Maria Rosaria Ferrante (University of Bologna)
Enrico Fabrizi (Department of Economics and Statistics of the University of Naples Federico II., Università Cattolica del S. Cuore)
Carlo Trivisano (University of Bologna)

Small Area Estimation of the Relative Median Poverty Gap

Topic 6 – Statistics closer to the ground

Keywords: lognormal distribution, Hierarchical Bayesian modeling, income, EU-SILC survey

Introduction

The reduction of poverty in Europe is a milestone in the Europe 2020 strategy. The most frequently used indicator, the At Risk of Poverty Rate (ARPT), measures the share of people below the poverty threshold but it has a serious shortcoming: it neglects how the left tail of the income distribution is shaped and thereby it misses to give indication of the depth of poverty.

The relative median poverty gap (RMPG) tries to complement the information provided by the APRT by measuring the relative distance between the median income of the poor and the poverty threshold. Specifically, it is defined as the difference between the poverty threshold and the median income of individuals below the threshold, relative to the threshold. RMPG = 0 if the income of all "poor" individuals is exactly equal to the threshold, and RMPG = 1 if the income of the poor is zero. Eurostat publishes estimates of the RMPG only at national level since the EU-Survey of Income and Living Condition (EU-SILC) doesn't reach a sample size large enough to produce reliable estimates at sub-national level. In this research we propose a small area estimation model for the RMPG. As far as estimation is concerned, we adopt a Hierarchical Bayesian approach implemented by means of a MCMC computational method. We discuss an application based on data from the Italian sample of the EU-SILC survey.

Methods / Problem statement

The literature of small area estimation of poverty indicators only recently has devoted attention to the estimation of the poverty gap. Molina et al. (2014), Rao and Molina (2016) adopt models for the estimation of mean poverty gap, specified at the unit level. When specified on a transformation of the income variable, as it is often the case, small area estimation based on these model requires unit-level auxiliary data for each non sampled unit in the population. We work within the "area level" model framework, where models are estimated starting from design based estimates, so they easily incorporate information on sampling design and on nonresponse adjustments.

We focus on the RMPG, that is based on median income of the poor, instead of the mean poverty gap, as the latter is more sensitive to extremely low and negative incomes. Moreover, the RMPG is one of the 'Laeken Indicators', a set of common European statistical indicators on poverty and social exclusion , selected by the European council in 2001. The design based estimates of RMPG at small area level are very imprecise. Denoting with na the small sample size in area a, these estimates are obtained using incomes of those who are poor among the na belonging to the sample, that can be very few in small in areas with a limited ARPR and small na. Moreover, in our experience the auxiliary information that can be used to explain RMPG has a really weak predictive power.

Results / Proposed solution

To overcome these problems we propose a model based on a suitable income distributional assumption with the RMPG expressed as functions of the parameters characterizing the income distribution. Specifically, we assume log-normality for the equivalized income. Under this distribution and under some approximations, we show that the RMPG can be expressed as a function of three parameters: the ARPR, the Gini index and the log income average. When deriving the small area estimator of the RMPG we did not make use of direct estimators of the same quantity.

These estimators are not only characterized by a very large variance for the reason discussed above but they are also biased since standard survey weighted estimators of the median are biased in small samples when the underlying distribution is skewed. Model based estimates of RMPG are obtained by using, in a MCMC framework, the posterior distributions of ARPR, Gini index and log income average at the area level as in Fabrizi et al. (2016). We checked the robustness of estimators obtained as posterior summaries with respect to the distributional assumption. The estimates obtained for RMPG in small areas are characterized by a very satisfying coefficient of variation.

The proposed estimators are also design consistent: the difference between model based and direct estimators go to zero as the domain specific sample size increases.

Conclusions

The RMPG is an important indicator to evaluate the depth of poverty since it provides indications on the amount of resources necessary to eliminate poverty, that is, the amount that one would have to transfer to the poor under perfect targeting to bring them all out of poverty.

To limit the attention to the ARPT can be misleading in this sense: ARPT can be unaffected by poverty relief policies that increase the incomes of all those in poverty without necessarily raising anyone above the threshold.

To further study the robustness to the method we propose with respect to the distributional assumption on the income variable, next to the lognormality assumption we are evaluating other assumptions and also to adopt a method for selecting the more suitable distributional assumption.