



**Title:** The Utility of Commercial Data for Survey Operations and Estimation: Examples from the National Survey of Family Growth

**Date/Time:** April 9, 2018 12:30 – 2:00 PM (ET)

**Speakers:** Brady West, University of Michigan

**Chair:** Tom Mule

**Sponsor:** WSS Data Collection

**Abstract:** The combination of declining response rates and increasing data collection costs has led to an increase in the use of responsive survey design strategies and consideration of alternative auxiliary data sources for improvement of survey data collection efficiency and post-survey nonresponse adjustments. Among the myriad sources of auxiliary data that might be appended to a sampling frame to support these efforts, commercial databases maintained by consumer marketing or credit scoring organizations have recently received some research attention, due to the rich amount of information that they make available for survey researchers. Unfortunately, the existing literature raises concerns about the quality of the information in these databases, and few studies to date have considered the ability of these auxiliary data sources to effectively predict survey outcomes, including household eligibility for a given survey, propensity to respond, and key survey measures of interest. Such studies are important to justify the costs of purchasing these commercial data for survey production purposes. In this seminar, I will examine the predictive ability of two alternative commercial databases offering different sets of auxiliary variables to survey researchers. Analyzing survey outcomes from the National Survey of Family Growth (NSFG), we find that these commercial variables improve the fits of models predicting survey eligibility and selected NSFG variables, but that neither data source substantially improves models of response propensity that already include selected NSFG paradata. As a result, the inclusion of the commercial variables in nonresponse adjustments does not result in substantial shifts in selected NSFG estimates. These results suggest that commercial data can be useful for selected survey operations (e.g., identification of eligible households during screening), but may not be useful for post-survey nonresponse adjustment. Suggestions for practice and directions for future research will be provided in conclusion.

**Location:** Bureau of Labor Statistics Conference Center, Room 9

To be placed on the seminar attendance list at the Bureau of Labor Statistics, you need to e-mail your name, affiliation, and seminar name to [wss\\_seminar@bls.gov](mailto:wss_seminar@bls.gov) (underscore after 'wss') by noon at least two days in advance of the seminar. Please bring a photo ID to the seminar. BLS is located at 2 Massachusetts Avenue, NE. Use the Red Line to Union Station. Parking in the area of BLS is available at Union Station. For parking information see <http://www.unionstationdc.com/parking>. No validation is available from BLS for reduced parking rates.

**WebEx:** WebEx event address:

<https://dol.webex.com/dol/j.php?MTID=m8a9d445ad56bfbaea5c1dd11d0a8cdc5>

**For audio:**

Call-in toll-free number (Verizon): 1-866-745-4894 (US)

Call-in number (Verizon): 1-517-233-8001 (US)

Attendee access code: 312 885 1

Note: Particular computer configurations might not be compatible with WebEx.