Small area estimation of poverty rate by model calibration and "hybrid" calibration

Risto Lehtonen[1] **and Ari Veijanen**[2]
[1] University of Helsinki, Finland
[2] Statistics Finland, Finland

Calibration techniques (Deville and Särndal 1992, Särndal 2007) are often used in design-based estimation of totals or means. In model-free (or linear) calibration, weights are calibrated to reproduce the known population totals of the auxiliary variables. Aggregate-level auxiliary data are sufficient for the method. The method does not involve explicit model statement. The same set of calibrated weights can be supplied to the diverse study variables of a sample survey. In official statistics production, these properties are sometimes considered a benefit.

In model calibration (Wu and Sitter 2001, Montanari and Ranalli 2005) representing a model-assisted technique, a model is specified and the weights are calibrated to the population total of the predictions derived via the model. Access to unit-level auxiliary data is assumed. A considerable modelling effort is often needed. A model calibration procedure presupposes the application of the method separately for each study variable. A benefit of the method is that models beyond the linear model can be specified such as logistic models and other members of the generalized linear models family.

In standard model calibration, consistency of estimated totals or means with published statistics is not guaranteed. Montanari and Ranalli (2009) introduced a variant of model calibration (called multiple model calibration) that retains the consistency property of estimates with published statistics, a key property of model free calibration. This is achieved by imposing into the calibration constraints the auxiliary x-variables for which the consistency requirement is needed. In practice this means that a part of the set of auxiliary variables is used for model calibration (to improve accuracy) and a part is used for reasons of consistency with published statistics. The two sets of x-variables can be distinct or they may overlap. We call this method "hybrid" calibration because the method aims at combining some of the favourable properties of model-free calibration and model calibration.

In the paper, alternative model calibration estimators introduced in Lehtonen and Veijanen (2012) are used for the estimation of at-risk-of poverty rate for population subgroups or domains and small areas. These methods include the semi-direct and semi-indirect model calibration methods. The hybrid calibration method is introduced and applied for small area estimation. The study variable describing poverty is binary and is modelled by a logistic mixed model. The calibration methods are compared with other design-based methods including direct Horvitz-Thompson type estimator and model-assisted logistic generalized regression estimator (Lehtonen and Veijanen 1998, Lehtonen, Särndal and Veijanen 2003). Statistical properties

(design bias and accuracy) of the methods are assessed by design-based simulation experiments using unit-level register data maintained by Statistics Finland.

Keywords: Poverty rate, Small area estimation, Calibration techniques, Generalized regression estimator.

References:

Deville, J.-C. and Särndal, C.-E. (1992). Calibration estimators in survey sampling. *Journal of the American Statistical Association*, 87, 376–382.

Lehtonen, R., Särndal, C.-E. and Veijanen, A. (2003). The effect of model choice in estimation for domains, including small domains. *Survey Methodology Journal*, 29, 33–44.

Lehtonen, R. and Veijanen, A. (1998). Logistic generalized regression estimators. *Survey Methodology Journal*, 24, 51–55.

Lehtonen, R. and Veijanen, A. (2012). Small area poverty estimation by model calibration. *Journal of the Indian Society of Agricultural Statistics*, 66, 125–133.

Montanari, G. E. and Ranalli, M. G. (2005). Nonparametric model calibration estimation in survey sampling, *Journal of the American Statistical Association*, 100, 1429–1442.

Montanari, G.E. and Ranalli, M.G. (2009). Multiple and ridge model calibration. Proceedings of Workshop on Calibration and Estimation in Surveys 2009. Statistics Canada.

Särndal, C.-E. (2007). The calibration approach in survey theory and practice. *Survey Methodology*, 33, 99–119.

Wu, C. and Sitter, R. R. (2001). A model-calibration approach to using complete auxiliary information from survey data. *Journal of the American Statistical Association*, 96, 185–193.