

Calibration is not good as the only weighting method

Seppo Laaksonen

University of Helsinki, seppo.laaksonen@helsinki.fi

Calibration is much used so that it is the only weighting adjustment method. Its good point is that the calibration margins are easy to create confidentially if these macro auxiliary variables are not concerning too small and sensitive items. This method does not exploit micro auxiliary variables that are available for individual units from registers, other administrative sources, even from social media and from interviewers. It means that the calibration only does not exploit micro auxiliary variables explicitly. Our recommendation thus is to use all existing auxiliary variables. Otherwise, it should have been used time and other resources to find and to compile such appropriate variables as much as possible during the survey process. It is possible in all countries to some extent although not done until now. There are some strategies to exploit micro auxiliary data but the most common one is response propensity weighting in which the binary model can vary, including symmetric probit and logit, and asymmetric log-log and complementary log-log. This paper presents principles with illustrative examples of response propensity weighting. We find that this method alone is not a good solution since it may include only simple calibration margins. In most surveys, it is possible to get more such margins and to exploit them in calibration. We recommend to use these margins after the response propensity weighting. In this stage, it is good to carefully think which margins are most useful so that they correct basic things, such as distribution by gender, age groups, region and education. They fortunately often reduce the bias due to unit nonresponse but it is not the main purpose of the method. It is good to recognize that some calibration methods, linear calibration particularly, do not work if too many calibration margins are tried. The reason is that some weights are below 1 or even negative. The empirical part of the presentation shows strategies for weighting adjustments. Some are only taken advantage of calibration, the others of both methods, first response propensity weights with macro and micro auxiliary variables and then raking ratio calibration that always leads to desirable weights. The data of the empirical part is artificial but downloaded from several sampling design data files (SDDF) of the European Social Survey. The reason here is that the country SDDF's are not publicly available.