HISTORY OF REINTERVIEW STUDIES AT NASS

George A. Hanuschak
Dale Atkinson
William C. Iwig
and
Victor Tolomeo

NATIONAL AGRICULTURAL STATISTICS SERVICE

U.S. DEPARTMENT OF AGRICULTURE

To be presented at the 1991
Joint Statistical Meetings
Atlanta, Georgia
August 22, 1991

HISTORY OF REINTERVIEW STUDIES AT NASS

George A. Hanuschak, Dale Atkinson, William C. Iwig, and Victor Tolomeo, USDA/NASS
George A. Hanuschak, USDA/NASS, 14th & Independence Ave, SW, Washington DC 20250-2000

KEYWORDS: Reinterview, response bias, reconciliation

Introduction: Over the years the National Agricultural Statistics Service (NASS) has conducted a series of reinterview surveys to validate inventory estimates from its agricultural surveys. Annual reinterview surveys targeting response bias for crop acreage, grain stocks, and hog inventories have been in place since December 1987. The Agency has also conducted several special-purpose reinterview surveys to address specific survey concerns. An important product of the reinterview surveys has been the identification of reasons for reporting errors. These include definitional problems, misinterpretation of questions and survey concepts, and simple reporting errors. Such cognitive information obtained from reinterviews has been valuable in survey instrument development, training, and the interpretation of survey results.

Reinterviews, in general, focus on quantifying response variance and bias. These components contribute to total survey error through the mean squared error (MSE) as follows:

MSE = sampling var. + response var. + covariance(s,r) + bias^2
where covariance(s,r) = covariance between the sampling and response deviations.

Response variance refers to the trial to trial response variability for a specific characteristic from a particular reporting unit. This variability may be inherent in the survey process or may be due to specific factors such as the enumerator or respondent. Response bias reflects the accuracy of the original survey response. To measure response bias, a "proxy to truth" must be obtained. The reinterview provides this benchmark through extensive probing or a reconciliation of differences between original and reinterview responses with a personal interview using an experienced interviewer.

Both response variance and bias contribute to nonsampling errors. Of the two contributors, response bias has generally been considered the more serious (and fixable) in NASS surveys. The objective of this paper is to review the current and historical use of reinterview surveys at NASS. Results of the reinterview program initiated in 1987 will be discussed and recommendations for future reinterview programs will be presented.

Background: The NASS probability survey program was in its early stages of development in the mid 1970's. Reinterviews were recognized even then as a valuable tool for determining the extent to which nonsampling errors contribute to total survey error. Both the reasons for reporting errors and the level of bias were of concern, and primary emphasis was placed on correcting the survey process as needed.

High quality survey results are highly dependent on correct association between sampling and reporting units. The best data are worthless if collected for and/or associated with the wrong unit. Communicating the NASS reporting unit concept (i.e., all agricultural activity on the total land operated by the selected unit) for probability agricultural surveys is an important step in the survey process. This involves verifying the name of the operation, identifying total land operated, and ensuring that all agriculture, and only agriculture, on those acres operated is reported. Only after establishing total land operated and ensuring the accuracy of the reporting unit, does the quality of reported survey item data come into play.

A review of early Agency studies indicates that various reinterview studies were conducted to identify problems with data reporting and reporting unit association. Some of the early results are presented in the ensuing paragraphs.

Bosecker and Kelly (1975) conducted a reinterview study in Nebraska to investigate how well reporting unit concepts were understood by the respondent. The original survey data had been collected by personal interview. Probing questions were asked on the personal reinterview to help the respondent more accurately answer the questions and to verify the accuracy of the reporting unit. The study found that approximately 30 percent of the respondents incorrectly reported total acres operated and 20-30 percent incorrectly reported specific livestock inventories. Respondents often reported livestock that they own, regardless of their connection to the total acres operated. One of the recommendations was that questionnaires should be more explicit about "the purpose of the land questions relative to livestock to be reported."

Hill and Rockwell (1977) conducted a reinterview study in Ohio and Wisconsin to determine how well the survey questionnaire obtained hog data for the correct reporting unit. A split sample approach was used in the study in which approximately half of the sample used the operational survey questionnaire. The other half used a test version with different wording that emphasized correct reporting unit

association. Both versions were mailed to the sample with a personal or telephone follow-up of nonrespondents. A personal reinterview and reconciliation was conducted within two weeks to arrive at the "true" values. Results for total hogs indicated that the operational estimate was biased upward 2-5 percent while the test version was biased upward 1-2 percent. While improper handling of partnership operations and reporting unit identification were major problems with both versions, the operation description section of the test version was recommended for operational use because of its increased clarity.

Hill and Farrar (1977) conducted a reinterview study "to identify nonsampling errors in entire farm acres and livestock and to report their effect upon the weights and weighted livestock expansions." The original survey data were collected by personal interview. Reinterview data were reconciled to establish the "true" values. Total farm acres were significantly biased downward about 6 percent on the original questionnaire, while reported cattle and hogs showed no significant bias. Bias in the total farm acres caused weighted estimates for cattle and hogs to be significantly biased upward. The major recommendation was "entire farm acres must be more accurately obtained from the operator if the current method of weighting is to be continued."

Nealon (1982) reported on a reinterview study in Minnesota, North Carolina, and Ohio "to determine if problems still exist with the reporting of farm acres operated." The parent survey data had again been collected by personal interview. The major objective was to measure response variance rather than response bias. Results indicated that about one-third of the respondents reported the same value, one-third differed by less than 10 percent, and one-third differed by more than 10 percent. The report stressed the need to improve the quality of reported total acres operated.

Reinterview Studies (December 1987-December 1990): Each year NASS conducts a series of multiple frame based agricultural surveys, collectively called the Agricultural Survey Program (ASP), to estimate specific agricultural commodities at the state and national levels. Reinterview studies designed to measure response bias in Computer Assisted Telephone Interviewing (CATI) collected ASP data were conducted in Indiana, Ohio, and Minnesota in December 1987, in Iowa, Nebraska, and Pennsylvania in March 1988, and in Indiana, Iowa, Minnesota, Nebraska, Ohio, and Pennsylvania in December 1988-1990.

The reinterview techniques used by NASS are similar to those used by the U.S. Census Bureau (Forsman, Schreiner, 1990); however, the NASS focus is on response bias rather than response variance or evaluating fieldwork. For the reinterview surveys, NASS used supervisory or experienced field interviewers for face-to-face reinterviewing of selected items from a subsample of ASP respondents. All reinterviews were conducted within 10 days of the ASP CATI interview. Any differences between original ASP and reinterview responses were reconciled. Considerable effort and resources were expended in procedural development, training and supervision to ensure that the reconciled values represent the best possible proxies to the "truth."

The reinterview samples targeted CATI respondents because CATI accounts for a large and increasing percentage of the ASP data collected, provides considerable control of the reinterview process, and affords flexibility in the computer generation of reconciliation forms. Parent survey (ASP) CATI interviews were completed from the state offices of the six states. A separate corps of supervisory and/or experienced field interviewers was used to conduct the follow-up face-to-face reinterviews. The objective of the reinterview was to obtain the best possible information about the subsampled operation by contacting the most knowledgeable person. It was preferred but not mandatory to contact the same respondent originally interviewed for the ASP.

Reinterview assignments, containing a reinterview questionnaire and reconciliation form, were mailed to the field interviewers the day after the CATI interviews were completed. After obtaining the reinterview response, the interviewer opened the sealed envelope containing the original responses and compared the reinterview responses to those from the ASP. When the ASP and reinterview response differed, the interviewer reconciled the difference. Interviewers were specifically instructed not to review the original answers until after the reinterview was completed, to maintain independence between the ASP and reinterview responses.

Interviewers were further instructed to complete the reinterview and reconciliation within 10 days of the original CATI interview to minimize recall problems. The average time between the original CATI interview and the reinterview ranged from 6.4 days in March 1988 to 5.9 days in December 1989.

Questionnaire wording in the reinterview was similar to that used in the ASP interview, but not all ASP questions were reasked on the reinterview. Estimates of response biases and their variances were based on a stratified sample design. For the i^{th} observation in stratum h, bias was measured as follows, with a negative bias indicating underreporting of the survey item:

$B_{hi} = O_{hi} - F_{hi}$ for stratum $h = 1,....,L$ and unit $i = 1,....,n_h$ where O_{hi} = original ASP CATI response and F_{hi} = final or reconciled value.

The Sample: The ASP which parented the reinterview studies is comprised of probability surveys based on a multiple frame survey design utilizing independent list and area frames. In the multiple frame context, the area frame is used to measure the incompleteness of the list and accounts for 10-20 percent of the survey estimates for major items. The types and mix of data collection methods used vary by state and to a lesser degree from survey to survey. Modes of data collection utilized are mail, CATI, non-CATI telephone interviewing, and face-to-face interviewing.

The reinterview subsamples were drawn from the portion of each state's ASP list sample completed on CATI. Samples eligible for reinterview were completed interviews, out-of-business operations, and interviewed operations that reported for some but not all items. Area samples, list interview refusals, and list operations not interviewed on CATI for the ASP were not eligible for reinterview. For the six reinterview states, approximately 45-50 percent of ASP list samples was completed on CATI.

Table 1 presents the reinterview sample sizes, subsampling rates and response rates for the six states in the December 1988-1990 reinterview surveys. For the sake of comparability, the results from the three state "pilot" reinterview study in December 1987 are not included.

Table 2 presents the resulting bias estimates (as a percentage of the CATI total) for selected grain stock items. Univariate test (H: Bias = 0) results are shown, with statistically significant biases for $\alpha = .05$ indicated by an asterisk. Levels of significance close to .05 are indicated in parentheses.

Negative biases were estimated for corn and soybean stocks, indicating that respondents contacted by CATI tended to underreport stock items. Notice, however, how the estimated bias decreased over the three year period. There appears to have been an interviewer conditioning effect in the later surveys reflecting knowledge of the reasons for differences from the early surveys which were emphasized in subsequent training sessions.

The precision of the percent bias estimates are shown in Table 2. The large standard errors (S.E.) indicate that, although biases do appear to exist, in general the precision of the estimates are very low. The low precision is due primarily to the distribution of individual reporting differences (i.e., differences between the initial CATI and reconciled responses) themselves. These distributions are highly kurtosed with a large spike at zero and pose a challenging estimation problem.

June 1990 Reinterview Survey: In December 1986 the NASS probability survey program was integrated and expanded to provide quarterly crop acreage indications. Previously, planted acreage estimates were based primarily on a probability area frame survey conducted each June. Nonprobability survey indications were provided for planting intentions in the spring, planted acreage in June and harvested acreage late in the year. Multiple Frame (MF) probability surveys replaced the nonprobability surveys in 1986, but from their inception they have produced upwardly biased crop acreage indications. The June 1990 Reinterview Survey was conducted in Indiana and Ohio to investigate three factors potentially contributing to this bias.

The first factor was correct identification of the reporting unit. The 1990 MF questionnaire had a single question asking the respondent how many total acres of land were in the operation. Specific types of land to be included such as land rented from others, or excluded such as land rented to others, were listed on the questionnaire but were not always read to the respondent. Previous studies (Bosecker and Kelly, 1975; Ford, 1975; Nealon, 1980a; Nealon, 1980b) had indicated that including the additional operation description questions did not affect the number of livestock reported. However, the effect on crop acreage was not addressed.

The reinterview questionnaire specifically asked for acres owned, acres rented from others, acres rented to others, and total acres operated. Reinterview acreage expansions for the two states were below parent survey expansions by about 14 percent for total land, 7 percent for cropland, 9 percent for corn, and 4 percent for soybeans. While some of this difference could be due to a personal reinterview versus a telephone interview, the reinterview data appeared to have less bias than the original MF data. The multi-question acres operated approach was adopted for the operational MF survey in December 1990.

The second factor investigated was whether or not farm operators accurately report their total crop acreage when reporting it on a total farm basis. The reinterview questionnaire asked the respondent to report parcel and field acreages for comparison to the farm level values. Results indicated that corn and soybean acreages reported at the farm level were approximately 3 percent larger than acreages reported at the parcel level. The average size of the difference increased as the number of parcels in the operation increased. There was very little difference between parcel and field level data. The results indicate that farm operators know their separate parcel acreage well, but overestimate when asked to provide a total farm acreage.

Nonresponse imputation was the third potential source of bias investigated. NASS has used an imputation procedure utilizing auxiliary control data and previous survey data (Atkinson, 1988) for acreage and grain stocks since June 1987. Reinterview data from nonrespondents to the June 1990 MF survey were collected to examine the assumptions of the imputation procedures and to

determine whether the procedures were introducing a bias. This is the first time in recent years that NASS has attempted to reinterview nonrespondents. The response rate was approximately 55 percent. Initial results indicate that the total cropland estimate is significantly biased upward (α = .05) by 3.5% due to imputation (Wesley, 1991).

Costs and Benefits of the Reinterview Survey Program: The Agency's reinterview survey program, as in place since 1987, has provided the following real benefits to the ASP:

1. Statistically representative measures of survey response bias have been provided in time for official Agricultural Statistics Board use, an extremely tight time frame of 30-40 days from the beginning of the data collection to Board publication.

2. Specific reasons for the biases have been provided. This information is valuable in directing future questionnaire designs and statistician and enumerator survey training sessions, to reduce bias levels at the state and regional/national levels. We have already observed significant reductions of bias in on-farm grain stocks in the reinterview states.

3. Another benefit is enhanced power of statistical testing procedures. Paired observations on identical reporting units provide an efficient design for detecting differences. When combined with hypothesis testing of two alternative questionnaire designs on the original interview, two designs can be compared not only to each other but also to a "proxy" best method. Thus, the combination of conventional split sample testing with a reinterview and reconciliation program can lead to more powerful and meaningful testing.

4. The reinterview program has provided the Agency with an independent methodology to evaluate its Agricultural Statistics Board's balance sheet approach. Balance sheet methods are used to compare current survey results with sources of administrative or check data over time. For example, current survey results for on-farm soybean grain storage are matched against soybeans utilized throughout the marketing year. Over several years, the balance sheet approach has indicated a downward reporting bias for on-farm soybean storage. This indicated bias, while variable from survey to survey and quarter to quarter, has usually represented between 10 and 15 percent of the estimate. The reinterview program provided an independent evaluation of the on-farm soybean storage indication from the current operational survey program. The reinterview results for December 1987, March 1988, December 1988 showed an underreporting bias of approximately 10-17 percent. Thus, an independent verification of an operational method used for many years was made possible.

In December 1989, the reporting bias in the six state study for soybeans became statistically insignificant. This may have been a result of conditioning the telephone enumerators with training that emphasized previous results. A corresponding, though somewhat smaller, reduction was reflected in the difference between the survey and balance sheet based board estimates for the reinterview states.

As with any program, there are costs incurred. These include:

1. The additional survey costs of personal interviews to get the proxies to "truth." Total out of pocket costs (training schools and enumeration costs) for the six-state reinterview program were $50,000 per survey. This cost included training schools for statisticians and interviewers.

2. An indirect cost is the additional respondent burden associated with recontacting farm operators with a personal interview within ten days of the original CATI interview. The refusal rate ranged from four to six percent on the reinterview surveys.

The authors contend that the reinterview program, per dollar spent, has yielded substantial statistical benefits to the Agency's ASP.

Recommendations for the Future: The Agency is at a crossroads concerning its reinterview survey program. There are several options for future courses of action for both the NASS operational survey and survey research programs. Based upon experiences to date, the authors list their prioritized recommendations for the operational and research programs. The list is ambitious, and implementation will take several years and be subject to changing budget conditions and overall Agency priorities.

Operational Program Options

1. Continue to use existing reasons for bias (with current questionnaire and CATI instrument design) to aid in the training of telephone and field enumerators. This was started prior to the December 1989 survey and bias levels have decreased in the reinterview states.

2. Develop an alternative questionnaire design and CATI instrument to be tested for on-farm grain storage. The reinterview studies demonstrated that the major reasons for bias were related to definitional issues. On-farm grain storage is a complex concept requiring a sizable list of inclusions and exclusions.

3. Expand, over the next several years, the reinterview program (including national and state training schools) to the national CATI domain. Reasons for this recommendation are the following:

a. The CATI domain itself is in a large expansion mode, with the installation of microcomputer local area networks (LAN's) in all State Statistical offices. This means that any bias measures can be applied to the largest domain of the ASP in the future.

b. The "mechanics and logistics" for conducting the reinterview survey program have been worked out for the CATI domain. To expand to the non-CATI domains (mail, non-CATI telephone and personal interviewing), will require new "mechanics and logistics". All other modes will be a greater challenge than CATI to meet the "within 10 days" reinterview requirement. It is anticipated that in several years the non-CATI domains in the ASP will be considerably smaller. Each of the subdomains will be very small; therefore, measuring a bias level with precision could be an even bigger problem than it is in the CATI domain.

4. Conduct research on developing procedures for the non-CATI domain. The "mechanics and logistics" of conducting a reinterview program when the initial interview is non-CATI need to be developed. A major issue is timing and making the "10 day rule" a reality for non-CATI interviews. A CATI interview is immediately in electronic form and has a partial edit built in, so that computer generated reconciliation forms with the original survey data can be printed the next day and air expressed or delivered to the assigned field enumerator. For the non-CATI domains, more time is required to get edited data into electronic form to computer generate reconciliation forms, making it a challenge to meet the "10 day rule."

5. Expand the reinterview program to the other major survey programs of the Agency. The authors think that it is better to complete recommendations 1-4 before taking on new surveys. Also, research on new surveys would need to be conducted first.

Research Program Options

1. Conduct research on alternative variance reduction methods for the bias estimates. This research has been started under a cooperative research project between the Agency and the Statistics Department at New Mexico State University. One method being evaluated is the use of a regression estimator. This estimator uses the entire original sample for the initial survey response and the reinterview subsample for the combined initial and reconciled responses. The result is a double sampling regression estimator that uses ASP data already available from the cheaper mode of data collection (CATI) in conjunction with the more costly reinterview data.

2. Conduct a test comparing the current questionnaire design and CATI instrument for on-farm grain storage with a newly developed questionnaire design and CATI instrument. Recall that the second priority recommendation for the operational program was to review existing reasons for the on-farm storage reporting bias and develop a new questionnaire and CATI instrument.

3. Aid in the expansion of the reinterview program to the national CATI domain. The following are unresolved issues associated with the proposed expansion:

a. As mentioned previously, new states are now being added to the CATI domain. As new microcomputer LAN's are installed, new states are being added to the program. These states need operational CATI survey experience before they become part of the reinterview program.

b. The sample size in any given state would probably be less than 300 for national CATI domain bias estimates. An alternative sampling procedure, such as some form of geographic clustering, should be considered for cost efficiency in enumerator training and data collection.

c. If on-farm grain storage is not the primary focus of future reinterview studies, procedures for other survey items and survey periods need to be developed. Procedures for hogs are already in place (McClung, Pafford, Tolomeo, 1990).

4. When doing split sample tests to compare two questionnaire designs, use a reinterview and reconciliation process to enhance the split sample test. For example, if one wants to compare two different telephone interview designs or CATI instruments using a split sample test, then a personal reinterview and reconciliation can provide benchmark values. Thus, in addition to comparing methods 1 and 2 to each other, one can also compare them to the "truth." This enhancement of split sample testing can become a routine improvement in the Agency's survey research program.

5. Conduct research on reinterview and reconciliation methods for the non-CATI domains of the ASP. When the Agency is ready to address the non-CATI domain, research into the potentially different "mechanics and logistics" will be necessary. For example, if the average time between the original interview and the reinterview and reconciliation is 6 days now but would become 9-12 days for the non-CATI domain, what effect would the 3-6 day time difference have on method comparisons?

6. Conduct research on using the reinterview survey procedures on other major Agency survey programs. When the Agency is ready to move on to other major surveys, research the mechanics, logistics, methodology and best questions for the proxy to true values for that survey program.

Table 1. Reinterview sample sizes and response rates, six states combined.

Year	Sample Size	Subsampling Rate (%)	Response Rates - Complete (%)	Refusal (%)	Inaccessible (%)
1988	1057	11.7	90.2	4.0	5.8
1989	1075	12.1	88.1	4.3	7.6
1990	1337	16.1	87.3	5.8	7.0

Table 2. Percentage bias estimates for grain stocks and storage capacity for the six states combined.

Year	Corn Stocks (% Bias)	S.E. of % Bias	Soybean Stocks (% Bias)	S.E. of % Bias	Storage Capacity (% Bias)	S.E. of % Bias
1988	-13.2 *	3.1	-16.8 *	7.1	2.15	.4
1989	-5.1 (.08)	3.0	4.3	3.5	6.0	4.7
1990	4.7	15.1	7.9	6.7	9.4	9.5

* indicates significance at $\alpha = .05$

ACKNOWLEDGEMENTS

The authors would like to acknowledge the extensive reinterview survey work by Brad Pafford and Gretchen McClung as reflected in their referenced reports. In addition, there was strong Agency managerial and technical support for the reinterview program by numerous individuals and units. Deserving special recognition for their work are the staffs of the six State Statistical Offices involved in the reinterview projects.

REFERENCES

Atkinson, Dale (1988). "The Scope and Effect of Imputation in Quarterly Agricultural Surveys," NASS, USDA.

Bosecker, Raymond R. and Kelly, William F. (1975). "Summary of Results From Nebraska Survey Concept Study," SRS, USDA.

Fecso, Ron and Pafford, Brad (1988). "Response Errors in Establishment Surveys with an Example from an Agribusiness Survey," Proceedings of the Section on Survey Research Methods, ASA, 1988.

Ford, Barry (1975). "The Effect of Land Questions on the Multiple Frame Hog Survey.", SRS, USDA.

Forsman, G., Schreiner, I (1990). "The Design and Analysis of Reinterview - An Overview," paper presented at the 1990 International Conference on Measurement Errors in Surveys.

Hill, George W. and Farrar, Martha S. (1977). "Impact of Nonsampling Errors on Weighted Tract Survey Indications," SRS, USDA.

Hill, George W. and Rockwell, Dwight A. (1977). "Associating a Reporting Unit with a List Frame Sampling Unit in Multiple Frame Sampling - Ohio and Wisconsin," SRS, USDA.

McClung, Pafford, Tolomeo (1990). "Estimating Response Bias Using Reinterviews," paper presented at the 1990 International Conference on Measurement Errors in Surveys.

Nealon, Jack (1980a). "The Effect of Omitting Acreage Questions and Modifying the Operation Description Section in Hog Surveys," SRS, USDA.

Nealon, Jack (1980b). "The Effects of Omitting Acreage Questions and Modifying the Operation Description Section in Cattle Surveys," SRS, USDA.

Nealon, Jack (1982). "Response Errors in the Weighted Estimator," SRS, USDA.

Warren, Fred B. (1991). "June 1990 Reinterview Survey: Part I, Effect of Alternative Acreage Operated Questions on Reported Acreages and Number of Hogs on Farms," NASS, USDA.

Warren, Fred B. (1991). "June 1990 Reinterview Survey: Part II, Effect of Parcel and Field Level Acreage Questions," NASS, USDA.

Wesley, Fatu (1991). "Evaluation of an Item Imputation Procedure for Agricultural Surveys," Proceedings of the Section on Survey Research Methods, ASA, 1991.