

SAMPLING DESIGN AND WEIGHTING IN THE EUROPEAN SOCIAL SURVEY

Seppo Laaksonen
University of Helsinki, Seppo.Laaksonen@Helsinki.Fi

The European Social Survey (ESS) has aimed to control the sample designs used by specifying sampling guidelines that should have been followed in each participating country. The main requirements are the use of probability sampling and the achievement of a minimum effective sample size that is determined by ineligibility rate, nonresponse rate, inclusion probabilities and clustering effects. Several sampling strategies have been used over the first seven rounds, but in most countries the design has not changed substantially.

The sampling requirements are not always well satisfied. In this paper we present key problems observed until now, and their trends. The following questions especially are considered and statistics over six rounds presented: (i) ineligibility has been growing and becoming more complex, (ii) nonresponse is a worsening problem and varies inside any country as well, (iii) design effects due to clustering vary by cluster size and intra-class-correlation (ICC); (iv) design effects due to variation in selection probabilities depend mainly on the type of sampling frame used. Since round 2, the prediction of effective sample size has improved as it has been possible for most countries to use parameter estimates from the previous rounds (ICC, eligibility rate, response rate, coefficient of variation of selection probabilities), but apparent changes over time in these parameters have caused difficulties.

The second section of the paper is to describe what happens after the fieldwork relating to sampling. First, the sampling design data file is created following the specific template. This file consists of all the gross sample units and its variables include those that give opportunity to create sampling weights, to analyse the survey quality, and to estimate. Its most important characteristics, including sampling design variables and weights, will be finally merged together with the real survey variables at respondent level, and then the survey analysis is ready to begin.

The ESS public use file currently includes the two types of weights, the analysis weights whose sum is the number of the respondents of each country (representing 15+ years old residents) for each country, and the country size weight. Their product multiplied by 10000 gives an ordinary sampling weight. There exists the two types of analysis weights since 2013; (i) the weights based on the design assuming that nonresponse is ignorable, and (ii) the raking ratio weights that adjust for nonresponse and in-eligibility to some extent. We thus see that the weighting can be further improved but a drawback is that the current sampling design files are not reasonably good in any country although many countries could do it better.

Fortunately, the minimum level of the sampling design file is achieved in each country so that the design based weights can be created. But even the raking-ratio weights are calculated separately taking benchmarking margins from an external source. The best situation would be such a file that consists of these variables and in addition, of a number of micro and macro auxiliary variables that give opportunity for creating a more sophisticated weight such as it is an appropriate combination of both a response propensity weight and a calibration weight.

References

Deville, J-C. & Särndal, C-E. (1992) Calibration Estimators in Survey Sampling. *Journal of the American Statistical Association*, 376-382.

Laaksonen, S. (2015). Sampling design data file. *New Techniques and Technologies in Statistics (NTTS)*. European conference. Brussels 10-13 March. <http://www.cros-portal.eu/sites/default/files//NTTS2015%20proceedings.pdf>

Laaksonen, S. & Heiskanen, M. (2014). *Comparison of Three Modes for a Crime Victimization Survey*, *Journal of Survey Statistics and Methodology* 2 (4): 459-483 doi:10.1093/jssam/smu018

Le Guenne, J. & Sautory, O. (2005). CALMAR 2 : Une Nouvelle Version de la Macro Calmar de Redressment D'Échantillon Par Calage. http://vserver-insee.nexen.net/jms2005/site/files/documents/2005/327_1-JMS2002_SESSION1_LE-GUENNEC-SAUTORY_CALMAR-2_ACTES.PDF

Lynn, P. & Gabler, S. & Häder, S. & Laaksonen, S. (2007). *Methods for Achieving Equivalence of Samples in Cross-National Surveys*. *Journal of Official Statistics*, 27, 1, 107-124.

European Social Survey website (europeansocialsurvey.org) including the sampling guidelines for the round 6: http://www.europeansocialsurvey.org/docs/round6/methods/ESS6_sampling_guidelines.pdf