CONTROLLING RESPONSE ERROR IN AN ESTABLISHMENT SURVEY

Chester H. Ponikowski, and Sue A. Meily, Bureau of Labor Statistics
441 G St. NW, Room 2821, Washington, DC 20212

I. INTRODUCTION

Response error may simply be defined as the difference between the value obtained from the survey and the desired or true value (Hansen, et al (1961)). The Bureau of Labor Statistics (BLS) has always recognized that response error is a potentially significant source of error in the results from the surveys that the Bureau conducts. In general, response error is difficult and expensive to control and measure. With the introduction of Computer Assisted Telephone Interviewing (CATI) into data collection, it has become more feasible to devise methods for the control and measurement of response errors.

Response error arises in the data collection phase of a survey. It may result from the failure of the respondent to report the true value, of the interviewer to record the value correctly, or of an instrument to measure the value correctly. Historically, the identification and control of response error in the Bureau's establishment surveys has been done through editing and screening of reported data. The editing and screening method cannot detect many definitional inconsistencies. The BLS has developed a record check technique, using CATI, to identify sources of such errors in one of its major establishment surveys and correct subsequently reported data.

This paper presents a discussion of the response error concepts, description of the record check technique used in an establishment survey, analysis of results from the application of the record check technique, and an indication of the future direction of the Bureau's record check technique and measurement of response error.

II. BACKGROUND

The BLS issues monthly payroll employment estimates for the nation from the Current Employment Statistics (CES) survey. The CES survey is a monthly establishment survey of approximately 300,000 nonagricultural units conducted by BLS in cooperation with State Employment Security Agencies. The CES survey provides data on employment, earnings, and hours by detailed industry and geography. The collection of the CES data is done by mail, CATI, and touchtone phone. For a more detailed description of the CES survey, refer to the BLS Handbook of Methods.

In 1984, a pilot test was launched to develop a record check technique for identifying the systematic and consistent error sources. The pilot test was conducted by personal visits to 400 establishments in Florida and Maine. The interview consisted of administering a structured record check questionnaire which obtained a detailed profile of the payroll recordkeeping systems within the establishment, including necessary adjustments to conform to the survey definitions. The results from the pilot test provided information on potential quality loss and options for controlling response error at the establishment level. However, the cost associated with the personal visits precluded the use of this technique in the CES survey.

To overcome the cost issue, the record check questionnaire was restructured and moved from the personal visit mode to the CATI collection mode. The error control features of the record check technique were maintained. A study of the CATI record check technique was conducted in 1985 on a subsample of about 500 CES reporters in Florida and Maine. The results demonstrated (Werking (1988) that the CATI record check technique can identify and rectify many of the systematic and consistent errors which traditional screens and edits miss. To address the concern that the CATI record check technique might yield different results from personal visit interviews, a personal visit reinterview was conducted on a subsample of CATI record check survey responders to identify any reporting discrepancies between the two modes of collection. The results from the reinterview showed no significant differences between the two modes of collection.

In 1988, the study of the CATI record check technique was expanded to nine States and conducted on random subsamples of CES units reporting by CATI and mail. Units collected by CATI are units which were late reporters for the preliminary estimates under mail. The results from this study are presented in the results section of this paper.

III. RESPONSE ERROR CONCEPTS

A. Definitions

Response errors are commonly identified as either response bias or response variance. The response bias consists of constant bias and variable bias. The constant bias is that component of bias

that affects all units in the sample with the same magnitude. That is, the expected value of the bias is the same for each unit. The variable bias, on the other hand, is the component of bias that varies with each unit and may be correlated with the true value. The variable bias in establishment surveys is primarily due to the establishment's recordkeeping system and the respondent's interpretation of the survey questions as they relate to the recordkeeping system. The deviation of the response on a given trial from the expected value for that unit is defined as the response deviation for the unit. The average of the squares of the response deviations summed over all trials and units is called the response variance. The response variance measures the variability in response for a given unit among trials.

B. Sources of Response Error

Sources of response error in establishment surveys arise primarily from the records used for responding, the questionnaire, the respondent, or the interviewer. Frequently, business establishment records used for responding are not consistent with specific survey definitions. Also, compiled records may be incomplete or inaccurate. Questionnaires may not always make it clear or easy for the respondent to determine the data requested. Question sequence can affect responses when it calls for data which are not readily available (OMB Statistical Policy Paper 15 (1988). Respondents contribute to response error by failing to report the correct figure. The failure of the respondent to report correct figures may be due to respondents inadequate knowledge of the establishment data desired, a confusing or lengthy questionnaire, or questions requiring extensive data gathering. Interviewers contribute to response error by failing to record responses correctly or to follow survey procedures. This failure is usually due to inadequate interviewer training, obscure interviewing procedures, too heavy workload, loss of interest in the survey, and inadequate verbal abilities. Other sources that may affect the accuracy of collected data include the mode of data collection (e.g., CATI or mail), the coders and analysts, the amount and type of publicity received by the survey, adverse legislation, and negative publicity or feelings about the survey organization.

C. Measurement of Response Error

A number of mathematical models have been developed for measurement of response errors. The model introduced by Cochran (1977) for the situation where errors of response can be assumed to be uncorrelated from one unit to another is the most appropriate model for establishment surveys. The model is

$$\bar{y}_{i\alpha} = u_i + B + B_i + d_{i\alpha} = u_i' + d_{i\alpha}$$

where

$\bar{y}_{i\alpha}$ is the value obtained in the αth repetition,
u_i is the true value for the ith unit,
B is the constant bias over all units,
B_i is the variable component of bias for the ith unit,
$u_i' = u_i + B + B_i$ is the average value given by the measuring process on the ith unit, i.e., the expected value over all trials on the ith unit,
$d_{i\alpha}$ is an error component which follows some frequency distribution. This component is called response deviation.

The mean square error (MSE) of the mean $\bar{y}_\alpha$ of a simple random sample of n establishments, under the assumptions that:

$E(B_i)=0$, $E(u_i')=u$, $E(d_{i\alpha})=0$, $E(d_{i\alpha}\, d_{j\alpha})=0$, and $V(d_{i\alpha})=\sigma_d^2$ is:

$$MSE(\bar{y}_\alpha) = 1/n(\sigma_d^2 + (1 - f)S_u'^2) + B^2$$

where $S_u'^2$ is the sampling variance, σ_d^2 is the simple response variance, and B^2 is the square of the bias.

The appropriateness of the above model to establishment surveys is due to the fact that responses are obtained from records primarily by self-filled mail questionnaires. In surveys with touchtone collection, data are self-reported by touchtone phone directly into the computer database. Thus in surveys taken from records and reported by mail or touchtone phone it is reasonable to assume that errors of measurement are uncorrelated among units. In surveys with CATI collection, the correlated response variance should be negligible because the respondent is only reading answers from a previously completed questionnaire.

A number of methods can be used to produce a quantitative estimate of the response error component of a sample estimate. The method of repeated measurement of subsamples is one of the most widely used methods. This method

involves selecting a subsample of an original sample and repeating the survey under some fixed set of identical conditions. Another widely used method is the method of interpenetrating subsamples. This method involves dividing at random a random sample of n units into k subsamples so that each subsample contains m = n/k units. The data collection and processing of the sample are done so that there is no correlation between two units in different subsamples. This method will also provide an estimate of the correlated component of the response variance. Other methods used to estimate response error are described in Cochran (1977) and Murthy (1967) including a combination of inter-penetation and repeated measurement.

No direct estimate of response error was obtained during the 1988 record check survey. However, during the 1984 record check survey, a variation of the repeated measurement technique was used. Respondents were asked to quantify the magnitude of the error whenever definitional differences were identified. This study provided a measure of response error on published estimates.

D. Control of Response Error

The measurement of errors is important in that it allows us to assess the overall quality of the reported data in a survey and identify the contribution of each component of error to the total error (sampling and nonsampling). However, control of response error may frequently be more useful than measure-ment of response errors in panel establishment surveys. This requires identifying problem areas and rectifying them, thus permanently reducing total survey error. To accomplish this may involve changes in the questionnaire design, data items definitions, data collection methodology, and data processing.

Screening and editing of reported data are the traditional techniques used to control response error in establishment surveys. These are particularly effective in panel surveys where data are compared over different reporting periods. For example, the CES uses internal consistency edits, range checks, and longitudinal edits.

Edits are designed to detect response error. The range checks are designed to detect response bias. However, range checks are usually too broad to detect many definitional inconsistencies. For example, a range check for average hourly earnings for production workers in a particular industry would most likely not detect the exclusion of vacation pay since vacation hours would likely be excluded too. The record check technique was developed for the CES survey to identify and control response errors that are not detected by screening and editing techniques.

IV. DESCRIPTION OF THE RECORD CHECK TECHNIQUE

The CES survey uses 11 different questionnaires, based on industry, to collect data from different types of establishments. Only one form is sent to a given establishment in the sample. Each CES form contains space to report data values to BLS and a set of definitions. The definitions include a series of inclusions and exclusions for each data item to aid the respondent in data preparation. However, the use and application of these definitions are not guaranteed through either mail or telephone collection of data.

The CES record check instrument is designed to compare the survey definitions to the establishment's recordkeeping system. The objectives are to identify definitional differences in recordkeeping, and to request that identified deviations be corrected in the future. The interviews are conducted by telephone using CATI. The questions are neutrally worded seeking "yes-no" answers. A sample question is: "Have temporary employees been included in the employment counts you have been submitting to us?". If the respondent's answer is "no", the respondent is asked if the temporary employees could be included in the employment counts reported to us in the future. If a category of employees is incorrectly included, the respondent is asked to exclude the group in the future.

To prolong correct reporting, a form is sent to the respondent listing adjustments to the reported data which the respondent agreed to make. A folder is provided to store both the data collection log and the list of adjustments for easy reference.

V. RESULTS FROM THE RECORD CHECK TECHNIQUE

The results obtained from the application of the CES CATI record check instrument provide information on the percentage of units needing adjustments and the percentage of units agreeing to adjust. The results are presented for the CES data items: employment, earnings, and hours. Results are compared across the two collection modes. Also, errors which occur most

frequently within each data item are included.

The sources of errors identified by the record check technique contribute to the constant bias and the variable component of bias in the Cochran error model. The results obtained on units needing adjustments are indirect measures of response error. Direct measures of error magnitude obtained from the 1984 record check survey are also presented.

Employment Data

The CES survey definition of employment appears to be straightforward, that is, persons who worked full-time or part-time or received pay for any part of the pay period including the 12th of the month. However, chart 1 shows that 59 percent of units in the sample were not adhering entirely to the survey employment definition, that is, they needed to make one or more adjustments to the reported employment data. About two-thirds of the reporters agreed to make one or more adjustments on future reports.

CHART 1

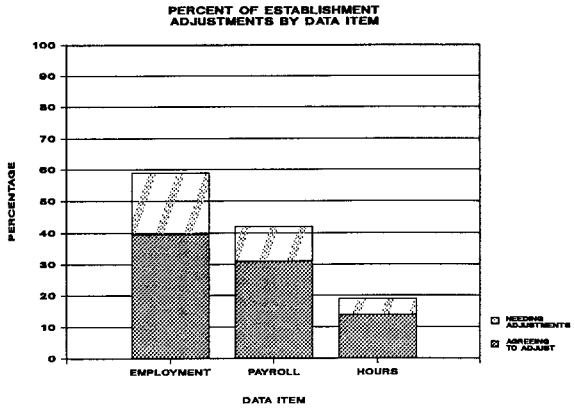

Chart 2 shows the breakdown for CATI and mail collection. Sixty-one percent of CATI units and 58 percent of mail units needed to make one or more adjustments to the reported data. The difference in the percentage needing to adjust between CATI and mail units is not significant at the .05 alpha level.

The most common errors in reported employment data are provided in Chart 3. Overall, more units needed adjustments to exclude employees on unpaid vacation than any other error type, but less were willing to adjust than the other error types shown.

CHART 2

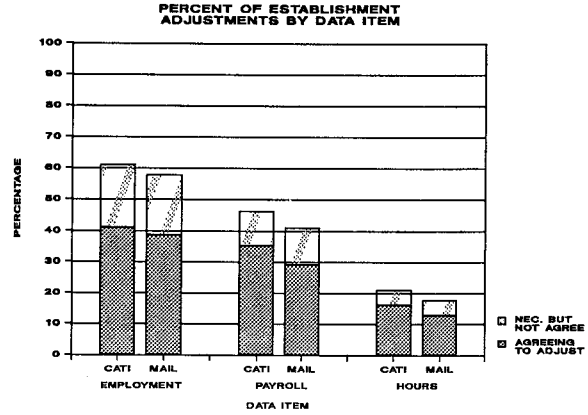

CHART 3

MOST COMMON ERRORS IN REPORTED EMPLOYMENT DATA PERCENT OF UNITS NEEDING/AGREEING TO ADJUSTMENTS

DESCRIPTION OF ERROR TYPE	NEEDING ADJUSTMENTS			AGREEING TO ADJUST		
	Total (%)	CATI (%)	Mail (%)	Total (%)	CATI (%)	Mail (%)
Included Empl. on Unpaid Vacation	11.2	11.6	11.0	55.7	43.5	62.5
Excluded Officers and Executives	5.7	6.9	5.0	80.7	75.7	84.3
Excluded Empl. on Paid Vacation	5.0	5.4	4.8	64.1	58.6	67.3
Excluded Vac. Empl. Paid in Advance	5.5	5.4	5.5	80.0	86.2	76.8

Sample Sizes: Total = 1553, CATI = 534, Mail = 1019

CHART 4

MOST COMMON ERRORS IN REPORTED EARNINGS DATA PERCENT OF UNITS NEEDING/AGREEING TO ADJUSTMENTS

DESCRIPTION OF ERROR TYPE	NEEDING ADJUSTMENTS			AGREEING TO ADJUST		
	Total (%)	CATI (%)	Mail (%)	Total (%)	CATI (%)	Mail (%)
Included Retroactive Pay	15.1	15.9	14.7	86.0	88.2	84.7
Included Nonregular Bonus	9.1	9.4	8.9	91.5	88.0	93.4
Advanced Vacation Pay in Wrong Pay Period	6.4	8.1	5.6	74.0	86.0	64.9
Included Contract Worker Pay	4.2	0.2	6.4	4.5	100.0	3.1
Included Routine Office Pay	3.2	1.5	4.0	75.5	87.5	73.2
Excluded Vacation Pay	3.9	4.5	3.5	40.0	41.7	38.9

Sample Sizes: Total = 1553, CATI = 534, Mail = 1019

CHART 5

MOST COMMON ERRORS IN REPORTED HOURS DATA PERCENT OF UNITS NEEDING/AGREEING TO ADJUSTMENTS

DESCRIPTION OF ERROR TYPE	NEEDING ADJUSTMENTS			AGREEING TO ADJUST		
	Total 1553 (%)	CATI 534 (%)	Mail 1019 (%)	Total 1553 (%)	CATI 534 (%)	Mail 1019 (%)
Included Vacation Hours in Wrong Pay Period	5.5	5.8	5.4	70.9	83.9	63.6
Excluded Vacation Hours	3.3	3.4	3.2	52.9	77.8	39.4

Sample Sizes: Total = 1553, CATI = 534, Mail = 1019

While the incidence of reporting errors seems large, the impact of these errors on the CES estimates of total employment has been found to be quite small. Many of the errors occur infrequently or affect only a small percentage of the employees at an establishment. Also, there is a cancelling effect at the aggregate since some of the error sources result in a positive bias, while others results in a negative bias. Some of the errors are consistent over time and would have little impact on over-the-month change estimates. A small sample record check survey conducted in 1984 found that the reporting errors would result in less than a one percent bias in total employment estimates.

Earnings Data

Chart 1 shows that 42 percent of sample units needed to make at least one adjustment to the reported earnings data. About three-quarters of the reporters agreed to make one or more adjustments on future reports. The difference in percentage needing to adjust between CATI and mail units shown in Chart 2 is not significant at the .05 alpha level.

The most common errors in reported earnings data are provided in chart 4. The majority of adjustments are needed when considering retroactive pay. Approximately 15 percent of the establishments are including retroactive pay when they should be excluding it. In both groups, CATI and mail units seem to be fairly even in percentages of agreeing to adjust except for the error in including contract worker pay. All of the CATI units agreed to adjust, while only 3.1 percent of the mail units agreed to adjust. This is deceiving, due to the small percentage of CATI units needing to adjust (0.2 percent).

The 1984 record check survey found that the effect of reporting errors would result in less than a three percent bias in earnings.

Hours Data

Chart 1 shows that 19 percent of sample units needed to make at least one adjustment to the reported hours data. About three-quarters of the reporters agreed to make one or more adjustments on future reports. The difference in percentage needing to adjust between CATI and mail units shown in Chart 2 is not significant at the .05 alpha level.

The most common errors in reported hours data are provided in chart 5. The inclusion of vacation hours in the wrong pay period and exclusion of vacation hours are only two error types that occur most frequently. Mail and CATI units showed comparable results when needing adjustments, while CATI units showed a higher percentage agreeing to adjust than mail units.

The 1984 record check survey found that the effect of reporting errors would result in less than a three percent bias in hours. Hours and earnings are published as ratios in the CES: average hourly earnings, average weekly hours, and average weekly earnings. It is difficult to gauge the effect of reporting errors on average hourly earnings, but the potential three percent error on hours and earnings should be relevant for average weekly hours and average weekly earnings.

VI. FUTURE DIRECTION

The results indicate that the record check technique can identify incidences of response errors present in the CES survey. Also, the results show some errors occur more frequently and may have a larger impact on total survey error or MSE. Future work in the area of detection and rectification of sources of errors in the CES survey will focus on modifying the record check instrument. The modified instrument will be much shorter and will have as its main objective the detection and rectification of only the most frequently occurring errors, and errors which may have a large impact on the quality of the CES estimates. The current long version of the record check instrument will be used continuously on a small sample of the CES reporters to identify new sources of errors that may arise as a result of changes in recordkeeping systems or corporate law. The use of the two versions will allow us to control and reduce response errors as much as possible at a more reasonable cost.

The modified record check instrument will be implemented on a large scale beginning in early 1990 as part of the CES Large Reporter Project. Over the next few years, large CES reporters, typically with 250 or more employees, will be shifted from mail to telephone collection: computer-assisted telephone interviewing, touchtone recognition, and voice recognition. As these reporters are shifted, the record check interview will be conducted. By improving the reported data from the largest reporters, the greatest reduction in response error of the published estimates is expected to be achieved.

Other future studies will include:

1. Obtaining values for adjustments which were identified and agreed to be made by respondents. This will allow us to evaluate the impact of the record check technique on the CES estimates.

2. Producing quantitative estimates of the response error component of the CES sample estimate using the method of repeated measurement and/or inter-penetrating subsamples. With the establishment of the Business Establishment List at the BLS in the near future it will be possible to produce estimates of response error for employment estimates using a repeated measurement method without additional cost to the CES survey.

3. Follow-up studies to determine when deterioration in reporting occurs and thus, how often record check reinterviews should be conducted.

VII. SUMMARY

The results from the CES record check study indicate survey researchers should take notice of the potential for deviations from survey definitions. The record check technique did not detect differences between modes of data collection in the incidences of response errors present in the CES survey. The CATI record check technique is a feasible and an effective method for detecting and controlling response errors in establishment surveys. The application of this technique is most useful in panel surveys. In a panel survey, response errors, once identified, could be reduced or eliminated in the future reports. Also, it is possible to monitor the impact of the changes on total survey error by continuing to evaluate the data after the changes have been implemented. We encourage the wider use of this technique in establishment surveys in addition to commonly used screening and editing techniques.

This study provides indirect measures of response error in the Current Employment Statistics Survey. The effect on published estimates are not as large as the measures of incidence, or indirect error, may imply. Many of the reporting errors occur infrequently or affect only a small percentage of employees at an establishment. A 1986 study found no significant difference in over-the-year change, except for earnings, between a group of respondents who had been interviewed with the record check instrument and a control group. Most importantly, the annual revisions to the published estimate of total employment, which are a measure of total survey error, have averaged only 0.2 percent over the last five years. The continued focus on controlling response error is intended to reduce further the magnitude of annual revisions.

REFERENCES

Cochran, W. G. (1977). Sampling Techniques, 2nd edition. John Wiley & Sons, Inc. New York.

Hansen, M. H., Hurwitz, W. N., and Bershad, M. A. (1961), Measurement Errors in Censuses and Surveys, Bulletin of International Statistical Institute, 38, pp. 359-374.

Murthy, M. N. (1967) Sampling Theory and Methods, Statistical Publishing Society. Calcutta-35, India.

United States Office of Management and Budget, Statistical Policy Working Paper 15 (1988): Quality in Establishment Surveys.

Werking, George S., Tupek, Alan and Clayton, Richard. (1988). CATI and Touchtone Applications for Establishment Surveys, Proceedings of the Census Bureau Annual Research Conference IV.

Werking, George S., Tupek, Alan R., Ponikowski, Chester H., and Rosen, Richard J. (1986). A CATI Feasibility Study for a Monthly Establishment Survey, ASA Proceedings of the Section on Survey Research Methods, 639-644.