SAMPLE REDESIGN FOR THE CURRENT EMPLOYMENT STATISTICS SURVEY

Shail Butani, George Stamas[1], Bureau of Labor Statistics; Michael Brick, Westat
Shail Butani, BLS, 2 Massachusetts Ave., NE Room 4985, Washington, DC 20212

Key Words: Multi-establishment firms; optimum allocation; collocated permanent random numbers; and atypical units.

I. Introduction

In this paper, we present some issues as well as empirical results pertaining to the sample redesign for the Bureau of Labor Statistics' (BLS) Current Employment Statistics (CES) Survey, commonly known as the payroll survey. CES is a survey of 390,000 businesses that is conducted monthly in a Federal/State cooperative environment to produce estimates of employment, payroll, and hours worked. The month-to-month movements in the CES series are closely followed by policy makers and forecasters as timely indicators of the overall strength and direction of the nation's economy.

We first give a brief background on the redesign; a detailed overview of the redesign is given in Werking (1997). In Section III we state the goals and in Section IV we give a description of the BLS' Business Establishment List (BEL) that serves as a sampling frame for the CES and as a source of "benchmark" or "known population" values for employment that lag in time. BEL is based on administrative records. Next, we give a profile of multi-site employers as they play an important role during sample allocation, selection, and estimation phase. In Section VI, we describe the sample design features such as sampling unit, stratification, allocation and selection cells, etc. The usage of permanent random numbers for frame maintenance, updating probabilities of selection, inclusion of new businesses (births) and removal of out-of-business units (deaths) and sample rotation is given in Section VII. Finally, Section VIII covers the transition phase followed by sections on performance measures of the new design and a brief summary.

II. Background

The CES is a quota sample whose inception over 50 years ago predates the introduction of probability sampling as the internationally recognized standard for sample surveys. Under the current design, the estimator for employment is: $\hat{E}_{h,t} = \frac{\sum_{i \in m \cap h} emp_{i,t}}{\sum_{i \in m \cap h} emp_{i,t-1}} \hat{E}_{h,t-1}$

where, m is the set of matched units i that responded in both month t and $t-1$ and the estimate is computed at the estimating cell level h. Estimating cells are usually defined by geography (national, state or metropolitan area) and industry. Some estimating cells for national estimates are defined by broad size class or geographic region in addition to industry. An estimate for a larger domain is the sum of estimates for the component estimating cells. With every March estimate, $\hat{E}_{h,t}$ is aligned with the "benchmark" level obtained from the BEL.

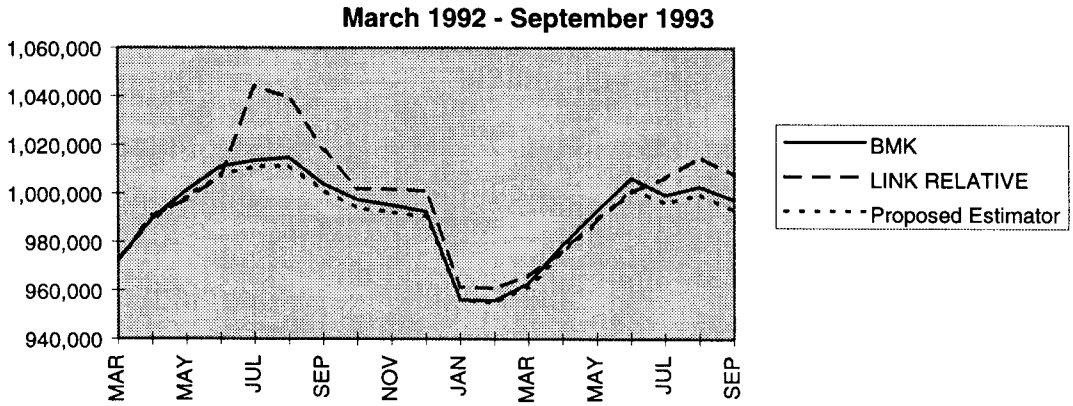
This estimator is known as the link relative. It implicitly adjusts for nonresponse and nonparticipation by assuming the trend is the same for respondents and nonrespondents. Quota samples as well as the link relative estimator are known to be at risk for potentially significant biases; these findings for CES were confirmed by recent research. In particular, the risk associated with the link relative is that often the units in estimating cells are not homogeneous and they are not weighted to compensate for differential probabilities of selection. The graph on the next page illustrates the bias that occurred in a probability based sample from using the link relative estimator. This bias is mainly arising from lack of using selection weights in the formula.

III. Goals of the Redesign

This paper describes the research conducted on a proposed probability based sample design for the CES. A few of the major goals of the redesign are as follows.

- The design should produce and maintain over time prespecified levels of reliability at the national, state, and Metropolitan Statistical Areas (MSA) levels by various levels of Standard Industrial Classification (SIC) code.

[1] At present, George Stamas is employed at Federal Aviation Administration.



- Users should have the flexibility to produce estimates at any geographic (e.g., county) and industry level (e.g., 4-digit SIC) and at any point in time with an estimate of its reliability.

- The total employment estimates (not seasonally adjusted) across various levels of geography and industry should be essentially additive. Under the current design, for various reasons, there is a wide gap between the sum of states and the national estimates.

- The annual revision to "known population values" at the total national level should be about two-tenths of a percentage point or less. A revision of one-half of one percent on a total nonfarm employment of about 120 million is considered to be very large in terms of over-the-year change that is generally of magnitude less than four million.

- The revisions between preliminary and final monthly estimates should be small, i.e. less than 50,000.

- Since CES estimates are typically released about 3 weeks after the reference period, it should be operationally feasible and relatively simple.

A major impediment to introducing probability sampling in CES has been issues associated with data collection. The Bureau during the last few years has invested a great deal of time and effort to improving response rates (Werking and Clayton, 1995) and obtaining timely reporting.

IV. Business Establishment List

The primary source of BLS' BEL is the quarterly contributions reports filed by employers for each unemployment insurance (U.I.) account with their state unemployment insurance agency. The data for both private and public sector workers are reported to BLS after they go through several stages of refinement by the employment security agencies of the 50 states and the District of Columbia as part of the Covered Employment and Wages, or ES-202, Program. Covered employment reported by these sources provide a virtual census (98 percent) of employees on nonfarm payrolls. Among other data elements, this rich and comprehensive database has about seven million records with SIC code, state, county, employment for each month, and total quarterly wages mostly broken down at the worksite or establishment level for each U.I. account.

The quarterly unemployment insurance files are generally transmitted five months after the end of the quarter by the states to BLS. It takes BLS an additional 3 months to process these files through various edits as well as perform record linkage to previous quarters before making it available as the sampling frame. The first quarter file of the current year, therefore, is generally available at the end of December or early January of next year for sampling and for "benchmarking" purposes. For example, in February of 1997, BLS can draw a sample from the first quarter 1996 file and at the same time ratio adjusts ("benchmark") to "known population values" the CES sample based estimates for the period April 1995 through March 1996; sample based estimates for April 1996 through January 1997 can then be recomputed using this updated information. At present, the benchmarking process is done with the release of May 1997 data in June.

V. Characteristics of the Multi-site Employers

As in most establishment surveys, multi-site employers (i.e., employers that own more than one establishment or worksite) have major implications for the CES redesign with respect to sampling unit, stratification, allocation, selection, data collection, and estimation. Data tabulated from the BEL indicate that:

- These employers comprise about 3.6 percent of all private in-scope U.I. accounts, 16 percent of all worksites (reporting units), and over 35 percent of the total private employment.

- About 93 percent of all private sector multi-site employers operate in only one 2-digit SIC and 90 percent are in one 4-digit SIC. Additionally, 99 percent operate in 3 or fewer in-scope 2-digit SICs.

- A total of about 6,000 U.I. accounts with 1,000 or more employees include 215,000 reporting units (sites) and about 17 percent of employment. There are 23,000 U.I. accounts with 250 or more employees and include 425,000 reporting units and about 25 percent of the total employment.

- On average, these accounts have sites distributed across 2 to 3 MSAs.

- There are no clear differences in the pattern of employment change between single and multi-site employers. Stratification by single and multi-site accounts does not appear to reduce the variance of the estimates.

VI. Sample Design

General parameters-- While both the MSA and industrial estimates are important, the consensus among the "workgroup" was that the reliability of the total employment estimates for states would be the top most consideration. Because of the large sample size, the national sample would be the sum of the state samples. Due to budget and workload considerations, the sample size for each state would not change from its present level; thus reliability would vary from state to state.

Sampling unit-- One of the first major decisions was to determine what the sampling unit should be for the redesign; especially for multi-site employers. Possibilities included reporting unit, MSA/U.I. account, or U.I. account. We selected the U.I. account to be the sampling unit because: (1) it is the unit of reporting by employers at the state level; (2) it facilitates tracking of breakouts and mergers; (3) it permits the greatest flexibility in terms of subsampling; and (4) it allows a mechanism for capturing opening of new worksites for multi-establishment firms.

Using BEL data, we tabulated over the year changes in employment broken down by single worksite and multiple worksite employers for the periods March 1990 to March 91, March 91 to March 92, March 93 to March 94, and March 94 to March 95. Even though the four periods cover both a recession and an expansion time, the employment from opening of new worksites accounted for 0.5 to 0.7 percentage point of the total employment. The most important reason for selecting the U.I. account, therefore, is point number four mentioned in the above paragraph; failure to capture this growth could lead to a gross underestimation that is well beyond the normal acceptable level of 0.2 percentage point.

Stratification-- For CES, the three primary variables available for stratification are SIC, employment size class, and MSA. General research conducted by BLS indicates that there is a great deal of variation by size class and by SIC and less so by geography within a state. For example, in general, construction industry is more variable than the manufacturing industry. Similarly, employers with 1-9 employees have different characteristics than those with 250 or more employees.

Frame construction-- A frame consists of a list of U.I. account numbers within state that had employment for any of the months April of the previous year through the latest available March. A 2-digit SIC code is assigned to the U.I. account after determining which 2-digit SIC has the most annual average employment across all of the reporting units under the account; MSA codes are assigned in a similar manner. Each U.I. account is placed into one of the following size classes based on its maximum employment for the year.

1-9; 10-19; 20-49; 50-99; 100-249; 250-499; 500-999; and 1,000 and more employees.

Allocation/stratification cells-- Using BEL data, we simulated several alternative sample designs for 11 states; in each design we chose to minimize the estimated variance for month-to-month change for total employment at the state level. Some of the designs we tried are: 2-digit SICs by 8 size classes; MSA by major industry divisions by 8 size classes; 2-digit SICs for services industries and major industry divisions for the remaining industries by 8 size classes; and probability proportion to size with the file sorted by MSA/4-digit industry/employment.

Based on the results of the simulations, characteristics of the multi-site employers, and in the interest of limiting number of stratification cells for purposes of allocation (e.g., many cells on the frame are empty or have only one U.I. account at a state/2-digit SIC/8 size class), we chose to stratify by industry division (Mining; Construction; Manufacturing--Durable goods; Manufacturing--Nondurable goods; Transportation and Public Utilities; Wholesale Trade; Retail Trade; Finance; Insurance; Real Estate; and Services) and by 8 size classes as defined above.

Certainty stratum-- In our original multi-year simulations, we let the optimum allocation determine the certainty stratum. In the large states, this was generally size class with 1,000 or more employees. The participating states wanted size classes with 500 or more employees to be absolute certainty because of the importance of large employers for MSA estimates. Results of the research indicate that, in general, states gained around six percent in efficiency (ratio of standard errors) as the certainty class cut-off moved from 500 to 1000, and around one additional percent as the cut-off is moved from 1000 to no cut-off. The benefit is, however, largely state-specific. Texas, for example, gains on average about 19 percent from moving the size cut-off from 500 to 1000 and an additional 5 percent from 1000 to no cut-off. Illinois, on the other hand, gains only about 2 percent from 500 to an unrestricted allocation. The final design has size class with 1000 or more employees as the absolute certainty stratum. If the optimal allocation for a stratum requires all of the units even if they have under 1,000 employees, they too will be sampled with certainty.

Cost-- At present, we do not have an estimate as to how much it cost to collect a multi-site U.I. account vs. a single site account. We have assigned a cost equal to the strata average number of UI/County combinations per U.I. number; this cost is utilized in the allocation formula.

Allocation method-- We used an allocation that minimizes the variance for the month-to-month change for total employment at the state level. For each allocation stratum, N_h is the number of U.I. accounts obtained from the current frame, s_h is the estimated average variance of month-to-month change based on several years of historical data, and the cost is as defined above. Additionally, we imposed the constraint that in each allocation cell at least two sample units are allocated provided the cell has two or more population units; cells containing one population unit are sampled with certainty.

Estimated variances are computed as follows. For each allocation stratum and for each of the six years from January 1989 to December 1995, we computed a variance of total employment for each of the 12 months and 12 month-to-month correlations (one for each pair of consecutive months). We then averaged them over the 12 months in a year, resulting in six estimates of stratum variances and six estimates of stratum month-to-month correlation. Within each allocation stratum, we then took the medians of these six estimated variances and correlations to compute the estimated average variance of month-to-month change.

Selection cells-- Because of the importance of MSA estimates, we are selecting the sample by MSA within each allocation cell at the sampling rate for the allocation cell. This provides near proportional allocation across MSAs within any allocation cell.

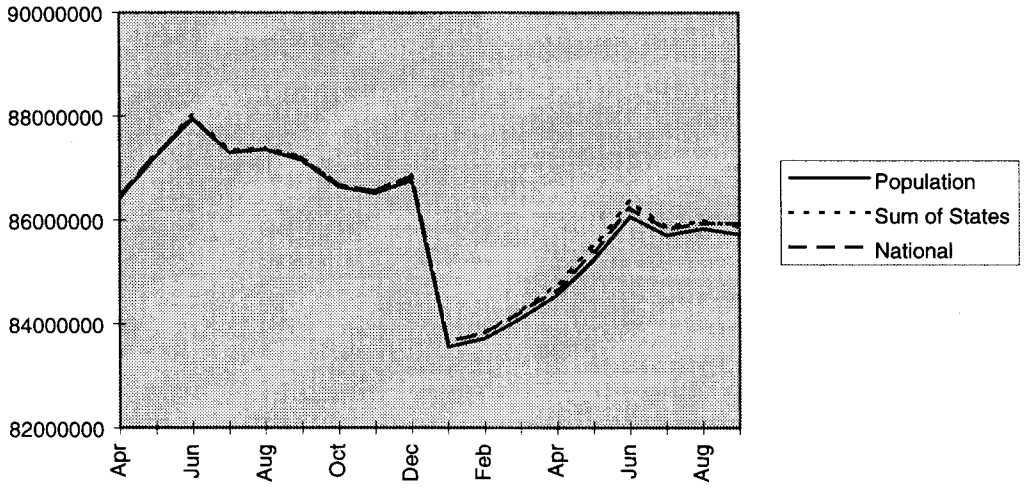
Selection weight-- While selection of continuing UI accounts is based on near optimum allocation, once the sample is selected, the weights are calculated based on the number of U.I. accounts actually selected within each allocation cell; weights are computed as $\frac{N_h}{n_h}$. It is important to note that selection of a U.I. account means all of its reporting units (sites) are selected and assigned the same weight. Method of selection is equivalent to simple random sample; the technique employed is collocated permanent numbers.

Estimation procedures--These procedures including identification and adjustment for outliers, definition of modeling industry, and estimators are described more completely in Butani, Harter and Wolter (1997). Our basic estimator is $\sum y_{it} * w_{it}^{aty} * w_{it}^{RAF}$. Where, the summation is over the appropriate domain as specified by the user; i is the reporting unit (site) level data including the MSA and SIC code; t is the time; y is the data type (all employees, women workers, production workers, hours and payroll); w^{aty} is the atypical weight (selection weight adjusted for outliers); and w^{RAF} is the raked ratio adjustment factor for state estimates. This adjustment factor is computed as: benchmark employment of all non-certainty reporting units divided by the sum of the selection weighted benchmark month employment for the non-certainty units in the sample. Two-dimensional raking is performed: state model cells by MSA industry division totals. For national estimates, w^{RAF} is calculated at the national model industries level with no geography component; hence the term ratio estimator. Thus, each reporting unit is

assigned an atypical weight, a raked ratio adjustment factor for state estimates, and a ratio adjustment factor for national estimates.

Since the sample size is quite large at higher levels of aggregation, the sum of the state estimates are virtually the same as independently produced national estimates. In the empirical study conducted on the simulations of the 50 states plus District of Columbia, the difference at the total private employment level was less than 0.05 percentage point as shown in graph below. This difference is expected to be much smaller than the difference arising from performing independent seasonal adjustments to national and states data.

National industry division employment against replicate 10 Sum of States



Estimates of average weekly hours are derived by taking the ratio of total hours to the total of production workers:

$$\sum y_{Hit} * w_{it}^{aty} * w_{it}^{RAF} / \sum y_{PWit} * w_{it}^{aty} * w_{it}^{RAF} .$$

Similarly, average hourly earnings is the ratio of total payroll to the total hours (Grden and Butani, 1997).

$$\sum y_{\$it} * w_{it}^{aty} * w_{it}^{RAF} / \sum y_{Hit} * w_{it}^{aty} * w_{it}^{RAF}$$

Variance estimation-- As of yet, we have not finalized the estimator of variances. A replication method that accounts for nonrespondents including nonparticipants and finite population correction factor is being considered.

VII. Frame Maintenance

Permanent random numbers-- We are using collocated permanent random numbers (PRN) to: update the probabilities of selection; include new U.I. accounts (births); remove old U.I. accounts (deaths); handle units that have changed size class, MSA, or SIC; and perform sample rotation. The main reasons for using PRN is to achieve the desired sample overlap from one year to the next for the CES sample and to minimize the sample overlap between CES and Bureau's Occupational Safety and Health and Occupational Employment Statistics Surveys. The collocation distributes the PRNs evenly on the interval 0 to 1. (Ohlsson, 1995)

We begin by assigning to each U.I. account on the frame a PRN; collocating PRNs at MSA/3-digit SIC/size class; the frame is then sorted by selection cells, h, and the first n_h units are selected within each stratum. With each quarterly update of the file, new U.I. accounts are identified and assigned PRNs; they are then collocated at the CES selection stratum level. Each quarter we supplement the existing sample with a sample of new U.I. accounts (births) and we also remove from the existing sample all the old U.I. accounts on the frame that cease to exist (deaths). Our plan is to perform frame maintenance once a year as was done in the simulations; once a year, we would resize all the units, reclassify them in their proper SIC and MSA, reallocate, and select the entire sample by shifting the starting point to achieve the desired overlap. For example, if we wanted to rotate one-third of the units in a stratum, then we would begin the selection after the first one-third n_h units. As described below, our plan is to rotate one panel at a time.

Rotation-- At present, the plan is to first phase-in the probability sample for all industries. Once the entire CES is on a probability sample, the plan is to rotate the sample in size class 1, 2, 3, and 4 (i.e., less than 100 employees) provided the sampling weight is greater than 3.000. A possible rotation procedure is to sort the file by state/MSA/size class (1-4), sequentially assign to each stratum numbers 1 to 36 beginning with a random start, and then rotate one panel per month. Rotating roughly 1/36 of the sample every month is subject to change depending on the cost of enrolling new units.

VIII. Transition Phase

The plan that is under consideration is to introduce one industry at a time beginning with wholesale trade industries. The initial plan is to have a production test that would last two years starting in July 1997; the purpose of this test is to work out operational procedures especially with respect to data collection and estimation systems. After that, we are planning to phase in the probability sample and rotate out the old sample of one or more industry divisions with subsequent years' benchmark releases. At benchmark time with March data, we would revise the employment of all the new and old series to align with the ES-202 data. As we rotate new industries, we would add the change from the new sample to the levels of the old sample for all estimates except employment. Once the probability sample is phased-in, we would take one break in series (e.g., average hourly earnings); this way users of the data will get a month-to-month change based on the probability sample but not perpetual break in the level of the series.

IX. Performance Measures

The following are some of the measures we plan to use in evaluating the new design.

- Initial enrollment response rates
- Ongoing response rates
- Target vs. actual relative standard errors
- Revisions between closings to assess imputation methods
- Annual benchmark revisions
- Disaggregation of benchmark revision into birth, death, and continuing units
- Accuracy of imputation methods and the effect on benchmark revision
- Effect of atypicals on the estimates, variances, and benchmark revisions
- Effect of response error
- Monitor edit and screening procedures
- Examine potential breaks in hours and earnings series
- Evaluate seasonal patterns among current CES, ES-202, new CES series

X. Summary

In summary, the Bureau has taken a major step to move one of its most important surveys to a probability basis. While we have attempted to build-in many of the operational aspects into our procedures, undoubtedly, we will be faced with several unforeseen issues. These issues and others that we have thought of but do not know how to handle without more data (e.g., seasonal adjustment) will be dealt with on an ongoing basis.

Any opinions expressed in this paper are those of the authors and do not constitute policy of the Bureau of Labor Statistics or Westat.

Bibliography

1. Butani, S., R. Harter and K. Wolter (1997), "Estimation Procedures for the Bureau of Labor Statistics Current Employment Statistics Program", *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

2. Grden, P. and S. Butani (1997), "A Comparison of Hours and Earnings Estimators", Internal Report, U.S. Bureau of Labor Statistics, Washington, D.C., (unpublished).

3. Ohlsson, E. (1995), "Coordination of Samples Using Permanent Random Numbers", In B.G. Cox et al., *Business Survey Methods*, New York: Wiley, pp. 153-183.

4. Werking, G. (1997), "Overview of the CES Redesign", *Proceedings of the Section on Survey Research Methods*, American Statistical Association.

5. Werking, G. and R. Clayton (1995), "Automated Telephone Methods for Business Surveys", in B.G. Cox et al., *Business Survey Methods*, New York: Wiley, pp. 317-337.