within regions makes it easier to allocate the sample size as desired. Another major advantage of independent regional samples is that it simplifies the construction of regional estimates – estimation of parameters as well as those of their sampling error. Commonly used variance estimation techniques (such as Taylor linearization, bootstrap, JRR) can be easily adapted for application at the regional level.

In fact, it is highly desirable to go beyond independent selection of regional samples, and treat regions as *sample design domains*. All aspects of the sample structure – sampling stages and clustering, sampling units, stratification, selection method, sample allocation, etc. – can be varied across regions to take into account specific conditions and requirements.

2.2 Improved estimates with techniques using auxiliary information

(1) Small area (or domain) estimation (SAE): Empirical Best Linear Unbiased Predictor

There is a wide variety of small area estimation (SAE) techniques available, and the field is rapidly expanding. The suitability and efficiency of a particular technique depends on the specific situation and on the nature of the statistical data available for the purpose. A standard reference on small area estimation methodology is Rao (2003). See also, among others, Gosh and Rao (1994) and Handerson (1950). It is, of course, not possible in this paper to develop and evaluate SAE models for diverse poverty and related indicators in the specific situation of individual EU countries. Among the various SAE techniques, the most used approach is the Empirical Best Linear Unbiased Predictor (EBLUP) estimator. [3]

The Empirical Best Linear Unbiased Predictor

In the EBLUP methodology, an intensive and small-scale survey (such as, for instance, the SILC for EU countries) provides direct poverty-related information at the micro (unit) level; this information can be aggregated to areas such as NUTS regions where the survey contains an "adequate" number of sample units and the areas identifiers are available in the micro data. On the other side, correlates of poverty-related characteristics of the areas can come from aggregated statistics (such as censuses or other administrative sources). The two sources can be combined to produce composite estimates, provided that (i) the survey data contain information for the

[3] Elbers, Lanjouw and Lanjouw (2003) developed an alternative approach, specifically aimed at 'poverty mapping'. See an application in Neri, Ballini and Betti (2005).

identification of the area to which each unit belongs, and (ii) the aggregate data on the correlates are available for all the areas in the population of interest. The approach can be to apply area level random-effect models relating small area direct estimates to domain specific covariates, considering the random area effects as independent. The basic area-level model includes random area specific effects, and the area specific covariates, $x_i = \left(x_{i,1}, x_{i,2}, \dots, x_{i,p}\right)$, i = 1…m, are related to the target parameters θ_i (totals, means, proportion, etc.) as $\theta_i = x_i \ \beta + z_i v_i$, where z_i are known positive constants, β is the regression parameters vector of dimension *px1*, v_i are independent and identically distributed random variables with 0 mean and variance $\hat{\sigma}_v^2$. The model assumes that the direct estimators $\hat{\theta}_i$ are available and design unbiased, in the form:

$\hat{\theta}_i = \theta_i + e_i$, where e_i are independent sampling errors, with zero mean and known variance ψ_i. The BLUP estimator is a weighted average of the design-based estimator and the regression synthetic estimator $\tilde{\theta}_i \ (\sigma_v^2) = \gamma_i \hat{\theta}_i + (1 - \gamma_i) x_i \ \hat{\beta}$, where $\gamma_i = \ \sigma_v^2 / (\psi_i + \sigma_v^2)$ is a weight (or 'shrinkage factor') which assumes values in the range [0-1]. This parameter measures the uncertainty σ_v^2 in modelling θ_i in relation to the total uncertainty including the variance of the direct estimator φ_i. The mean square error of the BLUP estimator depends on the variance parameter σ_v^2, which in practice is replaced by its estimator; hence the estimator obtained is called Empirical BLUP (EBLUP).

Merits and limitations of EBLUP

There are some serious limitations to the application of the SAE methodology described above in the context of regional estimation in EU-SILC. But let us first note some potential merits of the procedure (see Betti and Lemmi, 2013). SAE methods such as EBLUP make use of external data aggregated to NUTS2 (area level) only. R-codes are available under projects funded by the EU 7th Framework program (such as SAMPLE, AMELIE, etc…); estimates could be performed every year, given that such external sources are available.

There are three types of limitation to be faced.

(a) The first concern the lack of external data for the purpose of making SAEs. The methodology needs information from Census data, which are usually available every ten years in many countries. Often such external sources are not correlated sufficiently highly to the poverty measures under investigation. Also, most of the models assume the external data to be error-free, which is certainly not the case when the data come from large-scale field studies and surveys.

(b) The methodology tends to be complex and require specialised knowledge and software.

(c) The major concern in application to a multi-country undertaking such as EU-SILC is that the results may lack comparability. Generally, the procedures and application would have to be country-specific, and ensuring the application of common standards required for EU-SILC may be very difficult.

(2) Calibration

Calibration is another methodology that is useful for increasing efficiency through sample reweighting. The idea is summarised from Ardilly (2015).

Consider a sample s, the associated unbiased weights d_k and an auxiliary set of variables X_k known for every unit in frame U, so that $\sum_{k \in U} X_k$ is known. The aim is to seek new weights w_k as close as possible to d_k, so that the following calibration equation is valid:

$$\sum_{k \in s} w_k X_k = \sum_{k \in U} X_k$$

To do this the distant function $\sum_k D(w_k, d_k)$ is used and minimised under the above constrain.

The calibration result has two fundamental properties: the resulting calibrated estimator $\sum_{k \in s} w_k Y_k$ has no significative bias if the sample size is large. Calibrated estimator closely depends on the linear correlation between X_k and Y_k ; if it is high we can expect a large decrease in the variance compared to the variance of the initial unbiased estimator $\sum_{k \in s} d_k Y_k$. The method proposed by Ardilly marginalised the calibrated weights, to get calibrated weights for each region. The author notes that "the method is based essentially on: 1) the existence of 'enough explanatory' variables of the phenomena of interest and for which we can have regional margins; 2) the hypothesis that conditionally in these explanatory variables, the geography does not have impact anymore on the phenomenon of interest."

Calibration and weighting at the level of individual regions can improve the precision (reduced variance and/or reduced bias) of regional estimates. However, from the point of view of country-level estimation, the introduction of additional regional level calibration controls can be problematic. For instance, it may result in more extreme weights and hence an increased variance. More severe trimming of weights may have to be introduced to control this increase. Adding additional regional level constrains in the calibration process can even lead to non-convergence problems in the process itself.

2.3 Adjusting the reporting requirements to be less demanding in certain respects

This involves reducing the detail with which the results are reported: whether through cumulation (1) over time, and/or (2) over space.

(1) Cumulation of data or estimates over time, such as over annual waves of a EU-SILC survey

Cumulating data from consecutive waves of a panel survey is perhaps the most effective and tractable (practical) strategy for increasing the statistical precision of regional indicators. Below we briefly address some statistical aspects of the cumulation method for improving sampling precision of indicators for subnational regions (see also Verma et al., 2013). Two types of measure can be so constructed at the regional level by aggregating information on individual elementary units: average measures such as totals, means, rates and proportions constructed by aggregating or averaging individual values; and distributional measures, such as measures of variation or dispersion among households and persons in the region. An important point to note is that, more than at the national level, many measures of averages can also serve as indicators of disparity and deprivation when seen in the regional context: the dispersion of regional means is of direct relevance in the identification of geographical disparity.

Survey data such as from EU-SILC can be used in different forms or manners to construct regional indicators. When two or more data sources contain – for the same type of units such as households or persons – a set of variables measured in a comparable way, then the information may be pooled either (a) by combining estimates from the different sources, or (b) by pooling data at the micro level. Technical details and relative efficiencies of the procedures depend on the situation. The two approaches may give numerically identical results, or the one or the other may provide more accurate estimates; in certain cases, only one of the two approaches may be appropriate or feasible in any case.

Here our concern is with pooling of different sources pertaining to the same population or largely overlapping and similar populations. In particular, the interest is in pooling over survey waves in a national survey in order to increase the precision of regional estimates. Estimates from samples from the same population are most efficiently pooled with weights in proportion to their variances (meaning, with similar designs, in direct proportion to their sample sizes). Alternatively, the samples may be pooled at the micro level, with unit weights inversely proportional to their probabilities of appearing in any of the samples. This latter procedure may be more efficient (e.g., O'Muircheataigh and Pedlow, 2002), but may be impossible to apply as it requires information, for every unit in the pooled sample, on its probability of selection into each of the samples irrespective of whether or not the unit actually appears in the particular sample (Wells, 1998). Another serious difficulty in pooling samples is that, in the presence of complex sampling designs, the structure of the resulting pooled sample can become too complex or even unknown to permit proper variance estimation. In any case, different waves of a survey like EU-SILC do not correspond to exactly the same population. The problem is akin to that of combining samples selected from multiple frames, for which it has been noted that micro level pooling is generally not the most efficient method (Lohr and Rao, 1996). For the above reasons, pooling of wave-specific estimates rather than of micro data sets is generally the appropriate approach to aggregation over time from surveys such as EU-SILC.

Merits of the cumulation include the following: the implementation of cumulation with data such as EU-SILC data is quite easy to implement, especially for average measures; it is also possible to use it for distributional measures; the procedure has been already implemented in several EU and other countries.

The main limits of the cumulation are: yearly estimates of any variable would tend to be highly positively correlated, reducing the gain in precision from cumulation[4]. For instance, with a correlation of 0.8 between successive waves (and say 0.64 between two waves separated by one wave in-between), and a wave-to-wave sample overlap of 0.75 as in standard EU-SILC,[5] the variance of an estimate cumulated over three years is reduced to roughly 0.60 of that with single year data, rather than to 0.33 in the case of independent samples[6].

(2) Constructing fewer, larger reporting domains by grouping small regions

A grouping of regions could be applied in order to achieve more useful regional estimates. Estimates for a group are then used to represent the average conditions in the regions so grouped. As a rule, the procedure should be applied only to smallest regions, where essential and where separate reporting can be dispensed with, or is simply unachievable. For instance, this has been done in EU-SILC in Finland for the smallest region, which in fact is exceptionally small in population. This possibility is of course limited if there are many small regions so that the needed regional detail is largely lost as a result of grouping. Grouping of small regions, if permissible, does not require the overall sample size to be increased, and hence no extra sources are needed, but obviously, the disaggregation to full regional level is lost.

3. Issues in estimating sampling error at the regional level

This section discusses techniques which can help to improve the precision of sampling error estimates of regional statistics based on small but complex samples. First, it is useful to outline the basic variance estimation techniques suited for complex statistics from complex sample of reasonably large size.

3.1 Practical variance estimation techniques

(1) Relative merits and limitations of different variance estimation procedures

A diversity of variance estimation approaches has been developed for computation of sampling errors for complex statistics arising from complex samples of households and persons with complex designs, both general and some very specific for particular applications. Among the former, for the 'typical' social surveys based on reasonably large samples but with complex designs, the applicability of at least two broad approaches is generally well-established in the literature. These are the approaches based on (a) Taylor linearization and (b) on resampling such as the Bootstrap, Balanced repeated replication (BRR) and Jackknife repeated replication (JRR). For details on the two methodologies as applied to poverty and inequality measures, Taylor linearization and JRR, see Verma and Betti (2011). As noted by the mentioned authors, the two methods tend to give similar

[4] The cumulative measure is an average over the period of cumulation, and in this sense is not biased in itself.
[5] The design recommended by Eurostat has been developed and described in Verma and Betti (2006).
[6] The procedure for arriving at these figures has been described in Betti and Gagliardi (2017).

results and often the choice among them is dictated by practical considerations. Other recent contributions on variance estimation based on EU-SILC survey include Osier (2009), Osier et al. (2013), Goedemé (2013a, b).

The following provides a comparison between linearization and repeated-replication techniques, in particular JRR.

Linearization tends to be lighter in terms of the computational work involved, especially if the sample contains many PSUs. The main drawback of the linearization procedure is that different computational formulae have to be used for different types of statistics, the development of which can be complex and – what is more critical – may be intractable in some cases. Another complexity is the need for numerical evaluation of the slope at various points of the income distribution function. This can be problematic because of irregularities in the empirical income distribution based on sample survey data.

The JRR procedure, by comparison, is considerably heavier in terms of the computational work involved, especially if the sample contains many PSUs, though this disadvantage is becoming less important given the rapid increase in computing power. In addition, care is needed in the application of the JRR method to non-smooth statistics, such as the median or other quantiles of the income distribution, as it may not always provide a consistent variance estimator for such statistics. The same applies, but to a lesser degree, with regard to measures based on quantiles, such as the poverty rate defined with reference to the median income. For this reason, as well as to reduce computational work, the grouping of existing PSUs and strata is often needed to define new computational units. The relative advantages of the JRR method include the following. (1) The same variance estimation formula applies to different types of statistics. This permits the development of highly standardized software for its application. (2) The method can be extended to take into account the effect on variance of aspects such as imputation and weighting in the estimation process, insofar as those aspects can be repeated for each replication. (3) JRR methodology can be easily extended to longitudinal samples and to measures of change between different cross-sections. Once a common structure (PSUs and strata) is defined for all waves of a data set, the JRR can be used to estimate variance of any measure that incorporate measures from different waves. Incorporating such additional aspects into the variance estimation is generally much more complex, sometimes not possible, with the linearization technique. Extensions of the linearization method to deal with added complexity involved (longitudinal measures, aggregate measures of net changes and averages) are not available, but a replication-based approach such as the JRR may be extended more readily to deal with correlated samples. (4) Generally JRR procedure is well suited and stable to deal with cumulative measures such as Gini coefficient and percentile share ratios such as S80/S20. Finally, it should be noted that a very extensive use of JRR methodology is done at EU and OECD official level and also at single country level.

Considering the above merits, the JRR provides an attractive methodology for the present purpose.

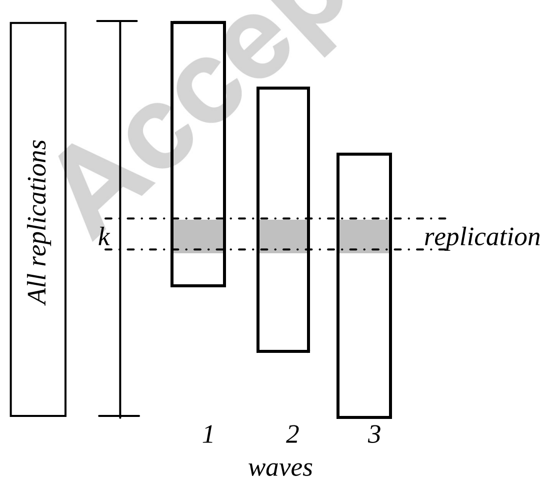
(2) Special issues involved in variance estimation of statistics based on cumulation of correlated samples

This section addresses some issues concerning pooling over correlated samples, such as over annual waves of longitudinal component of a EU-SILC sample. Let us consider variance estimation under three-year cumulation.

Consider that for each wave, a person's poverty status (poor or non-poor) is determined based on the income distribution of that wave separately, and the proportion of poor at each wave is computed. These proportions are then averaged over a number of consecutive waves. The issue is to quantify the gain in sampling precision from such pooling, given that data from different waves of a rotational panel are highly correlated. For this purpose, the JRR variance estimation methodology is very convenient. In brief, for the purpose computing variance of estimates pooled over waves of a rotational panel, the JRR methodology can be easily extended on the following lines. The total sample of interest is formed by the union of all the cross-sectional samples being compared or aggregated. Using as basis the common structure of this total sample, a set of JRR replications is defined in the usual way. Each replication is formed such that when a unit is to be excluded in its construction, it is excluded simultaneously from every wave where the unit appears. For each replication, the required measure is constructed for each of the cross-sectional samples involved, and these measures are used to obtain the required averaged measure for the replication, from which variance is then estimated in the usual way.

As an example, suppose we have a three consecutive year's dataset. In order to estimate the average of the three years, we proceed as follows.

[reformat the rest of the text to conform to the above paragraph format]



We first construct a common structure of strata and PSUs from the union of the sample for the three waves and assign to this common structure new weights equal to the average of the weights of the three years:

$w_t^{(Common)} = (w_t\ )^{Average} = (w_1 + w_2 + w_3)/3$.

For each year (t) and for each replication (k), we can estimate $y_k^{(t)}$ where t=1, 2, 3 and from this, the required statistic $y_k^{Average} = \sum_t a_t y_k^{(t)}$; that in our case is just $y_k^{Average} = (y_k^{(1)} + y_k^{(2)} + y_k^{(3)})/3$.

The variance estimate of this measure can be easily estimated applying the usual JRR for variance estimation as if the statistic were a common cross-sectional measure.

3.2 Special variance estimation problems arising from limitations of the information provided in EU-SILC public-use micro data files

(1) Lack of information on sample structure in EU-SILC data sets available in the public domain

Appropriate coding of the sample structure, in the survey micro-data and accompanying documentation, is an essential requirement for computing sampling errors taking into account the actual sample design. This is done through the definition of the sample structure. Information concerning the following three aspects must be available:

(1) Codes of the sample structure in the micro-data files.

(2) Detailed description of the sample design, for instance identifying features such as the presence of self-representing units, systematic selection, etc.

(3) Information connecting the sample structure codes in the micro-data with descriptions of the particular sample design features, so as to be able to identify the design features applicable to particular units.

Lack of information on the sample structure in survey data files is a long-standing and persistent problem in survey work. For EU-SILC, currently this information is not provided for all countries in the micro-data available to researchers. Presumably (and hopefully) it is available within each country for its own national survey. Indeed, the major problem in computing sampling errors for EU-SILC is the lack of sufficient information for this purpose. Actually the situation differs between the two versions of the dataset available for EU-SILC, namely the 'cross-sectional dataset' (providing separate information for each annual wave of the survey), and the 'longitudinal dataset' (for panel of individuals covering two to four survey waves). In the EU-SILC longitudinal dataset we have the necessary information available to construct the common structure; however, information has not been included in the dataset to enable us to construct the common sample structure in the cross-sectional case (households and individuals are not linkable in different years). The main variables necessary concern the identification of strata and PSUs. The variables corresponding to the stratification code have been suppressed in the publically available dataset. We have had to develop approximate procedures in order to overcome these limitations at least partially, and used them to produce useful estimates of sampling errors.

(2) On the specification of sample structure variables for computation of sampling errors

For the type of sample designs involved in EU-SILC, and in the practical procedures for variance estimation used, generally *all the necessary information about the sample structure can be provided, in addition to unit sample weights, in the form of two variables defined for each unit*: the 'computational stratum' and 'computational PSU' to which the unit belongs. Normally, we may expect the new variable 'computational stratum' to be related (and sometimes identical) to the explicit stratum in the survey; similarly concerning 'computational PSU' and explicit PSU in the survey. *However, very often the explicit variables reported in the survey require some redefinition before they can be used for the purpose of variance estimation.* Practical variance estimation methods need to make some basic assumptions about the sample design: the sample selection is independent between strata; two or more primary selections units are drawn from each stratum; these primary selections units are drawn at random, independently and with replacement; and the number of primary selections units is large enough for valid use of the in the variance estimation equations. The computational strata and PSUs have to be defined to meet, or at least approximate, these requirements.

In many practical situations, some aspects of sample structure need to be redefined to make variance computation possible, efficient and stable. *This requires a sufficiently detailed description of the sampling design, and on how that description relates to stratum and PSU codes in the data – information which is often lacking.* There are two major considerations involved in the redefinition of sample structure for the purpose of variance computation:

(i) The first concerns ensuring that each stratum has at least two sample PSU's – the minimum number required for the computation of variance.

It may be necessary to regroup ('collapse') strata so as to ensure this. In samples selected systematically, for instance, the implied implicit stratification is often used to define explicit strata formed by pairing or otherwise grouping of PSU's in the order of their selection from the systematic list, ensuring that each resulting computational stratum has at least two primary selections, assumed independent. A similar consideration applies to designs involving the selection of a single PSU per stratum. In addition, sometimes non-response can result in leaving fewer than two PSU's in some strata.

(ii) The second aspect concerns the aggregation of very small sample PSUs to construct larger and fewer computational units for variance estimation.

In a procedure like the JRR, the number of replications is equal or at least similar to the number of PSU's in the sample. In a large sample where elements (households, persons) have been selected directly, the number of replications which can be formed will be of the order of the sample size, normally running into thousands. This necessitates forming much fewer computational units, such as by creating 'pseudo-cluster' from random groupings of sample elements, and then randomly pairing of these 'clusters' to construct computational strata.

The above issue in fact arises in applying a procedure such as JRR in the case of any sample irrespective of its structure when we want to estimate not only variances but also design effects. The denominator of the design effect is variance under simple random sampling (SRS); that variance is estimated by assuming the sample to be SRS, which in turn necessitates forming fewer computational units from random regrouping of sample elements.

Sometimes there are additional consideration for using groupings of strata and PSUs. For instance, such grouping can help in preserving confidentiality of individual units in public-release micro data.

The above-mentioned problem arises more frequently and seriously when computing sampling errors for subclasses (subpopulations), especially for regions and other geographic subdivisions. The risk can be reduced by aggregating PSU's and strata to create fewer, larger computational units.

(3) Absence on information for the linking of units across cross-sectional samples

We need to address a special problem in EU-SILC that there is no information for the linking of units, i.e. of same households or persons, across survey waves in the publically available cross-sectional data sets (as distinct from longitudinal data sets where such linkage is possible).

Our proposed solution is to pass through the longitudinal dataset of EU-SILC to get an imputed measure of the correlation between the cross-sectional datasets. For example, if we want to produce estimates for the average of three years, we keep the longitudinal dataset for three years, say year 1, 2 and 3. We work separately with pairs of consecutive years, such as 1-2, 2-3 and 1-3. For each couple of years, the two datasets can be divided into three part: a) same individuals; b) same sample area, but different individuals; c) different individuals and different area. We have to work separately with each of these parts.

For part a) of the sample, same individuals, we apply exactly the same procedure as applies to the cross-sectional dataset and estimate: the variance of the first year $V(Y_t)$, the variance of the second year $V(Y_{t+1})$ and the variance of their average $V\left(\frac{Y_t+Y_{t+1}}{2}\right)$. From the relationship:

$$V\left(\frac{Y_t+Y_{t+1}}{2}\right) = \frac{1}{4}\left(V(Y_t) + V(Y_{t+1}) + 2\rho\sqrt{(V(Y_t)V(Y_{t+1}))}\right) \quad (2)$$

we can derive the correlation ρ, estimated for the three pairs of years, being, respectively, ρ_{12}, ρ_{23} and ρ_{13}.

For the correlation between adjacent years, a more stable estimate is provided by taking the average of the two estimates, $\rho_{adj}^{L} = \frac{\rho_{12}+\rho_{23}}{2}$ where L stands for longitudinal. For two waves separated by a year, we have $\rho_{13}^{L} = \rho_{13}$.

For part b) of the sample, same area but different individuals, again we apply the same procedure as above and we get $\rho_{12}^{A}, \rho_{23}^{A}, \rho_{13}^{A}$, where A stands for Area level. We use their average $\rho^{A} = \frac{\rho_{12}^{A}+\rho_{23}^{A}+\rho_{13}^{A}}{3}$ as a more stable estimate of the correlation at area level.[7]

For part c) of the sample, no overlap even at area level, we can take correlation between years to be zero.

Now we have all the information needed from the longitudinal dataset: ρ_{adj}^{L}, ρ_{13}^{L} and ρ^{A}.

We can return to the cross-sectional datasets, three sets in our case, and estimate at national level the three variances $V_1(Y)$, $V_2(Y)$, $V_3(Y)$. The general expression for variance of measures of average and net change taking into account the correlation between waves is:

$$V\left(\sum_i a_i Y_i\right) = \sum_i a_i^2 V(Y_i) + 2\sum_{j>i} a_j \rho_{ij}^{CS} \sqrt{V(Y_i)V(Y_j)} \qquad (3)$$

with parameters a_i=1/3 for the average measure.[8]

The main problem is to estimate ρ_{ij}^{CS} between the correlated cross-sectional samples from the correlations for the three types of sample overlap, a), b) and c), defined above. The former can be taken as a weighted sum of the latter since it is a mixture of the three types of overlap. With reference to a particular pair of waves (i, j), we take:

$$\rho^{CS} = w^{L}\rho_{adj}^{L} + w^{A}\rho^{A} + w_0 0 \qquad (4)$$

(for simplicity, subscript i,j has been dropped in the above).

The weights are determined by the following considerations:

$w^{L} + w^{A} + w_0 = 1$, since they define a weighted average;

$w_0 = \frac{1}{4}$ from the structure of EU-SILC;

$w^{L} = \frac{n_L}{n_c}$ with n_L is the sample size of the panel considered and n_c is the larger of the samples of the two cross-sectional dataset considered.

With the above-defined set of correlations and weights, the set of ρ_{ij} in equation (3) are defined as follows:

$\rho_{12} = (\rho_{adj}^{L} w_{12}^{L} + \rho^{A} w_{12}^{A})$, $\rho_{23} = (\rho_{adj}^{L} w_{23}^{L} + \rho^{A} w_{23}^{A})$, $\rho_{13} = (\rho_{13}^{L} w_{13}^{L} + \rho^{A} w_{13}^{A})$

3.3 Additional problems arising in estimating variances at the regional level

[7] The EU-SILC 'longitudinal dataset' is in fact more inclusive than covering only a fixed panel of individuals. It also covers all current household members of the panel individuals, and hence a changing set of persons in the same sample areas. This provides a basis for estimating ρ^{A} .

[8] For net change from year i to year (i+1), the corresponding coefficient would be a_i=1, a_{i+1}=-1.

Additional problems arise in applying the country-level variance estimation procedures to the regional level where generally sample sizes are much smaller. The approach involves two steps. We begin with decomposition of total variance into components. Firstly, this involves decomposition into variance under (hypothetical) simple random sampling and design effect accounting for the effect on variance of complexity (departures from simple random sampling) of the sample design, followed by decomposition of design effect itself according to the aspect of the design giving rise to it. The second step involves separation of components into two types: components which can be easily and directly estimated at the regional level despite the small sample sizes involved; and components which cannot be reliably estimated directly at the regional level but have to be derived ('ported') or even borrowed (copied) from the corresponding estimates from the bigger national sample or estimates averaged over regions.

(1) Decomposition of total variance into components

Variance estimates at regional level can be obtained directly by repeating the same procedure described above at regional level, or they can be constructed by borrowing some components from estimates from the national sample or estimates averaged over regions. Below we describe the second option. It involves decomposition of variance and design effects (see Verma et al., 2010).

From Kish (1965), the design effect is the ratio of the variance under the given actual sample design (V), to the variance under a simple random sample of the same size (V_{SRS}): $d^2=V/V_{SRS}$. This implies decomposition of variance into SRS variance and design effect: $V=V_{SRS}*d^2$.

At country level (C) the total design effect can be decomposed as follows

$$d^{2(C)} = d_W^{2(C)} \cdot d_H^{2(C)} \cdot d_D^{2(C)} \cdot d_X^{2(C)} \cdot d_R^{2(C)}$$

$d_W^{(C)}$ is the effect of unequal sample weights, which we have termed as the “Kish factor”. It can be approximated in terms of coefficient of variation of the sample weights: $d_W^{(C)\,2}$ = $(1+cv(w)^2)$

$d_H^{(C)}$ is the effect of clustering of persons within households and the quantity $d_H^{(C)\,2}$ can be estimated as the weighted mean household size.

$d_D^{(C)}$ is the effect of clustering of persons and households within dwellings. The effect of clustering of households within dwellings or addresses is absent (d_D =1) when we have a direct sample households or persons, or when such units are selected directly within sample areas - as is the case in most of the EU-SILC surveys. This effect is present when the ultimate units are dwellings, some of which may contain multiple households, but the effect tends to be small (i.e. $d_D \sim 1.0$) in so far as there is generally a one-to-one correspondence between addresses and households.

$d_R^{(C)}$ is one of the components of the design effect and it is the effect of correlation between dependent samples of waves in a panel. It can be calculated as:

$$d_R^{2(C)} = \frac{V\left(\sum a_i Y_i\right)}{\sum a_i^2 V\left(Y_i\right)} \qquad (6)$$

This is a function of rohCS defined in equation (4): with similar variance across waves, d2(C)R=1+rohCS. We refer to rohCS as "Kish correlation" or synthetic correlation (Kish, 1965 and 1995); it represents the ratio of variance of correlated samples over variance as if the sample were independent.[9]

Factor $d_X^{(C)}$ is the effect of multi-stage sampling, stratification and other design complexities, averaged over the three waves. It can be computed at the country level using the JRR procedure described above.

(2) 'Porting' estimates of some variance components from the country level to regional level

The next step is to separate variance components into two types: components which can be easily and directly estimated at the regional level despite the small sample sizes involved; and components which cannot be reliably estimated directly at the regional level but have to be derived ('ported') or even borrowed (copied) from the corresponding estimates from the bigger national sample.

The main difficulty in moving from national to regional level is the reduced sample size.

Our practical strategy in estimating variance and design effects at the regional level is to consider the different components involved (Verma *et al*., 2010):

$$V=V_{SRS}*d^{2(G)};\ \mathrm{d}^{2(G)}=\mathrm{d}_W^{2(G)}\times\mathrm{d}_H^{2(G)}\times\mathrm{d}_D^{2(G)}\times\mathrm{d}_X^{2(G)}\times\mathrm{d}_R^{2(G)} \qquad (7)$$

Quantities V_{SRS}, d_W, d_H and d_D do not depend on structure (especially clustering) of the sample, and can be normally estimated well from samples of elements (households or persons) at the regional level. With a national sample of several thousand units, for instance, most regional samples would contain hundreds of elements. The main difficulty arises with estimating quantities d_R and d_X directly and separately for each region. The reliability of the estimates depends on the structure of the sample, in particular on the number of sample PSUs available. This number may be small for individual regions: even when a national sample contains several hundred PSUs, most regional samples would contain merely tens of such units.

Concerning $d_R^{(G)}$, with the common structure of the panel in all regions of a country, it is reasonable to take $\mathrm{d}_R^{(G)} = \mathrm{d}_R^{(C)}$, the value already estimated at the country level on the lines described in the preceding subsection.

[9] There are alternative approaches; see for instance Tam (1984) or Berger (2004).

Concerning $d_X^{(G)}$, in many situations, depending on the relation between national and regional sample design, d_X^G can be reasonably defined as a function of already computed d_X^C, thus avoiding its re-estimation at the regional level. For instance, if the sample design in the region is the same or very similar to that for the country as a whole – which is quite often the case – we can take $d_X^{2(G)} = d_X^{2(C)}$. For models for other cases, see Verma *et al.* (2010); these models are essentially based on assumed identity or closeness of intra-class correlations between national and regional samples with similar designs.

4. Application of the procedure in Austria and Spain

The proposed procedure of variance estimation under three-year cumulation of SILC waves has been recently applied in a project funded by OECD (Piacentini, 2014). As shown in Table 1, poverty and inequality measures have been estimated at regional level in 28 OECD countries; in half of them the EU-SILC survey has been the main data source (countries in *italic*). The table lists the data sources used for the construction of indicators in OECD countries, and provides information on NUTS2 regions and minimum and maximum regional sample sizes.

and the described methodology has been implemented. Here results of the methodology are presented for two countries, Austria (AT) and Spain (ES) for three poverty and inequality measures: the at-risk-of poverty rate (ARPR) at 60% the national median income, the S80/S20 inequality index, and the Gini coefficient. The choice of these two countries has been determined by the fact that the necessary information on the sample structure was available to us only for those countries.

Table 2 provides a summary of the precision of the estimates. The results are based on the variance estimation methodology described in this paper, produced with the collaboration of the present authors. The table covers countries and regions for which it was possible to produce estimates of the variance for the three indicators: poverty rate, Gini and S80/S20. The table shows the median size of the 95% confidence intervals, both in relative and absolute terms, of regional estimates. For example, column 2 'relative poverty with poverty line defined at 60% the national median' shows that the difference between the upper and the lower bound of the confidence interval of the Australian estimates is 4.41 percentage points (as a median value across the 8 regions). Given that the absolute value of the size of the confidence interval is affected by the absolute value of the indicator (countries with higher poverty rates have, *ceteris paribus*, larger confidence interval for the poverty rate estimate), column 1 (and also 3 and 5) are reported as a way to improve the comparability of this summary measure across countries. Even for these relative summary measures, the cross-country comparability is possibly hampered by the fact that countries use different methodologies for calculating the confidence intervals.

As expected, the confidence intervals tend to be larger in countries where regional samples are smaller, such as Germany where in median the upper and lower limits of the confidence interval of the relative poverty rate at 60% the national median income are separated by 8.4 percentage points. For a similar annual sample size, the confidence intervals are smaller in Spain, where three-year averaged data were used, than in Italy, where the indicators refer to a single year. For the Gini index of disposable income, in around 7% of the regions the absolute size of the confidence interval (i.e. the difference between the upper and lower bound of the estimate) is higher than 10 percentage point; while for the poverty headcount based on the 60% the national median, the confidence interval is larger than 10 percentage points in around 10% of the regions.

Production of the estimates covered in Tables 1 and 2 was made possible by the necessary information on sample structures provided under special arrangements for the above-mentioned OECD-funded project, thus bypassing some of the difficulties noted in Section 3.2.

Table 3 presents more detailed results for individual regions. In contrast to the previous tables, the illustration in this table is confined to only two among all the EU-SILC surveys. The practical factor limiting the choice to these two particular countries is the general availability in these cases of the necessary information on the sample structure, which is not available in the EU-SILC public-use data files. In any case, this serves the current objective, which is mainly to illustrate the methodology developed in this paper. Numerical results are presented using EU-SILC for Spain and Austria, using pooled data over three annual waves of the rotating panel in each case, using the cumulation methodology detailed above. The table reports for the 'Headcount 60% national poverty line for disposable income', the point estimate and the lower and upper bound of its 95% confidence interval. Results are very good at national level, where the width of the confidence interval is only 1.6 for Austria and 1.8 for Spain. At regional level results are more heterogeneous, with a median width of the confidence interval of 4.2 in Spain and 5.2 in Austria. The minimum values are 1.9 in Pais Vasco and 2.8 in Oberösterreich, while the maximum are 10.7 in Aragón and 9.6 in Burgenland.

5. Concluding remarks

This paper has demonstrated how reasonable estimates of sampling error may be produced even when the available sample sizes are relatively small and full information on the sample structure is lacking. For estimation at the regional (subnational) level, the main approach proposed is to "borrow" parameters estimated from the bigger national sample, in so far as such parameters can be considered "portable". In particular, we have identified and used two parameters for this purpose: the intra-class correlation determined by clustering, stratification and other features of the sample structure; and the synthetic correlation coefficient determined by correlations among samples of waves of the survey.[10]

In conclusion we would like to emphasise a point of great practical concern. Assessment of sampling precision of the estimates, taking into account the actual structure of the sample on which the data are based, has an essential requirement: provision of codes describing the sample in the survey micro data itself, along with accompanying documentation describing the design and the code. Inadequate (or sometimes even absence of) information on sample structure in survey data files is a long-standing and persistent problem in estimation from sample surveys. Unfortunately, even outstanding and highly standardised multi-country surveys such as EU-SILC suffer from this sort of shortcomings, as demonstrated in the application in this paper.

Acknowledgements:

The estimates of income distribution at sub-national level for Austria and Spain were produced in the context of the OECD project "How's life in your region: Measuring regional and local well-being for policy making" carried out by the OECD Public Governance and Territorial Development under the guidance of the Working Party on Territorial Indicators (WPTI).

References

Ardilly P. (2015), Regional estimates of poverty indicators based on a calibration technique. *Eurostat Statistical Working Papers*, Luxembourg.

Berger, Y. G. (2004), Variance Estimation for Measures of Change in Probability Sampling, *The Canadian Journal of Statistics*, **32(4)**, pp. 451–467.

Betti G., Gagliardi F. (2017), Extension of JRR method for variance estimation of net changes in inequality measures, *Social Indicators Research*, **DOI:** 10.1007/s11205-017-1590-8.

Betti G., Lemmi A. (2013), Poverty and Social Exclusion in 3D: Multidimensional, Longitudinal and Small Area Estimation, in Betti G., Lemmi A. (eds.), *Poverty and Social Exclusion: New Methods of Analysis*. London and New York: Routledge, pp. 301-315.

Elbers, C., Lanjouw, J.O., Lanjouw, P. (2003), Micro-level Estimation of Poverty and Inequality. *Econometrica*, **71**, pp. 335-364.

Fay, R.E., Herriot, R.A. (1979), Estimates of income for small places: an application of James-Stein procedures to census data. *Journal of the American Statistical Association*, **74**, pp. 269-277.

[10] Of course, values of these parameters are in general specific to the variables being estimated.

Goedemé, T. (2013a), '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*, **110(1)**, pp. 89-110.

Goedemé, T. (2013b), *The EU-SILC sample design variables: critical review and recommendations*, CSB Working Paper Series, WP 13/02, Antwerp: Herman Deleeck Centre for Social Policy.

Gosh, M., Rao, J.N.K (1994), Small Area Estimation: An Appraisal (with discussion), *Statistical Science*, 9(1), pp. 55-93.

Handerson, C.R. (1950), Estimation of Genetic Parameters, *Annals of Mathematical Statistics*, **21**, pp. 309-310.

Kish, L. (1965), *Survey Sampling*, New York: John Wiley & Sons, Inc.

Kish, L. (1995), Methods for Design Effects, *Journal of Official Statistics*, **11(1)**, pp. 55–77.

Lohr, S.L., Rao, J.N.K. (2000), Inference from dual frame surveys. *Journal of American Statistical Association*, **95**, pp. 271-280.

Neri, L., Ballini, F., Betti, G. (2005), Poverty and inequality mapping in transition countries, *Statistics in Transition*, **7(1)**, pp. 135-157.

O'Muircheataigh, C., Pedlow, S. (2002), *Combining samples vs. cumulating cases: a comparison of two weighting strategies in NLS97*. American Statistical Association Proceedings of the Joint Statistical Meetings, pp. 2557-2562.

Osier, G. (2009), Variance Estimation for Complex Indicators of Poverty and Inequality Using Linearization Techniques, *Survey Research Methods*, **3(3)**, pp. 167-195.

Osier, G., Berger, Y., Goedemé, T. (2013), *Standard error estimation for the EU–SILC indicators of poverty and social exclusion*. Eurostat Methodologies and Working Papers 13-024, Eurostat, Luxembourg.

Piacentini, M. (2014), *Measuring Income Inequality and Poverty at the Regional Level in OECD Countries*, OECD Statistics Working Papers, 2014/03, OECD Publishing.

Rao, J.N.K. (2003), *Small Area Estimation*. Wiley, London.

Tam, S. M. (1984): On covariances from overlapping samples. *The American Statistician*, **38**, pp. 288–289.

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

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

Verma, V., Betti, G., Gagliardi, F. (2010), *An assessment of survey errors in EU-SILC*, Eurostat Methodologies and Working Papers, Eurostat, Luxembourg.

Verma, V., Gagliardi, F., Ferretti, C. (2013), *Cumulation of poverty measures to meet new policy needs*, in Advances in Theoretical and Applied Statistics. Torelli, Nicola; Pesarin, Fortunato; Bar-Hen, Avner (Eds.) 2013, XIX, Springer.

Wells, J.E. (1998), Oversampling through households or other clusters: comparison of methods for weighting the oversample elements. *Australian and New Zeeland Journal of Statistics*, **40**, pp. 269-277.

Table 1. Data sources used for the sub-national indicators

Country	Data Source and year	Regional level and number of regional units	Households in regional samples (min. -
Australia	2009-10 Survey of Income and Housing (SIH)	TL2, 8 regions	578 -3314
Austria	***EU-SILC, 3 year averages for 2008-2009-2010***	***TL2, 8 regions***	***207 -1315***
Belgium	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 3 regions*	*837 -3087*
Canada	Survey of Labour and Income Dynamics, 2011 reference	TL2, 10 regions	1766-18050
Chile	CASEN Survey, 2011 reference income	TL2, 15 regions	1588-5779
Czech Republic	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 8 regions*	*871 -1441*
Denmark	Danish Law Model System 2010	TL2, 5 regions	Register
Finland	*EU-SILC, 2012 wave (2011 reference income)*	*TL2, 4 regions*	*2298-2755*
France	ERFS, 2010 reference income	TL2, 21 regions	304 -8560
Germany	SOEP, 2011 wave (2010 reference income)	TL2, 16 regions	66-1789
Greece	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 4 regions*	*706 -2288*
Hungary	*EU-SILC, 2011 wave (2010 reference income)*	*NUTS1, 3*	*2932-5446*
Israel	Integrated Income Survey, 2011	TL2, 7 regions	2181-12213
Italy	*UDB IT-SILC, 2012 wave (2011 reference income)*	*TL2, 21 regions*	*344 -2031*
Japan	Comprehensive Survey of Living Conditions, 2009	TL2, 10 regions	729-3378
Mexico	Módulo de Condiciones Socioeconomicas, 2012	TL2, 32 regions	299 -2805
Netherlands	Income Panel Survey, 2010	TL2, 4 regions	9583-44587
New Zealand	Household economic survey, 2011 reference income	TL2, 2 regions	1134-2402
Norway	Income Statistics for Household, 2011 reference income	TL2, 7 regions	Register
Poland	*EU-SILC, 2011 wave (2010 reference income)*	*NUTS1, 6*	*1294-2651*
Slovak Republic	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 4 regions*	*611 -2099*
Slovenia	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 2 regions*	*4380-4859*
Spain	***EU-SILC, 3 year averages for 2008-2009-2010***	***TL2, 19 regions***	***113-1558***
Sweden	Income Distribution Survey, 2011 reference income	TL2, 8 regions	630 -3778
Switzerland	*EU-SILC, 2011 wave (2010 reference income)*	*TL2, 7 regions*	*266-1856*
Turkey	*Turkish SILC, 2011 reference income*	*NUTS1, 12*	*610 -2137*
United Kingdom	Households Below Average Income, average for 2010-	TL2, 12 regions	938-3842
United States	Current Population Survey, average for 2010-2012	TL2, 50 regions	2169-20056

Source: Authors' elaboration from Piacentini (2014).

Table 2. Median size of confidence intervals for poverty and income distribution indicators

Country	Relative poverty defined at 60% national median income		Gini disposable income		S80/S20 (quintile share ratio) disposable income	
	Relative median confidence interval	Absolute median confidence interval	Relative median confidence interval	Absolute median confidence interval	Relative median confidence interval	Absolute median confidence interval
Australia	0.18	4.41	0.11	0.03	0.34	1.80
Austria	0.42	5.17	0.16	0.04	0.23	0.86
Belgium	0.22	4.41	0.14	0.05	0.17	0.61
Chile	0.24	6.00	0.06	0.02	0.87	10.30
Czech Republic	0.31	3.51	0.33	0.08	0.15	0.52
Germany	0.42	8.45	0.17	0.04	0.22	0.88
Spain	0.23	4.30	0.09	0.03	0.23	1.40
Finland	0.21	3.39	0.08	0.02	0.11	0.40
Greece	0.27	6.00	0.11	0.04	0.25	1.56
Hungary	0.22	2.77	0.06	0.02	0.09	0.32
Israel	0.08	2.20	0.00	0.00		
Italy	0.33	5.60	0.15	0.04	0.25	1.10
Mexico	0.24	6.35				
Poland	0.23	4.43	0.09	0.03	0.15	0.73
Slovak Republic	0.31	4.60	0.14	0.04	0.19	0.72
Slovenia	0.17	2.57	0.11	0.03	0.09	0.31
Switzerland	0.28	4.81	0.33	0.10	0.16	0.65
Turkey	0.20	5.02	0.42	0.17	0.17	1.19
United States	0.19	4.10	0.08	0.03	0.17	1.20

Note: the relative median confidence interval is calculated as the median of the difference between the upper and lower bound of the 95% confidence interval of the indicator in each region, divided by the point estimate of the indicator. The absolute median confidence interval is the median absolute value of the difference between the upper and lower bound in each region. Source: Authors' elaborations from Piacentini (2014).

Table 3. Selected relative poverty indicators, percentages, forNUT2 regions of Austria and Spain: the point estimates (percentages) and the lower and upper bound of the 95% confidence interval

	Headcount 60% national poverty line for disposable income		
	Point estimate	Lower bound	Upper bound
Austria	13.8	13.0	14.6
Burgenland	13.8	9.0	18.6
Niederösterreich	10.8	9.2	12.3
Wien	19.0	16.4	21.7
Kärnten	18.6	14.3	23.0
Steiermark	13.8	11.6	15.9

Oberösterreich	9.2	7.8	10.6
Salzburg	11.7	9.1	14.2
Tirol	10.9	8.6	13.1
Vorarlberg	11.1	7.9	14.2
Spain	22.0	21.1	22.9
Galicia	18.4	16.8	20.1
Asturias	10.3	9.0	11.7
Cantabria	19.9	15.9	24.0
País Vasco	11.0	10.1	11.9
Navarra	9.5	8.3	10.7
La Rioja	25.0	23.1	26.9
Aragón	17.9	12.6	23.2
Comunidad de Madrid	16.4	14.6	18.2
Castilla y León	23.0	20.4	25.6
Castilla-La Mancha	31.8	29.5	34.1
Extremadura	35.2	30.6	39.7
Cataluña	17.1	16.0	18.2
Comunidad Valenciana	20.1	18.0	22.1
Baleares	18.6	15.0	22.2
Andalucía	32.0	30.1	33.8
Región de Murcia	24.7	22.1	27.3
Ciudad Autónoma de Ceuta	23.5	18.7	28.3
Ciudad Autónoma de Melilla	31.8	28.8	34.7
Canarias	32.0	29.9	34.1