Dissertation Defense

Adaptive Design to Adjust for Unit Nonresponse Using an External Micro-level Benchmark

Julia Lee

Dissertation Committee:

Steven Heeringa, Co-Chair, Trivellore Raghunathan, Co-Chair, Richard Valliant, Roderick Little

Monday, December 5th at 8:30am

368 ISR, University of Michigan
Via video to 1218 LeFrak, JPSM at University of Maryland

Traditional survey design draws a representative sample and implements post-survey weighting adjustments to compensate for nonresponse. When survey participation decline renders respondents nonrepresentative, the effectiveness of post survey weighting adjustment becomes uncertain. Recent developments to improve respondent representativeness via adaptive data collection design have delivered promising results on bias reduction.

This dissertation develops a new adaptive design to improve survey data quality, by capitalizing on a benchmark data which captures the target population. The basic idea is to adaptively draw samples that lead to representative respondents; and to compensate for nonrespondents by benchmarked imputation procedures. Respondent representativeness is enhanced by the sampling procedure as opposed to data collection, eliminating costs of nonresponse follow-up and inferential complexity due to varying data collection protocols.
The new adaptive design consists of benchmarked sequential sampling (BSS) and benchmarked multiple imputation (MI) procedures. The new design first improves respondent representativeness by BSS, which conforms either the frame variables alone (BSS-Z) or both frame and survey covariate information (BSS-X) to those of the benchmark. With improved respondent representativeness, the benchmarked MI recovers the population information, leading to better quality survey estimates that are less susceptible to the unknown nonresponse pattern. This design can be applied to surveys with rich micro-level data and surveys that use respondents of other surveys as sampling frame.

The BSS-Z method is demonstrated using the National Health Interview Survey and Behavior Risk Factor Surveillance System; the BSS-X and the benchmarked MI methods are demonstrated using the American Community Survey, the Current Population Survey, and the Census Planning Database.

An evaluation is done between the new design of adaptive sampling and imputation and the traditional design of fixed sampling and weighting (generalized regression estimator). To assess respondent representativeness, data from the new design is compared to those of the benchmark in marginal, conditional, and descriptive statistics. To assess the quality of the survey inference, a sample mean is calculated along with its root mean square error (RMSE), bias and coverage rate. To assess whether a design is of better value, a cost-effectiveness measure is derived from RMSE and a new cost model.