QUALITY ASSURANCE PROGRAM FOR THE AGRICULTURAL SURVEY PROGRAM

Brad Pafford, U.S.D.A., NASS
Room 4815 South Bldg., Washington, D.C. 20250

KEY WORDS: bias, reinterviews, reconciliation, CATI, response errors

INTRODUCTION

The Research Division of the National Agricultural Statistics Service (NASS) is developing a more statistically defensible quality assurance program for its surveys of the farm sector. This effort is not what might be thought of as a re-planning effort to fix specific problem areas as is often done to improve quality (Galvin, 1988; Linebarger, 1988). Instead, the research is focusing on a overall change in NASS philosophy on how to attack survey quality (Pafford et. al, 1988; Fecso et. al, 1988).

The reason for this re-direction is due to comments made from a number of sources both within NASS (Crop Reporting Board Standards, 1985) and from outside the agency. Among other things there is interest in increased standardization of NASS procedures, and less manual or "hands-on" control of the survey processes.

This paper will first introduce the Agricultural Survey Program (ASP). Second, the components of the quality assurance program will be briefly discussed. Finally, the heart of the paper discusses accuracy measurements, which is one component of the quality assurance program.

AGRICULTURAL SURVEY PROGRAM

The ASP is a series of surveys for producing national, regional, and state estimates of the number of livestock (e.g., cattle, hogs, and sheep), acreage and production of crops, and amount of on-farm grain storage. On-farm grain storage is farmer storage of grains on the land he/she operates, as opposed to off-farm grain stocks which are stocks of grains in commercial facilities.

Estimates for on-farm grain stocks are made on a quarterly basis, in March, June, September, and December. Livestock estimates are quarterly for hogs and bi-annual for other livestock.

The ASP uses multiple frame procedures, comprising a list frame of farm operators and an area frame of generally one square-mile land segments. The area frame is used in the case of grain stocks to estimate for incompleteness in the list. The ASP is also a panel survey with replicated samples rotated in and out, maintaining at least a 60% overlap from quarter to quarter. Finally, data is collected through four modes of data collection (mail, face-to-face, Computer Assisted Telephone Interviewing (CATI), and non-CATI). The percentage breakdown by mode varies by state. For example, there are only 14 out of 44 NASS State Statistical Offices with CATI capabilities.

QUALITY ASSURANCE PROGRAM

Verma (1987) and Fecso (1988) have outlined four important components of a quality survey. They are relevance, timeliness, resources, and accuracy. Overall survey quality can only be achieved as a mixture of these components. One asks: Is the survey relevant? Is the survey timely? Is the survey accurate? Do we have the resources? One example that is used to illustrate the interplay of these components is that even though a survey can be extremely accurate it may not be important for the data users. Therefore, one has not produced a quality survey. It is within this framework that the quality assurance program for the ASP is being developed.

The remaining parts of this paper focus on the accuracy component of a quality survey. In surveys, it is common to assess accuracy by looking at mean squared error (MSE). MSE is equal to variance plus squared bias (Cochran 1977). The bias component of MSE is presented.

MEASUREMENT OF BIAS

Reinterview procedures were used to approximate bias in grain stocks survey indications. Reinterview programs are well established in other federal agencies, providing useful tools to measure survey response errors (U.S. Census Bureau, 1987; Kosary and Sommers, 1987).

Specifically, the reinterview procedure involves personal reinterviews of a subsample of original sample units (farm operations) without either the interviewer or respondent's knowledge of the original responses. This is called an independent reinterview. Immediately after the reinterview is completed a form containing the original responses is opened and any differences between the original and reinterview responses are reconciled. In the reconciliation process, the correct response is obtained for each difference along with the reason for the discrepancy, and the source of the error (e.g., initial interviewer, initial respondent, reinterviewee, or reinterviewer). Experienced enumerators, preferably supervisors, are used for reinterview. Training schools are held for enumerators, covering reinterview techniques and survey procedures.

Commonly, reinterviews are used to measure one component of survey error, called simple response variance (SRV) (O'Muircheataugh, 1986; Bailar, 1968). This is the variability in responses over repeated surveys under essentially the same survey conditions each time. No assumption is made on how "biased" responses are; just on consistency in item reporting.

The focus of this paper is the use of reinterview procedures to obtain approximate measures of bias. Bias, is defined as $B = (Y - \mu)$, where Y is the estimate from the survey, and μ is the true value.

"Truth" is best obtained through access to good check data or, in general, through reliable information obtained outside the survey itself (Arends et. al (1973); Belloc (1954)). Remeasure by an independent method considered more accurate is sometimes used (Hansen, Hurwitz, and Bershad, 1961). It is with this technique of remeasure by an independent method that the reinterview procedures outlined in this paper adhere to. That is, approximation to the truth can be obtained by final reconciled values using reinterview trained supervisory enumerators who specifically attempt to contact the farm operators.

This reinterview technique was applied in December 1987 ASP in the states of Minnesota, Ohio, and Indiana with approximately 1,000 reinterviews conducted within 10 days after the original responses were collected. Original interviews were collected with the CATI mode of data collection. The interest was to measure the accuracy of <u>CATI</u> grain stocks data. The area frame component and consequently the multiple-frame estimate of the bias is not discussed in this paper.

ESTIMATORS FOR BIAS

Let, Y_{ij} be the ij-th observation for the CATI response, where $i = 1, 2, \ldots, L$; $j = 1, 2, \ldots, n_i$; i = the stratum identification, and j is the element within stratum identification.

Let, V_{ij} be the ij-th observation for the reinterview response, where i and j are defined as above.

Then, let T_{ij} be the truth measure,

$$T_{ij} = o_{ij}\ p_{ij}\ Y_{ij} + q_{ij}\ r_{ij}\ V_{ij} + C_{ij} ,$$

where,

$$o_{ij} = \begin{cases} 1 & \text{if operate in the CATI interview} \\ 0 & \text{otherwise (out-of-business),} \end{cases}$$

$$q_{ij} = \begin{cases} 1 & \text{if operate in the Reinterview} \\ 0 & \text{otherwise (out-of-business),} \end{cases}$$

$$p_{ij} = \begin{cases} 1 & \text{if the CATI response was correct,} \\ 0.5 & \text{if said "both estimates - unsure which is more correct"} \\ 0 & \text{otherwise} \end{cases}$$

$$r_{ij} = \begin{cases} 1 & \text{if the reinterview response was correct, and not equal to the CATI response,} \\ 0.5 & \text{if said "both estimates - unsure which is more correct"} \\ 0 & \text{otherwise} \end{cases}$$

$$C_{ij} = \begin{cases} \text{compromise value if one was reported} \\ 0 \quad \text{otherwise} \end{cases}$$

Then, the bias is, $B_{ij} = Y_{ij} - T_{ij}$, and estimated by

$$\hat{B} = \sum_{i=1}^{L} \sum_{j=1}^{n_i} w_{ij}\ z_{ij}\ B_{ij} ,$$

where, w_{ij} is the weight for the ij-th observation (one divided by the sampling fraction), and z_{ij} is the list adjustment factor (adjustment for duplication in the list).

The estimated variance of $\hat{B}$ is,

$$\hat{V}(\hat{B}) = \sum_{i=1}^{L} (n_i - 1)^{-1}\ n_i\ (1 - f_i) \sum_{j=1}^{n_i} (d_{ij} - d_{i.})^2 ,$$

where

$$d_{ij} = w_{ij}\ z_{ij}\ B_{ij} ,$$

$$d_{i.} = n_i^{-1} \sum_{j=1}^{n_i} d_{ij} ,$$

f_i is the sampling rate for the i-th stratum, n_i is the number of responses in stratum i, and z_{ij} and w_{ij} defined as before.

RESULTS - TOTAL BIAS

Table 1 presents the estimates of bias for the CATI collected December 1 grain stocks data.

Table 1. Estimates of bias in the CATI collected data - December 1987 Reinterview study.

Survey Item/State	(CATI - Recon.) 000 bu.	% of CATI
Corn Stocks		
Minnesota	-60,987	-10.5 *
Indiana	-48,289	-16.1 *
Ohio	-22,609	-12.0 *
Soybean Stocks		
Minnesota	-13,082	-14.7 *
Indiana	-3,012	-5.9
Ohio	-5,393	-13.6 *

* - Indicates significance at the α=.05 level.

As this table indicates there were significant biases for most stocks items (α=.05). The bias was in **under-reporting of stocks** for the CATI mode of data collection. Specifically, there were 10.5% to 16.1% more corn stocks found in reconciliation in the three states. Similarly, there were 13.6% and 14.7% more soybean stocks uncovered in Ohio and Minnesota (the 5.9% difference in Indiana not differentiable from zero). Finally, there was a significant 24.9% increase in wheat stocks uncovered during reconciliation in Minnesota, while a bias could not be detected for the other two minor wheat stocks states.

These bias levels, while only representing the CATI portion of the multiple frame estimate, are consistent with the correction the Agricultural Statistics Board (ASB) is making in the survey indications. The ASB's function is to take the survey indications along with any other available data and come up with a final estimate that is published. In the case of grain stocks, the ASB has good check data such as exports and crushings in order to arrive at an independent estimate- "balance sheet estimate."

Figure 1 presents a flow chart on the ASB could incorporate these bias estimates or accuracy measures into a comprehensive quality assurance plan, instead of a "hit-or-miss" nonsampling error research study.

Looking at this figure we first see that NASS needs to measure total bias in year one - the first year of a national reinterview program that involves reinterviews across all modes of data collection (telephone, both CATI and non-CATI, and personal interviews). With reasons for biases documented, NASS may want to eliminate the largest contributors to the bias. For example, this might be specifically adding a new question that asks the operator for reserve stocks that are stored on the land a farmer operates. Continuing along in figure 1, we see that once a bias is uncovered in year one the reinterview study must be conducted in year two to assess the bias. If they do not exist, then the reinterview study could be discontinued for a period of time. If they do exist one needs to ask whether the biases are constant or variable. If constant, periodic measures of bias would be needed to validate this constant bias. If the bias is variable, national measures of accuracy are needed on a yearly basis.

If a bias is not present in the first year, one follows down a different path of figure 1, however, similar concepts exist. Figure 1 provides answers then to the question, "How long do we need to do reinterviews?" It also provides a coherent and sound policy for the measures of accuracy in the ASP.

BIAS ESTIMATES BY REASON FOR ERROR

The reinterview-reconciliation procedure collects reasons for every difference between the original and reinterview responses. This information is helpful in detecting where and why errors are occurring. Bias estimates were generated for reason for discrepancy, with the reasons classified into "definitonal," "estimating," and "other" errors. Estimation/rounding occurred when the respondent felt that one or more of the responses were estimated, or differences were too small to be bothered with, which included rounding problems. The second classification is termed definitional. Here, the respondent gave specific mention to reasons that are directly attributable to a lack of understanding of what should have been included and excluded in the stocks. These cover a large spectrum ranging from confusion with government storage to confusion with stocks sold but still on the operation. The final category was the "other" class. These responses were nether definitional nor estimation related. They included responses such as "doesn't give out information over the phone," "doesn't know why there is a difference," and "mistake in addition."

Table 2 gives estimates of bias by reason for discrepancy for the bias in the corn stocks reporting.

Table 2. Bias estimates for corn stocks by reason for discrepancy - December 1987 Reinterview Study.

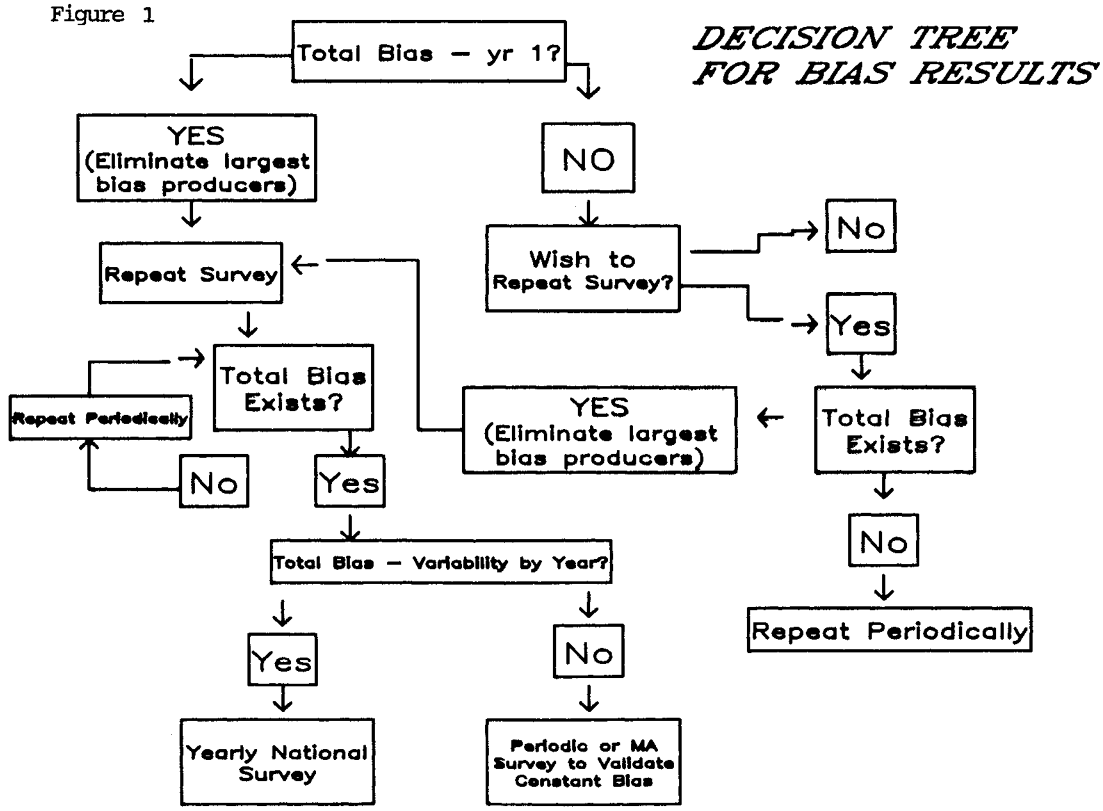
Reason for Discrepancy	Bias (000 bu.)	Percent of Total
Other	-57,006	43.3%
Estimating	-14,171	10.7%
Definitional	-60,709	46.0%
Total	-131,886	100.0%

The most controllable sources of bias creating agents comes from the definitional type errors. These we found contributed around 45% of the total bias. The ability to eliminate estimating/rounding errors will be minimal, and, in fact, these contribute less than 20% of the bias. The "other" errors, are also probably difficult to control. Some are due to the use of the telephone, such as "can't hear well on the phone," and "didn't take the time on the phone to add correctly." Some may be indirectly related to the use of the phone, such as "thought he/she reported this the first time," and "misunderstanding between enumerator and respondent. Finally, some "other" type errors will occur in any interview environment.

In order to control definitional errors that occur through CATI several alternatives are reasonable. First, a more detailed stocks section is needed. Questions should appear on the CATI screens that specifically ask the respondent if, for example, he/she has included stocks on all land operated. Another CATI screen would ask if they included grain stored on his/her operation belonging to someone else. This could be continued until the most important reasons for biases are addressed with the respondent during the interview. The sheer number and variety of definitional problems would make including every reason for difference almost impossible.

A second procedure to remove the definitional type biases is through improved telephone training and careful screening of potential telephone enumerators (they must be qualified). It is reasonable to suspect that some definitional problems are the result of enumerators lack of understanding.

Figure 1



CONCLUSIONS

Survey accuracy is just one component of survey quality. Considerations of relavance, timeliness, and resources are equally important. Survey accuracy can in part be dealt with by measures of mean squared error. The bias component is typically the most difficult to measure. The reinterview techniques developed by the Census Bureau provide a means to approximate the bias by using final reconciled responses as the "truth." While the results are encouraging for the ASP this technique cannot be blindly applied to all survey programs. In this application, approximate measures of bias are obtained through use of personal reinterviews and experienced supervisory enumerators.

Measurement of bias, and survey quality must be more than a set of "hit-or-miss" nonsampling error research projects. Figure 1 of this paper presents a procedure to incorporate measures of bias into a coherent framework to use in evaluating the accuracy of the ASP. Other measures of accuracy are important and are discussed in two NASS research proposals (Pafford et. al, 1988; Fecso et. al, 1988).

Finally, the Census Bureau's reinterview program is ideal for identifying why and where errors are occuring. Bias measures can be obtained for reasons for discrepancies. Then, those reasons that contribute most to the overall bias can be eliminated through the alteration of such things as survey procedures and questionnaires.

REFERENCES

Arends, William L., Robert L. Addison, Robert W. Young, and Raymond R. Bosecker (1973), "An Evaluation of Enumeration Techniques and Associated Response Errors and Biases," Staff Report, National Agricultural Statistics Service, U.S. Department of Agriculture, November.

Bailar, B. A. (1968) "Recent Research in Reinterview Procedures," Journal of the American Statistical Association, March, 41-63.

Belloc, N.B. (1954) "Validation of Morbidity Survey Data by Comparison with Hospital Records," Journal of the American Statistical Association, Vol. 49, pgs. 832-846.

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

Crop Reporting Board Standards (1985) A Report of the Crop Reporting Board Plicy and Procedures Working Group, July.

Fecso, Ronald (1988) "Survey Quality," Presented at the Quality Assurance in the Government Symposium, Dupont Plaza, Washington, D.C., May 24-25, 1988.

Fecso, Ronald, Toni Tremblay and Reed Johnston (1988), "Statistical Process Control of Objective Yield Surveys." Internal agency research proposal. Survey Research Branch, National Agricultural Statistics Service, U.S.D.A., Washington, D.C.

Galvin, John M. (1988) "Redesign of a Professional Work Process for Quality Improvement: A Team Project At the Bureau of Labor Statistics," Presented at the Quality Assurance in the Government Conference, May 24-25, 1988, Dupont Plaza, Washington, D.C.

Hansen, M.H., W.N. Hurwitz, and M. Bershad (1961) "Measurement Errors in Censuses and Surveys," Bulletin of the International Statistical Institute, Vol. 38, No. 2, pgs. 359-374.

Kosary, Carol L. and John P. Sommers, (1987) "The Effect of Computer Assisted Telephone Interviewing on Response Inconsistency." Bureau of Labor Statistics Report.

Linebarger, John S. "A Quality Improvement Team at the Bureau of the Census," Presented at the Quality Assurance in the Government Conference, May 24-25, 1988, Dupont Plaza, Washington, D.C.

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.

Pafford, Bradley, Gretchen McClung, and Victor Tolomeo (1987), "Process Control of the Quarterly Agricultural Survey (QAS)." Internal agency research proposal. Survey Research Branch, National Agricultural Statistics Service, U.S.D.A., Washington, D.C.

U.S. Census Bureau, (1987) Reinterview Manual, Available from the author.

Verma, Vijay (1988) "Some Less-Well Remembered Aspects of Quality of Statistical Data," Presented at the National Symposium on Sample Surveys in Indian Agriculture - Problems and Prospects, India.