Predicting response error in multiple survey items

Barry Schouten*
Statistics Netherlands and Utrecht University, the Netherlands - jg.schouten@cbs.nl

Frank Bais
Utrecht University, Utrecht, the Netherlands - f.bais@uu.nl

Vera Toepoel
Utrecht University, Utrecht, the Netherlands - v.toepoel@uu.nl

Many statistical institutes are exploring survey designs that combine the web survey mode with the telephone, paper and face-to-face survey modes. Partly this is because of cost reasons and partly this is to accommodate preferences that persons may have in responding to surveys and/or in answering survey questionnaires. Redesigns of surveys to mixed-mode designs lead to method effects in which mode-specific response error (or measurement error) may play an important role. At Statistics Netherlands, redesigns to mixed-mode surveys led to breaks in time series. Some of the breaks were not anticipated and some of the breaks were strongly conjectured to be caused by response error differences. For a number of surveys, experimental studies indicated that this conjecture is indeed true.

From the occurrence of method effects arose the need to be able, to the extent possible, to predict and explain response error differences between modes before and during redesigns. One of the methodological complexities is the multidimensionality of a survey questionnaire; questionnaires contain many survey items that are affected to different extents by the choice of survey mode. For this reason, attention was shifted to response styles, i.e. to deficiencies in the answering process that persist throughout a significant part of the questionnaire and affect multiple survey items simultaneously.

In the paper, we model answering behaviour through multilevel models in which both person and survey item characteristics are included. The levels in the model are formed by persons, surveys, and items within those surveys. Different forms of answering behaviour are deduced from paradata and response styles are viewed as small between variations for survey items (or large between variations for persons). For this purpose an experimental study was designed that combines survey data from the Dutch LISS panel, survey data from Statistics Netherlands, validation data from Dutch government registries and a rich set of paradata observations. In the paper, we focus attention on an important methodological research question: How to deal with missing survey item data due to unit-nonresponse, filter questions and item-nonresponse in the multi-level models?