

Impact of sampling on small area estimation in business surveys

Jan Pablo Burgard

University of Trier, Germany, jpburgard@uni-trier.de

Ralf Münnich*

University of Trier, Germany, muennich@uni-trier.de

Thomas Zimmermann

University of Trier, Germany, thzimmer@uni-trier.de

Modern Business statistics often faces the difficulty that an increasing demand of information on sub-levels defined by regions or cross-classifications of variables such as industry classes and business size can be observed, eg for measures of competitiveness by policy makers. In order to enable data producers to provide estimates on those sub-levels, sophisticated stratifications are implemented in the sampling design. These detailed stratifications may produce two difficulties. First, many strata contain only very few elements and, hence, make it difficult to derive optimal sample sizes. Second, statistical model building may suffer from the survey weights derived under these constraints. Additionally, optimization of sampling designs may have a strong impact on the accuracy of different estimation strategies. The aim of the paper is to evaluate different sampling designs in the context of estimation on sub-levels by regions and cross-classifications and their impact on these domain estimates. As estimators of interest, the Horvitz-Thompson- and generalized regression estimator as design-based methods as well as the Battese-Harter-Fuller-, the You-Rao-, and the augmented estimators are considered. The analysis is performed by means of a Monte Carlo study based on Italian business data.

Key words: optimal sampling design, model-based estimation, design-based estimation