

Handling nonignorable nonresponse using generalized calibration with latent variables

M. Giovanna Ranalli*

University of Perugia, Perugia, Italy giovanna.ranalli@stat.unipg.it

Alina Matei

University of Neuchâtel, Switzerland alina.matei@unine.ch

Andrea Neri

Bank of Italy, Rome, Italy, andrea.neri@bancaditalia.it

In this work we focus on the treatment of unit nonresponse in surveys when estimation of finite population totals and means of a set of variables of interest is of concern. Nonresponse is known to introduce bias in estimates. Successful reduction of such a bias may be achieved using powerful auxiliary information coupled with a well-specified model either for unit response probabilities (response model) or for the variables of interest (superpopulation models). Auxiliary information takes the form of a set of auxiliary variables whose values are known for the respondents and whose population totals (or means) are known or can be unbiasedly estimated via their Horvitz-Thompson estimator.

Calibration is a general tool to include auxiliary information at the estimation stage and to treat unit nonresponse. Here, we consider generalized calibration that, unlike other reweighting methods, allows to correct for nonresponse bias even when the variables that cause nonresponse are known only for the respondents. This is particularly relevant when nonresponse is nonignorable, i.e. when the topic of the survey and, therefore, the variables of interest influence the response probability of a unit. In fact, it is possible to introduce the variables of interest (known only of respondents) as instrumental variables for correcting this type of nonresponse.

In this work we move from noticing that response to a survey is the outcome of a complex process that involves several aspects, from the topic of the survey to some of the variables of interest or of the auxiliary variables. We assume that a part of such a process may be measured by an unobservable variable. Latent variable models can be employed to extract either a continuous construct (latent trait models) or a categorical one (latent class models) from a set of dichotomous/ordered manifest variables. This type of latent variables is particularly relevant, although not limited, to attitude and behavioral surveys and provide a measure of unobservable variables (like the "attitude towards politics" or other sensible topics) that likely influence the "willingness to respond" of a unit. We, therefore, propose to use such constructs as instrumental variables in the generalized calibration procedure. This allows to include variables of interest among the set of manifest variables and also for the construction of a single set of weights. The proposed methodology is first tested on a series of simulation studies and then applied to adjust estimates from the Italian Survey of Households' Income and Wealth.

Key Words: Finite population, auxiliary information, instrumental variables, item response theory models, latent class models.