EXTENDED ABSTRACT

Title: A nonstationary multidimensional estimation of poverty. The case of Morocco

Authors and e-mail of all:

Gianni Betti gianni.betti@unisi.it

Vasco Molini vmolini@worldbank.org

Lorenzo Mori lorenzo.mori7@unibo.it

Department:

Department of economics and statistics

\\

Department of Statistics P. Fortunati

University:

University of Siena

50x2030 Initiative

University of Bologna

Subject area: *(*Composite indicators and fuzzy indicators.*)*

Abstract: *(minimum1500 words)*

1.Introduction

In the last two decades Morocco has made progress in reducing economic inequalities (Boutayeb , 2006) and in the control of preventable childhood diseases (Achy, 2010). On the other hand, social inequalities and health inequities remain major problems for the third millennium (Achy, 2010). How can economic poverty be reduced while social inequalities increase? Answering this question is the principal aim of this work. In order to be able to do it at its best, it is necessary, from our point of view, to start with some definitions.

The status of poor defined as such on the basis of a single economic variable (generally monetary variable) and a national threshold (poverty line) hardly manages to capture all the nuances of a much more complex phenomenon such as the poverty one. One's poverty status is often the result of a plurality of simultaneous deprivations that go beyond the shortage of financial resources. For instance, a person who is not classified as income-poor can still suffer other deprivations, such as poor health, malnutrition, little schooling, or inadequate housing, all conditions that reduce well-being while increasing the risk of marginalization and social exclusion.

The multidimensional approaches to poverty are a well-known method for addressing such a complex issue. Ezzrari and Verme (2013) and El Bouhadi et al. (2012) present two multidimensional analysis of poverty in Morocco both based on multiple correspondence analysis (MCA). Despite this, they disagree about the results and whether or not poverty has increased in recent decades while agree that poverty is not constant over Moroccan's regions.

With the aim of combining a multidimensional approach and a high territorial disaggregation we propose to use a fuzzy measure of the incidence of relative poverty and deprivation and to estimate them at provinces level using EBLUP. With the former we can consider more than one dimension of poverty, not only the economical one, while the latter will allow us to obtain reliable estimates at provincial level taking account of the spatial correlation.

2. The fuzzy sets approach

Fifty years after the theorization of fuzzy sets, research and applications have been very active and have produced very important results in many fields, from decision making to engineering by away of medicine and economics. Several papers were published to evoke the concepts of fuzzy sets in the analysis of poverty and living condition, among them Betti et al. (2006), Betti and Verma (2008), Belhadj (2014). Cheli and Lemmi (1995) proposed "fuzzy monetary index" where the membership function was defined as:

$$\mu_i = FM_i = (1 - F_{(M),i})^{\alpha} = \left(\frac{\sum_{\gamma=i+1}^{n} w_{\gamma}|y_{\gamma}>y_i}{\sum_{\gamma=2}^{n} w_{\gamma}|y_{\gamma}>y_1}\right)^{\alpha}, i = 1, \dots, n \qquad (1)$$

where y_i is the income/expenditure of the i^{th} individual, $F_{(M),i}$ is the value of the distribution of income/expenditure for the i^{th} individual, $(1 - F_{(M),i})$ is the proportion of individuals less poor than the person i, w_{γ} is sample weight of person i with rank γ ordered distribution of income and α is a parameter.
In 1999, Betti and Verma replaced $(1 - \mathrm{F}_{(\mathrm{M}),\mathrm{i}})$ with $(1 - \mathrm{L}_{(\mathrm{M})\mathrm{i}})$ where $\mathrm{L}_{(\mathrm{M})\mathrm{i}}$ is the value of the Lorenz curve of income for the i^{th} individual. In the same year Betti and Cheli (1999) highlighted as α is an arbitrary parameter, but it can be chosen in the way of the mean of membership function will be equal to the Head Count Ratio (HCR) for the official poverty line. Betti, Cheli, Lemmi and Verma (2005, 2006) unified the two membership functions in the "Integrated Fuzzy and Relative" index that is the combination of them.

$$\mu_i = FM_i = (1 - F_i)^{\alpha-1}\big(1 - L(F_i)\big) \qquad (2)$$

In this approach the conventional poor/non-poor dichotomy is replaced by defining poverty as a matter of degree, determined by the place of the individual in the income distribution. The same methodology (Betti and Verma, 2008) facilitates the inclusion of other dimensions of deprivation into the analysis: by appropriately weighting indicators of deprivation to reflect their dispersion and correlation. This last indicator is known as Fuzzy Supplementary (FS) indicator. After that each question was converted into [0, 1] interval, the construction of FS on the basis of different items begins by identifying and investigating the latent dimension of this index. This latent dimension is expected to point to groups where by ‘group’ we mean a distinct set of items. Exploratory and confirmatory factor analyses allow us to achieve this objective. The procedure begins with an exploratory factor analysis to provide a preliminary outline of the groups. In some cases, the exploratory analysis confirms that all items within a single question are highly correlated to a specific dimension. In other cases, the exploratory analysis outlines that some of the items within a question do not have a good correlation with the dimension the question points to. From the latent factors observed in the exploratory analysis and after re-allocating some of the items to different groups on the basis of the correlations calculated among all items, a final list of k groups was proposed. Finally, a confirmatory factor analysis performed on these k groups was used to investigate the Goodness of Fit. After this analysis FM was computed following Betti and Verma (2008).

3. EBLUP with Additive Logistic Transformation and Nonstationary EBLUP

Small area estimation is any of several statistical techniques involving the estimation of parameters for small sub-populations. This technique is used to borrow strength from auxiliary variables to improve the effectiveness of a domain sample size. The term "small area" generally refers to a small geographical area such as a province or a specific sub-population group like people in a certain range of years (Rao and Molina, 2015). If a survey has been carried out for the population, the sample size within any particular small area may be too small to generate accurate estimates from data. Recently, a lot of studies have been carried out about small area techniques and many articles try to summarize the existing works. Between them, we remember Tzavidis et al. (2018), Guadarrama et al. (2016) and Sugasawa and Kubokawa (2020).

SAE based on the widely used area-level model proposed in Fay and Herriot (1979) assume that the area-level direct estimates are spatially uncorrelated. The classical Fay-Herriot (FH) method can be described as follows.

Let j index the J areas of interest, and let y_j be an unbiased direct survey estimator of an unobservable population parameter Y_j of a variable of interest y for area j. Let x_j be a vector of q auxiliary variables for area j that are related to the population mean Y_j . The FH model is then defined by the two equations:

$$y_j = Y_j + e_j \quad and \quad Y_j = \theta + x_j^T \lambda + u_j \tag{3}$$

where the first equation models the prediction error of the observed survey estimate y_j of the true population Y_j, while the second models the unobservable Y_j in terms of an overall mean θ, a linear combination of the components of the vector x_j, and specific area random effects u_j. The combination of these two equations leads to an area-level linear mixed model. Three assumptions are necessary:

- The area random effects u_j are independently and identically distributed;
- The area random effects u_j and the prediction errors e_j are assumed to be independent of each other within and across areas;
- The prediction variances σ_{ej}^2 are known. In practice σ_{ej}^2 are unknown and are replaced with s_{ej}^2.

Here we propose to use EBLUP with Additive Logistic Transformation and the nonstationary empirical best linear unbiased predictor (NSEBLUP) proposed by Chandra et al. (2015). The former, is a classical EBLUP where data are transformed with the Additive Logistic Transformation (ALT-EBLUP) by Aitchison (1986). While in the latter, authors replace the second equation in (3) with:

$$Y_j = \theta(loc_j) + x_j^T \lambda(loc_j) + u_j \tag{4}$$

Where loc_j is the centroid of the *j-th* area, i.e., is the spatial location of the area *j* expressed in coordinates. Under (3) the expected value of Y_j given x_j is the same at any two points in the study area that have the same set of values for this covariate. On the other hand, sing (4) the expected value of Y_j given x_j could be different at two points in the study area that have the same set of values for this covariate depending on the position of those points. The MSE estimation suggested by the authors is based on a parametric bootstrap procedure proposed by Gonzalez-Manteiga et al. (2008). The MSE estimator defined by this procedure is consistent provided the model parameter estimators are consistent.

4. Preliminary results and further development

Since 2002, as highlighted by Litvack (2007), the World Bank has been providing technical assistance to Morocco in the production and analysis of poverty maps. Between others the "*enquête nationale sur la consommation et les dépenses des ménages (2013/2014)*" collects data about household consumption and several variables that could be useful for the definition of the dimension of deprivation for the FS. The poverty line is defined equal to 5.5$ PPP per day per capita and we estimate the HCR at provincial level with direct estimator. The average HCR is 0.34 with an average coefficient of variation (CV) equal to 15.10 while the maximum is 83.99. Given these values and the covariates in our possession the CV of HCR is too high also after SAE (mean CV: 12.40, max: 69.70). For this reason, and for what said before we decide to use the FM. This result to be more precise than HCR with a mean CV of 8.24 and the maximum that is around 36. Using ALT-EBLUP the mean CV reduces to 6.38 with all value lower than 15. This is due to the fact that covariates work better with FM than with HCR given that these last have more heterogeneity than a dummy variable. For what concern the FS, the factor analysis points out 7 different groups:

- Personal satisfaction;
- Development and progress;
- Ordinary purchases;
- Extraordinary purchases;
- Environmental;
- Social life;
- Personal abilities and possibilities.

Next steps will be the direct and the NSEBLUP estimation using also meteorological auxiliary information.

References

Achy, L. (2010). "Morocco's experience with poverty reduction: Lessons for the Arab world". CEIP: Carnegie Endowment for International Peace.

Aitchison, J. (1986). The Statistical Analysis of Compositional Data. Springer Netherlands

Belhadj, B. (2014), "Employment Measure in Developing Countries via Minimum Wage and Poverty New Fuzzy Approach." OPSEARCH 52, no. 2: 329–39.

Betti G., Cheli B., Lemmi A. and Verma V. (2006), "On the construction of fuzzy measure for the analysis of poverty and social exclusion". Statistica & Applicazioni, 4(1), pp. 77-97

Betti, G. and Verma, V. (2008), "Fuzzy Measures of the Incidence of Relative Poverty and Deprivation: A Multi-Dimensional Perspective." Statistical Methods and Applications 17, no. 2: 225–50.

Boutayeb, A. (2006). "Social inequalities and health inequity in Morocco". International Journal for Equity in Health, 5(1), 1–6.

Chandra, H., Salvati, N., & Chambers, R. (2015). "A spatially nonstationary Fay-Herriot model for small area estimation". Journal of Survey Statistics and Methodology, 3(2), 109–135.

Cheli, B. and Lemmi, A. (1995), "A 'Totally' Fuzzy and Relative Approach to the Multidimensional Analysis of Poverty." Economic Notes, 24, 115-133.

El Bouhadi, A., Elkhider, A., & Kchirid, E. M. (2012). "The Multidimensional Approach to Poverty Measurement: Case of Morocco". Applied Econometrics and International Development, 12(2), 135–150.

Ezzrari, A., & Verme, P. (2013). "A multiple correspondence analysis approach to the measurement of multidimensional poverty in Morocco 2001–2007". In Poverty and social exclusion around the mediterranean sea (pp. 181–209). Springer.

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(366a), 269–277.

González-Manteiga, W., Lombardía, M. J., Molina, I., Morales, D., & Santamaría, L. (2008). "Bootstrap mean squared error of a small-area EBLUP". Journal of Statistical Computation and Simulation, 78(5), 443-462.

Guadarrama, M., Molina, I., & Rao, J. (2016). "A comparison of small area estimation methods for poverty mapping". Statistics in Transition New Series, 1(17), 41–66.

Litvack, J. (2007). "The poverty mapping application in Morocco". Using poverty maps to design better policies and interventions. Citeseer.

Rao, J. N., & Molina, I. (2015). Small area estimation. John Wiley & Sons.

Sugasawa, S., & Kubokawa, T. (2020). "Small area estimation with mixed models: A review". Japanese Journal of Statistics and Data Science, 3(2), 693–720.

Trani, J.-F., Bakhshi, P., Myers Tlapek, S., Lopez, D., & Gall, F. (2015). "Disability and poverty in Morocco and Tunisia: A multidimensional approach". Journal of Human Development and Capabilities, 16(4), 518–548.

Tzavidis, N., Zhang, L., Luna, A., Schmid, T., & Rojas-Perilla, N. (2018). "From start to finish: A framework for the production of small area official statistics". Journal of the Royal Statistical Society: Series A (Statistics in Society), 181(4), 927–979.

Verma, V. and Betti G. (2011), "Taylor Linearization Sampling Errors and Design Effects for Poverty Measures and Other Complex Statistics." Journal of Applied Statistics 38, no. 8: 1549–76.

Keywords: *(Fuzzy sets, Morocco, Multidimensional poverty, Nonstationary EBLUP, Small area)*

JEL codes: I32, C13 and E24