Unit level small area models for business survey data

C. Bocci[1] and P.A. Smith[2]

[1] University of Florence, Italy
e-mail: chiara.bocci@unifi.it

[2] University of Southampton, UK
e-mail: p.a.smith@soton.ac.uk

Abstract

Small area estimation (SAE) encompasses a wide range of methods (Rao and Molina, 2015) that have become a significant component in the official statistician's toolkit, with many applications to a wide range of variables and domains in many different types of surveys. Nevertheless, SAE is still uncommon in business surveys because of challenges arising from the skewness and variability of many size-related variables. SAE methods are generally based on mixed effects models which have assumptions of normal errors, but skewed variables violate this assumption. Moreover, the specific characteristics of sample designs used in business surveys (detailed stratification, non-negligible sampling fractions, large variations in estimation weights) affect the models on which small area estimates are based.

Several approaches have been suggested to deal with such skewed data in different ways: through transformation, by employing robust models to accommodate outlying tail observations, and by directly modelling the skewed distribution. Smith et al. (2021) examined a range of robust approaches, which reduce the impacts of observations in the tails of skewed distributions, in a dataset with known outcomes. Here we replicate this analysis with a second dataset of Italian retail businesses, and compare it with a second group of methods based on transformations of the initial data before modelling. The back-transformed predictions need bias adjustments to produce estimates with acceptable quality. We review the transformation-based methods which have been proposed in the literature and make an assessment of the best approaches to use for business surveys based on our repeated sampling simulation study.

Keywords: robust estimation, outliers, skewed distribution.

References

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

Smith, P.A., Bocci, C., Tzavidis, N., Krieg, S., Smeets, M.J.E. (2021), Robust estimation for small domains in business surveys. *Journal of the Royal Statistical Society Series C: Applied Statistics*, **70**, 312–334.