

Combining Probability Non-probability Samples: Theory and Practice

Michael Elliott

University of Michigan
E-mail: mrelliot@umich.edu

Abstract: Although probability sample designs remain a “gold standard” in survey research, demand for use of non-probability samples is increasing, due to, among other reasons, rising costs and falling response rates in probability samples and the availability of “big data” from administrative databases, social media users, and other sources. Design-based inference, in which the distribution for inference is generated by the random mechanism used by the sampler, cannot be used for non-probability samples. One alternative is quasi-randomization in which pseudo-inclusion probabilities are estimated based on covariates available for samples and nonsample units. Another is superpopulation modeling for the analytic variables collected on the sample units in which the model is used to predict values for the nonsample units. A third alternative is a model-assisted approach in which probability samples are used to develop calibration estimators. We will overview these approaches and discuss their unique advantages in different analytic and application settings.