

Poverty Indicators at Local Level: Definitions, Comparisons in Real Terms and Small Area Estimation Methods

Luigi Biggeri¹, Caterina Giusti², Stefano Marchetti² and Monica Pratesi²

¹*University of Florence and Interuniversity Centre ASES (Advanced Statistics for Equitable Sustainable Development) “Camilo Dagum”*

²*University of Pisa and Interuniversity Centre ASES (Advanced Statistics for Equitable Sustainable Development) “Camilo Dagum”*

Received: April 14, 2018 ; Revised: April 29, 2018; Accepted: May 08, 2018

Abstract

The importance of computing poverty measures at sub-national level is nowadays widely attested. Local poverty indicators are relevant both for a detailed planning of the policy actions against poverty and social exclusion, and for the citizens to evaluate their effects. However, there are still open problems to compute adequate sub-national poverty indicators. They refer to: 1) the definition of poverty lines; 2) the methods for accounting the spatial variation of the cost of living to make comparisons in ‘real terms’ between different areas; 3) the use of Small Area Estimation methods when the sample size is not enough to obtain accurate estimates of the indicators at local level. In this paper, we discuss the issues above by presenting some analyses on the impact of using different poverty lines on the value of the poverty rate for the 20 Italian Regions, which represent a planned domain of study in Italy. Then, we estimate the poverty rate for the 110 Italian Provinces, unplanned domains in Italy, by using specific parametric models and SAE methods. The key results highlight strong differences in the territorial distribution of the poverty rate by using national versus sub-national specific poverty lines. The effect of the heterogeneity of the general spatial price indexes on the poverty rates seems instead less important in comparison with the relevant territorial differences in the cost of housing. Moreover, the different methods of estimation of poverty rates at local level provides interesting first results and indicates the route for further research to improve the methods of estimation of poverty at the sub-regional level.

Key words: local poverty indicators; poverty lines; spatial price indexes, small area estimation.

1. Introduction

The important role played by poverty measures at sub-national and local level in setting policy actions against poverty and social exclusion is nowadays widely attested (World Bank Group, 2015). Particularly, local poverty indicators are relevant both for a detailed planning of policies actions and for the citizens to evaluate their effects (Simler, 2016). However, there are still open problems to compute adequate sub-national poverty indicators that ask for attention, verification and a deeper discussion.

The poverty indicators are many, also because they are faceted to highlight the various aspects of poverty multidimensionality (Betti and Lemmi, 2014). In this work, that is part of the deliverables of a large research infrastructure InGRID-2, financed by the European Commission under the H2020 framework, we only consider one relative monetary poverty indicator, the poverty rate.

To estimate this indicator at sub-national and local level, the main issues currently discussed at international and national level are the following: i) the choice of a national poverty line versus sub-national specific poverty lines; ii) the use of spatial price indexes, and in general of Purchasing Power Parities (PPPs) that account for the spatial variation in the cost of living, to make comparisons in ‘real terms’ between the different areas; iii) the use of Small Area Estimation methods when sample size is not enough to obtain accurate direct estimates of the indicator at local level and/or of the whole local distribution of consumption expenditure. These issues are relevant *per se*, but also because they can strongly affect the estimated poverty rates, and consequently the inferences drawn for policy evaluations.

In this paper we discuss these problems in a common framework, using two subsequent coherent steps related to the objectives to achieve. The analyses are conducted at two different territorial levels: larger areas for which it is possible to obtain statistically significant direct estimates of the poverty rate and for which spatial price indexes are available, and smaller areas for which the sample size is not enough to obtain accurate direct estimates and spatial price indexes are not currently available.

In Section we study the impact of using different poverty lines on the poverty rates computed at sub-national level, and we compare the poverty rates among sub-national areas in real terms. As areas of interest we consider the 20 Italian Regions, which represent the first sub-national administrative level in Italy. For these areas the sample size is large enough to compute direct accurate estimates of the poverty rate and of the empirical distribution of the consumption expenditures. The survey data we use come from the 2012 Household Budget Survey (HBS), designed and conducted in Italy by the Italian National Statistical Institute (Istat), and the values of regional Purchase Power Parities are those estimated by Istat for the year 2009, updated to 2012 as we explain in section 2.2.

Section 3 describes a first tentative estimation of the poverty rate and of the whole distribution of the consumption expenditures at sub-regional level, that is in areas where the sample size is not enough to obtain accurate direct estimates. In this case the areas of interest are the 110 Italian Provinces, which represent the second sub-national administrative level in Italy. According the “Nomenclature Territorial Units for Statistics” used by Eurostat, Italian Regions correspond to the NUTS-2 level while Provinces to NUTS-3 level. The poverty rates are estimated by using SAE methods and the estimation of the consumption expenditures distributions by using specific parametric models. In order to compare the poverty rates in real terms among the different provinces we use the cost of housing as a proxy of spatial price indexes.

2. Estimating Poverty Rates at Territorial Levels for which the Sample Size Allows Reliable Direct Estimates

2.1 The impact of the national poverty line vs regional-specific poverty lines on the estimation of the poverty rate

As it is well known, a common method used to measure monetary poverty is based on income or consumption as welfare indicators. An individual or household is considered poor when his income or consumption level is beyond a minimum level called poverty line (PL),

defined necessary to satisfy basic needs. This level varies in time and place, and each country uses poverty lines that are appropriate to its level of development, social organization and scale of values. In particular, we refer here to the at-risk-of-poverty rate or Head Count Ratio (HCR) computed by using consumption expenditures data. This is the simplest poverty indicator usually elaborated by most of the National Statistical Offices and, among them, by the Istat.

This issue has been discussed since long time and more recently by Jolliffe and Prydz (2015) and Ayala et al. (2014), who assert the inconvenience to use only one poverty line, the national one, as usually done by National Statistical Offices. The issue arises when within a country there are big differences in the per capita income or consumption. In Italy the differences in the consumption expenditure level per capita are very high, with a relative difference reaching the 50% from Northern to Southern Italian Regions (for example between Lombardia and Campania). It is therefore important to evaluate the impact of the use of different poverty lines in measuring the poverty rate and the consequent measure of the number of poor.

By using the micro-data of the Italian 2012 HBS, we estimated the empirical distributions of the household average monthly equalized consumption expenditure in the 20 Italian regions and for the whole country. In Fig.1a, all the regional empirical distributions of the household consumption expenditure (HCE) are reported, and in Fig.1b a focus on the distributions of the regions Lombardia and Campania in comparison with Italy is presented. From these figures the different characteristics of the distributions are evident and confirm the importance of considering regional poverty lines.

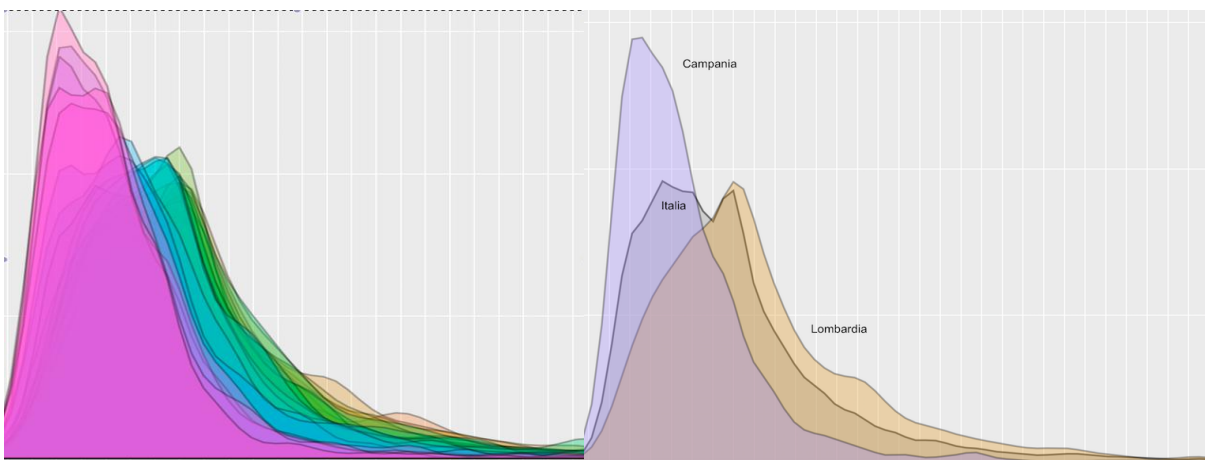


Fig. 1a: Estimates of the regional HCE distributions – 2102

Fig.1b: Estimates of the HCE distributions for two regions -2012

The Head Count Ratio (HCR) or at-risk-of-poverty poverty rate, for area $d=1, \dots, 20$, can be placed in a mathematical framework, the so-called FGT measures (Foster, Greer and Thorbeck, 1984):

$$P_{\alpha d} = \frac{1}{N_d} \sum_{i=1}^{N_d} \left(\frac{z - E_{id}}{z} \right)^{\alpha} I(E_{id} < z) \quad (1)$$

where N_d is the population size of the area d , E_{id} is the expenditure of the i th household in area d , z is the national poverty line and $I(E_{id} < z)$ is an indicator function (which is equal to 1 when expenditure is below the poverty line, and 0 otherwise). The HCR corresponds to $\alpha = 0$. When considering regional poverty lines z is replaced by z_d in equation (1).

The direct estimator of $P_{\alpha d}$ is defined as follows:

$$\hat{P}_{\alpha d} = \frac{1}{\sum_{i=1}^{n_d} w_{id}} \sum_{i=1}^{n_d} \left(\frac{z - E_{id}}{z} \right)^{\alpha} I(E_{id} < z) w_{id},$$

where w_{id} is the sample weight of household i in area d . Again, z can be replaced by z_d when considering regional poverty lines.

We computed the HCR for the 20 Italian regions, by using the National Poverty Line (NPL) and the Regional Poverty Lines (RPLs). To define the poverty thresholds we used the definition applied by Istat, setting them equal to the per capita expenditure for one household of two components, and then applying an equivalence scale for households with a different number of components (Istat, 2017).

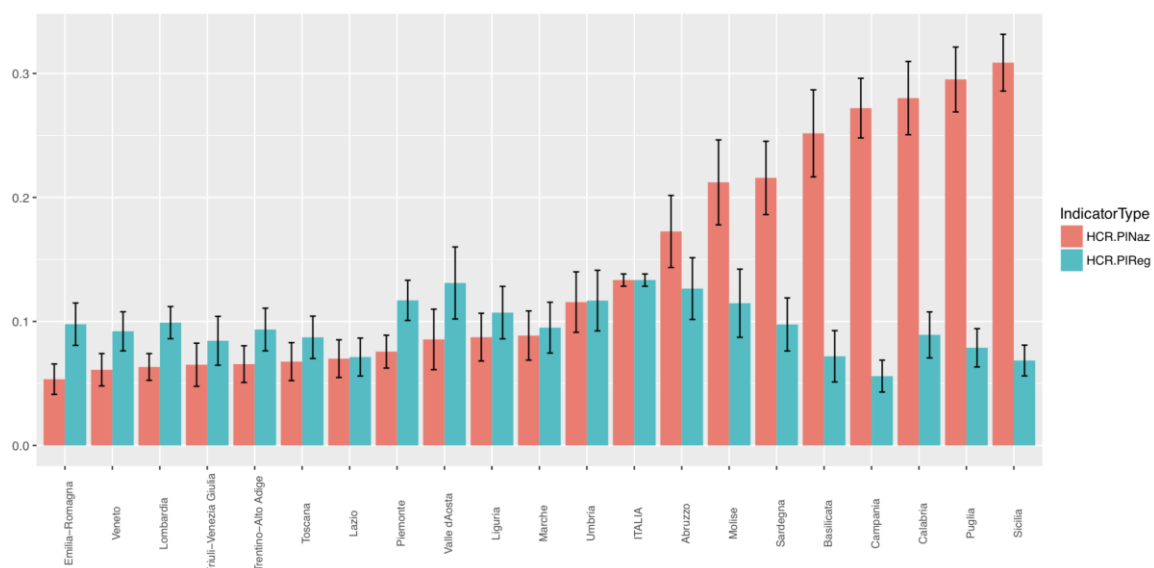


Fig. 2 Household HCR for Italian regions computed with National Poverty Line and Regional Poverty Lines – 2012

Fig. 2 shows for each region the HCR estimated with the national (z) and regional (z_d) poverty line. The results obtained clearly show that the variability of the spatial distribution of the HCRs is quite smaller by using the RPLs (the max and min values of the HCRs become 13.1% and 5.6% instead of 30.9% and 5.3% by using the NPL). Abandoning the reference to the NPL, the HCRs estimated with the RPLs are higher in 12 regions and lower in 8 regions and some of the differences in the ranking are striking. In Lombardia and Campania, the Italian largest regions in terms of population, the value of the poverty rate and the correspondent number of poor households (computed by using the total number of the households in the two regions) have relevant variations: Lombardia +3.8 points in incidence and + 148,000 poor households; Campania -21.6 points in incidence, and -445,258 poor households. Averaging, the national HCR using the RPLs instead of the NPL decreases of 4,4 points and the decrease of the number of poor households is about 1,100,000 units (from original 3,283,000).

The results certainly depend on the different levels of the household expenditure in each region, which are used to establish the RPLs. When the regional level is below the national poverty line, the regional line is of course lower. However, from Fig. 1a and 1 there is evidence that for every region the main part of the difference between the national and regional values of the poverty line is also due to the shape (skewness and kurtosis) of the

distribution of the expenditure in the region. This indicates that for policy decision making it is very important to carefully consider the left side of the distributions.

Apart from the difficulty of obtaining an exact measure of the poverty¹, it is clear that the use of different poverty lines has strong geographical implications in the evaluation and location of poverty. Surely, this topic needs to be discussed by policy makers. On the other hand, the choice of the poverty definition and of the poverty line depends on the level of analysis and on the kind of policy to be implemented (Kangas and Ritakallio, 2007). For comparing relative monetary poverty and the count and location of poor at regional level it seems justified the use of region-specific poverty lines (Mogstad et al., 2007) and, more generally, local-specific poverty lines when the economic situation of the sub-areas within a region is heterogeneous.

2.2 The impact of the regional cost-of-living differences on the measure of the poverty incidence

The importance to account for cost-of-living differences when comparing personal incomes and poverty indicators referring to different territorial areas (including urban, suburban, and rural areas) by using a spatial price index is nowadays widely recognized (Jolliffe, 2006a and 2006b; Aten, 2006; Ayala et al., 2014; Aten and Figueroa, 2015).

To assure that the poverty line(s) represent approximately the same standard of living across the different areas², there are two groups of indexes that are suggested and used at subnational level: Purchasing Power Parities (PPPs) and Cost of Housing, which have different background and justification (Renwick, 2009; Renwick et al., 2014)³.

At the international level, the use of PPPs computed by the International Comparison Program (ICP) of the World Bank (World Bank, 2014) is considered the most adequate spatial index to do poverty comparisons. However, the computed PPPs refer to the whole population and are not specific for the consumption or income of the poor. Some researchers, and in particular Deaton and Dupriez (2011), suggested to compute poor-specific PPPs by using the same vector of general average prices of groups of products and different weights, taking into account that the chosen baskets of goods and services are different at the different quintiles or percentiles of the consumption distribution. This is a first attempt to solve the problem, and the ICP-WB has included in his research agenda the study of poverty-specific PPPs. However, this is not enough. The consumer behavior of households and in particular of

¹ Within a country, the issue to have a good measure of the poverty and of the poor could be solved by computing the absolute poverty at very small areas level. In this case, it is possible to use on ad-hoc data collection where the basket of goods and services corresponding to the basic needs and its monetary value are monitored to capture the spatial and longitudinal variations in quantities, qualities and prices of the expenditures for consumption. Istat decided to follow this approach in 2009 (Istat, 2009) and is continuing to compute the absolute poverty in the same way. The baskets of goods and services are, by definition different based on many variables: the characteristics and typology of the household (number of components and their age, educational degree, etc.), the size of the municipality where the household is living and the geographical location of the municipality (Northern, Center or Southern part of Italy). The results of this computation are very interesting and useful, but for the time being the absolute poverty is not computed at the regions' level.

² The measures of absolute poverty, computed as mentioned in footnote 1, take implicitly into account of the different cost of living of the different areas.

³ Actually, in order to estimate consumer Sub-National PPPs alternative approaches have been proposed which are based on the Engel's Curve and/or a demand system model applied to data collected with sample surveys on household consumption expenditure. Instead of price data, this approach makes use of specific unit values obtained by the above-mentioned surveys. This approach was used for the first time by Coondoo et al (2004). More recent developments and applications have been done by Majumder, Ray and Sinha (2015). However, much of their work focuses purely on food price index numbers as the household expenditure surveys provide reliable unit value measures for food items. Wishing to use this approach, we have to bear in mind that the results here are based on unit value prices from household surveys and not on prices collected for the CPI.

the poor varies for quality of the commodities, channels of distributions, location of the markets and, above all, are likely to be lower in volume. It is known that the variability and the relative variation of prices (of elementary price indexes) by type of outlet and area are usually rather high (ISTAT, 2014). More difficulties emerge when sub-national PPPs are needed, mainly because of data collection complexity and, in fact, few countries are computing them.

For sub-national cost-of-living adjustments to compare poverty, also spatial indexes based on the cost of housing are used. In particular, these indexes are used in the USA because their variation across areas can be significant. The most used index refers to the average monthly rent for different type of houses. The hypothesis is that renting is the most important issue faced by poor, representing 40%-50% of their total consumption expenditure that is approximately the share of expenditure devoted to the cost of housing by the poor households. Moreover, it is important to consider that many policies to combat poverty are based on rental housing subsidies or on providing no housing cost to the poor.

To evaluate the impact of the regional cost-of-living differences on the estimation of the HCR in Italy in the year 2012 the question is: do we have the above-mentioned spatial price indexes, or can we compute them?

In Italy, there is only one computation of PPPs for the chief-towns of the regions that Istat computed in 2009 (Istat, 2010)⁴, following the ICP method. These PPPs have been updated by Marchetti and Secondi (2017) to the year 2012 and therefore they can be used as conversion factors to adjust the regional poverty lines, as they refer to the household expenditures.

As concerns the spatial indexes based on the cost of housing, since no results were available for Italy, we decided to estimate the regional median monthly expenses of the households to rent a house (RMHs) by using the HBS data.

The two available conversion factors have some limitations: the PPPs for chief-towns do not represent the level of prices for the whole Region, and on the other hand, the RMHs do not cover the poor that are owners of the house. Therefore, their use and interpretation should be done carefully. However, in order to have an idea of their impact, we applied the two conversion factors to the NPL to obtain new “Regional” Poverty lines. Then, we estimated new regional HCRs accounting for the spatial cost-of-living differences. As suggested by other researchers, we applied the MRHs to the 50% of the National Poverty Line⁵. The results are illustrated in Fig. 3.

⁴ These PPPs have been already used to deflate the regional disposal income and consumption (Pittau et al., 2011; Marchetti and Secondi, 2017).

⁵ On the basis of the HBS data the households belonging to the first quintile of the consumption expenditure distribution devoted, in average, about the 40% of the total expenditure to the housing cost.

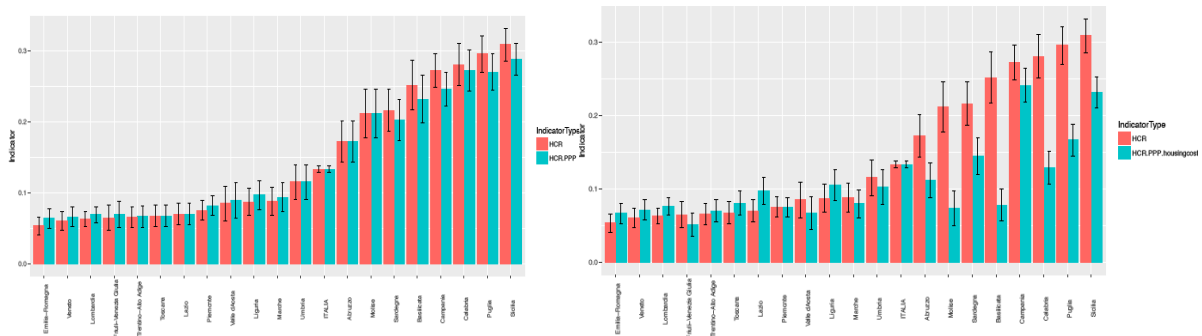


Fig. 3 Household HCR for Italian regions computed with NPL adjusted by PPPs (left) and RMHRs (right) – 2012

As we can see, the use of the conversion factors reshuffles in some way the territorial distribution of the HCRs values. More reshuffling effect is evident when the MHR is used. The range of the two spatial indexes is quite different: 13.2 percent points for PPPs and 68.8 percentage points for MHR. However, using the PPPs the resulting shift does not seem so radical in comparison with the changes obtained by using the RPLs.

Similar results were obtained by other researches (Ayala et al., 2014), and they indicate that the general PPPs are not the best conversion factors for both the poverty line and the expenses of the poor, also because they can be applied only to the value of the poverty lines and not to the different values of the distribution of the household consumer expenditure. Some researchers suggest that the best approach is using local poverty lines directly, because it is assumed that the regional standard better approximates the community standards for social standing. The within-country differences in poverty lines reflect in principles both difference in prices and needs, while the effect of the cost-of-living differences is only one explanation of the total difference highlighted by using the RPLs.

However, to correctly interpret the differences among the regional HCRs it would be necessary to analyze the effect of the different characteristics of the households (age of the components, profession, location and type of the house, etc.) under the poverty line, also to better define appropriate policy actions.

3. Estimating Poverty Rates at Territorial Levels for Which Sample Size is not Enough to Obtain Direct Accurate Estimates

As described in the introduction, we carried out a first tentative to face the issues described in section 2 in computing the poverty indicators at sub-regional level, that is for the 110 Italian Provinces, where the sample size of HBS is not enough to obtain direct accurate estimates. Obviously, the results obtained with the previous analyses guide us in the tentative estimation of the distribution of consumption expenditures and of the related poverty incidence at provincial level. In addition, the problem of measuring the different cost-of-living among the provinces is considered.

3.1 The estimation of the HCRs by using parametric distribution models and SAE methods

The use of parametric distribution models

The Log Normal and the General Beta2 models are frequently applied with good results to the income distributions, particularly the GB2 nests as special or limiting case the Weibull, Dagum and other important models for the income distribution (Bandourian et al, 2002, Graf and Nedyalkova, 2013; Marin and Molina, 2016).

We fitted the Log Normal and GB2 model separately for each Region, estimating the model parameters on the data of the equalized household consumption expenditure using the survey weights multiplied by the number of components of the household⁶. Our objective is to test the fitting of the two distributions at regional level, where the sample size is large, to choose one of them to mimic the distribution of consumption at provincial level.

To give an example of the results obtained, the fitted models are reported in Fig. 4, together with the empirical consumption expenditure distributions. For reasons of space we report here only the results for two regions: Campania and Lombardia. In Campania the distribution is shifted towards lower values of the consumption expenditure which have higher frequencies than in Lombardia.

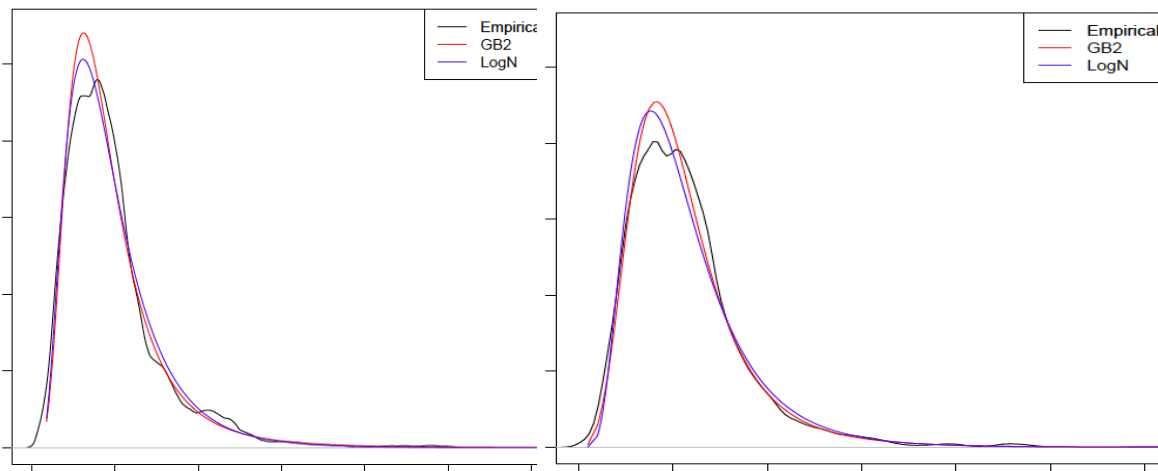


Fig. 4 Estimated and empirical distributions of HCE - 2012: Campania Region (left), Lombardia Region.

Both Lognormal and GB2 model fitted well the empirical distribution and particularly the left side of it and this is useful in practice when estimating the poverty rate at sub-regional level. However, the GB2 model represents better the data in most of the 20 Regions, and, therefore, it was chosen for the fitting at provincial level, even if the variance of the model parameter estimates is sometimes large. We computed then the provincial HCRs on the basis of the fitted distributions, using different poverty lines (national, regional, provincial). The results will be discussed at the end of this section.

The use of SAE methods

As a second approach to estimate the HCR at provincial level, we applied SAE methods. A wide range of methods have been proposed and used in literature to obtain reliable small-area estimates by model-based estimators (Pfefferman, 2013; Rao and Molina, 2015). More specifically, for the estimation of poverty indicators see among others Pratesi (2016), Molina and Rao (2010), Guadarrama et al. (2016), Marchetti et al. (2012).

Basing the analysis on the consumption expenditure distribution, we used the area-level approach proposed by Fay and Herriot, as the unit-level approach was impractical

⁶ Details on the estimated values of the parameters are available asking directly to the authors. To estimate the distribution of the consumption expenditure at provincial level we used the library GB2 of the R software (Graf and Nedyalkova, 2015), which explores the Generalized Beta distribution of the second kind. In particular, we used the function `mlfit.gb2()` that estimates the GB2 parameters by computing the maximum of the weighted likelihood. The log Normal distribution has been fitted by the R package `fitdistrplus` (Delignette-Muller and Dutang, 2015).

because of the lack of availability of auxiliary information (see also Marchetti and Secondi (2017)).

Let us assume that there are m small areas of interest and that $\theta_d, d=1, \dots, m$, represents the population characteristic of interest in area d , such as a mean, a proportion or a percentile, in our case $P_{\alpha d}$. Furthermore, assume that the vector of direct estimators $\hat{\boldsymbol{\theta}}$ is available for all or some of the areas of interest and design unbiased:

$$\hat{\boldsymbol{\theta}} = \boldsymbol{\theta} + \mathbf{e}, \quad (2)$$

where \mathbf{e} is a m -vector of $e_d \sim N(0, \psi_d)$, independent sampling errors with mean 0 and known variance ψ_d . Usually ψ_d is unknown and is estimated according to a variety of methods, including ‘generalized variance functions’, see Wolter (1985, Ch. 5) and Wang and Fuller (2003) for details.

A p -vector \mathbf{X}_d contains the population characteristics for area d , known without error and used as auxiliary information to reduce the mean squared error of the direct estimates.

The basic area level model assumes that an $m \times p$ matrix of area-specific auxiliary variables (including an intercept term), \mathbf{X} , is linearly related to $\boldsymbol{\theta}$ as:

$$\boldsymbol{\theta} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad (3)$$

where $\boldsymbol{\beta}$ is the p -vector of regression parameters, \mathbf{u} is the m -vector of independent random area specific effects with zero mean and $m \times m$ covariance matrix, $\Sigma_u = \sigma_u^2 \mathbf{I}_m$, with \mathbf{I}_m being the $m \times m$ identity matrix.

The combined model (Fay-Herriot, 1979) can be written as:

$$\hat{\boldsymbol{\theta}} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} + \mathbf{e}, \quad (4)$$

and it is a special case of linear mixed model.

Under this model, the Best Linear Unbiased Predictor (BLUP) is

$$\tilde{\boldsymbol{\theta}}_d^{FH} = \gamma_d \hat{\boldsymbol{\theta}}_d + (1 - \gamma_d) \mathbf{X}_d^T \tilde{\boldsymbol{\beta}}, \quad (5)$$

where $\gamma_d = \sigma_u^2 / (\sigma_u^2 + \psi_d)$. The predictor $\tilde{\boldsymbol{\theta}}_d^{FH}$ is a convex combination of the direct estimator $\hat{\boldsymbol{\theta}}_d$ and the predicted value $\mathbf{X}_d^T \tilde{\boldsymbol{\beta}}$ from the regression model. The extent to which it depends on the two terms is determined by the relative sizes of the model error variance σ_u^2 and the sampling error variance ψ_d . Using the restricted maximum likelihood, we can obtain an estimate of σ_u^2 ($\hat{\sigma}_u^2$), and then an estimate of $\boldsymbol{\beta}$ ($\hat{\boldsymbol{\beta}}$). Then the empirical best linear unbiased predictor is (Rao, 2003, see paragraph 6.2.4 page 100) $\hat{\boldsymbol{\theta}}_d^{FH} = \hat{\gamma}_d \hat{\boldsymbol{\theta}}_d + (1 - \hat{\gamma}_d) \mathbf{X}_d^T \hat{\boldsymbol{\beta}}$, where $\hat{\gamma}_d = \hat{\sigma}_u^2 / (\hat{\sigma}_u^2 + \psi_d)$. More details concerning analytic MSE estimation of the predictor $\hat{\boldsymbol{\theta}}_d^{FH}$ can be found in Rao (2003); Datta and Lahiri (2000); Prasad and Rao (1990).

The model relates the area target parameter to auxiliary variables that are known for each area and includes area effects to account for the between area heterogeneity. The auxiliary variables included are the provincial mean taxable income (per capita), computed by using data available from the Italian Ministry of Economy and Finance archives; the percentage of households in the province that own their own house, computed on the data of the Housing and Population Census 2011.

As a second approach, we obtained the HCR estimates using only the data coming from the HBS, by using the national poverty line. The variance of the estimates is sometimes large, but, in any case, we obtain useful information on the spatial distribution of the HCR. However, applying the Fay and Herriot model we obtain only which percentage of the distribution is below the poverty line and not the whole distribution of the consumption expenditures of households at local level.

Summary description of the results obtained

On the basis of the two presented approaches and methods, we computed the HCR for each of the 110 Italian provinces by using different poverty lines (national, regional, provincial).

The results obtained show that sometimes the variance of the estimations of the HCRs is rather large and we need to deepen the reasons of that analyzing and justifying one by one the value obtained, considering both the size of the areas involved and the data available. This will be done in a future work, focusing also on small area estimation of quantiles of the distributions.

In this first step of the research, we only investigate if both methods - the use of parametric distributions models and SAE methods - could justify the approaches that we are following to verify the impact of the choice of different poverty lines in the estimation of the poverty indicators at provincial level.

In Table 1 we report a summary description of the dispersion of the estimated HCRs (%) among the provinces, obtained by using different methods (SAE and GB2) and different poverty lines (NPL, RPLs and Provincial) PLs.

Table 1.: Summary statistics among the provinces of different estimators of provincial HCRs (in %).

	NPL		RPL		PPL
	SAE	GB2	SAE	GB2	GB2
Min.	0.00	0.00	0.00	0.00	0.00
1st Q.	6.35	6.35	6.61	7.16	7.74
Median	10.53	10.30	8.57	10.97	10.56
Mean	13.43	14.31	8.20	11.14	10.36
3rd Q.	20.55	20.50	9.99	13.82	12.86
Max.	33.98	42.15	15.09	31.09	17.98

As we can see, the result of the HCRs estimated are not so different in term of the min-max values obtained, while the heterogeneity of the HCRs among provinces became smaller passing from NPL to RPLs and PPLs, as expected. Indeed, for example, the interquartile range is 14.2 for SAEs with national poverty line, while it is 3.38 for SAEs with regional poverty lines. We did not compute HCR estimates using provincial poverty lines since the within provinces sample sizes are small and thus the estimated poverty lines are not reliable. As a future work a SAE model could be used to estimate reliable provincial poverty lines.

3.2 The impact of the provincial cost-of-living differences on the measure of the poverty rate

Indeed, PPPs conversion factors are not currently available for regions and provinces in Italy. Istat is now implementing a project to compute the PPPs for household consumption in provincial capital cities (Ferrante et al, 2014). For this reason, we estimated the conversion factor by using median monthly rent estimates for the household at provincial level (PMRHs). However, in some provinces the number of observations was too small to estimate the PMRH and in other provinces the variance of the estimates was large. However, the summary results of the spatial variability of the PMRHs normalized to Italy =1 (max =1.67; min = 0.32; CV = 0.3) and normalized to Region=1 (max=1.49; min= 0.45; CV= 0.2) show clearly the importance of the difference in the provincial cost-of-living measured by the housing cost proxy and the need for better measure of it.

The lack of adequate measure of the cost of living differences at provincial (local) level precludes the possibility to compute poverty rates in real term, which are comparable between the different territorial areas. A solution to this issue must be found and we are operating, in cooperation with Istat, to compute adequate spatial price indexes at lower territorial level.

4. Concluding Remarks

The paper presents the first results of an extended research group belonging to the Italian Inter-University Research Centre “Camilo Dagum”, involved in the European research project InGRID-2. There are many directions to continue the research, taking into account the first results presented in this paper.

The analyses conducted at regional level (that is for areas where the sample size allows reliable direct estimates) provided key results. The use of different poverty lines - national versus sub-national specific poverty lines - has strong geographical implications in the evaluation and location of poverty. Surely, this topic needs to be discussed among academic and official statisticians and at policy makers' level.

The impact of the regional cost-of-living differences on the measure of the poverty rates (HCRs) is also evident, especially considering the relevant territorial differences in the cost of housing. This is quite important because many policies against poverty are based on rental housing subsidies or on providing no cost house to the poor. Anyway, the analyses show the lack of adequate spatial price indexes at territorial level in Italy.

The analyses conducted at provincial level (that is for areas where the sample size is not enough) in order to obtain direct accurate estimates, highlighted interesting first results but also two critical issues. First of all, the variance of the estimates of the poverty rates, obtained by using parametric distribution models and SAE methods, is sometimes large. Second, the lack of adequate cost-of living correction factors (spatial price indexes).

Since taking into account the key role played by the poverty measures at sub-national and local level in setting policy actions against poverty and social exclusion is widely attested, it is necessary to try to face and to solve these two issues.

We could try to improve the methods of estimation of poverty rates at local level (provinces and sub-provinces areas), may be finding other auxiliary variables at area and household level, and by implementing the estimation of all the consumer expenditure

distribution at local level or at least focusing on small area estimation of quantiles of the distributions.

As far as the spatial price indexes, there is current research in Istat on the estimation of PPPs at provincial level, also by using scanner data on retail prices. However, the problem of identifying the prices paid specifically by the poor still exist and it seems that it will be not directly addressed. Thus, we are carrying out a project with the objective to estimate different spatial housing cost indexes at provincial and sub-provincial level. At this end we are using a very large and detailed administrative database available in Italy (Tax Agency of the Ministry of Economy and Finance) that contains, for more than 8.000 Italian municipalities, information on the rent and the cost of the houses purchased, including typology and surface of the houses and of their location in the different zones of the municipality.

Finally, to correctly interpret the reasons of the differences among the local HCRs it is also necessary to better investigate, with adequate methodology, the demographic characteristics of those individuals with consumption level below the poverty line, to better define appropriate policy actions.

Acknowledgments: This research was supported by InGRID-2 project (www.inclusivegrowth.eu) funded by the European Union's Horizon 2020 research and innovation programme under grant agreement No 730998. A preliminary version of this paper has been presented during the 61st ISI World Statistics Congress (WSC) held in Marrakech, Morocco from 16 to 21 July 2017.

References

- Aten, B. (2006) Interarea Price Levels: An Experimental Methodology, *Monthly Labor Review*, **129** (9), 47–61.
- Aten, B. and E. Figueroa (2015), Real Personal Income and Regional Price Parities for State and Metropolitan Areas, 2009–2013. *Survey of Current Business*, **96** (July).
- Ayala L., Jurado, A. and Pérez-Mayo, J. (2014), Drawing the Poverty Line: Do Regional Thresholds and Prices Make a Difference? *Applied Economic Perspectives and Policy*, **36** (2), 309-332.
- Bandourian R., McDonal J.B. and Turley R.S. (2002), A Comparison of Parametric Models of Income Distribution Across Countries and Over Time, Luxembourg Income Study, Working Paper n. 305.
- Betti G. and Lemmi A. (2014), *Poverty and Social Exclusion: New Methods of Analysis*, Ed., Wiley.
- Coondoo, D., Majumder, A., and Ray, R. (2004). A method of calculating regional consumer price differentials with illustrative evidence from India. *Review of Income and Wealth*, **50**(1), 51-68
- Datta, G. and Lahiri, P. (2000) A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica* **10**, 613 – 627.
- Deaton, A. and Dupriez, O. (2011), Purchasing Power Parity Exchange Rates for Global Poor, *American Economic Journal: Applied Economics*, **3**, 137-166.
- Delignette-Muller, M.L. and Dutang, C. (2015) fitdistrplus:an R Package for Fitting Distributions, *Journal of Statistical Software*, **64**, 1-34.
- Fay, R.E. and Herriot, R.A. (1979) Estimation of income from small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, **74**, 269–77.
- Ferrante, C., Occhiobello, R. and Polidoro, F. (2014), State of play of ISTAT project for compiling sub-national PPPs in Italy. Tech. Rep. Paper presented at the Workshop on

- Inter-Country and Intra-Country Comparisons of Prices and Standards of Living, September 1-3, Arezzo, Italy (provisional draft).
- Foster, James, J. Greer, and Eric Thorbecke (1984). A Class of Decomposable Poverty Measures. *Econometrica* **52** (3), 761–65.
- Graf, M. and Nedyalkova, D. (2013), Modeling of income and indicators of poverty and social exclusion using the Generalized Beta Distribution of the Second Kind, *Review of Income and Wealth*. DOI: 10.1111/roiw.1231.
- Graf, M. and Nedyalkova, D. (2015) Package GB2: Generalized Beta Distribution of the Second Kind: Properties, Likelihood, Estimation. Version 2.1, CRAN.
- Guadarrama, M., Molina, I. and Rao, J. (2016). A comparison of small area estimation methods for poverty mapping. *Joint issue Statistics in Transition new series & Survey Methodology* **17**, 41–66.
- Istat (2008). Le differenze nel livello dei prezzi tra i capoluoghi delle regioni italiane per alcune tipologie di beni Anno 2006, Roma (in Italian language).
- Istat (2009). La misura della povertà assoluta, Metodi e Norme n. 39, Roma (in Italian language).
- Istat (2010). La differenza nel livello dei prezzi al consumo tra i capoluoghi delle regioni italiane, Anno 2009, Roma (in Italian language).
- Istat (2014). La misura dell'inflazione per classi di spesa delle famiglie, Technical Report, Statistiche Flash, Istat Rome.
- Istat (2015). La misura dell'inflazione per classi di spesa delle famiglie, Flash, Luglio 2015, Roma (in Italian language).
- Istat (2017). La povertà in Italia, Roma, Report 2017, https://www.istat.it/it/files/2017/07/Report_Poverta_2016.pdf?title=La+poverta+in+Italia++13%2Fflug%2F2017++Testo+integrale+e+nota+metodologica.pdf, last access April 13, 2017.
- Jolliffe, D. (2006a). The Cost of Living and the Geographic Distribution of Poverty, USDA, Economic Research Report Number 26.
- Jolliffe, D. (2006b). Poverty, Prices, and Place: How Sensitive is the Spatial Distribution of Poverty to Cost of Living Adjustments. *Economic Inquiry* **44** (2): 296–310.
- Jolliffe, D. and Prydz, E.B. (2015). Global Poverty Goals and Prices, How Purchasing Power Parity Matters, Policy Research Working Paper 7256, Development Research Group Poverty and Inequality Team, World Bank Group, Washington DC.
- Kangas, O. and Ritakallio V. (2007). Relative to What? Cross National Pictures of European Poverty Measured by Regional, National and European Standards. *European Societies* **9** (2): 119–145.
- Majumder, A., Ray, R. and Sinha, K. (2015). Estimating purchasing power parities from household expenditure data using complete demand systems with application to living standards comparisons: India and Vietnam. *Review of Income and Wealth*, **61**(2), 302–328.
- Marchetti, S., Tzavidis, N., and Pratesi, M. (2012). Non-parametric bootstrap mean squared error estimation for M-quantile estimators of small area averages, quantiles and poverty indicators. *Computational Statistics & Data Analysis*, **56**(10), 2889-2902.
- Marchetti, S. and Secondi, L. (2017). Estimates of Household Consumption Expenditure at Provincial Level in Italy by Using Small Area Estimation Methods: “Real” Comparisons Using Purchasing Power Parities. *Social Indicator Research*, **131**(1), 215-234.
- Marin J.M. and Molina I. Graf M. (2016). A generalized linear mixed model for skewed distributions, paper presented at SAE 2016, Small Area Estimation Conference, August 17-19, Maastricht, NE (paper provided by the authors).

- Mogstad, M., A. Langørgen and R. Aaberge (2007). Region-specific versus Country-specific Poverty Lines in Analysis of Poverty. *Journal of Economic Inequality* **5** (1), 115–122.
- Molina, I. and Rao, J. N. K. (2010). Small area estimation of poverty indicators. *Canadian Journal of Statistics* **38**, 369-385.
- Phefferman, D. (2013), New important developments in small area estimation, *Statistical Science*, **28**(1), 40-68
- Pittau, M.G., Zelli, R. and Massari, R. (2011). Do spatial price indices reshuffle the Italian income distribution? *Modern Economy* **2**(3).
- Prasad, N. and Rao, J. (1990). The estimation of the mean squared error of small-area estimators. *Journal of the American Statistical Association* **85**,163–171.
- Pratesi, M. (2016). *Analysis of Poverty Data by Small Area Estimation*, Editor, Wiley Series in Survey Methodology, Wiley.
- Rao, J.N.K. (2003). *Small Area Estimation*, John Wiley & Sons, New York.
- Rao J.N.K. and Molina I. (2015). *Small Area Estimation*, Wiley Series in Survey Methodology.
- Renwick, T. (2009), Alternative Geographic Adjustments of U.S. Poverty Thresholds: Impact on State Poverty Rates. Working paper 2009-11, Washington DC: U.S. Department of Commerce, Census Bureau.
- Renwick, T., Aten, B., Figueroa, E. and Martin T. (2014). Supplemental Poverty Measure: A Comparison of Geographic Adjustments with Regional Price Parities vs. Median Rents from the American Community Survey, SEHSD Working Paper No. 2014-2, U.S. Census Bureau.
- Simler, K. (2016). Pinpointing Poverty in Europe: New Evidence for Policy Making. World Bank, Washington D.C.
- Wang, J. and Fuller, W.A. (2003): Mean Squared Error of Small Area Predictors Constructed With Estimated Area Variances, *Journal of the American Statistical Association*, **92**, 716-723.
- Wolter, K. (1985). *Introduction to Variance Estimation*. New York: Springer-Verlag.
- World Bank (2014). Purchasing Power Parities and the Real Size of World Economies: A Comprehensive Report of the 2011 International Comparison Program. Washington DC: World Bank.
- World Bank Group (2015). A measured Approach to Ending Poverty and Boosting Shared Prosperity, Concepts, Data and the twin Goals, Policy Research Report, Washington DC.