

Discussion: FROM FEASIBILITY TO IMPLEMENTAION October 2013

Edwin L. Robison¹
U.S. Bureau of Labor Statistics
2 Massachusetts Ave NE, Washington DC 20212

Invited Session: Research Developments in the Analysis of Repeated Survey Data

This paper summarizes my discussion of three papers presenting research on different aspects of weighting and using data from a repeated (periodic) survey: calibration, composite estimation, and small area estimation. The principal survey referenced by the papers is the Current Population Survey (CPS). The CPS is jointly sponsored by the U.S. Bureau of Labor Statistics (BLS) and the U.S. Bureau of the Census (Census). For each paper an abstract is given here that modifies the original abstract to better fit the presentation of this discussion.

Single-Stage Generalized Raking Weight Adjustments in the Current Population Survey (Eric Victor Slud and Christopher Gries, U.S. Bureau of the Census)

General Topic: Calibration

Abstract: This research concerns the adaptation to CPS of single-stage weight adjustment techniques to replace multiple stages of weight adjustment. The techniques involve weight optimization with respect to a loss function subject to population-control constraints. There are additive penalty terms for discrepancies between weight-adjusted survey totals and corresponding known or base-weighted estimated totals for certain survey attributes, and there is an additional nonlinear penalty term designed to force weights not to be too different from design weights scaled to the population total. The novel elements of the current research include: defining appropriate quadratic penalty terms corresponding to current CPS nonresponse adjustment and three-step second-stage ratio estimation; developing a methodology to define penalty multipliers by tracking properties of the current CPS weights across weighting stages; enforcing weight compression by a penalty term in place of the current CPS approach based on cell collapsing; and implementing the method on CPS data for detailed comparison with the weights as currently adjusted in CPS.

Composite Estimation in the Current Population Survey (Jun Shao and Zhou Yu, University of Wisconsin-Madison)

General Topic: Composite Estimation

Abstract: A composite estimator is currently used by CPS to estimate monthly totals of employment, unemployment, and other characteristics of the civilian non-institutional population. The current composite estimator has a non-negligible bias (unless some conditions are satisfied). The bias is studied as well as the variance. An adaptive method is developed to find the best values of the parameters used in composite estimation. Other issues, such as the use of administrative information and variance estimation, are also studied.

¹ Views expressed are those of the author and do not necessarily reflect the views or policies of the Bureau of Labor Statistics.

Analysis of Longitudinal Complex Survey Data Using Parametric Bootstrap (Snigdhanu Chatterjee, University of Minnesota and Partha Lahiri, University of Maryland College Park)

General Topic: Small Area Estimation

Abstract: A general theory is developed for a parametric bootstrap method, in the context of a general mixed model. Its usefulness in analyzing a variety of longitudinal complex survey data is illustrated. As a special case the problem is considered of estimating small area characteristics, using a time series and cross-sectional model that combines data from previous time points of a longitudinal survey and relevant auxiliary variables. The general methodology is used to demonstrate the flexibility of the parametric bootstrap method to produce highly accurate (low mean squared error) estimates of complex estimators with calibration and Winsorization adjustments.

Main Ideas Presented in the Discussion

The three papers cover research into possible calibration, composite estimation, and small area estimation methodologies for complex surveys that are regularly repeated, in particular the monthly Current Population Survey. For each paper this discussion presents an example is given of allied BLS research and outlines what is needed to move from feasibility to implementation.

Due to the importance of the CPS, the Bureau of Labor Statistics (BLS) is very careful and/conservative about making methodological changes. Once the feasibility of a methodology is established, it can be a “long way” to adopting the methodology for implementation, gauged either by the calendar time it takes or the amount of additional research work that is needed.

There are other considerations for implementation, but the limited scope or limited amount of data used for most feasibility research is emphasized. For a major program like CPS it is necessary to follow up feasibility study with research in more depth.

The Current Population Survey and its Importance

The CPS is a monthly survey primarily aimed at collecting labor force information for the United States and for each individual state. Two reasons are given here for considering CPS to be an “important” survey. First, CPS is the source of the official national unemployment rate, designated by the U.S. Office of Management and Budget to be a Principal Federal Economic Indicator (PFEI). Second, CPS data is a key contributor to measures that are used to determine the distribution of some federal funds to states and localities. What follows in this section is a brief summary of current CPS methodology; just enough to give some context to the areas studied in each of the three papers (Census, 2006).

Each month about 60,000 occupied housing units are eligible to be included in data collection, and about 55,000 of those respond. Each month detailed information including labor force data is obtained for about 105,000 adults. The CPS sample can generally be described as a two-stage stratified probability sample – large areas called Primary Sampling Units (PSUs) are selected, and then housing units are selected from those PSUs. In addition to national data reliability needs there are also state reliability

needs, so samples in less populous states are denser than samples in more populous states.

There are several stages in CPS weighting. A housing unit's base weight is the inverse of its probability of selection. There is a housing-unit-based nonresponse adjustment procedure; simple weighting class ratio adjustment is used. All subsequent weighting procedures are person-based, so persons in the same household can end up with different weights. Four procedures use forms of ratio estimation or benchmarking that force weighted CPS data to match population "control" figures provided by the Census Bureau (controls derived "externally" to CPS): first-stage ratio adjustment, national coverage adjustment, state coverage adjustment, and a raking procedure called second-stage ratio estimation. The second-stage procedure has three steps (national race, national ethnicity, and state) that are iterated 10 times to ensure that the number of weighted-up respondents can almost exactly match a large set of population controls. The symbol $y(t)^{SS}$ is used to indicate a simple weighted estimate for month t that uses second-stage weights, that is the weights that result after all of the weighting procedures through second-stage ratio estimation have been applied. A final weighting procedure, called composite weighting, is also a raking procedure with three steps (national race, national ethnicity, and state).

Note: The paper by Slud and Grieves looks at combining the nonresponse adjustment procedure and the second-stage ratio adjustment procedure. This discussion also refers to BLS work on combining weighting steps.

Composite estimation is a procedure that exploits month-to-month sample overlap in the CPS to lower variances on month-to-month change. The CPS sample is divided into 8 panels. Any given panel is interviewed four consecutive months, is dropped for 8 months, and then is brought back for another 4 consecutive months of interviewing (a total of 8 months in the sample). Each month two panels are rotated out (dropped) and two panels are rotated in, so 6 of the 8 panels are in common from one month to the next. Those 6 panels in common can be used to compute a month-to-month change estimate $\Delta(t-1,t)$ that can be used in the formula $y(t-1)^{SS} + \Delta(t-1,t)$ to update an estimate from the prior month to the current month. An example of a simple type of composite estimate is $y(t)^C = .6y(t)^{SS} + .4[y(t-1)^{SS} + \Delta(t-1,t)]$ which is a simple weighted average of two different estimates of the same quantity of interest. In a continuing system a more common form is $y(t)^C = .6y(t)^{SS} + .4[y(t-1)^C + \Delta(t-1,t)]$ where it is the composite estimator $y(t-1)^C$ from the previous month that is updated. Weighting coefficients can be optimized (.6 and .4 are just provided as an example). There are more complicated forms of the composite estimator with other parameters (CPS adds an extra term). Whether simple or complicated, knotty potential bias problems arise. For example, CPS has known month-in-sample bias where panel estimates consistently are different depending on how many times a particular panel has been included in the sample – and that bias affects all terms of the CPS composite estimator. (Composite weighting uses selected sets of composite labor force estimates as controls in the three raking steps.)

Note: The paper by Shao and Yu considers an adaptive method of optimizing composite estimation parameters. The discussion focuses on related BLS work.

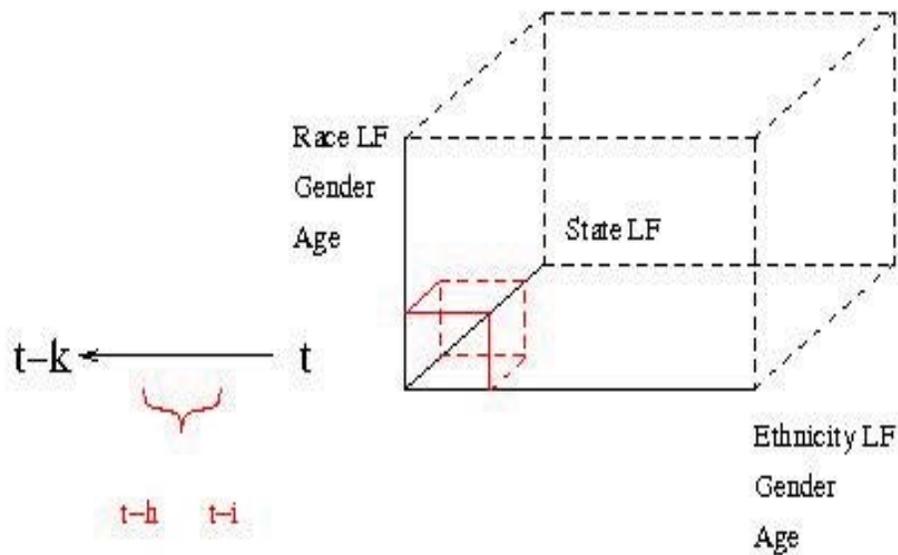
For subnational CPS data, small area estimation techniques are needed (BLS, 1997, Chapter 4). Even with specified minimum reliability criteria, states with the smallest populations have coefficients of variation for unemployment of 15% - 20%. "Noise" due to sampling error masks seasonality and economic trends/changes. BLS uses state-space

time series models to produce official employment and unemployment estimates for each state. Substate estimates are produced by the Local Area Unemployment Statistics program (LAUS) using a custom-designed methodology known as the “handbook” method.

Note: The paper by Chatterjee and Lahiri considers a parametric bootstrap method of small area estimation for complex periodic surveys. The discussion refers to related BLS work on time series models for producing state labor force estimates.

The 3-D Picture for Feasibility Studies

The following picture was used at the 2013 Joint Statistical Meetings to guide a general conceptual discussion of CPS feasibility studies. The larger box represents all data for CPS for a given month, the entire “data space”. Each axis of the larger box represents an important data product. Vertical: Race labor force data by gender and age is specified, but a lot of other related published data by race is not specified (examples: wages, education, and veteran status). Horizontal: Ethnicity labor force data by gender and age is specified, but a lot of other related published data by ethnicity is not specified. Depth: State labor force data. Do not construe any axis as a logical linear representation of data. Think of anything not on the axes, including the entire interior of the box, as unpublished data. The maximum data available for a feasibility study is shown as the latest month t going back to month $t-k$.



The smaller interior box indicates a limited scope of estimates that are examined for a feasibility study. Do not be literal in interpreting the smaller box. A feasibility study may be limited to partial data on one or two axes. For example, race data by gender may be looked at without considering age. A feasibility study may not examine any unpublished data at all. The bracketed segment from $t-i$ back to $t-h$ on the timeline indicates that a feasibility study may be limited to a slice of time. The discussion characterizes a feasibility study as research that homes in on promising methodologies based on limited data.

From Feasibility to Implementation

An important part of moving from feasibility to implementation is further study that conceptually fills up the bigger box. Look at other published data on the axes. If only two axes are in the feasibility study, move on to look at the third. Look at some unpublished data. A proposed methodology designed for a subset of variables (in a loose sense optimized for that data) may not do well when applied to other data. A proposed methodology designed when looking at “topside” data (large aggregates such as total employment) may not perform well for breakdowns to finer-level subpopulations.

Another important part of moving from feasibility to implementation is further study that extends the slice of time used. For periodic surveys, with their emphasis on measuring change across time, methodologies that are stable across time are desirable. What works during a stable economy may not do so well during a boom or a slump. Dynamic parameters may be needed for composite estimation or modeling to properly react to underlying changes in data over time.

There are other factors to take into consideration when moving toward implementation of a proposed methodology. (These were specifically mentioned in the discussion.)

- Efficiency, especially in terms of cost/variance tradeoffs
- Is one methodology best for all data?
- Is it sustainable in the future? (software; can the staff handle it)
- Breaks in series
- Impact on seasonal adjustment
- Selling the methodology to sponsors

Example 1, CPS Weighting

A BLS study (Cruz, Robison, and Zimmerman, 2006) has similarities to the presentation on single-stage generalized raking. Second-stage CPS weighting is a 3-step raking procedure that is followed by composite weighting, another 3-step raking procedure. To some extent composite weighting deconstructs second stage weighting. For any second-stage population control, composite weights can be used to make an estimate of the same population -- but usually the two will not be equal. The study established that a more complex raking could match all controls of both weighting steps.

Feasibility was demonstrated, but the procedure was not implemented. When “filling up the box” and looking at more data, systematic differences were found between the procedure’s results and official CPS estimates. The systematic differences had no clear statistical basis. Also, some relatively large (presumably undesirable) weight changes were noticed for not-in-labor-force persons for some unpublished demographic groups.

The work by Slud and Grieves is statistically more subtle than my 2006 work, and I think their work has a lot to offer. The weighting procedures they are combining are nonresponse adjustment that is household based and second-stage weighting. The research concentrates on topside labor force estimates. It is logical to extend the research to look at the impact on household data, but “filling up the box” is secondary for now -- since demonstrating feasibility is still underway. Software needs to be expanded to handle more controls, and evaluating the effect of intervening weighting steps is needed

(first-stage ratio estimation, national coverage adjustment, and state coverage adjustment).

Example 2, CPS Compositing

BLS has been pursuing work on generalized compositing (Erkens, 2012). The main benefit of compositing is that smaller variances can be obtained for estimates of month-to-month change. The main problem is that the current compositing procedure produces labor force estimates that are systematically different from “unbiased” second-stage estimates. (It is recognized that second-stage estimates are not “unbiased” but no credible statistically defensible explanation has been given for the systematic differences.) The BLS work thus far keeps the main benefit and eliminates the problem, but the work has been done only on topside data. Even though compositing for the work is done at a coarser demographic level than official CPS compositing, feasibility of the methodology has been demonstrated. The BLS work used over two decades of data, but it remains to “fill in the box”.

For the methodology presented by Shao and Yu, similar conditions apply when moving forward to implementation. However, more work is needed to demonstrate feasibility. Variance approximations are key to the presented results. Variance approximations are not used in the BLS work by Erkens; instead directly computed replicate variances are used. The BLS concern is that the advantages of composite estimation are subtle and that the form and parameters used in variance approximations could strongly influence results.

Example 3, Small Area Estimation -- CPS Models of State Unemployment and Employment

CPS monthly state samples are “small” in the sense that monthly estimates and estimates of change have relatively high standard errors, and that limits their use for economic analysis. Using CPS data, models are used by BLS to make official monthly estimates of employment and unemployment. The models have been through decades of development and official estimates are now made using “3rd-generation” models. Theory has been worked out for a 4th generation of models with better benchmarking and better information on trend and error components (Tiller and Pfeffermann, 2010). Initial programming has demonstrated the feasibility of 4th generation methodology using the entire time series of data back to 1976. That is, “the box is filled in” for this project, so why not implement the methodology right away? Other considerations come into play. The methodology is complex and needs further validation. Work on ease of running the software and future sustainability is needed since each state runs the software. State users need to be educated on the interpretation of historical time series changes. The impact on seasonal adjustment needs assessment. The methods need to be sold to the BLS program office that is the sponsor.

The paper by Chatterjee and Lahiri on small area estimation presents a very different parametric bootstrap methodology. The method has both time series and cross-sectional elements. For this methodology more work is needed to demonstrate feasibility, with “filling in the box” and extending the time frame. Even if feasibility is demonstrated, an important consideration is the following: for state labor force estimates this methodology would be in direct competition with a well-entrenched BLS methodology. However, it is indicated in the paper that there is potential to extend the work to substate areas.

References

Cruz, Samantha, Edwin L. Robison, and Tamara S. Zimmerman. 2006. Iteration of Second-Stage and Composite Procedures in the Current Population Survey. *JSM Proceedings*, Section on Survey Research Methods. Alexandria, VA; American Statistical Association.

Erkens, Greg. 2012. Changes in Panel Bias in the U.S, Current Population Survey and its Effects on Labor Force Estimates. *JSM Proceedings*, Section on Survey Research Methods. Alexandria, VA; American Statistical Association.

Tiller, Richard and Danny Pfeffermann. 2010. Time Series Modeling of Official Statistics with Two-Stage Benchmarking. *JSM Proceedings*, Section on Survey Research Methods. Alexandria, VA; American Statistical Association.

U.S. Census Bureau. 2006. *Current Population Survey Design and Methodology, Technical Paper 66*.
(www.census.gov/cps/methodology/techdocs.html)

U.S. Bureau of Labor Statistics. 1997. *BLS Handbook of Methods*.
(www.bls.gov/opub/hom; also see www.bls.gov/lau/laumthd.htm)