

Matthias Till
Johannes Klotz
Bernhard Meindl

Enlarging Small Domains: An Application of EU-SILC Regression Parameters to the LFS for Obtaining Disaggregated Poverty Estimates

Social inclusion in EU-Member States is primarily measured by the EU-SILC instrument. Interventions against poverty on a sub-national level call for a targeted monitoring of regional patterns and vulnerable groups. Our paper presents a straightforward combination of the strengths of the two main flagship data sources of official social statistics in the European Union. The LFS with its large sample size gives a reliable representation of small domains in the population. On the other hand, EU-SILC captures a broad range of living conditions and income poverty.

Structural patterns which are observable in EU-SILC are condensed into a logistic regression model. Its parameters yield a predicted propensity for a certain living standard for any combination of characteristics in the model. Upon this model the risk of poverty for specific groups determined for example by household composition, age, sex, citizenship and activity status can be estimated. A table of the regression parameters is usually considered an exhaustive presentation since estimated risks are identical within the same combinations of characteristics. Alternatively, risk estimates can be interpreted as a threshold for a uniformly distributed random variable for the imputation of a dichotomous poverty status on each record in a micro dataset. An intuitive estimate for poverty will then be presented in a simple frequency table.

Regression parameters are themselves estimates and hence subject to sampling variance. Further, bias may result from specification problems where omitted variables are of relevance. Finally, the composition of a small subgroup in a sample may be subject to large sampling variance or even bias. This is where a fusion of data sources and a projection of estimates appear particularly useful. Instead of imputing the poverty status within the same EU-SILC data set we use the LFS to borrow from its greater statistical strength. A bootstrapping procedure provides a first assessment of relative variances and mean squared errors.

Given the existence of harmonized core variables the approach presents a practicable alternative to more advanced techniques of small area estimation. This contribution takes a comparative view between the EU-SILC and LFS Data, with findings for Austria, Italy and the Netherlands.