

Selected ICES-IV Presentations from the Newly Released Journal of Official Statistics Special Issue on Establishment Surveys

Date and Time: Wednesday, February 4th, 2015
1:00 - 4:30 p.m.

Chair: Darcy Miller, NASS

Location: Bureau of Labor Statistics Conference Center

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 (underscore after 'wss') by noon at least 2 days in advance of the seminar, or call 202-691-7524 and leave a message. 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.

Sponsor: Methodology Section

WebEx event address for attendees:

<https://dol.webex.com/dol/j.php?MTID=m94030e09869538075cfe95f5d8a2b7fb>

For audio: Call-in toll-free number (Verizon): 1-866-747-9048 (US)
Call-in number (Verizon): 1-517-233-2139
(US) Attendee access code: 938 454 2

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

Schedule

Time	Speaker	Affiliation	Point of Contact
1:00	Polly Phipps	BLS	Phipps.Polly@bls.gov
1:10	Richard Sigman	Westat	Richard.sigman@westat.com
1:40	Morgan Earp	BLS	Earp.Morgan@bls.gov
2:10	Mary Mulry	Census	mary.h.mulry@census.gov
2:40	Intermission		
3:00	MoonJung Cho	BLS	Cho.Moon@bls.gov
3:30	Vanessa Torres van Grinsven	Utrecht University & Statistics Netherlands	V.TorresvanGrisven@cbs.nl
4:00	Daniell Toth	BLS	Toth.Daniell@bls.gov

Title: Does the Length of Fielding Period Matter? Examining Response Scores of Early Versus Late Responders

Abstract:

This paper discusses the potential effects of a shortened fielding period on a large, Federal employee survey's item and index scores and sample demographics. Using data from the U.S. Office of Personnel Management's 2011 Federal Employee Viewpoint Survey, we investigate whether early responding employees differ from later responding employees on their key policy-relevant survey item scores. Specifically, we examine differences in scores on items and indices relating to conditions conducive to employee engagement and global satisfaction. We define

early responders as those who responded in the first two weeks of the fielding period. We also examine the extent to which early versus late responders differ on certain demographic characteristics such as grade level, supervisory status, gender, tenure with agency, and intent to leave or retire. Our analysis focuses on large and independent Federal agencies so as to eliminate agencies with smaller sample sizes. Our findings provide insight about how a shorter fielding period and thus lower response rates (i.e. by including only early responders) affects sample characteristics and resulting estimates of Federal employee surveys.

-Richard Sigman¹, Taylor Lewis², Naomi Dye Yountr¹, Kimya Lee²

¹Westat, 1600 Research Blvd Rockville, MD 20850

²U.S. Office of Personnel Management, 1900 E Street, NW, Washington, DC 20415

Title: Modeling Nonresponse in Establishment Surveys: Using an Ensemble Tree Model to Create Nonresponse Propensity Scores and Detect Potential Bias in an Agricultural Survey

Abstract:

Increasing nonresponse rates in federal surveys and potentially biased survey estimates are a growing concern, especially with regard to establishment surveys. Unlike household surveys, not all establishments contribute equally to survey estimates. With regard to agricultural surveys, if an extremely large farm fails to complete a survey, the United States Department of Agriculture (USDA) could potentially underestimate average acres operated among other things. In order to identify likely nonrespondents prior to data collection, the USDA's National Agricultural Statistics Service (NASS) began modeling nonresponse using Census of Agriculture data and prior Agricultural Resource Management Survey (ARMS) response history. Using an ensemble of classification trees, NASS has estimated nonresponse propensities for ARMS that can be used to predict nonresponse and are correlated with key ARMS estimates.

-Morgan Earp¹, Melissa Mitchell², Jaki McCarthy², & Frauke Kreuter³

¹Bureau of Labor Statistics

²National Agricultural Statistics Service

³Joint Program in Survey Methodology, University of Maryland

Title: Detecting and Treating Verified Influential Values in a Monthly Retail Trade Survey

Abstract:

In survey data, an observation is considered influential if it is reported correctly and its weighted contribution has an excessive effect on a key estimate, such as an estimate of total or change. In previous research with data from the U.S. Monthly Retail Trade Survey (MRTS), two methods, Clark Winsorization and weighted M-estimation, have shown potential to detect and adjust influential observations. This paper discusses results of the application of a simulation methodology that generates realistic population time-series data. The new strategy enables evaluating Clark Winsorization and weighted M-estimation over repeated samples and producing conditional and unconditional performance measures. The analyses consider several scenarios for the occurrence of influential observations in the MRTS and assess the performance of the two methods for estimates of total retail sales and month-to-month change.

-Mary H. Mulry, Broderick E. Oliver, Stephen J. Kaputa, U.S. Census Bureau

Title: Analytic Tools for Evaluating Variability of Standard Errors in Large-Scale Establishment Surveys

Abstract:

Large-scale establishment surveys often exhibit substantial temporal or cross-sectional variability in their published standard errors. This article uses a framework defined by survey generalized variance functions to develop three sets of analytic tools for evaluation of these patterns of variability. These tools are for (1) identification of predictor variables that explain some of the observed temporal and cross-sectional variability in published standard errors; (2) evaluation of the proportion of variability attributable to the predictors, equation error and estimation error, respectively; and (3) comparison of equation error variances across groups defined by observable predictor variables. The primary ideas are motivated and illustrated by an application to the U.S. Current Employment Statistics program.

-MoonJung Cho, John Eltinge, Julie Gershunskaya, Larry Huff, Bureau of Labor Statistics

Title: In Search of Motivation for the Business Survey Response Task

Abstract:

Increasing reluctance of businesses to participate in surveys often leads to declining or low response rates, poor data quality and burden complaints, and suggests that a driving force, that is, the motivation for participation and accurate and timely response, is insufficient or lacking. Inspiration for ways to remedy this situation has already been sought in the psychological theory of self-determination; previous research has favored enhancement of intrinsic motivation compared to extrinsic motivation. Traditionally however, enhancing extrinsic motivation has been pervasive in business surveys. We therefore review this theory in the context of business surveys using empirical data from the Netherlands and Slovenia, and suggest that extrinsic motivation calls for at least as much attention as intrinsic motivation, that other sources of motivation may be relevant besides those stemming from the three fundamental psychological needs (competence, autonomy and relatedness), and that other approaches may have the potential to better explain some aspects of motivation in business surveys (e.g., implicit motives). We conclude with suggestions that survey organizations can consider when attempting to improve business survey response behavior.

-Vanessa Torres van Grinsven,¹ Irena Bolko,² and Mojca Bavdaž²

¹Utrecht University and Statistics Netherlands

²University of Ljubljana

Title: Data Smearing: An Approach to Disclosure Limitation for Tabular Data

Abstract:

Statistical agencies often collect sensitive data for release to the public at aggregated levels in the form of tables. To protect confidential data, some cells are suppressed in the publicly released data. One problem with this method is that many cells of interest must be suppressed in order to protect a much smaller number of sensitive cells. Another problem is that the covariates used to aggregate and level of aggregation must be suppressed before the data is released.

Both of these restrictions can severely limit the utility of the data. We propose a new disclosure limitation method that replaces the full set of micro-data with synthetic data for use in producing released data in tabular form. This synthetic data set is obtained by replacing each unit's values with a weighted-average of sampled values from the surrounding area. The synthetic data is produced in a way to give asymptotically unbiased estimates for aggregate cells as the number of units in the cell increases. The method is applied to the U.S. Bureau of Labor Statistics Quarterly Census of Employment and Wages data, which is released to the public quarterly, in tabular form, aggregated across varying scales of time, area, and economic sector.

-Daniell Toth, Bureau of Labor Statistics