

## **Title: Non-probability Samples**

**Date and Time:** September 9th, 2015  
12:30pm to 3:30pm

**Moderator:** Donsig Jang, Mathematica Policy Research

**Location:** Offices of Mathematica-MPR 1100 1st Street NE, 12th Floor, Washington DC 20002. Once in the building, inform the receptionist at the first floor lobby that you are visiting Mathematica for a WSS seminar. Then, take the elevators to the 12th floor and tell the Mathematica receptionist that you are attending the WSS seminar. Please call Mathematica's main office number (202 484-9220), if you have trouble finding the building.

**By Metro:** Take the Red Line to either the NoMa-Gallaudet U (used to be called New York Ave) Station or Union Station. From the NoMa-Gallaudet U Station, follow signs to exit at M Street. Then walk 1 block west on M street and 2 blocks south on 1st Street NE (the building will be on your right). From Union Station, walk north along 1st Street NE for about 4-5 blocks until you reach L Street (the building will be on your left after crossing L street).

**By Car:** Pay parking is available in the building parking garage, which is located 1 block east of North Capitol on L Street NE.

**Guest List:** To be placed on the attendance list for webinar, please RSVP to Alyssa Maccarone at amaccarone@mathematica-mpr.com or (202) 250-3570 at least 2 days in advance of the conference. Provide your name, affiliation, and contact information (e-mail is preferred). Once on the attendance list, you will be provided with information about webinar.

**Sponsor:** WSS Methodology Section

**Abstract:** This Washington Statistical Society conference on non-probability samples follows upon the 2014 WSS President's Invited Lecture given by Mike Brick who discussed the findings of the American Association for Public Opinion Research (AAPOR) task force on conditions under which survey designs that do not use probability samples might still be useful for making inferences to a larger population. The speakers of this conference will discuss applications of using non-probability samples and their implications.

**Following the seminar, snacks and refreshments will be served, encouraging the attendees to continue questions and discussions on the talks.**

### **Schedule**

Time	Speaker	Affiliation	Point of Contact
12:30	Donsig Jang	Mathematica Policy Research	DJang@Mathematica-Mpr.com
12:35	Mike Brick	Westat	mikebrick@westat.com
1:00	Michael Sinclair	Mathematica Policy Research	MSinclair@mathematica-mpr.com
1:25	Jill A. Dever	RTI International	jdever@rti.org
2:00	Intermission		
2:10	Scott Keeter	Pew Research Center	SKeeter@PewResearch.org
2:35	Richard Valliant	Universities of Michigan and Maryland	rvallian@umd.edu

## **Title: Non-Probability Sampling Assumptions and Methods**

### **Abstract:**

Non-probability sampling must rely on modeling to make inference because design-based theory is impossible without known selection probabilities. The typical opt-in, non-probability survey applies weights to the sample observations in ways that mimic methods from probability samples. This weighting process implies a set of assumptions about the distributions of the variables and the selection mechanisms. In this talk, we begin by reviewing some of these weighting methods and the implied or explicit model assumptions. Next, we discuss whether some of the lessons learned from probability sample designs can be employed to improve the accuracy of non-probability samples.

**-Mike Brick**

Vice President and co-Director of the Survey Methods Unit at Westat

## **Title: Non-Probability Samples and Panel Surveys: A look at Two Strategies for Blending Panel Surveys with Traditional Probability Samples**

### **Abstract:**

With increasing survey costs, reduced participation rates, and a greater need for timely and domain specific estimates to inform treatment and policy decisions, interest has grown in the use of panel surveys and other forms of extant data to augment traditional survey methods. Likewise, while the direct use of panel surveys and nonprobability based samples have been widely accepted for polling and marketing research studies, the authors suspect that a broader array of researchers may begin to look at methods to answer questions where traditional survey methods are dominant. Given concerns with the representative nature of panel data, researchers may require some form of validation provided by an independent and smaller probability based data collection. To cover both of these situations, we present two approaches for blending data from panel or non-probability sample collections with probability samples. The first approach is based on the work of Valliant, Dorfman and Royall (2000) coupled with the use of a composite estimator that we suggest may be most useful when a traditional probability based survey data collection is used to validate a larger panel study. The second approach explores the use of a matching process based on Guo and Fraser (2010) to supplement a traditional survey with panel survey data to produce estimates for smaller geographical area, for cases with rare characteristics or to create survey estimates between study cycles to generate more timely information. We will also explore the use of Bayesian statistical methods to capture the variability in the blending procedures based in part on the work of Zheng and Little (2003) and Zangeneh and Little (2012).

**-Michael Sinclair**

Senior Fellow at Mathematica Policy Research

## **Title: Can Estimated-Control Calibration Reduce Bias in Estimates from Nonprobability Samples?**

### **Abstract:**

Nonprobability (or design-free) surveys are becoming more prevalent because they offer both increased speed in obtaining data on emerging issues (e.g., an opt-in web survey) and decreased costs compared with probability-based surveys. However, evaluation studies have shown that many nonprobability estimates are biased because of errors associated with coverage, selection, and model misspecification.

Calibrating design-based survey weights to control totals estimated from other surveys has been implemented for years. Referred to as estimated control (EC) calibration, this technique has been shown to reduce bias for design-based estimates beyond levels seen when calibrating to typical controls alone (e.g., demographic characteristics, geographic location). By comparison, propensity score adjustments (PSA) are used to calculate estimates from nonprobability surveys, and may include questionnaire items (e.g., webographic questions) as logistic model covariates. However, research to date on PSA shows mixed results with bias reduction.

This presentation begins with a brief background on PSA and EC calibration techniques. Next, results are discussed from an empirical study to compare bias reduction obtained from the two methods for nonprobability samples. The presentation concludes with future research.

**-Jill A. Dever**

Senior Survey Statistician, RTI International

**Title: An Exploration of Threats to Inference with Nonprobability Samples: Bias, Homogeneity, Both or Neither?**

**Abstract:**

Aside from coverage problems, online panels based on nonprobability samples may suffer from large but unknown biases in the types of internet users who participate. These samples may be significantly less diverse than in a typical random sample of the public. This presentation will compare probability samples and nonprobability samples from several vendors, focusing on both the comparison of probability versus nonprobability samples and the differences among the nonprobability samples on a set of variables that may be related to the propensity to take part in a survey. Among the topics addressed will be the impact of adjustments and weighting on the bias including relationships among variables and subgroups and the effective sample size relative to the cost per interview.

**-Scott Keeter**

Director of Survey Research, Pew Research Center

**Title: Inferential Problems with Nonprobability Samples**

**Abstract:**

This talk will briefly review some of the basic issues in making inferences to populations using samples where the investigator has limited control over which units appear in the sample. Repeated sampling (design-based) inference cannot be used, so if estimators have any justification, it must be model-based. Finding a model(s) that can be used to project a sample to a population is, thus, the critical step. Approaches that have been proposed are (1) use of models to calculate of pseudo inclusion probabilities to use in quasi-randomization inference, (2) projection using population structural models for analysis variables, and (3) combinations of (1) and (2). Combining probability samples (reference surveys) with nonprobability samples is one line of attack for (1). Calibration estimation (e.g., raking or poststratification) has been used for (2). Validating the models and the procedures used by different organizations is difficult. Some options for testing the methods will be discussed.

**-Richard Valliant**

Research Professor, Universities of Michigan and Maryland