Nikos Tzavidis (University of Southampton)
Natalia Rojas-Perilla (Freie Universität Berlin)
Timo Schmid (Freie Universität Berlin)
Soeren Pannier (Freie Universität Berlin)

The use of transformations in model-based small area estimation

Topic 2 – Learning more from what we already know

Keywords: MSE estimation, power transformations, Johnson transformations, poverty, inequality

Introduction

Eradication of extreme poverty is one of the most important Millennium Development Goals to be addressed in this century. To achieve this, it is crucial to have detailed knowledge about the spatial distribution of poverty.

One approach to studying the spatial distribution of poverty is by using small area model-based methods. Small area models make use of combined survey and Census data for producing estimates of poverty indicators at geographical levels where direct estimation is either not possible due to the lack of sample observations or very imprecise.

Among many, one approach for poverty mapping is the Empirical Best Predictor (EBP) approach proposed by Molina & Rao in 2010, which -in its current form- relies on the use of Gaussian assumptions for the model error terms. The research questions in this work are motivated by methodological challenges faced by practitioners.

Methods / Problem statement

A real-world question users of small area estimation SAE methods often ask is whether data transformations can help improve estimation and whether the use of data transformations reduces the need for using semi-parametric and non-parametric methods.

The issue of data transformations is addressed in the current SAE literature in a fairly ad-hoc manner. Contrary to standard practice in applied work, recent empirical work indicates that using transformations in SAE is not as simple as transforming income by computing its logarithm. This paper addresses the issue of data transformations systematically.

Results / Proposed solution

The objective of the present work is twofold. First, we analyze how the performance of the EBP method is affected by departures from normality and explore the use of alternative transformations for achieving normality.

For this purpose, we explore model-based estimation under various transformations, the estimation of transformation parameters and how such transformations can assist with improving the validity of the model assumptions and the precision of small area predictions. Second, we adapt the mean squared error (MSE) estimator to account for the additional uncertainty due to the estimation of transformation parameters.

Conclusions

The methods are illustrated by using survey and Census data for estimating deprivation and inequality in Mexico and results from model and design-based simulation. Our results show that using transformations can improve the precision of small area estimates. Moreover, emphasis is placed on the development of easy to use tools that will enable users of SAE methods to select and use appropriate transformations for their data.