

Modelling final outcome and length of call to improve efficiency in call scheduling

Gabriele Durrant, Olga Maslovskaya, and Peter W.F. Smith

Southampton Statistical Sciences Research Institute, University of Southampton

For interviewer administered surveys many survey agencies nowadays routinely collect call record data. Examples of such data may be recordings of the day, time and outcome of each call or visit and, for face-to-face surveys, information collected by the interviewer such as physical and social characteristics of the selected housing unit and neighbourhood. Researchers have increasingly become interested in how best to use and analyse such information. It is hoped that a better understanding of the calling patterns and the mechanisms leading to particular call sequences will help to improve data collection.

More specifically, for statistical agencies, investigating time and effort into repeated calls and follow-ups is very resource-intensive. From a survey management perspective, it seems desirable to avoid long unsuccessful call sequences to improve efficiency. The aim then is to identify early on in the data collection process cases prone to long and unsuccessful call sequences. This paper models call record data predicting final call outcome and length of a call sequence early on in the data collection process. Separate logistic and joint multinomial models for the two outcomes are considered and the models account for the clustering of sample cases within interviewers. Of particular interest is to identify good explanatory variables that predict final outcome and length of a call sequence, in particular characterising long unsuccessful call sequences. Findings so far indicate that outcomes of the initial call attempts are highly predictive but also other variables such as interviewer observations. Further research questions that we aim to address in this study are: how can predictors best be incorporated into the model (e.g. as summary statistics or as individual outcomes?); how predictive are the models?; does their ability to predict the variable of interest improve if more and more call record data are available (e.g. for later calls; or for later waves in a longitudinal study?); how can these models best be used in adaptive and responsive survey designs?

The study uses data from a large-scale longitudinal survey in the UK, Understanding Society. Implications for survey practice and fieldwork procedures are discussed.