RESPONSE ERRORS IN ESTABLISHMENT SURVEYS WITH AN EXAMPLE FROM AN AGRIBUSINESS SURVEY

Ron Fecso and Brad Pafford, National Agricultural Statistics Service
Ron Fecso, USDA/NASS, Room 4801-S, Washington, D.C. 20250

INTRODUCTION

This paper is part of a session organized to present topics from the Statistical Policy Working Paper 15 - "Quality in Establishment Surveys". Response errors in establishment surveys are discussed and illustrated with agribusiness examples.

DEFINITION OF RESPONSE ERROR

Response error, which occurs in the data collection phase of a survey, may simply be thought of as the difference between the value collected during the survey and the correct value. Response errors may result from (1) the failure of the respondent to report the correct value (respondent error), (2) the failure of the interviewer to record the value correctly (interviewer error), or (3) the failure of an instrument to measure the value correctly. Although the concept of "correct value" is often simple and well defined, its measurement is often difficult and may result in response error. Survey researchers commonly identify response errors as either response deviation or response bias, which is made up of constant bias and variable bias. Constant bias, when it occurs, is a difference between the correct value and the recorded value, the difference being evident over all units in the sample. Variable bias is a change in the difference between correct and reported values for different reporting units. The change in bias may be correlated with the correct value. Response deviation is the component of error associated with differences in the response over repeated measurements of an individual element of the sample. Response deviation is often caused by factors which are unique to the specific interview times, such as, the respondents attention or the interviewer's actions.

Examples

In an agricultural establishment survey, a farmer may report that 160 acres (a quarter of the square mile section which is a common ownership size in the Midwest) are planted in corn when in fact only 154 acres are planted--the remaining 6 acres being roads, streams, irrigation ditches, and the like. This is an example of a respondent error. However, had the enumerator observed the crop growing in the quarter section and recorded 160 acres, the error would be an interviewer error. If interviews at another time or by another interviewer would have resulted in a 154 acre response, the 6 acres would be a response deviation and possibly variable bias. If farmers would always reply 160 acres, the 6 acres are a constant response bias for 160 farms and might be expected to be a correlated response error with, for example, a 12 acre bias for a 320 acre farm.

Response deviation may occur when several persons who are allowable respondents for the establishment have differing knowledge of the value to be reported. For example, although either spouse is often an allowable respondent for family businesses, one may provide more accurate answers than the other (Pafford, 1988). Thus reported values may depend on which spouse is actually contacted. This may also be considered a specification error. In establishment surveys, interviews prior to or after completing tax forms may result in response deviations for these data items since the respondent may have more complete financial knowledge after doing taxes.

The simplest example of response bias is when a measurement instrument is miscalibrated. If the error is constant, it would result in a constant response bias. When the error is proportionate to the measurement, there is a variable response bias which is correlated with the correct value.

SOURCES OF RESPONSE ERROR

The sources of response error in establishment surveys discussed here are grouped into three categories: task error, respondent error, and interviewer error (Bradburn, pp. 289-328 in Rossi, Wright and Anderson, 1983). If an error source is mentioned in only one category, it is done for ease of discussion, and does not imply that sources do not belong in more than one of the categories. Bradburn notes that although "much of the research on response effects has focused on interviewer and respondent characteristics...the characteristics of the task are the major source of response effects and are, in general, much larger than effects due to interviewer or respondent characteristics."

Task Error

The task is the process of obtaining information. It includes what is measured and how it is measured. The formulation of the task often interacts with the enumerator or respondent to contribute to differences in probing, interviewer or respondent behavior, memory, etc.

A questionnaire of excessive length can cause errors resulting from fatigue or boredom of the respondent or the interviewer. Question sequence can affect the responses when it affects recall or creates confusion.

Questionnaire requirements can also contribute to response error. As mentioned previously, permitting multiple respondents can result in respondents with different knowledge of the desired value and thus contribute to response deviation and/or bias. In situations where

multiple respondents are required to complete a questionnaire, the interaction of the group of respondents can cause differences in the reported values.

Records error is a task error which arises from inaccuracy in the records used for responses. Typical causes include inaccurately or incompletely compiled data, the use of inaccurate or out-of-date administrative data, and unavailable or inaccessible records.

Respondent Error

Respondent error, the failure of the respondent to report the correct value, has many causes. The error may be deliberate or may not be deliberate as in the case where the respondent does not have adequate knowledge of the establishment data desired. Confusing or lengthy questionnaires or questions requiring extensive data recall or records gathering can also cause respondent error. The burden of reporting is especially worrisome for small establishments that already suffer considerable time loss completing required tax, employment, and other government program forms.

The timing of an interview can also impact respondent error. Interviews soon after the end of a business cycle, tax preparation, or other reporting period may improve recall, while interviews during busy times may result in rushed responses.

Memory problems may occur. Two causes of memory errors are age and the respondent not considering the requested information to be important. An excessive number of inapplicable questions may cause even the relevant data to suffer. Recall problems include the omission of events or details and telescoping (the inclusion or exclusion of events which are beyond the survey's frame of reference). In establishment surveys in which the respondent is often expected to provide records, the problem may be less severe.

The willingness of the respondent to cooperate also affects the accuracy of responses. This may be influenced by the sensitivity of the information, any sense of possible loss of prestige associated with a response, use of the data for taxation or entitlement programs, the respondent's mood, interest in the survey, level of fatigue, available time, sense of burden resulting from repeated visits, and provisions for a tangible or intangible reward for cooperating.

When responses are gathered using a measurement instrument, response errors have been called measurement errors especially in industrial quality control applications. An inaccurate counter, a faulty scale, or poorly calibrated equipment may cause measurement errors. Sometimes weather conditions such as extreme cold, heat, or humidity, as well as physical conditions such as inadequate work areas contribute to measurement errors.

Events that may increase response errors include negative presurvey publicity, adverse legislation or low prices in the establishment's industry, and negative feelings about the survey organization.

Interviewer Error

Interviewer error, the failure of the interviewer to record responses correctly, commonly results from poor interviewer training or ambiguous guidelines. Deviation from survey procedures is another type of interviewer error. Too heavy a workload may contribute to interviewer error, as does loss of interest in the survey, discomfort with prescribed probing techniques, a negative attitude, fatigue, and inadequate verbal abilities. These factors can cause interviewer error or may result in an interaction with the respondent that promotes respondent error.

The interaction of the respondent with the interviewer or the survey instrument may cause conditioning errors or changes in the response because the respondent perceives a desired answer or realizes that the interview could be shortened.

CONTROL OF RESPONSE ERROR

The most common approach to controlling response error is that reflected by O'Muircheartaigh (U. S. Bureau of the Census, p. 209, 1986): "While it is important to assess the overall quality of the data in a survey, it is frequently a greater concern to identify particular problem areas. Some variables will be more susceptible to unreliability in reporting than others, and some classes of respondents will be less consistent than others in their responses. It would be useful to identify these variables and these types of respondents and to examine the reasons for the lower quality of data they provide.

"Having identified problem areas the next stage should be to change the survey procedures to take the problems into account and if possible to overcome them. This might involve changes in the definitions of, and questions for, the constructs being measured and/or changes in the field work strategy and execution. Such changes are more appropriate in the context of a continuing survey (or of a program of related surveys) than in a single ad hoc survey. In a continuing survey it is possible to monitor the impact of the changes by continuing to evaluate the data after the changes have been introduced."

Some techniques for controlling the previously mentioned sources of response error in establishment surveys are discussed next.

Task Error

Some basic methods used to control questionnaire misspecification include studying establishment recordkeeping practices prior to designing the survey forms, attempting to understand how respondents interpret the questions and answer

them, and using questionnaire pretests. Working Paper 10 (Statistical Policy Office, 1983) provides detail about controlling questionnaire misspecification. Techniques used include: individual and group interviews, interview observations, formal testing, and post survey evaluation.

Studies to check records and to eliminate nonmeasurable data items from the survey or to improve collection methods are useful ways to control records error in establishment surveys.

Respondent Error

A simple method of controlling respondent error in establishment surveys is to check responses against administrative data when they exist. An analyst familiar with the industry may be able to spot responses which are uncharacteristic of establishments in the industry with similar administrative data.

Where respondents must provide data in repetitive contacts, personal contact with the respondents whose data often contain problems may help improve responses. Finally, a computer edit which utilizes all reasonable relationships within the record is essential, as are effective followup procedures.

Recently, techniques from cognitive psychology have been used to study sources of respondent error. Discussions by Loftus and DeMaio, et al. in U. S. Bureau of the Census (1986) are good beginning references.

Interviewer Error

The control of interviewer error starts with detailed and understandable training and procedural guidelines for the interviewers. The management aspects of a survey--recruitment, training, and supervision of the enumerators--must receive proper attention. Testing and well defined, relevant selection criteria during interviewer recruitment can control interviewer error. Supervision practices will vary with the survey conditions such as telephone vs. personal interviews or number of interviewers supervised. Developing good supervisory practices is essential because the supervisors are often the first level at which problems are recognized or corrected. Supervisors can help interviewers understand their job better, provide additional training, and assure that workload does not impact the quality of the work. Field editing may be useful, or when using telephone interviews, on-line monitoring is useful. A reinterview of a sample of the interviewer's work is also a commonly accepted practice.

MEASUREMENT OF RESPONSE ERROR

Since the sources of response error are extremely diverse, the techniques for measuring it are also diverse. Measurement studies have been conducted to:

(1) estimate the precision of survey results,
(2) identify specific survey problems,
(3) identify improvements in the survey methodology, and
(4) monitor the impact of changes to the survey methodology.

The following is a generalization of some of the measurement approaches taken in studies of response error.

The measurement of response errors requires that they be represented by a mathematical model. A number of alternative models have been proposed, often to accommodate special situations. Most sampling textbooks provide an example of an error model and further references. To illustrate, a general response error model (similar to Cochran 1977) is

$$y_{ij} = x_i + e_{ij}$$

$$= x_i + b + b_i + d_{ij}$$

where y_{ij} is the value obtained from the i^{th} element in the j^{th} repetition,

x_i is the correct value,
e_{ij} is the error of measurement,
b is the constant bias term of e_{ij}, if any,
b_i is the variable component of bias which may be correlated with x_i, and
d_{ij} is the fluctuating component of error from repetition j which follows some frequency distribution.

The variations in the response error models which have been developed depend upon the survey itself, the error sources assumed to be a problem in the survey, and the assumptions made about e_{ij}. Survey factors which must be considered by the model formulation include (1) the existence of, or ability to obtain, "correct" values for units in the survey, (2) the complexity of estimation given the sample design, (3) the ability to make remeasurements under reasonably fixed conditions, one of the most difficult conditions to achieve, (4) the ability to randomize work assignments, and (5) budget constraints for these costly measurement studies.

The predominant method of measuring response error involves formulating a response error model, postulating that the survey is repeatable under some fixed set of identical conditions, and measuring the components of variability (response variance) among the repetitions. Interpenetration and reenumeration (or a combination of the two) are commonly used to measure the response variance. Fellegi (1964) presents a framework for the joint application of these techniques while Cochran (1977), Wright (1983), Zarkovich (1966), and the U. S. Bureau of the Census (1985) provide numerous references to approaches taken in different circumstances. A discussion of reinterview methods, sometimes called response analysis surveys, can be found in Working Paper No. 10 (Statistical Policy Office, 1983).

Measurement techniques can also be used as a control method. This approach involves controlling the survey estimates by adjusting the survey estimate to counteract the bias.

Zarkovich (1966) recommends double-sampling approaches which estimate response bias. Basically, this approach consists of selecting a subsample of the original sample, collecting "correct" values for these responses, and forming a difference estimator using the original responses. A limitation is that the "correct value" which is necessary for the approach often can not be obtained. Examples of double sampling can be found in Tenebein (1970), Ostry and Sunter (1970), and Fecso (1986).

Measurement techniques include both indirect and direct measurement techniques. These techniques are discussed next.

Indirect Techniques

Indirect measurement of response error involves examining the information related to response error. This includes the usual survey practice of computing edit failure rates and interviewer error rates. This type of information does not measure the response error, but does provide a reasonable idea of the magnitude of the error. Feedback sessions with respondents and/or interviewers may also help find sources of response error. Questionnaire pretests and cognitive studies, which among other things can help determine whether different word meanings are assumed by different respondents or how recall methods affect response, also provide clues concerning the magnitude of response errors.

Direct Techniques

Direct measurement of response error requires a designed study. The study may be as simple as a records check or may be a detailed content or reinterview study that attempts to control causes of error. Interviewer and respondent variation studies often assume that an identical set of survey conditions have occurred during repeated or randomized assignments of data collection by the interviewer or in repeated inquiries from the respondent. Under such conditions the contribution to error from interviewers or respondents can be measured.

AN EXAMPLE

Statistical Policy Working Paper 15, part of which formed the basis of this paper, presents profile information on the use of various measurement and control procedures for response error in government establishment surveys. Briefly, editing, analyst review and the production of edit failure rates were used by the majority of surveys reviewed by the subcommittee report. Other valuable techniques such as reinterviews, recordkeeping studies and cognitive studies were not widely used. Further detail can be found in the working paper. This section presents an example of a reinterview study in an agricultural survey. A challenging aspect of this type of study is gaining agency momentum to do it. Once done, we found that agency personnel, across the board, found the study reasonably enjoyable, very informative, and a positive experience. Further, we hope the data presented can convince others of the value of the information provided and thus begin the planning for similar studies in other surveys.

A Reinterview Study from Agribusiness

As an example of measuring response error in establishment surveys we present results from a reinterview study used to measure the bias of Computer Assisted Telephone Interviewing (CATI) methods in a National Agricultural Statistics Service (NASS) survey. As part of its estimating program the NASS publishes quarterly estimates of crop acreage intentions and plantings, harvested acreage, stocks of grains, and livestock numbers. These estimates are derived from a multi-purpose, multi-frame survey.

Because of the detailed nature of acreage, stocks and livestock inventory items, the NASS has relied primarily on personal interviews in the past to get the most accurate answers from the farm population. For example, in collecting on-farm grain stocks data, farmers may store these grains in multiple bins on property they own and/or rent. In addition, farmers often are involved in multiple operating arrangements involving their own grains, those of landlords, and those where formal and informal partnerships exist. Correct accounting of these grains is extremely important because of their effect on commodity trading.

Recently, the NASS has made extensive use of telephoning, including CATI to collect these data. The primary reasons for this are the reduced federal budget and the need to reduce the time between initial data collection and publication. The difficulty in using the telephone for collecting some of these quarterly survey data has been suspected, yet no attempts have been made to apply reinterview survey methods to check for response errors. Obtaining accurate responses is considered a problem not only because of the detailed nature of these data, but also because the centralized state telephoning crews lack farm experience and familiarity with farm terms.

Use of reinterview methods is well recognized in the literature for measurement of simple response variance (Bailar, 1968; O'Muircheartaigh, 1986), and correlated response variance (Groves and Magilavy, 1986). The focus in this response error study was to measure the bias by treating the final reconciled response between the CATI and independent personal reinterview response as the "truth". In order to obtain "truth" measures, experienced supervisory field enumerators were used in reinterviewing approximately 1,000 farm operations for the December 1986 Agricultural Survey. It is these results that are reported in the following tables, specifically for the grain stocks items (corn and soybean stocks).

As Table I indicates the difference in the CATI and final reconciled responses, the bias, was significant for all but one item (soybean stocks in Indiana). The direction of the bias indicates

Table I. Estimates of Bias in CATI Collected Responses 1/

State	Corn Stocks % of 1,000 Bu. CATI	Soybean Stocks % of 1,000 Bu. CATI
Minnesota	60,711 (10.4%) *	13,230 (14.9%) *
Indiana	53,218 (17.9%) *	2,998 (5.9%)
Ohio	22,711 (12.0%) *	5,434 (13.7%) *

1/ The sampling design in the original CATI sample was stratified simple random sampling. The reinterview sample was a random sample of CATI respondents within strata. The bias was approximated by expanding the difference in reconciled and CATI responses at the sample unit level.

* - Indicates the CATI and final reconciled responses were significantly different at α = .05.

that the CATI data collection mode tends to underestimate stocks of corn and soybeans.

In the process of reconciliation, the reasons for differences were collected. A summary given below in Table 2 indicates that an overwhelming percent of differences, 41.1% could be related to definitional problems (bias related discrepancies), and not those of simple response variance (random fluctuation). Examples of these definitional problems are rented bins not included, confusion with reporting government reserve grains, failed to include grain belonging to someone else, and bins on son's farm mistakenly included. While not shown, these discrepancies contributed primarily to the large bias with approximately 25 out of the 37 having a relative difference (the reconciled responses minus the CATI response divided by the CATI response) more than 25% or less than -25%. In contrast, the differences due to rounding and estimating contributed little to the overall bias. Approximately 26 of the 28 responses in this category had relative differences of between -25% and 25%.

These results suggest that the bias in the survey estimate generated from the CATI telephone sample might be reduced through a revised questionnaire design, improved training, or a shift in mode of data collection back to more personal interviews. Considering the constraints of time and budget the change to additional personal interviews is unlikely. Thus, the alternative is to use reinterview techniques to monitor this bias over time to determine whether the bias has been reduced through improvement in questionnaires and/or training. If large discrepancies continue the estimates for grain stocks can be adjusted for bias through a continuing reinterview program. If the bias stabilizes, even at zero, periodic reinterview studies are recommended to validate the "constant" bias adjustment which could be used in interim periods.

Table 2. Reason for Differences in CATI and Reinterview Responses for Corn Stocks in Minnesota - December 1987.

REASON	NUMBER	PERCENTOF TOTAL
ESTIMATED/ROUNDING	28	31.1%
"DEFINITIONAL"	37	41.1%
operation crossing state boundaries		
misunderstanding in what to include/exclude		
failed to report reserve corn		
failed to include grain belonging to someone else		
confused with reporting that stored in town		
confusing with reporting landlord corn on farm		
didn't ask/forgot to include last year's crop		
included rented storage on other farms not his		
not all bins included		
rented bins not included		
bins on son's farm mistakenly incl.		
didn't ask about ear corn/only reported shelled corn		
confusion with reporting high moisture corn		
forgot to include grain purchased for feed/seed		
forgot to include storage on the ground		
gave all whole grain/not just corn		
included cracked corn in silage		
didn't include storage used by another		
confusion with storage on acres operated versus where lives		
only included stocks of one partner/the one interviewed		
included only new crop in storage		
thought CATI enum wanted only gov. Program corn stored		
original figure did not reflect Dec. 1 stocks		
OTHER	25	27.8%
TOTAL	90	100.0%

REFERENCES

Bailar, B. A. (1968) "Recent Research in Reinterview Procedures." JASA 63:41-63.

Cochran, William G. (1977) Sampling Techniques. 3rd Edition, John Wiley & Sons.

Fecso, Ron. (1986) "Sample Survey Quality: Issues and Examples from an Agricultural Survey." Proceedings of The Section on Survey Research Methods, American Statistical Association.

Fellegi, I. P. (1964) "Response Variance and Its Estimation." JASA 59:1016-1041.

Groves, Robert M. and Lou J. Magilavy. (1986) "Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys." Public Opinion Quarterly, Vol. 50: 251-266.

O'Muircheartaigh, Colm A. (1986) "Correlates of Reinterview Response Inconsistency in the Current Population Survey." Second Annual Research Conference, Bureau of the Census, March 23-26, 1985 in Reston, Va.

Ostry, Sylvia, and Alan Sunter. (1970) "Definitional and Design Aspects of the Canadian Job Vacancy Survey." JASA 65:1059-1070.

Pafford, Brad. (1988) "Use of Reinterview Techniques for Quality Assurance: The Measurement of Response Error in the Collection of December 1987 Quarterly Grain Stocks Data Using CATI." National Agricultural Statistics Service Research Report, USDA.

Rossi, Peter H., James D. Wright and Andy B. Anderson. (1983) Handbook of Survey Research. Orlando, Academic Press.

Statistical Policy Office. (1983) Approaches to Developing Questionnaires. Statistical Policy Working Paper 10. Washington, D.C.: Office of Management and Budget.

Tenebein, Aaron. (1970) "A Double Sampling Scheme for Estimating from Binomial Data with Misclassification." JASA 65:1350-1461.

U. S. Bureau of the Census. (1986) Proceedings of the Second Annual Research Conference. Reston, VA. U. S. Department of Commerce.

U. S. Bureau of the Census. (1985) Proceedings of the First Annual Research Conference. Washington, D.C.: U. S. Department of Commerce.

Wright, Tommy, ed. (1983) Statistical Methods and the Improvement of Data Quality. Orlando, Fla.: Academic Press.

Zarkovich, S. S. (1966) Quality of Statistical Data. Rome, Food and Agricultural Organization of the United Nations.

ACKNOWLEDGEMENTS

Thanks to members of the Nonsampling Errors Research Section and Bob Tortora for reviewing early drafts and to Patsy Nelson for her word processing wizardry.