Can a neighbour region influence poverty? A fuzzy and longitudinal approach

Può una regione confinante influenzare la povertà? Un approccio longitudinale e sfocato

Gianni Betti, Federico Crescenzi and Francesca Gagliardi

Abstract One of the most important goals of the 2030 UN Agenda for Sustainable Development is to *"…eradicate poverty, in all its forms and dimensions …"*. In order to give a comprehensive answer to such needs, in this paper we propose to adopt a longitudinal measure recently proposed by Verma *et al.* (2017), which is based on the fuzzy set approach to multidimensional poverty: the "Fuzzy At-persistent-risk-of-poverty rate"; then we propose to estimate this measure at regional level via small area estimation techniques, by introducing a spatial correlation model. In this way we are able to take into account whether a neighbour region can influence poverty in all its forms and dimensions, namely, the multidimensional dimension, the regional dimension and the longitudinal dimension.

Abstract Uno dei più importanti *goal* dell'Agenda 2030 per lo Sviluppo Sostenibile delle Nazioni Unite è *"…eradicate poverty, in all its forms and dimensions …"*. Con lo scopo di tentare di rispondere a questa necessità, nel presente lavoro proponiamo di utilizzare una misura longitudinale recentemente proposta da Verma *et al.* (2017), basata sull'approccio multidimensionale e sfocato per la misura della povertà: il cosiddetto "*Fuzzy At-persistent-risk-of-poverty rate*"; inoltre, proponiamo di stimare tale misura a livello regionale, tramite l'introduzione di un modello con correlazione spaziale tra gli errori. In questo modo prevediamo di catturare l'influenza che ha una regione confinante nella misura della povertà, in ogni sua dimensione, ovvero quella multidimensionale, quella regionale, ed infine quella longitudinale.

Key words: small area estimation, fuzzy longitudinal poverty, SILC, Spain.

[1] Gianni Betti, Department of Economics and Statistics, University of Siena, email: gianni.betti@unisi.it

Federico Crescenzi, PhD student, University of Bologna; email: federico.crescenzi2@unibo.it

Francesca Gagliardi, Department of Economics and Statistics, University of Siena, email: gagliardi10@unisi.it

1 Introduction

One of the most important goals of the 2030 UN Agenda for Sustainable Development is to *"…eradicate poverty, in all its forms and dimensions …"* (UN, 2015). This is particularly necessary after the global crisis started in 2008 and the failure of meeting the Millennium Development Goal of halving extreme poverty in the world by 2015.

The need of reducing poverty and launching anti-poverty programmes and policies has also been expressed by the European Union by the Europe 2020 Strategy (European Commission, 2010). This consists in a series of policy objectives called "headline targets", which should be reached by 2020. Among these targets, there is the reduction of the at-risk-of-poverty rate (ARPR, known in the literature as head count ratio, or FGT(0) in the family of Foster *et al*., 1984), and of the at-persistent-risk-of-poverty rate in a longitudinal context for monitoring poverty over time. Moreover, poverty measures are most useful to policy-makers and researchers when they are finely disaggregated, that is when they want to represent geographic units smaller than whole countries; this is exactly the purpose of DG Regional Policy of the European Commission, aiming to use sub-national/ regional level data (NUTS 2) for the social indicators used for monitoring the "headline targets" at the regional level.

In order to give a comprehensive answer to the needs reported above, in this paper we propose to adopt a longitudinal measure recently proposed by Verma *et al.* (2017), which is based on the fuzzy set approach to multidimensional poverty: the "Fuzzy At-persistent-risk-of-poverty rate"; then we propose to estimate this measure at regional level via small area estimation (SAE) techniques, by introducing a spatial correlation model. In this way we are able to take into account whether a neighbour region can influence poverty in all its forms and dimensions, namely, the multidimensional dimension, the regional dimension and the longitudinal dimension.

2 Longitudinal measures of fuzzy and multidimensional poverty

In this section we describe the construction of the fuzzy longitudinal poverty measures, which aim at estimating occasional, persistent or chronic concepts of poverty. In the fuzzy literature, these measures have been defined as: i) anytime, for those individuals belonging to fuzzy set poverty for at least one out four years; ii) continuous, for those belonging to all four years; moreover, we adopt a very recent definition proposed by Verma *et al.* (2017), namely iii) the "fuzzy at-persistent-risk-of-poverty", which refers to those individuals belonging to the fuzzy set in the most recent year, and to two or three years in the last three years: this measure is the fuzzy counterpart of the Eurostat "at-persistent-risk-of-poverty rate", one of the most important Laeken indicators.

From a mathematical point of view, let μ_t be the series of T = 4 membership functions over the four periods, we can define the anytime fuzzy measure as the fuzzy union over the periods, which consists in the maximum of the T values:

$$\mu_{any} = \max(\mu_1, \mu_2, \mu_3, \mu_4) \tag{1}$$

In the same way, the continuous fuzzy measure is defined as the fuzzy intersection over the periods, which consists in the minimum of the T values as:

$$\mu_{cont} = \min(\mu_1, \mu_2, \mu_3, \mu_4) \tag{2}$$

The definition of the fuzzy at-persistent-risk-of-poverty is much more complex, and we suggest to read Verma *et al.* (2017) for a full and detailed description.

3 Model based small area estimation

Sample sizes of surveys like EU-SILC, designed to be representative at national level, are frequently too low to get efficient estimates of indicators at small area level, like NUTS 2. In other words, it means that the measures calculated from such small sub-samples – termed in the related literature as “direct estimators” – have too large variances. Small area estimation theory is concerned with resolving these problems.

This classic EBLUP model proposed by Fay and Herriot (1979) can be extended by considering that the vector of errors follows a Simultaneously Autoregressive Process (SAR) with spatial autoregressive coefficient ρ and proximity matrix $\boldsymbol{W}$ (Cressie, 1993). In this way, the model with spatially correlated random effects is:

$$\hat{\boldsymbol{\theta}} = \boldsymbol{X\beta} + \boldsymbol{Z}(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{u} + \boldsymbol{e}. \tag{3}$$

The estimator is unknown because it depends on some unknown parameters, such as ρ. By substituting them with consistent estimators, a two stage estimator is obtained which can be referred to as a Spatial EBLUP. (see Pratesi and Salvati, 2007, for further details).

In order to estimate the Spatial EBLUP models it is necessary to have the standard errors of the direct estimator $\hat{\theta}$. Since the poverty measures adopted in the present paper are quite complex (such as, for instance, the “fuzzy at-persistent-risk-of-poverty rate”), which are calculated on the basis of a very complex survey such as EU-SILC, we estimate their standard errors by means of the Jackknife Repeated Replication (JRR) in the version of Verma and Betti (2011)[1].

[1] Verma and Betti (2011) demonstrate how a variant of the JRR method can fit better in case of “complex measures”; moreover, Betti *et al.* (2018) show that the JRR variant of Verma and Betti (2011) is particular adapt for estimating variance of fuzzy poverty measures.

4 Empirical analysis

The reference data for the present work are based on a subset of micro-data from the EU Statistics on Income and Living Conditions (EU-SILC) survey, which is the major source of comparative statistics on income and living conditions in Europe.

Generally, the EU-SILC national surveys are designed with focus on the production of reliable estimates at the national level. In fact, although EU-SILC survey has a very large sample in Spain (13,109 households and 34,756 individuals for 2011), the regional sub-samples are very heterogeneous in size, so that in some NUTS 2 regions estimates are not significant.

From the Spain EU-SILC 2011 Intermediate Quality Report (INE, 2012) we have a detailed description of the sample design that is important both for understanding whether regions form independent sampling domains, and for the construction of the ‘computational’ PSUs and strata, needed for the estimation of JRR standard errors.

Using such numerical data, here we present analysis of the direct estimates and their relative sampling errors for poverty and deprivation variables. The separate results for each of the 19 regions of Spain allow the input to the Fay and Herriot (FH) and the Spatial EBLUP models. The following two statistics are considered in turn in a longitudinal context: fuzzy monetary poverty rate (FM); and fuzzy supplementary deprivation rate (FS). For each one of these statistics, the longitudinal measures are those described in section 2: any-time poverty, continuous poverty, and at-persistent-of-risk rate.

From the results concerning standard errors of FM fuzzy at-persistent-risk-of-poverty rate, we can appreciate that both FH and SEBLUP are lower than the direct estimates. In general, we have a mean reduction of the standard error for FH of 18%, and for SEBLUP of 26%. The largest reduction standard errors are clearly found in regions with small sub-sample sizes, such as Melilla, Ceuta and Rioja.

However, for the main purpose of this paper, it is quite interesting to observe the larger gain in spatial EBLUP over FH; the geographical information of the w matrix of vicinity clearly supplies an evident added value, being clear that an increase in poverty in a neighbour region can affect the region under investigation as well.

From the results concerning standard errors of FS fuzzy at-persistent-risk-of-poverty rate, the reduction is smaller compared to the corresponding FM measure: 16% for FH and 13% for SEBLUP. In this case, the effect of the geographical information of the matrix of vicinity does not supply an added value.

Analyzing the FM anytime poverty rates, the highest values of are found in Extremadura, Castilla-La Mancha and Andalucia, while the lowest are in Navarra, Aragon and Baleares. The average gains in standard errors for FH and SEBLUB are very similar to the one found for FM at-persistent-risk-of-poverty rate. Again the largest reductions are found in the smallest regions of Melilla and Ceuta.

The measures of the FS anytime poverty rates show that the highest values are found in Galicia, Andalucia, Canarias and Murcia, while the lowest are in Melilla, Navarra, Aragon and Pais Vasco. Again the average gains in standard errors for FH and SEBLUB are very similar to the one found for FS at-persistent-risk-of-poverty rate. The largest reduction is found in the smallest region of Ceuta.

Table 1 reports the FM continuous poverty rates; the highest values of are in Extremadura, Ceuta, Castilla-La Mancha, Andalucia, Murcia and Rioja, the lowest are in Navarra, Pais Vasco, Asturias, Madrid and Melilla. In this case, the average gain in standard error is larger than ones found in all other measures, with mean reduction of about 23% for SEBLUP.

Table 1: FM continuous poverty rates

Region	Direct	se	SEBLUP	se	Gain
Galicia	7.91%	1.14%	8.10%	1.05%	91.59%
Asturias	5.11%	1.03%	5.70%	0.94%	90.98%
Cantabria	7.01%	1.46%	6.49%	1.21%	83.18%
País Vasco	4.43%	1.17%	3.75%	1.06%	91.12%
Navarra	2.13%	0.76%	2.63%	0.73%	96.64%
Rioja	11.24%	3.09%	7.44%	1.58%	51.10%
Aragón	4.77%	1.02%	4.86%	0.93%	91.21%
Madrid	5.59%	0.87%	5.17%	0.88%	101.11%
Castilla y León	8.50%	2.07%	9.04%	1.29%	62.42%
Castilla - La Mancha	13.63%	1.96%	13.31%	1.37%	70.07%
Extremadura	19.49%	1.95%	17.07%	1.55%	79.32%
Cataluña	6.18%	0.92%	6.24%	0.87%	94.36%
Comunitat Valenciana	7.86%	1.17%	8.16%	1.04%	88.80%
Balears	6.04%	1.87%	6.98%	1.59%	85.02%
Andalucía	13.80%	1.58%	13.82%	1.33%	83.91%
Murcia	11.08%	2.96%	11.79%	1.64%	55.29%
Ceuta	14.43%	7.55%	11.83%	2.26%	29.98%
Melilla	5.76%	6.01%	7.68%	2.07%	34.52%
Canarias	9.19%	2.03%	9.02%	1.67%	81.97%
					76.98%

5 Concluding remarks and further considerations

In this paper we propose a series of mathematical procedures to properly estimate longitudinal poverty and deprivation at regional level. First of all, we consider direct estimates of poverty and deprivation measured by means of fuzzy sets theory, and in particular we take into account the new measure "fuzzy at-persistent-risk-of-poverty rate", recently proposed by Verma *et al.* (2017); then we implant for the first time the procedure of Jacknife Repeated Replications in the version of Verma and Betti (2001) for estimating standard errors of such fuzzy direct estimates.

The primary result obtained is the extension of variance estimation to beyond measures of monetary longitudinal poverty, specifically to fuzzy formulation of those measures and, as a corollary, to multidimensional measures of longitudinal deprivation, which by their very nature are a matter of degree i.e. are fuzzy. To our knowledge, in the literature, no such extension has been published. Moreover, we propose to utilise SAE techniques, such as spatial EBLUP, to further reduce the variability of fuzzy poverty measures at regional level.

Overall, we can conclude that both FH and SEBLUP are able to reduce standard errors by 20-30% in average, with picks of 70% for regions where sample sizes are particular small. Moreover, the larger gain in spatial EBLUP over FH is evident only for FM longitudinal measures, while the gain is practically absent in FS ones, for which the geographical information of the w matrix of vicinity does not supply an added value. So, in conclusion, neighbour region can affect poverty only when we adopt a monetary measure, while it seems to unaffected in the case of a multidimensional or non-monetary measure: further research is necessary to understand reasons of such phenomenon.

References

1. Betti G., Gagliardi F., Verma V. (2018), Simplified Jackknife Variance Estimates for Fuzzy Measures of Multidimensional Poverty, International Statistical Review, doi:10.1111/insr.12219.
2. Cressie N. (1993), Statistics for Spatial Data, New York: Wiley.
3. European Commission (2010), Communication from the Commission. Europe 2020. A strategy for smart, sustainable and inclusive growth. Brussels, 3.3.2010 COM(2010) 2020.
4. Fay R.E., Herriot R.A. (1979), Estimates of income for small places: an application of James-Stein procedures to census data, Journal of the American Statistical Association, 74, pp. 269-277.
5. Foster J.E., Greer J., Thorbecke E. (1984), A class of decomposable poverty measures, Econometrica, 52, pp. 716-766.
6. Instituto Nacional De Estadistica (INE), (2012) Intermediate Quality Report, Survey on Income and Living Conditions Spain (Spanish ECV 2011).
7. Pratesi M., Salvati N. (2007), Small Area Estimation: The EBLUP model based on spatially correlated random effects, Statistical Methods and Applications, 17(1), pp. 113-141.
8. United Nations (2015), Transforming our World: The 2030 Agenda for Sustainable Development, A/RES/70/1, UNITED NATIONS.
9. Verma V., Betti G. (2011), Taylor linearization sampling errors and design effects for poverty measures and other complex statistics, Journal of Applied Statistics, 38(8), pp. 1549-1576.
10. Verma V., Betti G., Gagliardi F. (2017), Fuzzy Measures of Longitudinal Poverty in a Comparative Perspective, Social Indicators Research, 130(2), pp. 435-454.