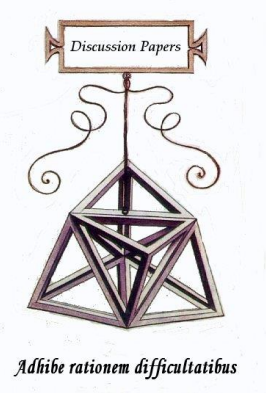
Discussion Papers
Collana di
E-papers del Dipartimento di Economia e Management – Università di Pisa



Bertarelli G. Biggeri L. Giusti C.
Marchetti S. Pratesi M.

Intra-Country comparisons of Poverty Rate

Discussion Paper n. 260

2020

Discussion Paper n. 260, presentato: **Ottobre 2020**

Indirizzo dell'Autore:

Gaia Bertarelli, Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy. Email: gaia.bertarelli@ec.unipi.it

Luigi Biggeri, Centro Interuniversitario Camilo Dagum, via Ridolfi 10, 56100 PISA – Italy. Email: luigi.biggeri@unifi.it

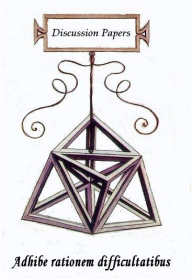
Caterina Giusti, Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy. Email: caterina.giusti@unipi.it

Stefano Marchetti, Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy. Email: stefano.marchetti@unipi.it

Monica Pratesi, Dipartimento di Economia e Management, via Ridolfi 10, 56100 PISA – Italy. Email: monica.pratesi@unipi.it

Si prega di citare così:
Bertarelli, Biggeri, Giusti, Marchetti, Pratesi (2020), "Intra-Country comparisons of Poverty Rate", Discussion Papers del Dipartimento di Economia e Management – Università di Pisa, n. 260 (http://www.ec.unipi.it/ricerca/discussion-papers).

Discussion Paper
n. 260



Bertarelli G. Biggeri L. Giusti C. Marchetti S. Pratesi M.

Intra-Country comparisons of Poverty Rate

Abstract

Reducing the inequality between member states is a target the European Union (EU) has set itself in its treaties and monitors through its cohesion reports. This has become a much debated and researched issue over the last decade. One result emerges clearly in the debate: this goal is far to be reached without a deeper studying of inequalities within each Member States. This paper describes a general approach to the previous issue proposing a set of statistical methods that can be applied in the almost totality of the EU countries. It is based on data from European current sample surveys on consumption expenditure (Household Budget Survey) and on income and living conditions (EUSILC - European Survey on Income and Living Conditions). It uses the most popular poverty indicator from the Laeken set, the At Risk of Poverty Rate (ARPR) or Head Count Ratio (HCR). The examples are built on Italian data. The sub national level used is defined on the basis of the NUTS classification used by Eurostat.

Keywords: Inequality, Small Area Estimation, Laeken indicators

JEL: C18, C80, C83

Intra-Country comparisons of Poverty Rate

Gaia Bertarelli[1,2], Luigi Biggeri[2,3], Caterina Giusti[1,2],
Stefano Marchetti[1,2], Monica Pratesi[1,2]

[1] Department of Economics and Management, University of Pisa
[2] Tuscan Universities Research Centre Camilo Dagum on Advanced Statistics for the Equitable and Sustainable Development
[3] Department of Statistics, Computer Science, Applications "G. Parenti", University of Florence

Abstract

Reducing the inequality between member states is a target that the European Union (EU) has set itself in its treaties and monitors through its cohesion reports. This has become a much debated and researched issue over the last decade. One result emerges clearly in the debate: this goal is far to be reached without a deeper studying of inequalities within each Member States. This paper describes a general approach to the previous issue proposing a set of statistical methods that can be applied in the almost totality of the EU countries. It is based on data from European current sample surveys on consumption expenditure (Household Budget Survey) and on income and living conditions (EUSILC - European Survey on Income and Living Conditions). It uses the most popular poverty indicator from the Laeken set, the At Risk of Poverty Rate (ARPR) or Head Count Ratio (HCR). The examples are built on Italian data. The sub national level used is defined on the basis of the NUTS classification used by Eurostat.

1 Introduction

Reducing the inequality between member states (i.e. convergence) is a target that the European Union (EU) has set itself in its treaties and monitors through its cohesion reports. This has become a much debated and researched issue over the last decade (OECD, 2015). One result emerges clearly in the debate: this goal is far to be reached without a deeper studying of inequalities within Member States. Building on this, the objective of this paper is to provide a statistical methodology to exploit current European official data sources by Eurostat and other institutions to provide a set of comparable local indicators to study inequality and make intra countries comparisons.

This objective brings us face-to-face with a complex pattern of possible issues and dimensions of inequality, which can be measured in different ways. Inequality exists regarding different characteristics such as income, wealth or life expectancy between different entities such as persons, households, gender, labour and capital, regions or countries. Again we have to make a decision and limit our field of work to the poverty, to the relative poverty expressed by its economic dimension measured by the monetary poverty indicators, adopted by Eurostat and named Laeken indicators. Among these, At Risk Of Poverty or social Exclusion, abbreviated as AROPE, corresponds to the sum of persons who are either at risk of poverty, or severely materially deprived or living in a household with a very low work intensity. Persons are only counted once even if they are present in several sub-indicators. The AROPE rate, the share of the total population which is at risk of poverty or social exclusion, is the headline indicator to monitor the EU 2020 Strategy poverty target. As it is well known, the indicator can be obtained from data on income or consumption expenditure. Individuals or families are poor if their income or consumption level is below a minimum level (called poverty line, PL) defined necessary to satisfy basic needs. This level varies in time and place, and the countries use poverty lines which are appropriate to their level of development, social organization and scale of

values. Eurostat defines this PL as the 60% of median income, some member states define it as the consumption level of a household with two members. Whenever it is defined the threshold represents a limit used to define poor and to identify them at the territorial level chosen for the analysis. In the economy of this work ,we do not calculate the AROPE rate but we limit the attention to the At Risk of Poverty Rate, the first Laeken indicator, that is internationally named also Head Count Ratio and correspond to the percent of households (or individuals) beyond the defined PL. Strictly speaking, this indicator does not measure wealth or poverty, but low income in comparison to other residents in that country, which does not necessarily imply a low standard of living.

As it concerns the definition of the geographical scale of the study to do meaningful intra country comparisons, we say that the interest here is in doing studies at local level, that is the "places where people live". Many questions arise on the definition of the most adequate geographical areas to translate in practice the concept of "local". Should they be cities, urban and/or not urban areas? Should we use the administrative borders used by local governance? Aggregations of these administrative areas, say for instance the Local Labour Systems? We do not enter here in this controversial discussion. The debate risks to be endless. In the economy of this paper we decided to refer to the Local Administrative Units (LAU) defined by Eurostat in the NUTS definitions[1]. The sub national level used is defined on the basis of the Nomenclature of territorial units for statistics, abbreviated NUTS (from the French version Nomenclature des Unités Territoriales Statistiques). It is a geographical nomenclature subdividing the economic territory of the European Union (EU) into regions at three different levels and two Local Administrative Unit level (NUTS 1, 2 and 3 respectively, and LAU 1 and 2, moving from larger to smaller territorial units).

There are other examples of studies at NUTS level (see Simler (2016); Azevedo, J.P. (2018)). But these previous studies do not approach the problem as we do here, providing solutions to four emerging issues recalled in the vast literature on the problem. These issues are relevant because their solution sometimes strongly affects both the final values of the poverty indicators and the number of the poor, impacting on the policy actions.

They are:

- the geographical level of the poverty lines used for the analysis;
- the spatial variation of cost of living as measured by the spatial price indexes, and in particular by the Purchasing Power Parities (PPPs) for comparisons in real terms between different areas;
- the use of Small Area Estimation methods when sample size is not enough in the data sources, in order to obtain accurate estimates of the indicators and price indexes at local level and/or of the whole local distribution of income and/or consumption expenditures in nominal and real terms.
- the problems related to cross-border statistics in Shengen Area where free movements of people, goods, services and capital make the national borders a very fuzzy limit.

The paper describes a general approach to the previous issues proposing a set of statistical methods that can be applied in the almost totality of the members states. It is based on data from European current sample surveys on consumption expenditure (Household Budget Survey) and on income and living conditions (EUSILC - European Survey on Income and Living Conditions). It uses the most popular poverty indicator from the Laeken set, the At Risk of Poverty Rate (ARPR) or Head Count Ratio (HCR). The examples are built on Italian data. The sub national level used is defined on the basis of the NUTS classification used by Eurostat. Many of the results can be applied to other geographical classifications as the more recent DEGURBA, referring ot the degree of urbanization (Eurostat, 2018).

The paper is organized as follows. In Section 2 we discuss the issue of the geographical definition of the threshold of poverty. It has been discussed since long time and more recently by Jolliffe and Prydz (2015) and Ayala et al. (2014), which assert the inconvenience to use only one poverty line (the national), as usually made the National Statistical Offices. The issue arises when there are large differences in the values of per capita income or consumption among the different areas to be compared.

[1]More information is available at `https://ec.europa.eu/eurostat/web/nuts/background`

Section 3 is dedicated to the spatial distribution of the cost of living. The need to account for the cost-of-living differences in the comparison of poverty among different territorial areas (including urban, suburban, and rural areas) by using a spatial price index is recognized everywhere (Jolliffe, 2006). To assure that the poverty line(s) represent approximately the same standard of living across the different areas, there are two groups of indexes that are used at sub-national level: Purchasing Power Parities (PPPs), and Cost of Housing, which have different background and justification (Renwick, 2009). Our effort is finalized to study the spatial distribution of poverty specific PPPs.

Section 4 describes our solution when the data sources do not offer enough observations to guarantee the statistical significance of the poverty indicators at the geographical level (called small area) chosen for the analysis. A wide range of methods have been proposed and used in the literature to obtain reliable small-area estimates (mostly model-based estimators) (Pratesi, 2016). We follow here what already done by Marchetti and Secondi (2017), which, taking into account the availability of the data, used the area-level approach proposed by Fay and Herriot. The area-level estimator is a linear combination of the small area direct estimator and a predicted component based on a linear mixed model. The model relates the parameter of interest (e.g. the HCR) to auxiliary variables that are known for each area, and includes area effects to account for the between area heterogeneity.

The summary of our findings and a list of recommendations for future analyses at European level are the content of the concluding Section 5.

2 The impact of the regional-specific vs national-specific poverty lines on the evaluation of the poverty incidence

The choice of the poverty definition and of the PL depends on the level of the analysis and the kind of policy to be implemented (Kangas and Ritakallio, 2007). For comparing relative income poverty at regional level, it seems justified the use of region- specific PLs (Mogstad et al., 2007). To compute the relative poverty incidence the poverty line is set at the 60% of the national (for the National PL) or regional (for the regional PL) median equivalised disposable income. National Poverty Lines (nPLs) allow us to establish a general scheme of how regions compare with national standards. However, the resulting regional rankings are conditioned by their relative wealth. Regional Poverty Lines (rPLs) allow us to gauge intraregional poverty. Considering the same nPL for each of the regions implies an equity concept in which individuals with equal income are assumed to have similar wellbeing regardless of the region where they live. Prior research has shown that relative differences are greater within nations than between nations (Kangas and Ritakallio, 2007).

To show the relevance of using region-specific Poverty Lines (PLs) instead of an unique National PL (nPL) in order to provide useful data for policy interventions against poverty at sub-national level, we present an example referring to the estimation of Italian households' Head Count Ratio (HCR) in the 20 Italian regions (NUTS-2 level) using Italian sample of the 2017 European Union Survey on Income and Living Conditions (EUSILC 2017). EUSILC 2017 survey has as reference population all private households and their current members residing in Italy at the time of data collection and the sample size is about 22 thousand households.

The use of different PLs has strong geographical implications in the evaluations of Italian household's poverty. The Italian nPL in 2017, set at 60% of the national median equivalised income, is 9994.2 euros. Table 1 shows rPLs in Italy for the same year. It is possible to observe that rPLs are heterogeneous from one region to another. In some regions the rPL takes on a lower value than the national one, in others a greater one. This has a great implication in the analysis of the HCR.

Figure 1 shows the HCR values for Italian regions computed with nPL and rPL and reports the corresponding confidence intervals. If the rPLs are used, it appears to be no statistically significant differences in the estimates of the regional HCR to measure relative poverty. On the contrary, when the HCR is calculated based on the nPL, regional HCRs seem to be different.

When using the nPL to calculate the regional HCR, it is possible to observe geographical inequalities: southern Italy seems to be much poorer than the north (Figure 2 (a)). Figure 2 (b) does not

Region	rPL	Region	rPL
Sicilia	6988.80	Campania	7139.20
Calabria	7442.40	Basilicata	8057.40
Molise	8344.00	Puglia	8697.60
Abruzzo	8828.67	Sardegna	9333.60
Lazio	10114.88	Umbria	10274.80
Marche	10677.60	Piemonte	10681.20
Valle d'Aosta	10698.00	Toscana	10888.57
Liguria	11116.80	Veneto	11259.20
Friuli V.G.	11400.00	Lombardia	11483.04
Trentino A.A.	11985.14	Emilia Romagna	12120.40

Table 1: Italian Regional Poverty Line from EUSILC 2017.

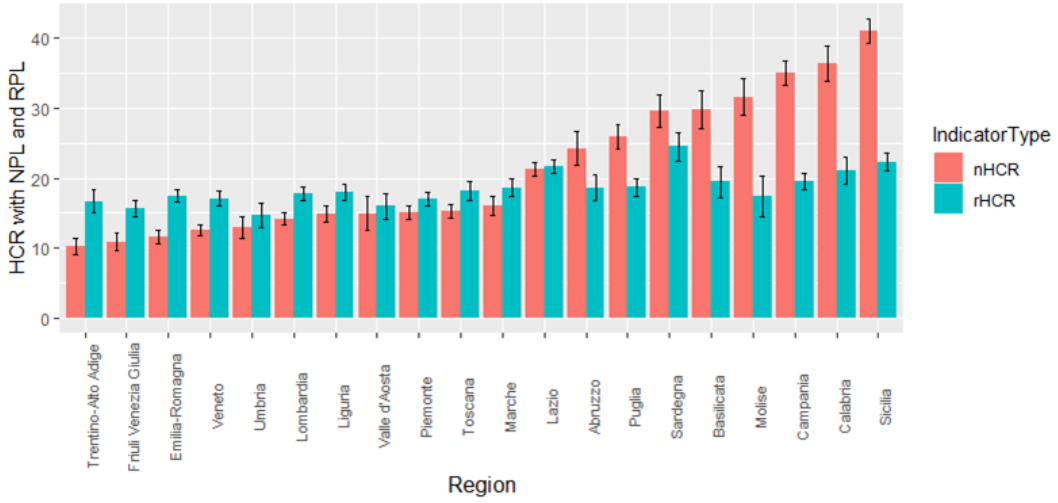


Figure 1: Household HCR for Italian regions computed with nPL and rPLs using EUSILC 2017

highlight geographical differences.

One of the main implications of using alternative regional poverty lines is the extent to which patterns of poverty might change. The number of poor families varies according to the poverty line used to compute the HCR (Table 2) and identifying the households with higher probabilities of being poor has become an important element in the monitoring and evaluation of targeted anti-poverty policies. This is especially relevant in the case of decentralized policies, where an adequate assignment of public resources to the poorest households has been a subject of increasing concern for both voters and policy-makers (Ayala et al., 2014). The effectiveness of poverty intervention largely depends on the links between the strategies implemented and the actual distribution of poverty across households. In terms of the sensitivity of poverty to alternative regional thresholds, it seems reasonable to test the extent to which poverty profiles vary as different lines are used.

A very straightforward way of dealing with this issue is to look at the results of poverty regression models, where the rationale is to model the probability of being or not being poor defined using different thresholds. To test different poverty pattern among the use of nPL and rPL we estimate logit models for each of the poverty rates with alternative thresholds (the best set of covariate selected using AIC criteria)

Those characteristics usually considered as being relevant factors in the relative risk of being poor have been chosen as explanatory variables. In addition to regional information, they include household characteristics, educational attainment, and employment status. The householder's gender (male as

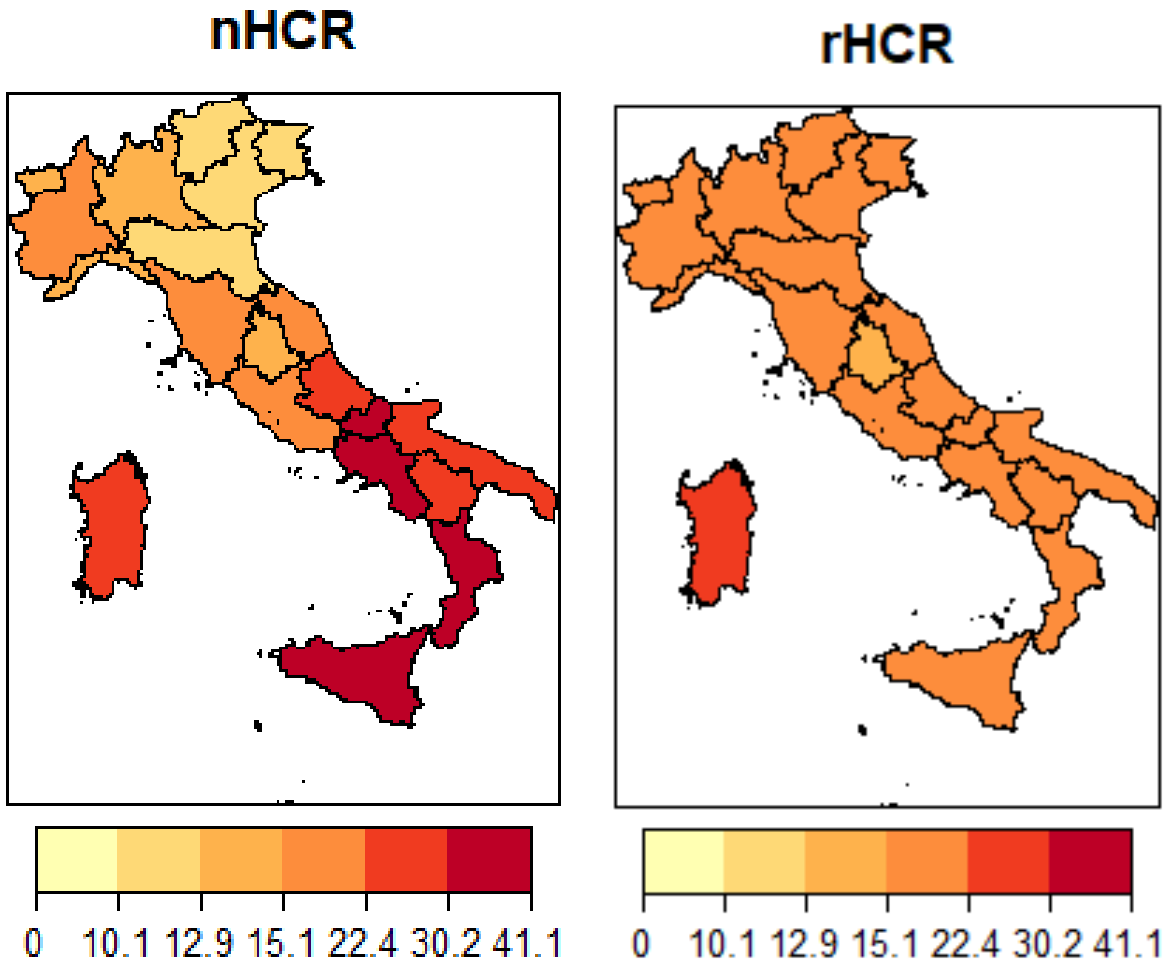


Figure 2: Household HCR for Italian regions computed with nPL (a) and rPLs (b) using EUSILC 2017.

reference category) and age, the household size, the number of household member less than 17 years old, and the marital status of the householder (5 dummy variables: never married, married, separated, widow and divorce with married as reference category) and the nationality of the householder (italian nationality as reference category) are included among the first of these. Educational and labor characteristics are key measures for explaining household living conditions. Eight dummy variables (no title, primary school, junior high school, high school diploma (not allow to attend university), high school diploma (allow to attend university), post graduate non-university diploma, university degree and PhD or post-graduated specialization) reflecting the educational attainment of the householder are included with high school diploma that allow to attend university as the reference category. In addition, four dummy variables are considered to reflect employment status (employed, unemployed, retired and other), with employment being the reference category.
The sample distribution of the socio-economic factors are:

- Gender of the householder: 64% male, 36% female;
- Age: median 59 years old, mean 60 years old;
- Household size: 29% one component, 36% two components, 18% three components, 17% more than three;
- Number of household members younger than 17: 75% zero, 16% one, 8% two, 1% more than

Region	numb. of Poor nPL	numb. of Poor rPL
Piemonte	224	253
Valle d'Aosta	62	68
Lombardia	308	406
Trentino	93	137
Veneto	203	275
Friuli	112	164
Liguria	156	192
Emilia	177	262
Toscana	236	287
Umbria	99	107
Marche	155	181
Lazio	380	390
Abruzzo	119	89
Molise	132	75
Campania	373	197
Puglia	232	165
Basilicata	124	74
Calabria	218	127
Sicilia	387	216
Sardegna	159	135

Table 2: Number of poor Italian Households at Regional level using EUSILC 2017 using nPL (left) and rPL (right).

two;

- Nationality: 89% Italian, 11% others;
- Work Condition: 50% employed, 4% unemployed, 38% retired, 4% other inactivity;
- Marital Status: 23% never married, 51% married, 4% separated, 18% widowed, 5% divorced;
- Highest level of education of the householder: 4% no primary school, 18 % primary school, 27% junior high school, 7% high schoolwithout the possibility to attend university, 29% high school with the possibility to attend university, 1% post graduate without a university diploma, 11% university, 2% PhD.

Table 3 presents estimates of the probability of being in poverty for each of the two defined thresholds (nPL and rPLs). Most of the effects are significant and appear with the expected signs. Household size and the number of children in a household seem to have a significant effect of poverty, both using nPL or rPL. Household head's age seems to have a statistically significant effect of poverty regardless to the thresholds too, but it is relatively small. The higher the age and the size are, the lower the probability of being poverty is. Furthermore, the number of children in a household has a significant positive effect on poverty. It is possible to noticed that household headed by women exhibit a higher probability of being impoverished. If the household's head is unemployed, retired or inactive the probability of being poor is significantly higher than if he/she is employed. Furthermore, if the household's head is single, separated or divorced, the probability of being poor is significantly higher than if he/she is married. No significant effect if he/she is widower/widow. Table 3 shows that the previous comments apply to all models, regardless the poverty line used. Nationality and education of the household's head seem to have a significant effect on poverty. If he/she is not Italian the probability of being poor is higher. Moreover, it is possible to observe that University education has a strong negative effect on the likelihood of being poor. At other extreme, a level of education attainment equivalent to primary or lower education notably increases the provability of poverty.

	nPL		**rPL**	
(Intercept)	-1.23^{***}	(0.15)	-1.34^{***}	(0.15)
Age	-0.02^{***}	(0.00)	-0.01^{***}	(0.00)
Components	-0.28^{***}	(0.03)	-0.30^{***}	(0.03)
Minors	0.58^{***}	(0.04)	0.64^{***}	(0.04)
Female	0.16^{**}	(0.05)	0.14^{**}	(0.05)
Unemployed	2.17^{***}	(0.09)	2.29^{***}	(0.08)
Retired	0.47^{***}	(0.07)	0.53^{***}	(0.07)
Other inactive	1.45^{***}	(0.07)	1.50^{***}	(0.07)
Never married	0.28^{***}	(0.06)	0.29^{***}	(0.06)
Separated	0.69^{***}	(0.10)	0.72^{***}	(0.10)
Widow	-0.12	(0.08)	-0.08	(0.07)
Divorced	0.51^{***}	(0.10)	0.50^{***}	(0.10)
Not italian	1.05^{***}	(0.06)	1.18^{***}	(0.06)
No title	0.99^{***}	(0.20)	0.65^{**}	(0.22)
Primary school	0.93^{***}	(0.11)	0.71^{***}	(0.12)
Junior high school	0.97^{***}	(0.07)	0.93^{***}	(0.07)
High school not univ.	0.69^{***}	(0.05)	0.69^{***}	(0.05)
Post-grad not univ.	0.22^{*}	(0.09)	0.26^{**}	(0.08)
University	-0.88^{***}	(0.21)	-0.88^{***}	(0.21)
PhD	-0.17	(0.39)	-0.34	(0.40)

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 3: Logit Estimates for HCR (Average Marginal Effects) using EUSILC 2017.

Regional variables appear to produce the most significant variation 4. Both signs and statistical significance change with the different poverty lines. Results change considerably when region-specific lines are used to calculate poverty indices. Moving from national to regional poverty lines yields remarkably different results in terms of the geographical distribution of the probability of being poor.

	nPL		**rPL**	
Piemonte	-0.51^{***}	(0.10)	-0.39^{***}	(0.10)
Valle d'Aosta	-0.58^{***}	(0.16)	-0.50^{**}	(0.16)
Lombardia	-0.57^{***}	(0.09)	-0.23^{**}	(0.09)
Trentino	-0.67^{***}	(0.13)	-0.21	(0.12)
Veneto	-0.56^{***}	(0.10)	-0.20^{*}	(0.09)
Friuli V.G.	-0.89^{***}	(0.12)	-0.44^{***}	(0.11)
Liguria	-0.38^{***}	(0.11)	-0.14	(0.11)
Emilia R.	-0.69^{***}	(0.11)	-0.21^{*}	(0.10)
Toscana	-0.31^{**}	(0.10)	-0.08	(0.09)
Umbria	-0.52^{***}	(0.13)	-0.46^{***}	(0.13)
Marche	-0.36^{**}	(0.11)	-0.18	(0.11)
Abruzzo	0.21	(0.13)	-0.25	(0.14)
Molise	0.64^{***}	(0.13)	-0.23	(0.15)
Campania	0.59^{***}	(0.10)	-0.48^{***}	(0.11)
Puglia	0.43^{***}	(0.10)	-0.08	(0.11)
Basilicata	0.68^{***}	(0.13)	-0.13	(0.15)
Calabria	0.79^{***}	(0.11)	-0.13	(0.13)
Sicilia	0.99^{***}	(0.10)	-0.04	(0.11)
Sardegna	0.42^{***}	(0.12)	0.13	(0.13)
Num. obs.	22227	22227		

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$

Table 4: Logit Estimates for HCR (Average Marginal Effects) using EUSILC 2017- Regional variables.

An alternative way of checking the possible differences in the effects of the chosen poverty characteristics is to look at the shape of the probability distributions resulting from the estimated models with the nPL and rPLs. Figure 3 does not report differences in the probability distributions.

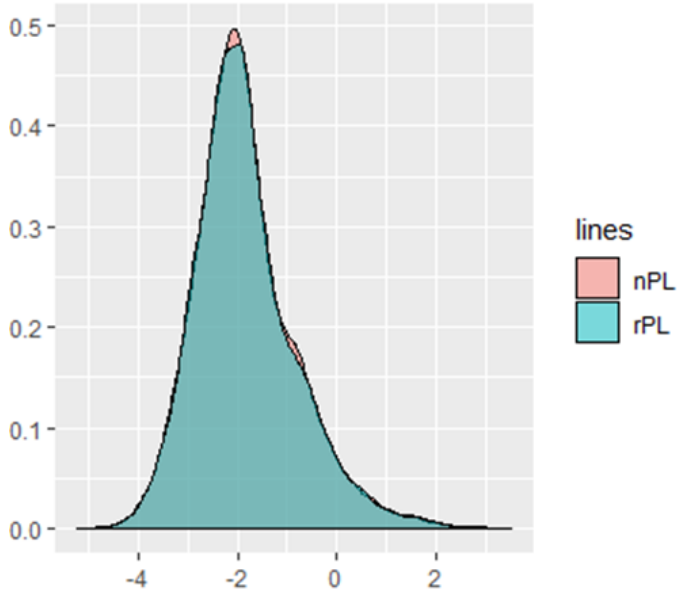


Figure 3: Estimated probability distributions of household HCR for Italian regions computed with nPL and rPLS using EUSILC 2017.

As in Ayala et al. (2014) almost all the variables (covariates) show significant effect on the probability of being poor, with expected sign. The results of the estimates of the two logit models using the two different thresholds (nPL and rPLs) are very similar. As a consequence we can state that the pattern of poor families is not influenced by the choice of poverty lines in Italian region in 2017.

Our results provide general support to the notion that regional levels of poverty change with each threshold, and regional rankings do not remain robust to the choice of poverty line. A second important finding is that poverty profiles in the 20 Italian regions in 2017 using EUSILC do not vary as different lines are used.

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

In the previous section we focused on the estimation of the HCR in Italy using data from the EUSILC survey. In Italy, ISTAT, the National Statistical Office, also estimates the HCR and the incidence of absolute poverty using data form the Household Budget Survey (HBS). In the Italian HBS data are collected on the basis of a two-stage sample design where the first stage are the municipalities and the second stage are the households. The regions (NUTS-2 level) are the finest geographical level for which direct estimates of the target indicators are usually reliable.

According to ISTAT, the Head Count Ratio (HCR), a relative measure of poverty incidence, is computed using HBS consumption data by defining for each household an indicator variable which takes value 1 if the Monthly Consumption Expenditure (MCE) of the household is less or equal the poverty line, value 0 otherwise. The values are then averaged by using the sample weights. To compute the HCR values, it is thus necessary to first compute the poverty line. At national level, the poverty line for households of two components is set equal to the per-capita mean MCE at country level:

$$nPL = \frac{\sum_{i=1}^{m}\sum_{j=1}^{n_j} CE_{ij}w_{ij}}{\sum_{i=1}^{m}\sum_{j=1}^{n_j} a_{ij}w_{ij}} \tag{1}$$

where CE_{ij} represent the Consumption Expenditure, w_{ij} the survey weight and a_{ij} the household size of household j living in area i, with $i = 1, \ldots, m$ and $j = 1, \ldots, n_j$. To take into account the

existence of economies of scale in consumption within households, the poverty line is then adjusted by using the Carbonaro scale (Istat, 2010). In this way, household expenditures can be directly compared with those of households composed of two members. The value of the HCR_{ij} is thus computed for each household as

$$HCR_{ij} = I(CE_{ij} \leq PL \cdot s_{ij}) \tag{2}$$

where s_ij represents the values of the Carbonaro scale, a specific coefficient depending on the household size. Specifically, according to the Carbonaro scale $s_ij = 0.66$ for households with $a_{ij=1}$, $s_{ij} = 1.33$ for a household with $a_{ij} = 3$, $s_{ij} = 1.63$ when $a_{ij} = 4$, $s_{ij} = 1.90$ when $a_{ij} = 5$, $s_{ij} = 2.16$ when $a_{ij} = 6$ and $s_{ij} = 2.40$ for households with 7 members or more. The *HCR* of a given area i computed by using the national poverty line PL is then computed as

$$HCR_i = \frac{\sum_{j=1}^{n_j} HCR_{ij} w_{ij}}{\sum_{j=1}^{n_j} w_{ij}}. \tag{3}$$

A corresponding measure of variability can be computed to derive the coefficient of variation and the confidence intervals for the *HCR* estimates. We computed direct estimates using the *sae* package that is available in *R* (Molina and Marhuenda, 2015). Under a different approach to that currently employed by Istat, the poverty line can be defined at the regional level as:

$$rPL_i = \frac{\sum_{j=1}^{n_j} CE_{ij} w_{ij}}{\sum_{j=1}^{n_j} a_{ij} w_{ij}} \tag{4}$$

and the corresponding *HCR* value is computed for each household j as

$$HCRrPL_{ij} = I(CE_{ij} \leq rPL_i \cdot s_{ij}). \tag{5}$$

Again, as in 3 the *HCR* of each area i, HCR_rPL_i, is then computed by averaging the HCR_rPL_{ij} in area i using the sample weights.

We consider as areas the 20 Italian regions (NUTS-2 level): in this case the intervals of estimation of the regionals $HCRs$ are not too large, and thus no specific model-based adjustments are needed. Table shows the values and the corresponding lower and upper estimated interval values of the HCR using the nPL and the $rPLs$ computed using Italian HBS data 2017.

Table 5 shows in the third column the HCR values obtained using the national poverty line nPL=1102.5 Euros, together with the corresponding interval lower and upper bounds (columns 4 and 5). For a clearer representation, the values of the HCR are also plotted in Figure 4. As expected, a clear north/south dived can be depicted when using the nPL. The values of the HCR ranges from 4.6% for the Emilia-Romagna region, in the north-east of Italy, to the 34.8% of the Calabria region, in the south of Italy. These strong geographical differences are also depicted in the values of the rPLs, presented in column 6, as the rPLs are estimated as weighted means of the CE levels. The rPLs values range from 1380.69 Euros, the rPL of Valle d'Aosta, located in the north-ovest of Italy, to 740.82 Euros, the rPL of Sicilia, an island in the south of Italy. In general, the rPLs are higher than the national PL for the north-west and north-east regions (Piemonte to Emilia-Romagna), the rPLs are closer to the nPL for the central regions (Toscana to Lazio), while the rPLs of southern regions (Abruzzo to Calabria) and the islands (Sicilia and Sardegna) are significantly lower than the nPL. Computing the corresponding HCR values, first of all, a clear decrease of the HCR values can be observed. The HCR values using the rPLs range from 5.71% of the region Marche, located in central Italy, to the 14.47% of the region Molise, i the south of Italy. From a geographical point of view, a more mixed situation is depicted with respect to the HCR values using the nPL: Figure 4 maps on the right the HCR computed using the rPLs.

Under a different approach aiming at taking into account the different price levels in each area, the national poverty line nPL can be modified using area-specific Purchasing Power Parities (PPPs), so that the rPLs represent approximately the same standard of living across the different areas. At the international level, the use of PPPs computed by the International Comparison Program (ICP) of the World Bank (World Bank, 2017) is the most adequate spatial index to do poverty comparisons. More

Table 5: HCR computed using Italian HBS 2017 data: HCR using the National Poverty Line (HCR), HCR lower bound (HCR_{lb}), HCR lower bound (HCR_{ub}), Regional Poverty Lines (rPL), HCR using rPLs (HCR rPL), HCR rPL lower bound (HCR rPL_{lb}), HCR lower bound (HCR rPL_{ub}). The estimates refers to the Regions, situated in the following areas of the country: North-Ovest (NO), North-Est (NE), Center (C), South (S) and Island (I).

Region	Area	HCR	HCR_{lb}	HCR_{ub}	rPL	HCR rPL	HCR rPL_{lb}	HCR rPL_{ub}
Piemonte	NO	7.28	5.50	9.07	1221.50	10.12	8.06	12.17
Valle d'Aosta	NO	4.60	2.58	6.62	1380.69	9.55	6.78	12.32
Lombardia	NE	5.62	4.53	6.71	1357.06	10.89	9.45	12.32
Trentino-Alto Adige	NE	5.30	3.25	7.35	1312.14	9.17	6.77	11.58
Veneto	NE	6.27	4.74	7.80	1166.68	7.81	6.10	9.51
Friuli-Venezia Giulia	NE	7.26	4.79	9.73	1205.81	8.81	6.14	11.49
Liguria	NO	8.54	5.99	11.08	1217.35	11.66	8.82	14.49
Emilia-Romagna	NE	4.81	3.43	6.20	1335.31	9.83	7.92	11.74
Toscana	C	6.06	4.29	7.83	1265.79	11.01	8.67	13.34
Umbria	C	12.61	8.91	16.31	1012.63	10.63	7.10	14.15
Marche	C	8.72	5.74	11.70	973.07	5.71	3.25	8.16
Lazio	C	8.43	6.85	10.01	1218.76	11.28	9.49	13.07
Abruzzo	S	15.96	12.21	19.70	906.29	7.16	4.40	9.92
Molise	S	21.97	18.43	25.50	894.90	14.47	11.57	17.37
Campania	S	25.16	22.43	27.88	782.00	8.69	6.91	10.47
Puglia	S	22.42	19.82	25.01	837.41	9.50	7.67	11.34
Basilicata	S	22.36	18.52	26.19	831.17	9.59	6.77	12.41
Calabria	S	34.80	30.62	38.97	740.82	13.63	10.41	16.86
Sicilia	I	30.00	26.77	33.23	780.87	11.68	9.44	13.93
Sardegna	I	17.73	14.10	21.36	918.23	9.44	6.82	12.07

difficulties emerge when sub-national PPPs are needed, mainly because of data collection complexity and in fact few countries are computing them.

A first tentative to get sub-national PPPs in Italy was done by Istat (2010). Istat computed PPPs for consumer prices for 20 cities (the regional capitals), following the procedure for international comparison, characterised by the principle of strict comparability of the products. For this reason, an analysis of data collected for Consumer Price Indices (CPIs) was carried out to check if the characteristics of products included in the divisions Food and beverages, Clothing and footwear and Furniture for the home (which represent about 34% of the total consumer expenditures) were the same in the different cities. Regarding Food and beverage products, Istat was able to use the CPI data (of about 1300 products) after having achieved comparability, while for Clothing and Furniture products it was necessary to carry out an *ad hoc* survey in order to collect new data for strictly comparable products in the 20 cities. The PPPs were computed for the basic headings of each mentioned division of expenditure and the results referring to year 2009 were published in 2010. Giving that PPPs were computed only for regional capital cities, that they were related to year 2009 and that they were built using about the 34% of the total consumer expenditures, we judged that they are not suitable to adjust the national poverty line computed using HBS 2017 data. Therefore, we present here a methodology that is based on the use of the data the cost of housing available from the HBS 2017. Indeed, for sub-national cost-of-living adjustments to compare poverty, also spatial indexes based on the cost of housing are currently used, in particular in the United States, because their variation across areas can be significant.

Specifically, in this section we propose the use of PPPs defined as Spatial Housing Price Indexes (SHPI) estimated using HBS data. As the HBS survey collects information on the rent payed by households for their house, together with the main characteristics of the house, we decided to use the classical approach based on computing price indexes based on geographical rent differentials. Indeed, the price payed by households to rent the house usually represent and important share of the total household consumption expenditure, as the house is a basic need. Therefore, the aim is to compute

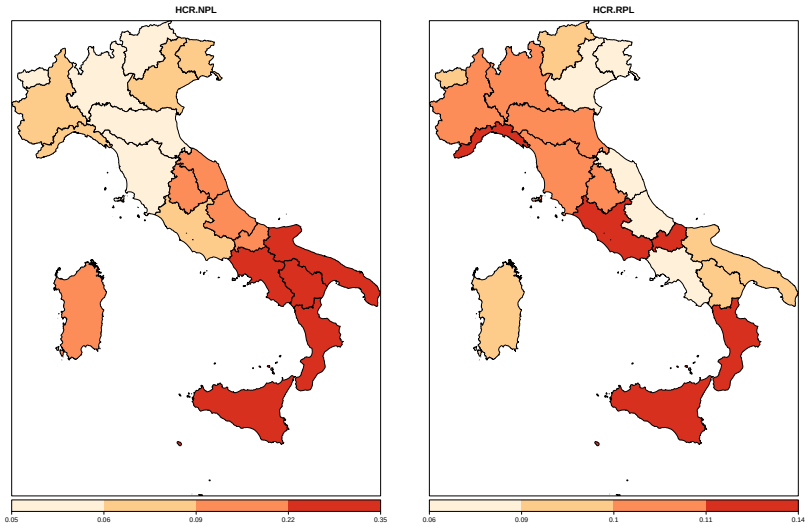


Figure 4: HCR values estimated with HBS 2017 data using the National Poverty Line (left) and Regional Poverty Lines (right).

rPLs that adjust the nPL using SHPI, instead of simply computing the PL at regional level as in the previous analysis.

Specifically, to estimate the SHPI of region r we use a hedonic price method. The hedonic price method is basically a regression of the price of the house (rent) against known relevant determinants (characteristics of the unit) that indirectly affect the price. A classical hedonic equation is as follows:

$$\ln p_{ir} = \alpha_0 + \sum_{r=1}^{R} \alpha_r D_r + \sum_{k=1}^{K} \sum_{h=1}^{H} \beta_{kh} C_{kh} + \epsilon_{ir} \tag{6}$$

where p_{ir} is the rent cost per square meter of house i in region r, D_r is a vector equal 1 if house i is in area r and 0 otherwise, α_r is the area r price, C_{kh} is the characteristic k and classification h of the house i with β_{kh} its regression coefficient, also called characteristic shadow price, and ϵ_{ir} is the error term for house i in area r, which should satisfy the standard assumptions of the multiple linear regression model (normal distributed error with constant variance). As usual, by imposing the constraint $\alpha_1 = 0$, then α_r is the difference of (fixed) effects connected with the area r compared with the base area 1. To use as a reference Italy instead of area 1, the coefficient α_r has been adjusted following Suits (1984). In this way, α_r represent the fixed effect of area r compared to Italy. Thus, the quantity $\exp(\alpha_r)$ represents the Spatial Housing Price Index in area r (SHPIr) with respect to Italy, and it is also called purchasing power parity of area r (PPPr).

To control for the characteristics and classification of the house, we use the following variables: municipality type (metropolitan area, suburbs of metropolitan area and municipality with more than 50,000 inhabitants, municipality with less than 50,000 inhabitants), presence of small kitchen (no, yes), kitchen included in the leaving room (no, yes), heating (centralized, single, district heating, single heating machine, other, no heating), building age (after 2009, 2000-09, 1990-99, 1980-89, 1970-79, 1960-69, before 1959), satellite TV (no, yes), garden (no, yes), dish-washer (no, yes), broadband (no, yes), rent type (free, regulated, "transitorio", for students or others), number of rooms and surface (in square meters, included in the model as a degree two polynomial). Parameters have been estimated using weighted least squares to account for the presence of heteroscedasticity.

Table 6 shows the estimated coefficients of the regression fitted only on the data of households with a rented house, that is without considered imputed rents. As we can see, almost all the regional effects are significant, meaning that, controlling for the house characteristics, there are residual spatial differences in rent prices. When including also imputed rents, more significant covariates are

obtained, as shown in Table 7. However, these may depend on the fact that imputed rent values are already estimated by Istat using an edonic regression. In any case, it is evident from both models that the regional effect are positive for regions located in the north of the country, negative for those located in the south (from Abruzzo).

Table 6: Model1 - Edonic regression model using only rents.

Variable / Base level	Effect	Estimate	p-value
Regions	Piemonte	0.085	0.078
Base: Italy	Valle d'Aosta	0.331	0.000
	Lombardia	0.144	0.018
	Trentino-Alto Adige	0.362	0.000
	Veneto	0.085	0.221
	Friuli-Venezia Giulia	0.197	0.018
	Liguria	0.222	0.009
	Emilia-Romagna	0.191	0.004
	Toscana	0.240	0.001
	Umbria	0.052	0.561
	Marche	0.086	0.371
	Lazio	0.231	0.001
	Abruzzo	-0.228	0.070
	Molise	-0.374	0.001
	Campania	-0.056	0.402
	Puglia	-0.102	0.126
	Basilicata	-0.689	0.000
	Calabria	-0.494	0.000
	Sicilia	-0.193	0.030
	Sardegna	-0.090	0.326
Municipality type	Suburbs with $>$ 50000 inhabitants	-0.138	0.005
Base: Metropolitan area	Suburbs with $<$ 50000 inhabitants	-0.344	0.000
Small kitchen, Base: No	Yes		
Heating	Single	-0.099	0.024
Based: centralised	District heating	-0.113	0.404
	Single heating machine	-0.189	0.002
	Other	0.556	0.000
	No heating	-0.142	0.034
Building age	2000-2009	-0.060	0.102
Base: After 2009	1990-1999	-0.146	0.006
	1980-1989	-0.306	0.000
	1970-1979	-0.197	0.000
	1960-1969	-0.251	0.000
	Before 1960	-0.060	0.163
Garden, Base: No	Yes	0.094	0.008
Dish-washer, Base: No	Yes	0.071	0.024
Broadband , Base: No	Yes	0.135	0.000
Type of rent contract	Based on special agreements	-0.220	0.000
Base: 4+4 years contract	Temporary (Max18 months)	-0.167	0.089
	Students/Regulated	-0.559	0.000
	Other	-0.183	0.068
Surface	Squared metres (Sqm)	0.012	0.000
	Sqm^2	-2.89E-05	0.000

Table 7: Model2 - Edonic regression model using rents and imputed rents

Variable / Base level	Effect	Estimate	p-value
Regions	Piemonte	0.012	0.423
Base: Italy	Valle d'Aosta	0.193	0.000
	Lombardia	0.133	0.000
	Trentino-Alto Adige	0.346	0.000
	Veneto	0.124	0.000
	Friuli-Venezia Giulia	0.110	0.000
	Liguria	0.001	0.966
	Emilia-Romagna	0.123	0.000
	Toscana	0.185	0.000
	Umbria	-0.017	0.461
	Marche	0.067	0.002
	Lazio	0.143	0.000
	Abruzzo	-0.055	0.025
	Molise	-0.223	0.000
	Campania	-0.075	0.000
	Puglia	-0.130	0.000
	Basilicata	-0.362	0.000
	Calabria	-0.302	0.000
	Sicilia	-0.306	0.000
	Sardegna	0.032	0.180
Municipality type	Suburbs with $>$ 50000 inhabitants	-0.263	0.000
Base: Metropolitan area	Suburbs with $<$ 50000 inhabitants	-0.446	0.000
Small kitchen, Base: No	Yes	-0.038	0.001
Water in the house, Base: No	Yes	0.056	0.031
Water from the well, Base: No	Yes	0.059	0.022
Heating	Single	-0.050	0.001
Based: centralised	District heating	-0.026	0.576
	Single heating machine	-0.133	0.000
	Other	-0.036	0.726
	No heating	-0.102	0.000
Building age	2000-2009	-0.024	0.050
Base: After 2009	1990-1999	-0.035	0.006
	1980-1989	-0.056	0.000
	1970-1979	-0.076	0.000
	1960-1969	-0.092	0.000
	Before 1960	-0.113	0.000
Satellite TV, Base: No	Yes	0.031	0.000
Balcony, Base: No	Yes	0.041	0.000
Garden, Base: No	Yes	0.053	0.000
Connection to the gas netweork, Base: No	Yes	0.060	0.000
Landline, Base: No	Yes	0.029	0.000
Washing machine, Base: No	Yes	0.085	0.001
Dishwasher, Base: No	Yes	0.096	0.000
Broadband , Base: No	Yes	0.070	0.000
Type of rent contract	Based on special agreements	-0.282	0.000
Base: 4+4 years contract	Temporary (Max18 months)	-0.098	0.187
	Students/Regulated	-0.602	0.000
	Other	-0.226	0.002
Box	Free	-0.002	0.940
Base: Owned	Rented	0.000	0.996
	Other	-0.050	0.000
Cellar	Free	-0.057	0.041
Base: Owned	Rented	-0.066	0.224
	Other	-0.038	0.000
Number of rooms	Count	0.027	0.000
Surface	Squared metres (Sqm)	0.138	0.000
	Sqm^2	-2.40E-05	0.000
	Sqm^3	2.52E-08	0.000

We also decided to run the two regression models - only using rents or both rents and imputed rents - focusing on a more homogeneous kind of house, apartments located in a large housing complex (with more than 10 flats). The results are shown in Table 8 and 9. As we can see, the two models do not show major differences in the estimated models parameters.

Table 8: Model3 - Edonic regression model using only rents, large housing complex (more than 10 flats).

Variable / Base level	Effect	Estimate	p-value
Regions	Piemonte	0.033	0.591
Base: Italy	Valle d'Aosta	0.351	0.004
	Lombardia	0.148	0.061
	Trentino-Alto Adige	0.400	0.000
	Veneto	0.000	0.997
	Friuli-Venezia Giulia	0.136	0.205
	Liguria	0.188	0.024
	Emilia-Romagna	0.193	0.047
	Toscana	0.263	0.020
	Umbria	0.075	0.597
	Marche	0.192	0.265
	Lazio	0.079	0.371
	Abruzzo	0.007	0.981
	Molise	-0.311	0.098
	Campania	0.138	0.158
	Puglia	-0.261	0.015
	Basilicata	-0.561	0.025
	Calabria	-0.705	0.005
	Sicilia	-0.043	0.731
	Sardegna	-0.320	0.123
Municipality type	Suburbs with $>$ 50000 inhabitants	-0.122	0.018
Base: Metropolitan area	Suburbs with $<$ 50000 inhabitants	-0.251	0.000
Water in the house, Base: NO	Yes	0.594	0.032
Heating	Single	-0.093	0.039
Based: centralised	District heating	-0.090	0.372
	Single heating machine	-0.219	0.033
	Other		
	No heating	-0.447	0.000
Building age	2000-2009	-0.193	0.000
Base: After 2009	1990-1999	-0.332	0.000
	1980-1989	-0.465	0.000
	1970-1979	-0.309	0.000
	1960-1969	-0.065	0.352
	Before 1960	-0.085	0.193
Balcony, Based: No	Yes	0.143	0.011
Broadband , Base: No	Yes	0.160	0.000
Type of rent contract	Based on special agreements	-0.300	0.000
Base: 4+4 years contract	Temporary (Max18 months)	-0.206	0.173
	Students/Regulated	-0.736	0.000
	Other	-0.418	0.001
Surface	Squared metres (Sqm)	0.008	0.000

Table 9: Model4 - Edonic regression model using rents and imputed rents, large housing complex (more than 10 flats).

Variable / Base level	Effect	Estimate	p-value
Regions	Piemonte	0.126	0.000
Base: Italy	Valle d'Aosta	0.178	0.000
	Lombardia	0.086	0.000
	Trentino-Alto Adige	0.423	0.000
	Veneto	0.042	0.190
	Friuli-Venezia Giulia	0.035	0.317
	Liguria	-0.107	0.000
	Emilia-Romagna	0.074	0.014
	Toscana	0.152	0.000
	Umbria	-0.119	0.006
	Marche	0.043	0.323
	Lazio	0.283	0.000
	Abruzzo	0.057	0.154
	Molise	-0.169	0.000
	Campania	0.003	0.923
	Puglia	-0.171	0.000
	Basilicata	-0.214	0.000
	Calabria	-0.336	0.000
	Sicilia	-0.188	0.000
	Sardegna	0.056	0.193
Municipality type	Suburbs with $>$ 50000 inhabitants	-0.257	0.000
Base: Metropolitan area	Suburbs with $<$ 50000 inhabitants	-0.357	0.000
Heating	Single	-0.054	0.000
Based: centralised	District heating	-0.016	0.637
	Single heating machine	-0.158	0.000
	Other	-0.039	0.144
	No heating	-0.324	0.000
Building age	2000-2009	-0.054	0.004
Base: After 2009	1990-1999	-0.087	0.000
	1980-1989	-0.092	0.000
	1970-1979	-0.087	0.000
	1960-1969	-0.074	0.000
	Before 1960	-0.030	0.175
Garden, Base: No	Yes	-0.063	0.008
Landline, Base: No	Yes	0.033	0.005
Dishwasher, Base: No	Yes	0.047	0.000
Broadband , Base: No	Yes	0.076	0.000
Type of rent contract	Based on special agreements	-0.374	0.000
Base: 4+4 years contract	Temporary (Max18 months)	-0.150	0.169
	Students/Regulated	-0.808	0.000
	Other	-0.453	0.000
Box	Free	-0.040	0.230
Base: Owned	Rented	0.001	0.983
	Other	-0.029	0.024
Cellar	Free	-0.112	0.002
Base: Owned	Rented	-0.049	0.377
	Other	-0.051	0.000
Surface	Squared metres (Sqm)	0.008	0.000
	Sqm^2	-1.09E-05	0.002

To better depict the differences among the four models, Table 10 shows some summary statistics of the estimated SHPI, computed by taking the exponential of the regional estimated parameters, using the four regression models by grouping the regions in three main areas (north, center, south/islands). As we can see, the results confirm that for the islands and for the regions located in the south almost

all the SHPI are below 1 under all the models, with the only exception for the largest values. These means to lower hosing price values - in terms of rents - with respect to the overall Italian territory. On the opposite for the northern and central regions almost all the values are close or over 1, indicating that in these case rent values are higher even after controlling for the house characteristics.

Table 10: Summary of estimated SHPI, by model and area.

Area	Model	Min	Q_1	Median	Mean	Q_3	Max
North Ovest/Est	Model1	1.09	1.14	1.21	1.23	1.28	1.44
	Model2	1.00	1.09	1.13	1.15	1.16	1.41
	Model3	1.00	1.12	1.18	1.21	1.26	1.44
	Model4	0.90	1.04	1.08	1.12	1.15	1.53
Center	Model1	1.05	1.08	1.18	1.17	1.26	1.27
	Model2	0.98	1.05	1.11	1.10	1.17	1.20
	Model3	1.08	1.08	1.15	1.17	1.23	1.30
	Model4	0.89	1.01	1.10	1.11	1.20	1.33
South/Islands	Model1	0.50	0.67	0.81	0.77	0.91	0.95
	Model2	0.70	0.74	0.84	0.84	0.93	1.03
	Model3	0.49	0.69	0.75	0.80	0.97	1.15
	Model4	0.71	0.82	0.84	0.89	1.02	1.06

Having estimated the SHPI values, we then used them to adjust the national poverty line (nPL) to obtain price-adjusted regional poverty lines (rPL_PPPs). Extending the idea in Renwick et al. (2014) we decided to adjust the nPL using for each region r the $SHPI_r$ value opportunely weighted:

$$rPL_PPPs = nPL \cdot (\lambda_r \cdot SHPI_r + 1 - \lambda_r) \tag{7}$$

where λ_r indicates the estimated share of household expenditure for the house in region r. The shares λ_r have been estimated using again HBS data. Table 11 shows the estimated λ_r values by region. As we can see, the shares ranges from 0.1535 (region Basilicata, located in the south) to 0.2367 (region Lazio, located in the centre). In this case, a clear north/south divide cannot be observed.

Table 11: Estimated share of household expenditure for the house using HBS data, by region.

Region	λ_r
Piemonte	0.1922
Valle d'Aosta	0.1889
Lombardia	0.1929
Trentino-Alto Adige	0.2194
Veneto	0.2021
Friuli-Venezia Giulia	0.1979
Liguria	0.2129
Emilia-Romagna	0.1958
Toscana	0.2053
Umbria	0.2033
Marche	0.2070
Lazio	0.2367
Abruzzo	0.2008
Molise	0.1783
Campania	0.2098
Puglia	0.1994
Basilicata	0.1535
Calabria	0.1755
Sicilia	0.1831
Sardegna	0.2088

Having computed the price-adjusted regional poverty lines, we then computed the regional HCR values. These are mapped in Figure 5, when using the SHPIs values estimated from Model1 and

Model2, and in Figure 6 using the SHPIs values estimated from Model3 and Model4. With respect to the strict north/south divide obtained using the nPL, the figures highlights some differences even if, adjusting the nPL only by the share presented in Table 11, approximately around the 20%, no major deviations can be observed. Specifically, the four figure together show that, independently from the chosen mode, the regions in the north central part of the country are always in the lowest quartile of the HCR values (lightest color), while several regions in the south and the island Sicilia are always in the higher quartile (darkest color). The changes in the other regions suggest that this methodology can be used to derive region-specific poverty lines adjusted for price differences. Moreover, the methodology can also be used extended by adjusting the poverty lines using other household costs - such as, for example, food costs - and to derive local poverty lines at a more detailed geographical level. In the next sections we focus on the use of the same methodology at provincial level.

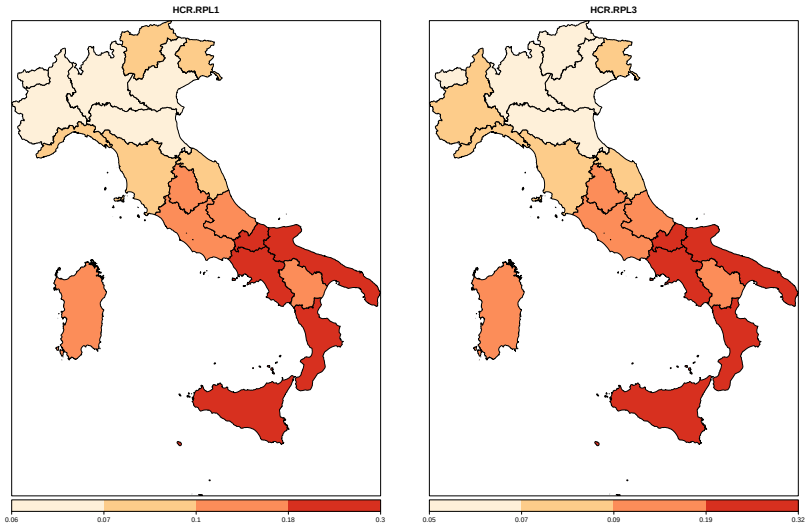


Figure 5: HCR values estimated with HBS 2017 data using the price-adjusted regional poverty lines using the SHPI values from model1 (left) and model2 (right).

4 The Estimation of Poverty Rates using Small Area Estimation Models

In the previous section we focused on the estimation of the HCR in Italy at national and regional level using data from the HBS survey. HBS data provides sound estimates at the national and regional level. However, sub-regional estimates, such as provincial estimates, can help the decision makers to plan and implement local welfare policies. Usually, such local estimates are unreliable because of the small sample size of the HBS at local level. For example, for the HBS 2017 the national sample was composed by 16496 households, with provincial sample sizes ranging from 20 to 1036, with a median value of 125 households per province. Thus, for many provinces direct estimates – i.e. estimates obtained using only province-related data – are unreliable. To obtain province-level reliable estimates we can resort to small area estimation methods (Rao and Molina, 2015). A collection of examples of usage of SAE for analysing poverty rates is in Pratesi (2016). They use auxiliary variables to build model-based estimates with increased efficiency with respect to direct estimates. The increasing of the estimates efficiency is related to the goodness of fit of the model used. We can distinguish between two main model-based small area methods: those based on unit-level models and those based on area-level models. Unit-level model-based estimates are a linear

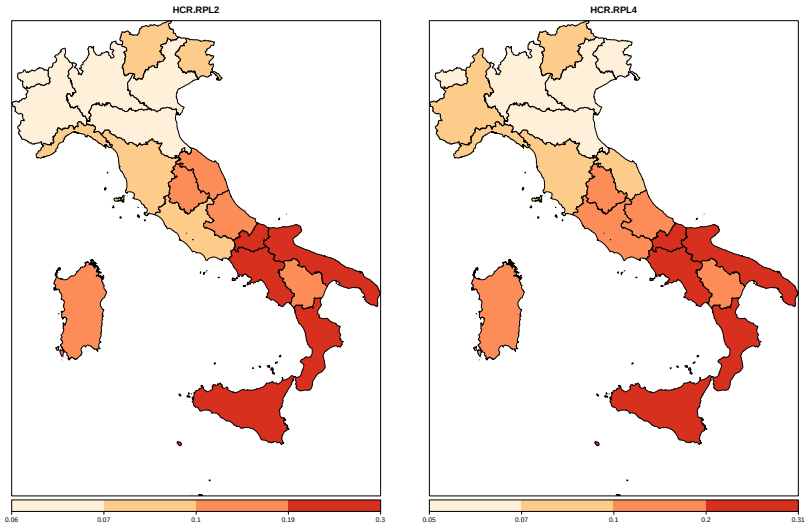


Figure 6: HCR values estimated with HBS 2017 data using the price-adjusted regional poverty lines using the SHPI values from model3 (left) and model4 (right).

combination of the direct information and a regression synthetic prediction of non-sampled units. The fixed part of the model links the target values to some known auxiliary variables, for each units belonging to the larger area to which the small areas of interest belong to. The area specific random effects is instead introduced in order to take into account the correlation among the units with each small area (between area variation). Area-level model-based estimates are a linear combination of the area direct estimator and a predicted component based on a linear mixed model. The model relates the parameter of interest to known auxiliary variables for each of the areas that constitute the partition of the whole population. An effect to account for (within) area homogeneity is included in the model. To obtain small area estimation of HCR at provincial level we use area-level model-based estimator because we do not have access to unit-level data for non-sampled units.

As already underlined in the previous section, when the focus of the HCR are regions, provinces or municipalities, the national (monetary) poverty threshold can be misleading because the price levels within the country are not equal, as noted by Biggeri et al. (2018). A purchasing power parity (PPP) for each province could handle this problem because we could adjust the national poverty line for each province taking into account the real purchasing power of the currency in each province, an issue that was already addressed in the previous section at regional level. Unfortunately, the PPPs at province level are not available nor usually computed by national statistical offices, mainly because it is very expansive in term of time and money to compute them.

Therefore, we opt to adjust partially the poverty line, using data available from the HBS 2017, as already suggested in the previous section. We propose to adjust the relative national poverty line (nPL) at the province level taking into account the Spatial Housing Price Index (SHPI), limited to the housing rents as collected by the Italian HBS. The rent cost represent about 20% of total expenditures of households.

To obtain estimates of HCRs at provincial level, we first compute spatial housing price indexes that are then used to partially adjust the national poverty line at provincial level, and, finally, we estimate the provincial HCRs using area-level model-based estimators. In each province we adjust the national poverty line proportionally to the average weight of rent expenditure on the total consumption expenditure.

4.1 Estimation of Spatial Housing Price Indexes

As already done at regional level, to estimate the Spatial Housing Price Indexes (SHPI) we use the data coming from the HBS, which collects information on the rent payed by the occupants and on the main characteristics of each house, as well the characteristics of the area where the house is located. Taking into account the data available, we use a hedonic price method to estimate the SHPI at provincial level. The hedonic price method is basically a regression of the price of the house (rent) against known relevant determinants (characteristics of the unit) that indirectly affect the price. A classical hedonic equation is as follows:

$$\log p_j i = \alpha_0 + \sum_{i=1}^{m} \alpha_i D_i + \sum_{k=1}^{K} \sum_{h=1}^{H} \beta_{khji} C_{khji} + \epsilon_{ji}, \tag{8}$$

where p_{ji} is the rent cost per square meter of house j in province i, D_i is an indicator vector being equal 1 if unit (house) j belong to area i and 0 otherwise, α_i is the area i price, C_{khji} is the characteristic k and classification h of house i in area j, with β_{khij} its regression coefficient, also known as characteristic shadow price, and ϵ_{ij} is the error term for house j in area i, which should satisfy the standard assumptions of the multiple linear regression model. The model as specified in (8) is not identified, therefore, as usual, we impose the constraint $\alpha_1 = 0$. Then, α_i is the difference of (fixed) effects connected with the area i compared with the base area 1. To use as a reference Italy instead of area 1, which make clearer the interpretation of these coefficients and easy to adjust the national poverty line, the coefficients α_i's have been adjusted following Suits (1984). In this way, α_i represent the fixed effect of area i compared to Italy. Thus, the quantity $\exp(\alpha_i)$ represents the SHPI in area i with respect to Italy.

To control for the characteristic and classification of the house we use the following variables: municipality type (metropolitan area, suburbs of metropolitan area and municipality with more than 50,000 inhabitants, municipality with less than 50,000 inhabitants), presence of small kitchen (no, yes), kitchen included in the leaving room (no, yes), heating (centralized, single, district heating, single heating machine, other), building age (after 2009, 2000-09, 1990-99, 1980-89, 1970-79, 1960-69, before 1959), satellite TV (no, yes), garden (no, yes), dish-washer (no, yes), broadband (no, yes), rent type (free, regulated, transitorio, for students or others), number of rooms and surface (in square meters, included in the model as a degree two polynomial). Parameters have been estimated using weighted least square to account for the presence of heteroscedasticity. The estimated model is summarised in table 12.

Table 12: Hedonic regression model for house rent.

		Estimate	pval
	(Intercept)	5.1274	0.0000
Provinces			
Base: Italy	TO	0.1442	0.0014
	NO	0.1798	0.0647
	CN	-0.1709	0.0328
	AT	0.1389	0.3257
	AL	-0.1463	0.2048
	AO	0.2772	0.0001
	IM	0.4845	0.0001
	SV	-0.0983	0.5579
	GE	0.2729	0.0000
	SP	0.2013	0.1482
	VA	-0.0179	0.8364
	CO	0.1349	0.3985
	SO	0.5743	0.0225
	MI	0.2637	0.0004
	BG	0.0963	0.2549
	BS	0.1003	0.2834
	PV	-0.1877	0.0572
	CR	0.2795	0.1314
	MN	-0.0598	0.6959
	BZ	0.6616	0.0000

TN	0.1288	0.0900
VR	0.1858	0.0352
VI	0.0619	0.5948
BL	0.3740	0.0573
TV	0.0572	0.5023
VE	-0.0970	0.4892
PD	0.0169	0.9116
RO	0.0756	0.7038
UD	0.1959	0.0355
GO	-0.0330	0.8743
TS	0.0859	0.4037
PC	-0.0124	0.9379
PR	0.1551	0.1131
RE	0.1364	0.2489
MO	0.0571	0.6570
BO	0.3949	0.0000
FE	0.1012	0.5097
RA	0.0971	0.4314
FC	0.0427	0.7351
PU	0.3944	0.0017
AN	-0.1589	0.1928
MC	0.0031	0.9869
AP	0.0977	0.7228
MS	-0.3141	0.0192
LU	0.0035	0.9913
PT	0.3817	0.0001
FI	0.4739	0.0000
LI	0.0555	0.7852
PI	0.0372	0.8533
AR	0.0527	0.6552
SI	0.4859	0.0277
GR	0.1342	0.5395
PG	-0.0088	0.9171
TR	0.0093	0.9471
VT	0.1861	0.0729
RI	-0.3846	0.3513
RM	0.3097	0.0000
LT	0.0628	0.6990
FR	-0.0269	0.8590
CE	-0.0053	0.9475
BN	-0.0110	0.9586
NA	0.1330	0.0403
AV	-0.2589	0.0333
SA	-0.2201	0.1985
AQ	-0.2843	0.1337
TE	0.1207	0.5158
PE	-0.1654	0.7055
CH	-0.3336	0.0119
CB	-0.4497	0.0000
FG	-0.4536	0.0006
BA	-0.1057	0.1870
TA	0.0477	0.5586
BR	-0.0745	0.6259
LE	-0.1985	0.1941
PZ	-0.6052	0.0000
MT	-0.6849	0.0000
CS	-0.4906	0.0000
CZ	-0.4412	0.0006
RC	-0.5791	0.0007
TP	-0.5104	0.0100
PA	0.1282	0.1778
ME	-0.0379	0.8102
AG	-0.3520	0.0420
CL	0.1190	0.1930
EN	-0.4975	0.2069
CT	-0.0139	0.9117
RG	-0.4459	0.0018
SR	-0.1673	0.4243

	SS	-0.5183	0.0173
	CA	0.0939	0.5408
	PN	0.1429	0.3776
	IS	-0.2634	0.0330
	OR	-0.2498	0.2243
	BI	0.2483	0.1268
	LC	-0.3500	0.2573
	LO	0.1398	0.4156
	RN	0.4582	0.0008
	PO	0.5717	0.0000
	KR	-0.0737	0.8190
	VV	-0.6244	0.0092
	OT	0.2352	0.0446
	VS	0.0157	0.9559
	CI	0.0292	0.8841
	MB	0.2905	0.0014
	FM	0.0601	0.7483
	BT	-0.3147	0.0360
Zone			
Base: Metropolitan area city center	Municipalities or suburb of metropolitan area with more than 50000 inhabitants	-0.0136	0.7109
	Municipalities or suburb of metropolitan area with less than 50000 inhabitants	-0.1206	0.0042
Heating			
Base: Centralised heating	Single	-0.0796	0.0035
	District heating	-0.1282	0.1117
	Single heating machine	-0.2025	0.0000
	Other	0.6176	0.0000
	No heating	-0.2480	0.0000
Construction period			
Base: After 2009	2000-2009	-0.0961	0.0004
	1990-1999	-0.2095	0.0000
	1980-1989	-0.3924	0.0000
	1970-1979	-0.2243	0.0000
	1960-1969	-0.1343	0.0008
	Before 1960	-0.0613	0.0601
Rent type			
Base: unconstrained rent	Regulated	-0.2596	0.0000
	“Transitorio”	-0.1599	0.0275
	Students rent	-0.6607	0.0000
	Other	-0.3383	0.0000
House dimension			
	Rooms	-0.0319	0.0434
	Surface	0.0148	0.0000
House facilities			
	Stallite TV	0.0506	0.0842
	Garden	0.0432	0.1773
	Dish-washer	0.0607	0.0092
	Broadband	0.1306	0.0000
	Presence of small kitchen	0.0503	0.0947
	Kitchen included in the living room	0.0849	0.0022

The estimated SHPIs, which are equal to the exponential of the estimated province coefficients are summarized in table 13, where provinces are grouped by Italian geographic repartitions (North, Centre, South). From table 13 we can see that the distribution of SHPI among provinces of north and central Italy are very similar, even though SHPI in the northern provinces are a little bit higher than those in the centre. The SHPI values in southern provinces are smaller, with almost 75% of the provinces with an index below 1, the SHPI value of Italy.

Table 13: Distribution of estimated province SHPI grouped by geographical repartition

Repartition	Min	1st Q.	Median	Mean	3rd Q.	Max
North	0.7047	1.0436	1.1444	1.1675	1.2818	1.9378
Centre	0.6807	1.0032	1.0595	1.1494	1.3234	1.7712
South	0.5042	0.6390	0.8025	0.8292	0.9919	1.2651

4.2 Small area estimation of poverty rates at provincial level using adjusted poverty lines

Once we have estimated the SHPI at provincial level we have to adjust the national poverty line (nPL) at provincial level to get the estimates of the HCRs. As already explained in the previous section, the national poverty line in Italy for a household of two components is set by Istat as the mean per-capita consumption expenditure. For different household sizes this threshold is adjusted according to the Carbonaro equivalence scale. We recall that in 2017 the nPL for two components is estimated equal to 1101.52 euros.

We decided to adjust the nPL for each province using the SHPI values opportunely weighted (adapting the idea in Renwick et al. (2014)):

$$nPL_i^* = nPL \times (\lambda_i SHPI_i + 1 - \lambda_i) \tag{9}$$

where nPL_i^* is the adjusted poverty line for province i, λ_i is the estimated share of expenditure for the rent in province i. The quantities λ_i's are estimated from the HBS 2017 as the provincial mean of the ratios between the rent expenditure and the total consumption expenditure:

$$\lambda_i = \frac{1}{\sum_{j=1}^{n_i} w_{ij}} \sum_{j=1}^{n_i} \frac{p_{ij}}{t_{ij}} w_{ij}, \tag{10}$$

where n_i is the sample size in province i, w_{ij} is the survey weight of household j in area i, p_{ij} is the rent price of household j in area i and t_{ij} is the total consumption expenditure of household j in area i. The survey weights have been calibrated to sum to the total households at provincial level.

Although the λ_i's are estimated at the provincial level – thus possibly unreliable because of small sample size – we judge the direct estimates suitable for our purpose. Indeed, about half of the provinces have a 95% confidence interval for λ_i%'s direct estimates that is less than 4% and it is less than 5% for about 75% of the provinces[2]. In table 14 we show the distribution over provinces of the λ_i's grouped by the main Italian geographic areas, which is similar among provinces in the north, center and south of Italy.

Table 14: Distribution over provinces of the λ_i%'s grouped by Italian geographic repartitions

Repartition	Min	1st Q.	Median	Mean	3rd Q.	Max
North	15.66	17.97	19.37	19.55	20.92	24.57
Centre	14.34	18.84	20.53	20.21	21.77	25.13
South	13.68	17.34	18.71	18.98	20.63	25.48

The adjusted nPL varies between 1000 euros in Reggio di Calabria province, south of Italy, to 1330 euros in Bolzano province, north of Italy. Moreover, 61 of the 107 Italian provinces have an adjusted nPL greater than the national reference (nPL), that is 1102.5 euros. We show the level of the adjusted nPLs in figure 7. We can see that a lot of provinces in the north and centre of Italy have an adjusted nPL above the national reference and only few of them are below, while for the provinces in the south is the contrary.

[2]Standard error of λ_i's are obtained ignoring the design effect at the province level.

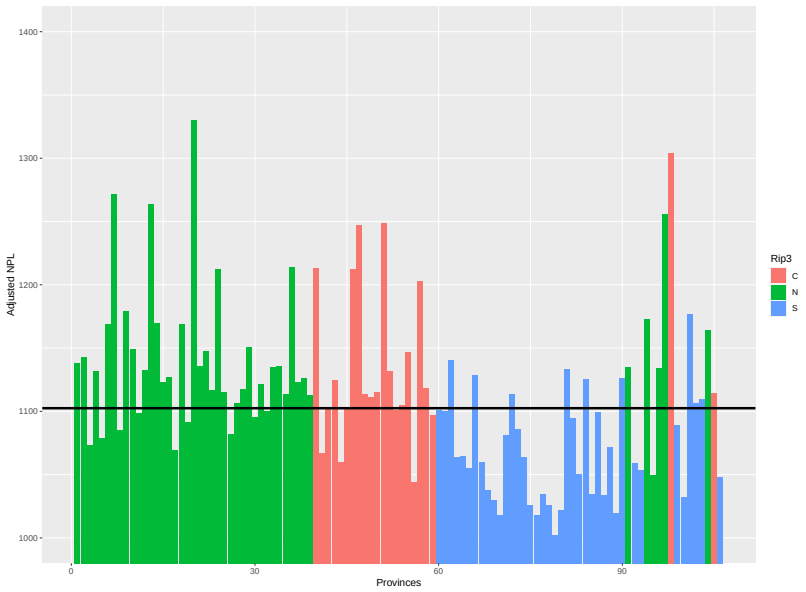


Figure 7: Adjusted nPL for Italian provinces. Black line indicates the national reference (nPL).

Once the adjusted nPLs are computed we can obtain related direct estimates of poverty rates. We compute direct estimates using the `direct` function of the R (R Core Team, 2019) package `sae` (Molina and Marhuenda, 2015). The variability of the direct estimates we obtained is judged to be too high, in particular to carry out comparisons between provinces. About half of the provinces have a 95% confidence interval length grater than 6% and about one third of them greater than 9%. Looking at the coefficient of variation (CV) of the direct estimates, we obtained in about half of the provinces a CV greater than 30% and in about 25% of the provinces a CV greater than 45%. Therefore, we decide to resort to small area estimation method to try to improve the efficiency of the poverty ratio estimates.

The availability of data lead us to the choice of an area-level model. Many proposal of this approach are described in literature, see for example Rao and Molina (2015); Pfeffermann (2013). However, giving the focus on cross-sectional data and the absence of spatial correlation in the conditional distribution of the target given the auxiliary variables, we decide to use the basic area-level model, also know as Fay-Herriot model (Fay and Herriot, 1979). In what follows a short description of the method is given.

Let us assume that there are m small areas of interest and that θ_i represents the target parameter of the area i, such as a mean or a proportion. A survey provides a direct estimator $\hat{\theta}_i^{dir}$ of θ_i for some or all of the small areas. As usual, we assume that under the sampling design $E[\hat{\theta}_i^{dir}] = \theta_i$. A p-vector $\mathbf{X}_i$ contains the auxiliary data sources of population characteristics for area i.

Let us assume that the auxiliary variables $\mathbf{X}_i$ are known exactly. The FH model is as follows

$$\hat{\theta}_i^{dir} = \mathbf{X}_i^T \beta + u_i + e_i \quad i = 1, \ldots, m, \tag{11}$$

where $u_i \overset{iid}{\sim} N(0, \sigma_u^2), i = 1, \ldots, m$ are the model errors and $e_i \overset{ind}{\sim} N(0, \psi_i^2), i = 1 \ldots, m$ are the design errors, with e_i independent from u_j for all i and j. It is assumed that the quantity of interest in area i is $\theta_i = \mathbf{X}_i^T \beta + u_i$.

Under the assumption of normality of both the errors (model and sampling design), the best linear unbiased predictor of θ_i is

$$\tilde{\theta}_i^{FH} = \gamma_i \hat{\theta}_i^{dir} + (1 - \gamma_i) \mathbf{X}_i^T \tilde{\beta}, \quad \gamma_i = \frac{\sigma_u^2}{\sigma_u^2 + \psi_i^2}, \tag{12}$$

where $\tilde{\beta}$ is the Best Linear Unbiased Estimator of β. The predictor $\tilde{\theta}_i^{FH}$ is a convex combination of

the direct estimator $\hat{\theta}_i^{dir}$ and of the predicted value $\mathbf{X}_i^T\tilde{\beta}$ from the regression model. The extent to which it depends on the the direct estimator or on the predicted value for the area is determined by γ_i and hence by the relative sizes of the model error variance σ_u^2 and the sampling error variance ψ_i^2.

According to the theory of small area estimation (Rao and Molina, 2015), the parameters β and σ_u^2 are unknown and must be estimated, while ψ_i^2 is assumed to be known. The estimators of the ψ_i^2s are often smoothed, and the smoothed estimators are treated as if they were the true sampling variances (Datta et al., 2005).

Estimators of β and σ_u^2 can be obtained using the restricted maximum likelihood from the marginal distribution $\hat{\theta}_i^{dir} \sim N(\mathbf{X}_i^T\beta, \sigma_u^2 + \psi_i^2)$ (Rao and Molina, 2015, see paragraph 6.2.4). By plugging in the estimates of β and σ_u^2 into equation (12) we obtain the empirical best linear unbiased predictor (EBLUP)

$$\hat{\theta}_i^{FH} = \hat{\gamma}_i\hat{\theta}_i^{dir} + (1-\hat{\gamma}_i)\mathbf{X}_i^T\hat{\beta}, \quad \hat{\gamma}_i = \frac{\hat{\sigma}_u^2}{\hat{\sigma}_u^2 + \psi_i^2}. \tag{13}$$

The terms $\hat{\gamma}_i$ are commonly known as *shrinkage* factors.

When all the parameters (σ_u^2, β) are known the mean squared error (MSE) of the estimator (12) is

$$MSE(\tilde{\theta}_i^{FH}) = E[(\tilde{\theta}_i^{FH} - \theta_i)^2] = \gamma_i\psi_i^2 = g_{1i}. \tag{14}$$

When the parameters in (12) are estimated we obtain the estimator (13) that has the following MSE

$$\begin{aligned} MSE(\hat{\theta}_i^{FH}) &= \gamma_i\psi_i^2 + (1-\gamma_i)^2\mathbf{X}_i^T V(\hat{\beta})\mathbf{X}_i + \psi_i^4(\psi_i^2+\sigma_u^2)^{-3}V(\hat{\sigma_u^2}) \\ &= g_{1i} + g_{2i} + g_{3i}, \end{aligned} \tag{15}$$

where g_{2i} is the contribution to the MSE from estimating β and g_{3i} is the contribution to the MSE from estimating σ_u^2. In (15) $V(\hat{\beta})$ and $V(\hat{\sigma_u^2})$ are the asymptotic variances of an estimator $\hat{\beta}$ of β and an estimator $\hat{\sigma_u^2}$ of σ_u^2, respectively. An estimator of (15) is as follows

$$mse(\hat{\theta}_i^{FH}) = \hat{g}_{1i} + \hat{g}_{2i} + 2\hat{g}_{3i}, \tag{16}$$

where $\hat{g}_{1i} = \hat{\gamma}_i\psi_i^2$, $\hat{g}_{2i} = (1-\hat{\gamma}_i)^2\mathbf{X}_i^T[\sum_{i=1}^m \mathbf{X}_i\mathbf{X}_i^T/(\psi_i^2+\hat{\sigma}_u^2)]^{-1}\mathbf{X}_i$, $\hat{g}_{3i} = \psi_i^4(\psi_i^2+\hat{\sigma}_u^2)^{-3}2[\sum_{i=1}^m 1/(\hat{\sigma}_i^2+\psi_i^2)^2]^{-1}$. More details concerning analytic MSE estimation for area level model can be found in Rao and Molina (2015); Datta and Lahiri (2000); Prasad and Rao (1990).

By applying the described method we get EBLUP of poverty rate at provincial level in Italy using the adjusted poverty lines (used to compute direct estimates). As auxiliary variables we used the ratio between number of taxed persons over the population, and the ratios between the number of persons with *i.* income coming from salary, *ii.* income coming from pensions and *iii.* income lower than 10,000 euros per year, over the number of taxed persons. These data come from the Italian tax agency database 2017.

The normality assumption of area-level errors of the Fay-Herriot model has been tested graphically, see figure 8, and by Shapiro normality test that result in a p-value of about 0.03. Therefore, considering also the robustness of the model to this assumption (Datta et al., 2005), we think the normality assumption is reasonable for our model.

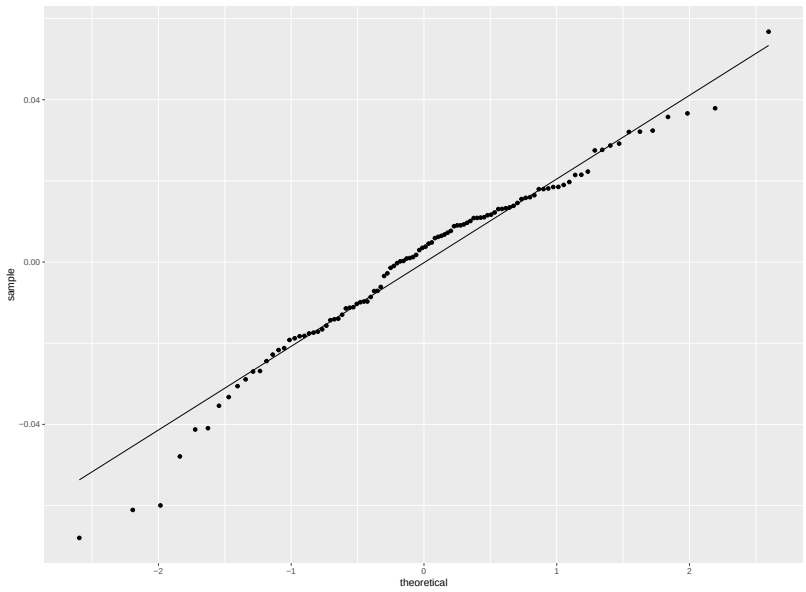


Figure 8: QQ-plot of estimated area-level error of Fay-Herriot small area model.

The EBLUPs show a gain in efficiency with respect to direct estimates. We obtained a CV smaller than 16% in 37 provinces, while half of the provinces have a CV smaller than 20%. The gain in term of variability is shown in figure 9 where we can see that the EBLUP is more efficient than the Direct estimator in all the provinces and the gain in efficiency is greater in those areas where the sample size is smaller, as expected.

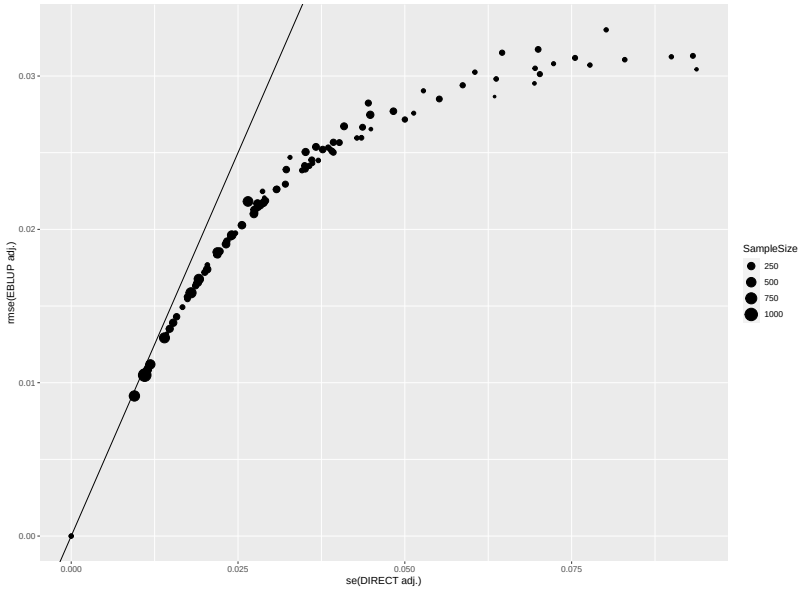


Figure 9: Estimated standard error of the Direct against estimated Root MSE of the EBLUP

We also computed the EBLUP without any adjustment of the poverty line to compare the poverty ratio with and without adjustment. Unadjusted EBLUPs have been obtained using the same model as for adjusted EBLUP[3]. The results of the comparison are showed in figure 10, which show that the

[3]The adjustment is referred to the nPL

HCR of southern provinces estimated using the adjusted poverty line is lower than the HCR obtained using the nPL, while for central and northern provinces is the contrary, as we expected given the perception that in the south of Italy food goods and house rents are cheaper.

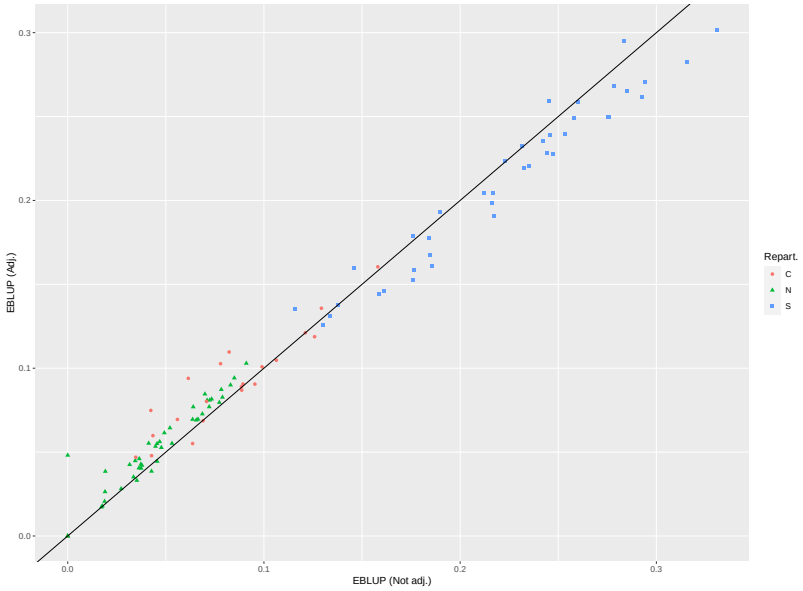


Figure 10: Poverty rate at provincial level in Italy: EBLUP not adjusted against EBLUP adjusted.

5 Summary and Outlook

In this deliverable we addressed the issue of intra-country comparisons of poverty indicators using a statistical methodological approach.

We first considered the estimation of a relative poverty indicator, namely the Head Count Ratio or At Risk of Poverty Rate at regional level (NUTS-2) in Italy. At this level the sample sizes of the EUSILC and HBS surveys are large enough to allow the estimation of reliable indicators.

When measuring the poverty incidence, the use of a national poverty line allows to establish a general scheme of how regions compare with national standards. However, considering the same poverty line for each of the regions implies an equity concept in which individuals with equal income are assumed to have similar wellbeing regardless of the region where they live. The use of regional poverty lines allows to gauge intra-regional poverty, which can be important for planning regional policies.

Apart from the value of the HCR itself, it is important to understand if and to what extent the patterns of poverty change when using regional poverty lines instead that a national poverty line. Identifying the households with higher probabilities of being poor is indeed an important issue especially when planning decentralized policies. Our results using 2017 EUSILC data provide general support to the notion that regional levels of poverty change with each threshold, and regional rankings do not remain robust to the choice of poverty line. A second important finding is that poverty profiles in the 20 Italian regions do not vary as different lines are used.

Under a different approach aiming at taking into account the different price levels within the country, the national poverty line can be modified using area-specific Purchasing Power Parities (PPPs), following the methodology currently applied at international level for international comparisons among different countries. In this work we proposed the use of regional PPPs defined as Spatial Housing Price Indexes (SHPI) estimated using hedonic regressions applied to HBS 2017 data. Adjusting the national poverty lines using the SHPIs values - properly weighted - allowed to derive regional poverty lines that represent approximately the same standard of living across the different areas. Given the promising results obtained at regional level, the analysis was repeated at provincial

level (NUTS-3). To obtain reliable estimates for the 110 Italian provinces using HBS data we resorted to Small Area Estimation (SAE) methods, that use auxiliary variables to build model-based estimates with increased efficiency with respect to direct estimates (i.e. estimates based only on survey data). By estimating Spatial Housing Price Indexes at provincial level, provincial poverty lines representing the same standard of living were estimated.

The results obtained at regional and provincial level suggest that the methodology can be extended to include other Spatial Price Indexes, therefore adjusting the national poverty line with other components of households' consumption expenditure. Indeed, our results suggest that the price payed by households to rent the house represent and important share of the total household consumption expenditure, but that in Italy this share is approximately equal only to the 20%. Therefore, by including other consumption expenditure components, such as for example the expenditure for food, the national poverty line could by adjusted in a more complete manner. The study of the patterns of poverty - already applied to "standard" regional poverty lines - could then be extended to the case of subnational cost-adjusted poverty lines. It is important to underline that, being based on EUSILC and HBS data, the proposed analyses - applied here only to Italian data - could be extended to other European countries provided that survey data at subnational level are available.

References

Ayala, L., A. Jurado, and J. Pérez-Mayo (2014). Drawing the poverty line: do regional thresholds and prices make a difference? *Applied Economic Perspectives and Policy 36*(2), 309–332.

Azevedo, J.P. (2018). *Small Area Estimation of Poverty in the EU: Lessons Learned*. Workshop on Small Area Methods and living conditions indicators in European poverty studies in the era of data deluge and Big Data. Pisa, 8 May 2018: World Bank Group.

Biggeri, L., C. Giusti, M. Pratesi, and S. Marchetti (2018). Poverty indicators at local level: Definitions, comparisons in real terms and small area estimation methods. *Statistics and Applications 16*, 351–364.

Datta, G. and P. Lahiri (2000). A unified measure of uncertainty of estimated best linear unbiased predictors in small area estimation problems. *Statistica Sinica 10*, 613–627.

Datta, G., J. N. K. Rao, and D. Smith (2005). On measuring the variability of small area estimators under a basic area level model. *Biometrika 92*(1), 183–196.

Eurostat (2018). *Methodological manual on territorial typologies*. Luxembourg: Publications Office of the European Union: Eurostat.

Fay, R. and R. Herriot (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.

Istat (2010). *La differenza nel livello dei prezzi al consumo tra i capoluoghi delle regioni italiane*. Roma, Italy: Istat, Italian national statistical office.

Jolliffe, D. (2006). 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 E. B. Prydz (2015). *Global poverty goals and prices: how purchasing power parity matters*. The World Bank.

Kangas, O. and V. Ritakallio (2007). Relative to what?: Cross-national picture of european poverty measured by regional, national and european standards. *European Societies 9*(2), 119–145.

Marchetti, S. and L. Secondi (2017). Estimates of household consumption expenditure at provincial level in italy by using small area estimation methods: 'real' comparisons using purchasing power parities. *Social Indicators Research 131*(1), 215–234.

Mogstad, M., A. Langørgen, and R. Aaberge (2007). Region-specific versus country-specific poverty lines in analysis of poverty. *The Journal of Economic Inequality 5*(1), 115–122.

Molina, I. and Y. Marhuenda (2015, jun). sae: An R package for small area estimation. *The R Journal 7*(1), 81–98.

OECD (2015). *In It Together: Why Less Inequality Benefits All*. Paris: OECD Publishing.

Pfeffermann, D. (2013, 02). New important developments in small area estimation. *Statist. Sci. 28*(1), 40–68.

Prasad, N. and J. Rao (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*. Wiley.

R Core Team (2019). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rao, J. and I. Molina (2015). *Small Area Estimation*. Wiley Series in Survey Methodology. Wiley.

Renwick, T. (2009). Alternative geographic adjustments of us poverty thresholds: Impact on state poverty rates. *Washington, DC: US Census Bureau*.

Renwick, T., B. Aten, E. Figueroa, and T. Martin (2014). Supplemental poverty measure: A comparison of geographic adjustments with regional price parities vs. median rents from the american community survey. Technical report, Bureau of Economic Analysis.

Simler, K. (2016). *Pinpointing poverty in Europe: New evidence for policy making*. World Bank.

Suits, D. (1984). Dummy variables: Mechanics v. interpretation. *Review of Economics and Statistics 66*, 177–180.

World Bank (2017). *Purchasing Power Parities and the Size of World Economies. Results from the 2017 International Comparison Program*. Washington, DC: World Bank Group.

Discussion Papers
Collana del Dipartimento di Economia e Management, Università di Pisa

Email della redazione: lfanti@ec.unipi.it