Small area poverty indicators adjusted using local price indexes

F. Schirripa Spagnolo[1,2] G. Bertarelli[1,2] C. Giusti[1,2] M. Pratesi[1,2] S. Marchetti[1,2]

[1]Dept of Economics and Management, University of Pisa, Italy

[2]Tuscan Universities Research Centre 'Camilo Dagum' on Advanced Statistics for the Equitable and Sustainable Development

ESRA 2019
Zagreb, 15-19 July 2019

Aims of the presentation

- Estimate Italian households' At-Risk-of-Poverty-Rate (ARPR) in the Italian provinces (NUTS-3 level) in order to provide useful data for policy interventions against poverty at sub-national level in Italy.

- Take into account the local cost of living using alternative data sources (*Big Data*).

The use of consumption data to estimate the ARPR

- We use consumption expenditures data from the Italian Household Budget Survey (HBS) 2017 to estimate the ARPR in the 110 Italian provinces.
- Italian HBS is carried out by the National Institute of Statistics (Istat) and it is used to compute the relative and absolute poverty incidence in Italy.
- HBS does not allow for reliable provincial estimates, therefore it is necessary to use small area estimation techniques.

Spatial variation of cost of living

- Classical approach: individuals or families are poor if their income or consumption level falls below a minimum level (called poverty line, PL).
- Increasing attention to the the use of spatial price indexes for accounting the spatial variation of cost of living in order to make comparisons in real terms between the different areas (Biggeri and Pratesi, 2017; Giusti et al., 2017; Marchetti and Secondi, 2017; Laureti and Rao, 2018).
- Spatial price indexes are measures of differences in price levels across areas, essential to compare economic well-being indicators.

Spatial variation of cost of living (cont.)

- Italy is a Country characterized by strong geographical differences in term of cost of living.
  - In 2009, ISTAT estimated sub-national Purchasing Power Parities (PPPs) for Italian regional capital cities showing that differentials in consumer price levels between the Italian cities are significant.
- Poverty thresholds should be adjusted for price differences across geographic areas in order to represent approximately the same standard of living across the different areas.

Provincial poverty lines

Adjustments in the national poverty line

$$PPL_j = NPL \cdot (\alpha_j \cdot HI_j + \beta_j \cdot GI_j + (1 - \alpha_j - \beta_j))$$

- PPL_j: Adjusted PL in the province j (Provincial PL);
- NPL: national PL;
- HI: Rent Index;
- α_j: rent expenditure weight;
- GI: Grocery Index;
- β_j: grocery expenditure weight.

Numbeo

- The values of the Rent and Grocery Index are computed using data from Numbeo.
- *Numbeo is the world's largest database of user contributed data about cities and countries worldwide. Numbeo provides current and timely information on world living conditions including cost of living, housing indicators, health care, traffic, crime and pollution..*
- It is a free and crowd-sourced global database of reported consumer prices. To collect data Numbeo relies on user inputs and manually collected data from authoritative sources.
- New source of data: big amount of data stored; daily data; heterogeneous data (different sources).

Numbeo (cont.)

- For our purpose we use the a tools that allow us to compare cost of living for two cities.
- Indexes for a city are provided in comparison to an other city. We decide to use Rome (the Capital of Italy located in the Central Italy) as benchmark and we assume that the price measured in the provincial capital is fixed over the entire province.
- Rent Index is an estimation of prices of renting apartments in each provincial capital compared to Rome (if Rent index is 80, Numbeo estimates that price of rents in the provincial capital is on an average 20% less than the price in Rome).
- Groceries Index is an estimation of grocery prices in each provincial capital compared to Rome.

Estimation of rent/grocery weights

- Direct estimates of the weights (α_j, β_j) at provincial level can be considered reliable.
  - In SAE, provinces are in a *grey zone*: the sample size is not so small (the minimum sample size is around 20) and sometimes direct estimates can be statistically sound (CVs not so high: no area has a CV greater than 33.3%; 15% of areas have a CV higher than 16.6%).
- PL_j are not SAE estimates.
- For the three out-of-sample provinces, we use the mean of the Region.
- Adjusted Provincial Poverty Lines (PPLs) range from 1131.97 in the province of Milan (Lombardy) to 853.42 in the province of Catanzaro (Calabria). The NPL is 1102.52.

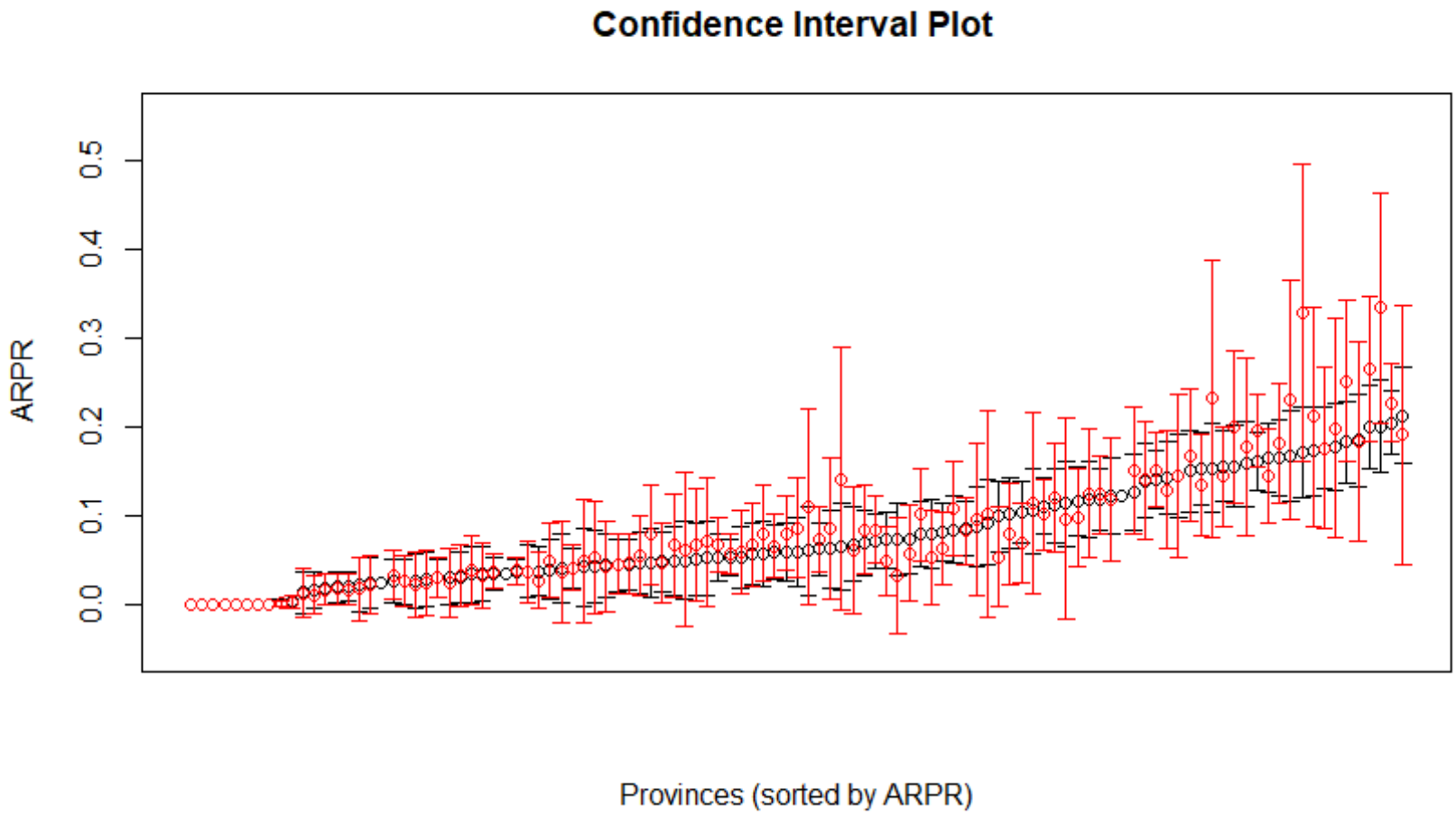
FH model ARPR in Italy using NPL and PPLs

- Direct estimates of the ARPR at province level using both NPL and PPLs are not reliable.
- In order to reduce the variability of the direct estimates we employ a FH model (Fay and Herriot, 1979).

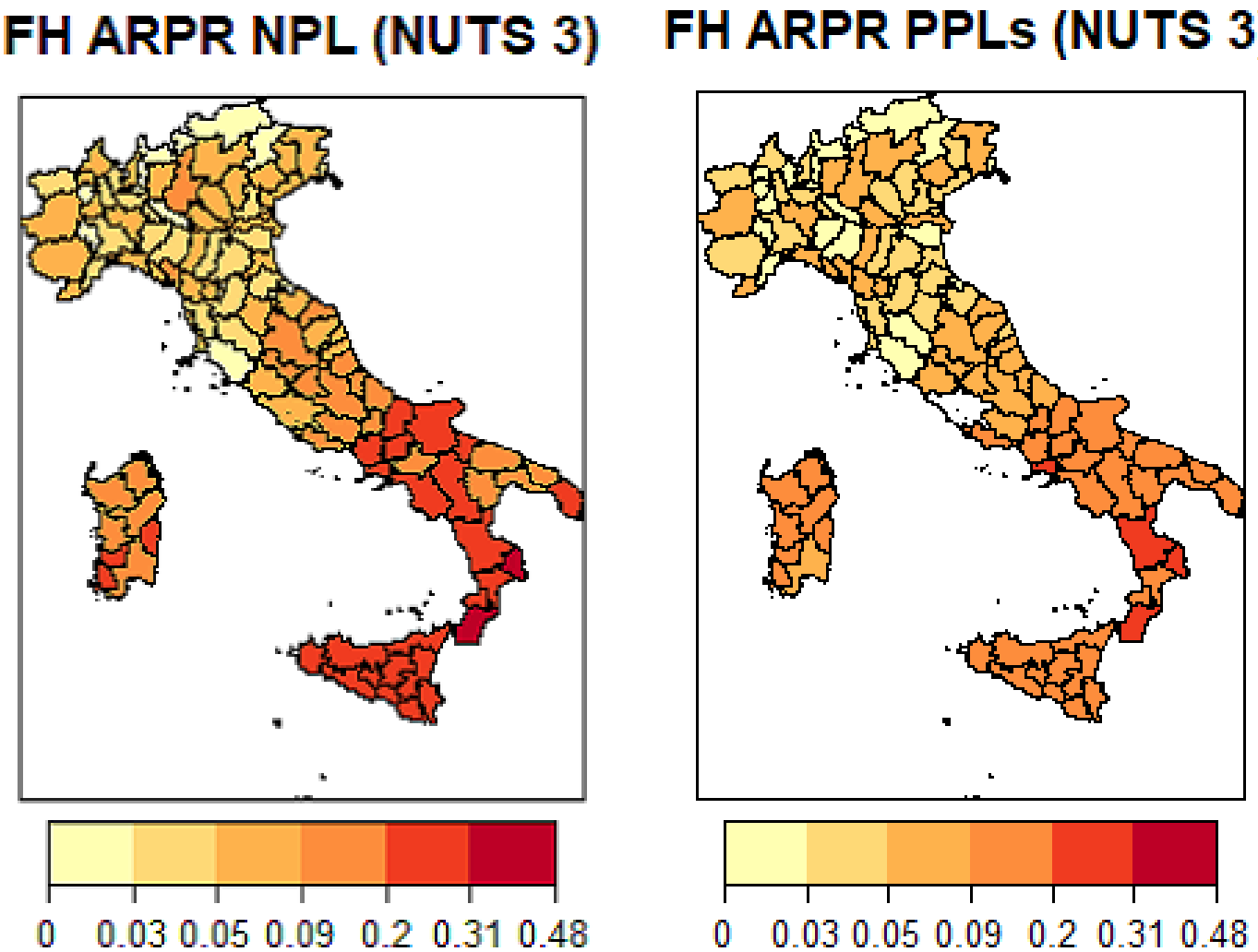
  FH model for the ARPR using PPLs (auxiliary variables come from the Tax agency administrative database).

Variable	Estimate	p-values
Intercept	0.403	0.000
Tax payers (%)	-1.369	0.000
Income of employees (%)	0.864	0.012
Income 0-10000 (%)	0.691	0.000
Retirement income (%)	0.665	0.064

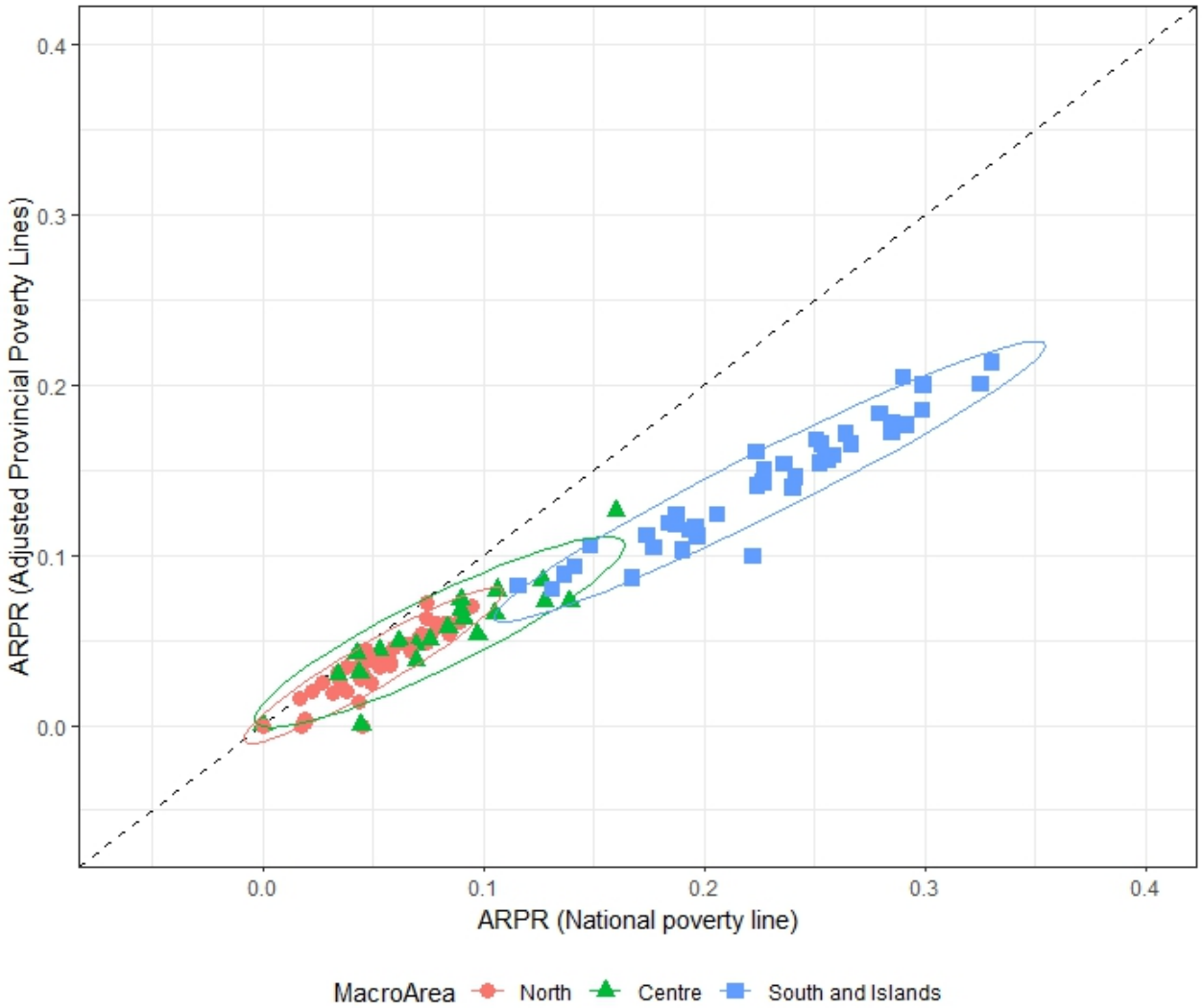
Confidence Interval Direct and FH Estimates (PPLs)



ARPR at provincial level with NPL and PPLs



ARPR at provincial level with NPL and PPLs (cont.)



Final remarks

- Estimates that include the cost of living can be useful for making comparison in real terms.
- Take into account the rent and the grocery prices, the poverty rates tend to be lower than the ones obtained with a fixed national poverty line.
- Cost of living in the provinces has an impact on the ARPR.

Future work

- The aim is to extend the present work by using two alternative new data source: the scanner data (sub-regional retail volumes and price for food and grocery) to estimate spatial price indexes for grocery and the transaction prices of houses made available by the Revenue Agency (OMI) to compute spatial price indexes for housing.
- In particular, scanner data will be provided us by Istat and these stored unit level transaction belonging to the 16 most important retail chains (95% of modern retail chain distribution in Italy).
- Grocery products: five divisions of the ECOICOP.
- These data might allow us to compute poor-specific price indexes and study in depth the living conditions of the poor population.

Essential Bibliography

Biggeri, L. and Pratesi, M. (2017). Monetary poverty indicators at local level: definitions, methods of estimations and comparisons in real terms. *Proceedings of the 60th World Statistics Congress of the International Statistical Institute, 2017*.

Fay, R.E., and Herriot, R.A. (1979). Estimation of income from small places: An application of James Stein procedures to census data. *JASA* 74, 269–277.

Giusti, C., Pratesi, M. and Masserini, L. (2017). Local Comparisons of Small Area Estimates of Poverty: An Application Within the Tuscany Region in Italy. *Soc Indic Res* 131(1), 235–254.

Laureti, T. and Rao, D.S.P. (2018). Measuring Spatial Price Level Differences within a Country: Current Status and Future Developments. *Estudos de Economia Aplicada* 36(1), 119–148.

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. *Soc Indic Res* 131(1), 215–234.

Francesco Schirripa Spagnolo

Department of Economics and Management
University of Pisa

francesco.schirripa@ec.unipi.it

This work is carried out with the support of the project MAKSWELL and of the project PRA_2018_9 'From survey-based to register-based statistics: a paradigm shift using latent variable models'