Title: FCSM/WSS Workshop on Sensitivity Analysis with Integrated Data

Date/Time: June 10, 2019, 1:00 to 4:30 PM

Location: Bureau of Labor Statistics Janet Norwood Conference Center and WebEx

Registration: The workshop is free, but you have to register on EventBrite HERE.

Sponsor: WSS Methodology

Abstract:

This half-day workshop will inform efforts by the interagency Federal Committee on Statistical Methodology (FCSM) to improve information shared with the public about how agencies combine data from multiple sources to produce estimates, analyses and other data products. The workshop will provide insight into a critical aspect of integration processes, namely, identifying and evaluating critical assumptions made during integration, and the possible impact on results when those assumptions are changed.
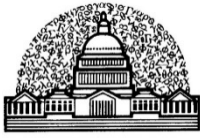
At this workshop, we will explore existing examples of such work, toward creating a broader understanding of techniques and issues involved. We will also emphasize a set of broader reporting issues and recommendations to help the public understand the implications of the different assumptions that are involved when integrating data from multiple sources into a single product. We will cover issues with evaluation and reporting of situations where 1) all sources have completely overlapping variables and we have to select a source, 2) the sources have some overlap on common variables but information has to be filled in across cases or domains with imputation and other techniques like statistical matching, 3) the sources are all providing unique information that has to be filled in for cases or domains not common to the initial data sources – again, relying on imputation, statistical matching, etc. – and 4) the sources have similar but not non-identical variables that must be expressed with a common metric (sometimes referred to as data harmonization). Integration procedures, sensitivity analyses, and reporting standards will vary depending on the combinations of data sources: multiple sample surveys, sample surveys with non-probability data, sample surveys and administrative data, and non-probability data with administrative data.

Discussion topics may also include what to report about how key methods for combining data are selected and evaluated: procedures and models for imputation, weighting for probability and non-probability samples, propensity score methods, techniques for exact matching and statistical matching, and considerations of remaining incomplete cases or coverage after integration is completed.

Session 1: 1:00-2:40
Understanding Sources of Uncertainty with Integrated Data
- Introductory Remarks, Joe Schafer, U.S. Census Bureau (10 min)
- "Measuring Uncertainty with Multiple Sources of Data," Sharon Lohr, Arizona State University (50 min)
- Discussant: John Eltinge, U.S. Census Bureau (15 min)
- Questions and floor discussion (25 min)

Session 2: 3:00-4:30

Perspectives on Sensitivity Analysis from Health-Outcomes Research

- "Measurement error correction and sensitivity analysis in longitudinal dietary intervention studies using an external validation study," Juned Siddique, Northwestern University (30 min)
- "Sensitivity Analysis for Unobserved Effect Modification when Generalizing Findings from Randomized Trials to Target Populations," Benjamin Ackerman, Johns Hopkins University (30 min)
- Floor discussion (30 min)

WebEx: WebEx event address:
https://dol.webex.com/dol/j.php?MTID=m193633665799a4135d9f2ba1387e1129

For audio:
Call-in toll-free number (Verizon): 1-866-865-9536 (US)
Attendee access code: 744 124 3