IMPLEMENTATION OF THE CURRENT EMPLOYMENT STATISTICS REDESIGN: DATA COLLECTION

Richard L. Clayton, Bureau of Labor Statistics
2 Massachusetts Avenue, NE, Suite 4840, Washington, D.C. 20212

Key Words: CASIC, Employment, Redesign, Touchtone Data Entry

I. Introduction

The Current Employment Statistics (CES) survey is undergoing a complete redesign. Since its inception, over 50 years ago, the CES has been a panel survey. The CES Redesign (CES-R) is instituting a probability sample design, see Werking (1997). The new sample design holds new challenges for the data collection activity of the survey. This paper profiles the existing and planned data collection environment and discusses the major issues involved in implementing the new sample design.

II. The Current Employment Statistics (CES) Survey

The CES is an ongoing monthly survey of about 390,000 non-agricultural establishments. The CES preliminary estimates, released on the first Friday of each month, which receive vast media coverage, are considered among the most influential data series for economic policy purposes, and are a driving force behind the financial markets. For over 50 years, the CES has been conducted in a cooperative Federal-State system under which BLS provides specifications and methodology which are implemented by the States. States have solicited new respondents, collected and edited the monthly data and made the State and MSA estimates. The same edited microdata are transmitted to BLS for compilation into the national estimates. The data collection environment allows only about 2 weeks to collect data before the preliminary estimates.

Existing Data Collection Environment

For most of the history of the CES, data collection was conducted via a mailed form shuttled between each of the States and respondents. The form provides 12 blank rows for each month of the year. Each month, the respondents fills in the 5 to 6 data items and mails the completed form to the state for key entry, editing. The form was then returned to the respondent for the next month.

This process, while simple, yielded very low response rate for preliminary estimates, about 50%, which lead to frequent, and often large, revisions to the preliminary estimates. To address this problem, the CES has invested heavily in the automation of data collection since 1984 to address long-standing problems of slow reporting and high costs.

The CES is well-suited for CASIC methods because it has a very short collection period (between 2 to 2.5 weeks) and collects a small number (5 or 6) of numeric, commonly available payroll-related data items. Since 1982, the CES staff has researched and implemented Computer Assisted Telephone Interviewing (CATI), Computer Assisted Personal Interviewing (CAPI), Touchtone Data Entry (TDE), Voice Recognition (VR), and Electronic Data Interchange (EDI) as replacements for traditional mail collection, see Werking and Clayton 1995 for a complete description of these methods and their role in changing CES data collection. At this writing, about 250,000 respondents reported monthly using TDE with thousands more in transition. The largest, multi-unit companies report through EDI. Under this array of methods, the average response rates for preliminary estimates have been raised by 20 percentage points and the average monthly revisions to the preliminary estimates have been reduced by 38 percent.

Over the ensuing 15 years, the CES has developed a dynamic collection scheme incorporating CATI, TDE, FAX, EDI. The basic approach was to convert mail respondents to CATI for a short period. During the CATI collection period, usually 6 months, interviewers educated respondents on the importance of the CES, its timeliness requirements and reviewed data quality definitions. Then, respondents were converted to the most appropriate ongoing collection mode. Over 90 percent of respondents are converted to TDE, about 7 percent to FAX and the remaining were returned to mail reporting if no automated method was acceptable. This mixed mode system raised response rates for the preliminary estimates from about 50 percent to 80 percent for the treated units. Revisions to the preliminary estimates were reduced by 39 percentage points.

Each of the automated methods represented a separate set of systems and procedures, targeting different sample groups based on the number of reports, the available technology and their willingness.

III. Profile of Old and New CES

Sampling Unit - UI versus Individual Location: Under the old CES, sample units were generally individual business locations. Under the CES-R, the sampling unit is an individual Unemployment Insurance (UI) number. Every employer covered by UI must apply for a UI number from the respective state. Employers can retain a UI number for the life of their business or sometimes they will find it useful or necessary to change their UI. Common reasons for changing UI numbers are to incorporate, in acquisitions of other businesses, or other changes in ownership, such as adding "& sons". Each such change presents challenges for data collection.

Specific rules have been designed for the range of known UI changes. Interviewers are trained in tracking down a knowledgeable respondent. However, often our respondent does not know their UI number, the scope of the business or other information regarding the UI system. This severely complicates the interaction and our ability to correctly identify and track the UI numbers selected for the sample.

Implications of the UI as a Sampling Unit: The UI number as a sampling unit poses significant challenges for data collection. The most difficult is identifying and capturing the creation of a new expansion unit and the surge of employment in a new location. Thus, it is crucial to immediately identify the existence of the new location. Only if the new site can be readily identified upon opening its doors, or even earlier for hiring and training, can the CES hope to accurately capture the new employment and thus depict the turning points in the economy.

The timely and accurate capture of the new unit, under any collection mode, is based on the respondents' actual knowledge of the event. We cannot assume that our respondents, typically payroll clerks, know about every opening, especially in large multi-unit firms.

It is important to note that the CES estimates are benchmarked annually to the entire universe count derived from the UI system. The quarterly UI filing process includes 3 monthly employment figures. The total of these figures across the US provides a universe count of UI covered employment. This annual event can also result in significant revision to the level of employment. One method of reducing such differences is to obtain the monthly CES data from the same respondents. Thus, the enrolling interviewers are targeting the UI respondent as the likely best sources of consistent monthly data. In a previous study (targeting units which some differences between CES and UI data), we have determined that existing CES respondents are the UI reporters in about 50% of the cases.

Under CATI, interviewers can ask each respondent each month whether a new unit was opened, its employment and insure that the new location is actually under the selected UI number. As already mentioned, many respondents have this knowledge, but some do not.

Importance of Expansion Units: There are two major sources of employment growth in the economy; business births and the expansion of existing businesses (Getz, et al, 1997). Capturing new births will be handled by direct solicitation of new UI accounts or by modeling. As for capturing the expansion of business expansion, there are two sources of expansion. First, individual locations can grow. This growth is relatively easy to capture for units participating in the survey. As they grow, their employment figures reflect this growth.

Second and much more difficult, is the expansion of businesses into new locations. The new employment created at the new location must be captured in the first month it is opened, or even earlier if hiring and training precede its opening. Thus, the data collection environment must specifically focus attention to the monthly potential of each sampled business, as a selected UI number, to add new units. Just because a business has not added units in months or years does not preclude it from expanding in any particular month. It is estimated that 1.5% of existing units open a new location in any single month.

Identifying and Collecting Expansion Units: Each of the CES collection modes must allow for the timely capture of new units. Each method has strengths and weaknesses in this regard. It is important to maximize the effectiveness of each of the modes to avoid a bias due to the ability or inability to capture new units. Under CATI, interviewers must simply ask whether new units, under the targeted UI, have opened or whether new units are being planned. Also, when editing for increases in employment, respondents should be asked whether there the increase is due to new units.

TDE reporting presents the most difficult environment. The TDE script will be expanded asking each unit each month whether a new unit was added, or closed. The recorded answer (1 for yes, 0 for no) will cause an interviewer to initiate a call to verify that the new unit is in the same states as the parent UI number, is under a CES-R sample UI number (and thus is in the same state as the UI number), and actually has employment. Since the strength of TDE is

very inexpensive data collection, the addition of these follow-up calls may dramatically change the cost-effectiveness of this method. We will strive to find ways to improve the cost effectiveness of this feature, such as automating these follow-up questions or only asking the "expansion unit questions" of a subsample of units.

WWW reporting (Clayton and Werking, 1998) is well-suited to this task, or virtually any other special task. Its interactive properties will allow prompting of respondents each month and on-line follow up questions. EDI is the best suited for capturing new locations. The data files received via EDI usually contain records for all units under that business, thus new locations tend to automatically be included.

We expect a few residual units reporting by mail. Again, forms will be modified to accommodate these additional questions, and positive responses will cause telephone follow-up to categorize the special events at the business.

Reporting Level: The level at which a respondent agrees to report also plays into our ability to capture new expansion units. Multi unit reporters can agree to participate in the CES for all of their individual locations under the targeted UI(s) either individually or at some aggregated level, such as a county or state level. Such aggregations not only reduce reporting burden and costs, but help insure that new expansion units are automatically captured in the report. State wide reporting may adversely affect the accuracy of the MSA estimates in that the new employment may not be accurately attributed to the correct county and thus the MSA. Currently, about 4 percent of CES units are reported at either the county-wide or state-wide level.

When initially solicited or enrolled in the CES, interviewers will have to balance reporting burden, costs and the willingness of respondents to provide potentially large amounts of data if there are many individual locations.

Solicitation: The solicitation or enrollment function has traditionally been the domain of each of the states. Under the CES-R, BLS will conduct enrollment through the production tess, about the first 18 months of implementation. Thus, BLS had to develop and test enrollment methods for the first time.

CATI is used as the primary solicitation vehicle. CATI will be used for a period of a few, generally 6 months until the timeliness of monthly data reporting is well established, the maximum number of data items have been obtained and the respondent is fully aware of the CES and its role in the national economic is well-linked to monthly publications and the media uses. A two year test of methods of methods was conducted with the assistance of the University of Michigan Center for Survey Methods. These tests lead to some significant improvements to the original enrollment methodology. The implementation of the CES-R will continue to include ongoing research capabilities including both Michigan and Westat.

Ongoing Collection: Based on the technology available to the respondent, the number of reports and respondent preferences, ongoing collection will be either by TDE, FAX, mail or EDI. For ongoing collection, attrition reduction will be essential. The basic TDE non-response prompting methods involving carefully timed fax and postcard prompts will likely be enlarged to include targeted CATI prompts for cases meeting predictive criteria for likely attrition.

Conclusion: The CES Redesign poses significant challenges to the existing series of available data collection modes and procedures. Changes to sampling unit, a variety of collection vehicles and the increased rigor implicit in probability sampling has caused a complete review of existing methods, including a complete re-engineering of the existing processes, including systems, forms, and each opportunity for respondent contact. Based on the wide array of collection modes available and features available by integrating the strengths of each, the redesigned CES will remain a state of the art collection environment.

References:

Getz, P., Kropf, J, and Strifas, S. (1997), Measuring the Contribution of Business Births and Deaths to Overall Employment Movements, ASA Proceedings of the Section on Survey Research Methods, in print.

Stamas, G., Levin, K., Cantor, D, and Goldenberg, K. (1997) Sampling For Employment at New Establishments in a Monthly Business Survey, ASA Proceedings of the Section on Survey Research Methods, in print.

Clayton, R., and Werking, G., "Business Surveys of the Future: The World Wide Web as a Data Collection Methodology", Monograph for International Conference on Computer Assisted Survey Information Collection, San Antonio, Texas, December 1996.

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

Werking, George S. (1997), "Overview Of the CES Redesign", *Proceedings of the Section on Survey*

Research Methods, American Statistical Association, in print.