Small area poverty indicators adjusted using local price indexes

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

STATISTISCHE WOCHE 2019
Trier, 10–13 September 2019

Overview

Introduction

Estimation of Head Count Ratio (HCR) at provincial level (NUTS 3) in Italy

Estimation of HCR at provincial level considering the different purchasing power

Results

Conclusions and future work

Aims of the presentation

- *Head count ratio* (HCR) at provincial level (NUTS 3) in Italy
  - Useful for policy interventions against poverty at sub-national level
  - Need of efficient estimates to discriminate among provinces
- HCR is a measure of monetary poverty incidence
- Money do not have the same purchasing power *within* the country
- HCR should consider this

Consumption data to estimate HCR

- HCR is usually based on income or consumption expenditure
- We use consumption expenditures data from the Italian Household Budget Survey 2017 (HBS)
- HBS is carried out by Istat with a sample survey of about 22 000 households
- Used to estimate the official HCR in Italy
- Also used to estimate the *absolute poverty incidence*

Consumption data to estimate HCR

- HBS is designed to get sound estimates at regional level (NUTS 2)
- HBS sample size at provincial level: 20 – 1024 units (median 125)
- Depending on the target, estimates at provincial level may results unreliable in some provinces
- Model-based small area estimators are used in these cases

HCR: definition

- HCR: proportion of persons whose consumption expenditure is below a given threshold
- Consumption expenditure is observed at household level
- Specific threshold for each household size
- All the persons in the same household are considered poor if the household consumption expenditure is below the specific threshold

HCR threshold(s)

- Threshold for household with *two persons*: the mean per-capita consumption expenditure (at national level)
- This threshold is called *(national) poverty line* (NPL)
  - Thresholds for different household sizes are obtained adjusting the NPL with a specific scale
  - *One person* household: NPL$\times$0.60
  - (*Two person* household: NPL$\times$1)
  - *Three person* household: NPL$\times$1.33
  - *Four person* household: NPL$\times$1.63
  - *Five person* household: NPL$\times$1.90
  - *Six person* household: NPL$\times$2.16
  - *Seven or more person* household: NPL$\times$2.40

Purchasing power within the country

- Take into account different purchasing power allows to make comparisons in real terms among areas
- Sub-national Purchasing Power Parities (PPPs) measure the differences in price levels across areas and are essential to compare monetary-based well-being indicators
- Italy is characterized by strong geographical differences in term of cost of living
  - In 2009 Istat estimated sub-national PPPs for regional capital cities
  - Differentials in consumer price levels between these cities were significant

Poverty thresholds should be adjusted for price differences across areas

- Some references: Biggeri and Pratesi, 2017; Giusti et al., 2017; Marchetti and Secondi, 2017; Laureti and Rao, 2018

Purchasing power within the country

- As already discussed (Biggeri and Pratesi, 2017; Giusti et al., 2017) to account for the cost of living differences in the comparison of poverty between different sub-national areas, it should be necessary to compute and use **poor-specific sub-national PPPs**
- Until now, no country has computed those indexes
- The second best could be, as done in USA by Bishop et al. (2017) and Renwick et al. (2014), to use a proxy based on the combination of a general sub-national PPPs and a cost of housing index, on the hypothesis that this last cost is the most important issue to be faced by poor
- We have planned to do it by using the scanner and CPI data provided by Istat and computing a cost of housing index by using the HBS data

Purchasing power within the country: a provisional solution

- Unfortunately, due to confidentiality restrictions, Dagum researchers had not access to Istat elementary prices data in time for this presentation
- Therefore, we decide to conduct an exercise (experiment) by using alternative sources of data
- We consider the indexes provided by the very large **database NUMBEO**, referred to many cities in various contries, including Italy

Numbeo

As stated by Economist Intelligence Unit, within the organisation which publish "The Economist" newspaper

> *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*

- In particular they computed the following indexes, which are of interest for us:
  - Cost of Living index
  - Cost of Living Plus Rent Index
  - Rent index
  - Groceries Index
  - Local Purchasing Power

Numbeo database drawbacks

- The Numbeo database source is surely interesting at our ends
- However, the database does not include information on metadata to evaluate the statistical quality of the computed indexes
- Some information on the methods followed by Numbeo can be found in `www.numbeo.com/common/motivation_and_methodology.jsp`
- Numbeo says to use heuristic technology to get "data quality" and that it "periodically discards data which most likely are incorrect statistically"
- But many procedures are unclear (i.e. it is unclear how Numbeo selects goods and how it collects prices, and in which way the collected data guarantee comparability among countries and regions)

Numbeo database drawbacks

- The lack of the necessary metadata may depend on the Numbeo commercial goal: to provide to companies and business man cost of living indexes (also tailor-made) for cost of living allowances
- In any case, with the information available, it is impossible to do now a detailed quality analysis of the Numbeo spatial price indexes
- However, considering the description of the indexes it is clear that neither the cost of living indexes, nor the rent index and the local purchasing power can be used at our aim
- We conducted an external evaluation of the various indexes comparing them with the PPPs computed by Istat in 2009 for the regional capital cities. We found an hight correlation between the PPPs and the Numbeo Groceries Index

Numbeo Groceries Index

- Numbeo Groceries Index is an estimation of grocery prices in the different cities
- Considering Rome as base $= 1$ the computation with the Numbeo data for 2018/19, show that:
  - The minimum and the maximum for the regional capitals are 0.600 (Catanzaro, Calabria) and 1.266 (Aosta, Valle d'Aosta)
  - The minimum and the maximum for the provincial capitals are the same
- To carried out the exercise we have in mind, we consider the groceries indexes as a proxy (also if inaccurate) of the sub-national PPPs

A preliminary estimation of rent spatial indexes

- For the time being, an House Price spatial index (HPSI) or better a spatial rent index are not published in Italy
- Therefore, we tried to carried out a preliminary estimation of the spatial rent indexes, by using the data collected with the HBS conducted annually by Istat
- The HBS collects information on the house occupied by each household, referred to the location of the building, and to many characteristics of the building and of the apartment (size, number of rooms, etc.)

A preliminary estimation of rent spatial indexes

- To make preliminary estimates of the rent spatial indexes, we use a classical hedonic equation as follows (Hill, 2012):

$$P = f(S, N, E)$$

  - P is the price rent of the house/apartment
  - S is the structural characteristics of the house
  - N is the locality/neighbourhood
  - E is the environmental characteristics

A preliminary estimation of rent spatial indexes

- We consider a sample of n independent observations of houses' rents r_{ij} $(i = 1, 2, \ldots, n)$ in the different areas (regions or provinces) j $(j = 1, 2, \ldots, M)$
- The semi-log formulation is as follows:

$$\ln r_{ij} = \sum_{j=1}^{M} \alpha_j A_j + \sum_{k=1}^{K} \sum_{h=1}^{H} \beta_{kh} C_{kh} + \epsilon_{ij}$$

  - A_j is a vector of geographical areas dummies
  - α_j is the vector of area prices
  - C_{hk} is the matrix of the characteristics (with $k = 1, \ldots, K$) and their classifications $(h = 1, \ldots, H)$
  - β_{kh} is the matrix of hedonic regression coefficients called also characteristic shadow prices
  - ϵ_{ij} is the error terms, that satisfy the standard assumption of a multiple regression model

A preliminary estimation of rent spatial indexes

- Once the α_j parameters are estimated, the HPSI of the area j with respect to the base area is given by

$$HPSI_j = \exp(\alpha_j)$$

- The computation of the above indexes, by using 2017 HBS data and considering Rome as base $=1$, shows that:
  - The minimum and the maximum for the capital regions are 0.502 (Potenza, Basilicata) and 1.361 (Bolzano)
  - The minimum and the maximum for the provincial capital are 0.445 (Enna, Sicilia) and 1.361 (Bolzano)

Estimation of the adjusted provincial poverty lines: Results

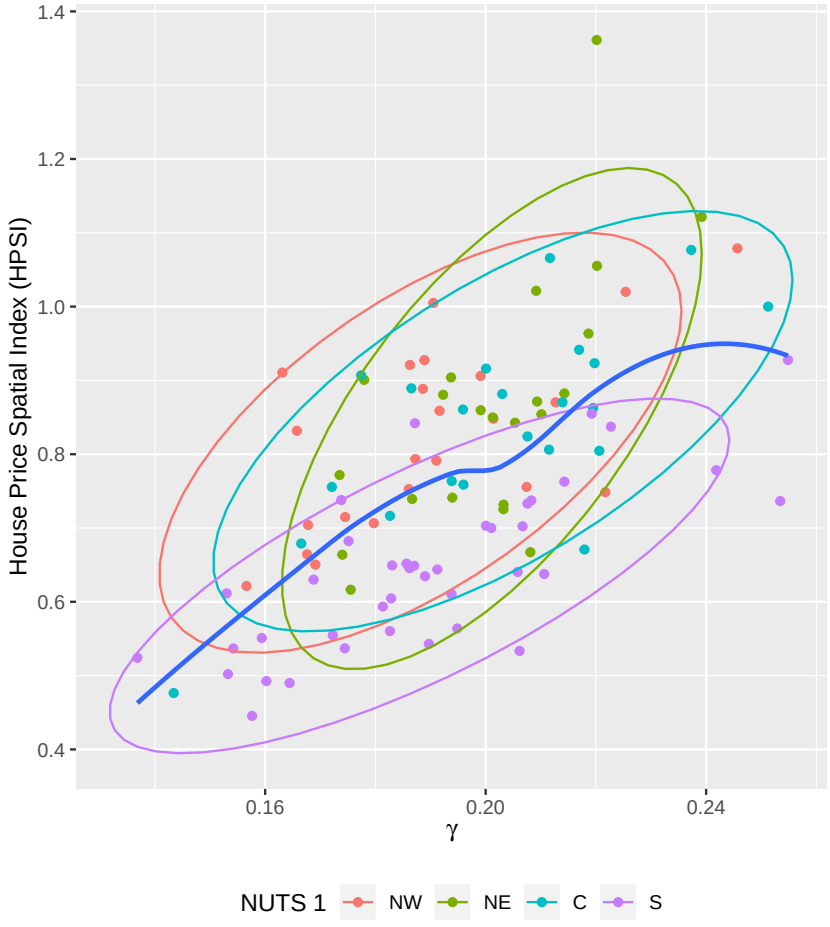
$$PPL_j = NPL \times \left(\gamma_j \times HPSI_j + \delta_j \times GI_j + (1 - \gamma_j - \delta_j) \times 1\right)$$

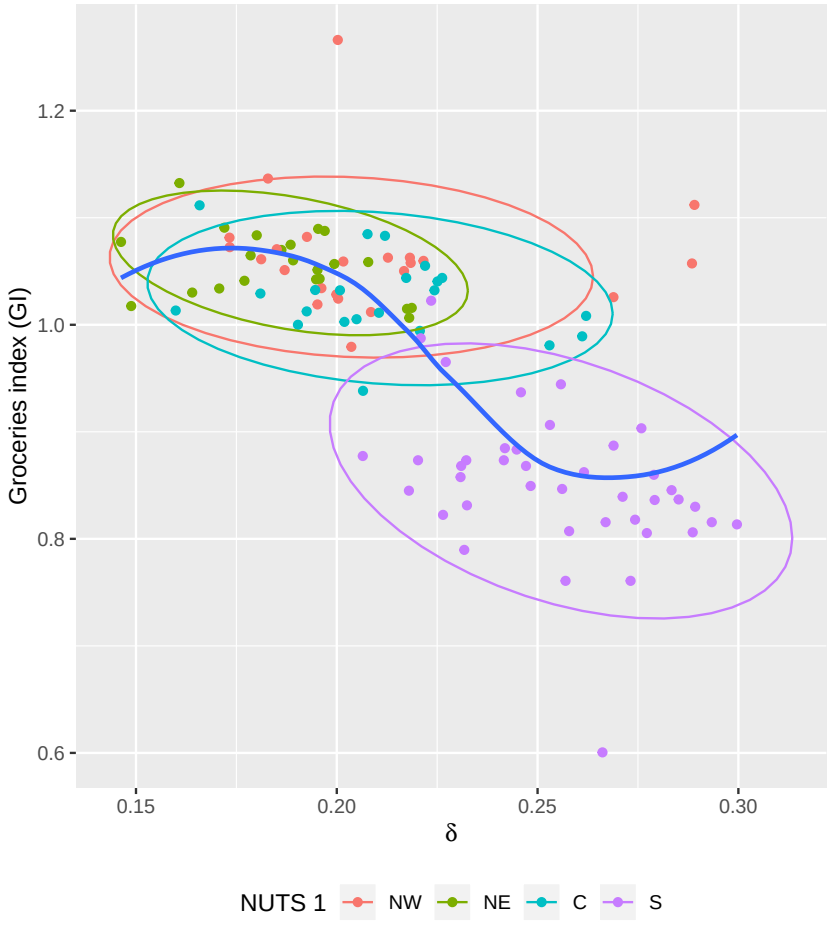
- PPL_j: Adjusted NPL in province j
- $HPSI_j$: Spatial price hedonic index in province j
- γ_j: Rent expenditure weight
- GI_j: Groceries index in province j
- δ_j: Groceries expenditure weight in province j

Estimation of the adjusted provincial poverty lines: Results

- Direct estimates of weights at provincial level have been judged reliable
  - In SAE, Italian provinces are often in a grey zone
  - In the case of share of consumption expenditure for house and groceries there is no need to resort on SAE
  - No area has a CV greater than 33.3% and only 15% of areas have a CV greater than 16.6%
- Main results:
  - PPLs depend on the choice of the base for the groceries and the house price spatial indexes, which is Rome
  - PPLs range from 1202.75 euro in the province of Bolzano to 891.60 euro in the province of Catanzaro (Calabria)
  - NPL is 1102.52 euro (PPL of Rome)

Indexes and weights



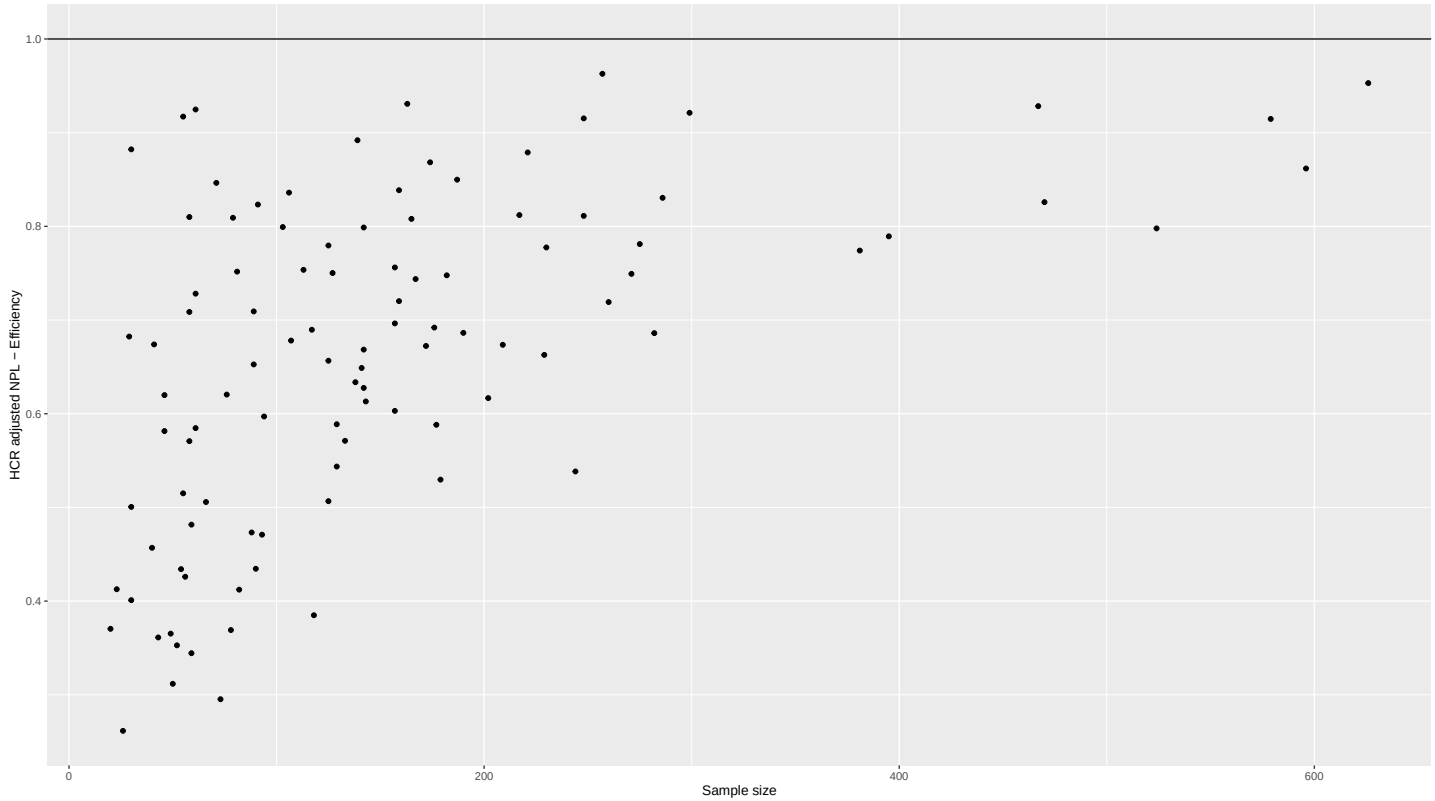


Estimation of Provincial HCR in Italy using NPL and PPLs

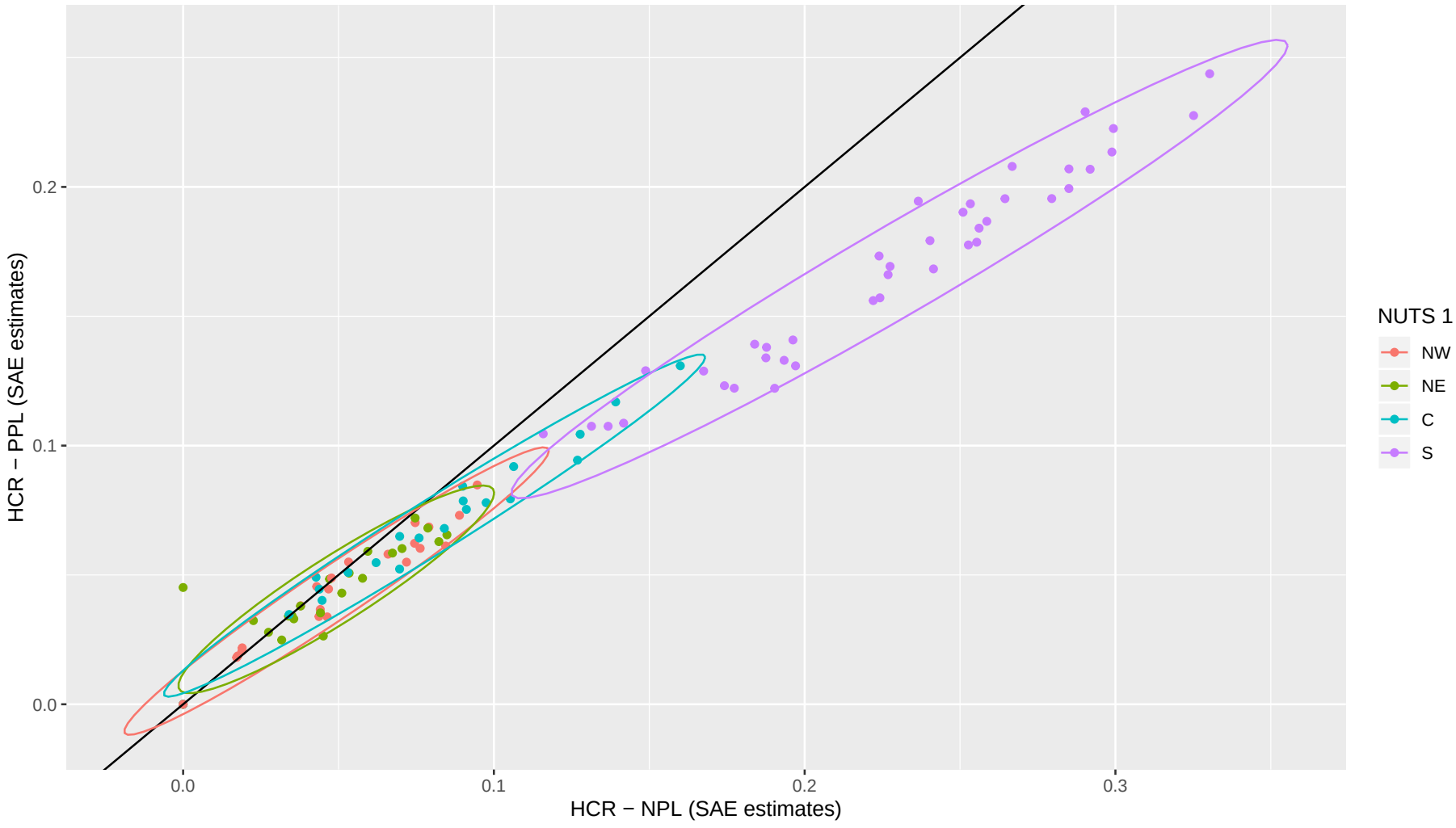
- HCR direct estimates at province level using both NPL and PPLs are not reliable
- In order to reduce the variability of the direct estimates we employ a basic FH model (Fay and Herriot, 1979).
  - FH model for the HCR using PPLs (auxiliary variables come from the Tax agency administrative database 2017).

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

Efficiency of small area estimates



HCR at provincial level with NPL and PPLs



Final remarks

- Cost of living can be useful for comparisons in real terms
- Cost of living in the provinces has an impact on the HCR (monetary-based measure)
- Also if the data are approximated, adjusting for rent and grocery prices causes substantial changes of the values of HCR
- There are spatial patterns:
  - Northern provinces: HCR general increasing
  - Souther provinces: HCR general decreasing

Future work

- Extend the present work in the MAKSWELL project
- *By using*:
  - Scanner data and CPI to estimate sub-national PPPs
  - Transaction prices of houses (available by the Revenue Agency) to compute spatial price indexes for housing
- *By improving*:
  - Hedonic regrerssion for spatial price indexes using HBS data
- There is work-in progress jointly with F. Polidoro (Istat's Head of Unit of integrated system on economic conditions and consumer prices) to use traditional and scanner data on price under an agreement for the Makswell Project

Final goal: adjust the NPL considering a poor-specific cost of living spatial index

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*.

Bishop J.A., Lee, J. and Zeager L.A. (2017). Incorporating spatial price adjustments in U.S. public policy analysis, ECINEQ WP n. 438, May 2017, `www.ecineq.org`.

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.

Hill, R. J. (2012). Hedonic price indexes for residential housing: A survey, evaluation and taxonomy, *Jour of Econ Surv*, 27 (5), 879–914.

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.

Renwick, T., Aten, B., Figueroa, E. and Martin T. (2014). Supplementary poverty measure: A comparison of Geographic Adjustments with Regional Parities vs. Median Rents from the American Community Survey, SEHSD Working Paper N. 2014–22.

MAKSWELL European H2020 Project

MAKing Sustainable development and WELL-being framework work for policy analysis

Proposes to extend and harmonise the indicators able to capture the main characteristics of the beyond-GDP approach proposing a new framework that includes them in the evaluation of the public policies

- Building up a database for a wide set of EU countries that selects and harmonizes the national framework on well-being as well as the available SDG indicators
- Improving the database both in relation to the timeliness and to the integration with big data measures and the methodologies able to reach these extensions
- Extending the geographical dimension especially focusing on the possible estimates of people vulnerability either looking at the measurement of poverty and inequality or to the measurement of regional inflation
- Using the extended database for policy evaluation at macro