THE MEASUREMENT OF RESPONSE BIAS IN MARCH 1988 QUARTERLY ON-FARM GRAIN STOCKS DATA

Gretchen McClung
Vic Tolomeo
Brad Pafford

THE MEASUREMENT OF RESPONSE BIAS IN MARCH 1988 QUARTERLY ON-FARM GRAIN STOCKS DATA. by Gretchen McClung, Vic Tolomeo, and Brad Pafford. Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250. July, 1990. NASS Research Report No. SRB 90-09.

ABSTRACT

A reinterview survey was conducted in March 1988 in Iowa, Nebraska, and Pennsylvania. A subsample of respondents interviewed by CATI (Computer Assisted Telephone Interviewing) was reinterviewed in order to measure the quality of the Agricultural Survey data. Underreporting of on-farm corn and soybean stocks was found to be significant. Larger biases were found when someone other than the operator reported. Definitional reasons for differences in the two interview responses accounted for most of the bias (61%). At the three-state level, there were no significant differences for on-farm grain storage capacity, on-farm wheat stocks, land in farms, or total cropland acres. Except for wheat stocks, results were consistent with those of the first reinterview study conducted in December 1987 in three different states.

```
*****************************************************************
*   This paper was prepared for limited distribution to    *
*   the research community outside the U.S. Department      *
*   of Agriculture.                                         *
*****************************************************************
```

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

SUMMARY

The reinterview survey was conducted in March 1988 in Iowa, Nebraska, and Pennsylvania. This was the second in a series of reinterview studies conducted to provide a measure of quality of the Agricultural Survey data by estimating response bias. A subsample of Computer Assisted Telephone Interviewing (CATI) respondents was selected for reinterviewing. Enumerators used face-to-face interviews to reask a portion of the questions asked over the telephone. Wording and sequence of the reinterview questions were the same as for the original interview. After the reinterview questionnaire was completed, the original and reinterview answers were compared, and any differences were reconciled. The final reconciled value was considered to be a proxy for the true value for that item, and it was used to measure bias in the CATI reported values.

Bias estimates were generated for on-farm corn stocks, soybean stocks, all wheat stocks, grain storage capacity, land in farms, and cropland acres. For corn stocks, significant differences were found for Iowa, Nebraska, and the 3 states combined. Soybean stocks differences were significant in Iowa and Nebraska, and at the 3-state level. The only other significant difference was for Nebraska in grain storage capacity. All of the significant differences were negative, which means that these items were underreported on the original CATI interviews.

Respondent categories were established based on the original respondent and the reinterview respondent. The operator was the respondent for both interviews almost 81% of the time. Reasons for the differences were grouped into 3 categories - estimating reasons, definitional reasons, and other reasons. A significant relationship was found between respondent category and relative bias, with larger biases associated with respondents other than the operator. There was also a significant relationship between the reason for the difference and the magnitude of the bias, with estimating reasons resulting in smaller biases than definitional and other reasons.

Estimates were generated for cross classification domains of respondent category by reason for difference. Definitional reasons accounted for most of the bias in corn and soybean stocks. Some specific definitional reasons for the corn stocks differences and their actual contribution to the bias were presented.

Bias estimates by size of operation were also generated. Using both land in farms and cropland acres as the size classification variable, the average unexpanded corn stocks bias increased as the size of the operation increased.

Most of the results for this March study were comparable to those for the December 1987 reinterview study even though the survey period and the states included were different. Future reports on the December 1988 and 1989 reinterview studies will be able to make more direct comparisons because the same 6 states were included in the later surveys.

INTRODUCTION

The National Agricultural Statistics Service (NASS) has conducted a series of reinterview surveys to measure response bias in Agricultural Surveys (AS) data collected by Computer Assisted Telephone Interviewing (CATI). These studies were designed to provide a quantitative measure of quality for the Ag Surveys. The first reinterview survey was conducted in December 1987 in Indiana, Minnesota, and Ohio. The second study, which included Iowa, Nebraska, and Pennsylvania, was conducted in March 1988.

Reinterviews were completed on a subsample of CATI respondents to the March AS. CATI was chosen for the study because of the large percentage of AS data accounted for in this way and because the data collected is easily accessible for use in computer generation of reconciliation forms containing the original responses. Experienced and/or supervisory enumerators were used to conduct face-to-face reinterviews. After the reinterview questionnaire was completed, the enumerator opened an envelope containing the original responses. These were compared with the reinterview responses. When a difference occurred between the original and reinterview responses, a specific reconciliation procedure was used to resolve the difference and determine, if possible, a reason for the difference. Enumerators were instructed not to review the original responses until after the reinterview questionnaire was completed.

Results for on-farm grain stocks, grain storage capacity, land in farms, and cropland acres for the March 1988 reinterview study are presented in this paper. The first report on reinterviewing, 'USE OF REINTERVIEW TECHNIQUES FOR QUALITY ASSURANCE: The Measurement of Response Bias in the Collection of December 1987 Quarterly Grain Stocks Data Using CATI', by Pafford (1989), provides background material on the motivation and reasons for this reinterview research. The authors of this report assume that the reader is familiar with this first research paper.

METHODS

Survey Procedures

The survey procedures followed those used for the December 1987 reinterview study. Copies of the March 1988 reinterview questionnaire and reconciliation form for Pennsylvania can be found in Appendix A. For details on the manuals, training, forms, enumerator procedures, editing, and data entry, see Pafford (1989).

The Sample

The reinterview sample was a subsample of the CATI list sample in the three states. Only completed CATI samples, including those coded out-

1

of-business, were subsampled. For the March 1988 reinterview survey, a total of 1160 samples were assigned for reinterviewing. Reinterview strata sample sizes were allocated in the same proportion as the CATI operational survey. Table 1 shows the sample sizes and response rates by state. The completion rate ranged from 74 to 89% at the state level, with an overall completion rate of 82%. The refusal rate was low, ranging from 3 to 5%. An operation was coded inaccessible if a reinterview was not completed within 10 days of the original CATI interview. Inaccessibles occurred for many reasons, including bad weather and length of time for delivery of the reconciliation forms. Some enumerators commented that, on occasion, forms mailed from Headquarters using one or two day delivery service were not received until several days later. The inaccessible rate was 14% for the 3 states combined, with a range of 8 to 21%. For the December 1987 reinterview survey (Pafford 1989), the completion rate was higher (86% overall) and the inaccessible rate was lower (10%), while the refusal rate was the same.

Table 1. Sample sizes and response rates for the 3 states in the March 1988 reinterview study.

	State			
Response	Iowa	Nebraska	Pennsylvania	Total
Completed	85%	74%	89%	82%
Refusal	4%	5%	3%	4%
Inaccessible	11%	21%	8%	14%
Total	100%	100%	100%	100%
	(n=357)	(n=483)	(n=320)	(n=1160)

Statistical Measures

For purposes of this study, the final reconciled value was assumed to be the proxy to the true value for use in estimating the bias. The assumption is that the process of reconciliation and the use of supervisory and experienced enumerators to conduct the reinterviews provides the best answers possible. In general, a final value was calculated in cases where the reinterview response differed from the original CATI response and the difference was reconciled. If the original interview and the reinterview generated different responses and the respondent did not know which answer was correct, the final "true" value was missing and the observation was not used for that particular item. If, however, the respondent said that either answer could be right because both answers were estimates, the average of the two values was considered to be the final value. If the same response was given on the original interview and reinterview, this response was considered the final or "true" value. A limitation of this procedure occurs when the same incorrect answer is given both times, in that there is no way to measure a bias.

2

Two measures of response bias were generated. One was a measure of the bias in the original CATI responses. This was calculated using the differences between the original CATI responses and their corresponding reconciled responses for each item of interest. The other measure of bias was estimated using the SPS edited CATI data, which was the CATI data that had been edited using the operational Survey Processing System (SPS). The difference in the SPS edited CATI data and reconciled responses is a measure of the effect of the reconciled data if it had been used in the operational estimating program. A comparison of the two measures of bias also provides an indication of the effect of the operational editing procedures.

Frequency distributions of the respondent combinations (original-reinterview respondent) and the reasons for the differences are presented. Crosstabulations of the relative magnitude of the bias by respondent categories and reason categories were also calculated. Bias estimates were also calculated based on respondent and reason categories. Finally, bias estimates were computed based on the size of the operation.

Formulas used to generate the bias and variance estimates were based on a stratified sample design (see Pafford, 1989). For the ith observation in stratum h, bias was measured as

$$B_{hi} = X_{hi} - T_{hi} \qquad \text{stratum } h = 1,\ldots,L \text{ and unit } i = 1,\ldots,n_h$$

where X_{hi} = original response
T_{hi} = final or reconciled value

RESULTS

Response Bias

Bias estimates for corn stocks, soybean stocks, wheat stocks, and grain storage capacity are shown in Table 2. These estimates, for both the original CATI and SPS edited CATI data, are given at the state level and the 3-state combined level. The biases as a percent of the original CATI and SPS edited CATI total estimates are also shown, and significant differences determined from univariate t-tests are indicated.

For corn stocks, Iowa, Pennsylvania, and the 3-state differences were significant for both original CATI and SPS edited CATI data. For soybean stocks, Iowa, Nebraska, and the 3-state differences were significant. There were no significant differences for wheat stocks at any level. The only significant difference for grain storage capacity was for the original CATI data in Nebraska.

Table 2. Bias estimates for stocks and grain storage capacity for the March 1988 reinterview study.

Survey Item / State	Original CATI - reconciled		SPS edited CATI - reconciled	
	1000 bu.	% of orig. CATI	1000 bu.	% of SPS edited CATI
Corn Stocks				
Iowa	-80,119	-14.2 *	-79,296	-14.3 *
Nebraska	-13,232	-2.8	-10,768	-2.2
Pennsylvania	-4,769	-10.9 *	-4,960	-11.5 *
Total	-98,120	-9.1 *	-95,024	-8.8 *
Soybean Stocks				
Iowa	-9,271	-11.3 *	-9,192	-11.4 *
Nebraska	-1,278	-7.8 *	-1,278	-7.8 *
Pennsylvania	+27	+2.0	+27	+2.1
Total	-10,522	-10.6 *	-10,443	-10.6 *
Wheat Stocks				
Iowa	-137	-34.5	-137	-34.5
Nebraska	-2,510	-7.1	-2,507	-7.1
Pennsylvania	-62	-14.3	-209	-73.5
Total	-2,709	-7.4	-2,853	-7.9
Grain Storage Capacity				
Iowa	-40,914	-4.0	-28,759	-2.7
Nebraska	-37,929	-3.8 *	-23,570	-2.4
Pennsylvania	+23,078	+14.3	+22,318	+14.1
Total	-55,765	-2.6	-30,011	-1.4

* indicates significance at α= .05 (unadjusted univariate tests)

Original CATI and SPS edited CATI data produced similar results except in two cases. For capacity in Nebraska, the original CATI bias was significant but the SPS edited CATI bias was not. This was mostly due to two operations where capacity was zero in both the original interview and the reinterview. However, capacity was entered during the operational editing process and this resulted in a smaller overall

state level bias. The other case was for wheat stocks in Pennsylvania, where the bias levels were quite different, although neither was significant. This was due to one operation where the original CATI reported wheat stocks were edited out during the operational edit.

In general, the direction of the estimated biases was downward. All of the significant biases were negative, with the corn and soybeans stocks biases (at the 3-state level) at about 10%. A negative or downward bias means that the operational CATI survey data underestimated the true level of the item. These results agree with the findings of Pafford (1989) in the December 1987 reinterview study and support the Agricultural Statistics Board in its consistent upward adjustment of the multiple frame stocks indications.

Table 3 shows the bias estimates for land in farms and cropland acres. The differences ranged from -2.8% to +5.7%, but none of these mostly small differences were significant (the smallest p-value was .19, with the remainder ranging from .37 to .99). The large difference in cropland acres between the original CATI and SPS edited CATI data was due to Iowa, where cropland acres were entered into the SPS edited CATI data for two operations where the original interview and reinterview both obtained responses of no cropland.

Table 3. Bias estimates for land in farms and cropland acres
          for the March 1988 reinterview study.

Survey Item / State	Original CATI - reconciled		SPS edited CATI - reconciled	
	1000 bu.	% of orig. CATI	1000 bu.	% of SPS edited CATI
Land in farms				
Iowa	-378	-1.7	-273	-1.2
Nebraska	1,901	5.3	2,074	5.7
Pennsylvania	-126	-2.1	-121	-2.0
Total	1,397	2.2	1,680	2.6
Cropland acres				
Iowa	48	.3	222	1.3
Nebraska	-189	-1.2	-118	-.7
Pennsylvania	-107	-2.8	-107	-2.8
Total	-248	-.7	-3	<-.1

Respondents

In order to analyze bias with respect to type of respondent, a respondent combination was generated for each reinterview sample unit based on who responded to the original (CATI) interview and to the reinterview. For example, the respondent combination 'partner - same partner' means that the same partner responded in both interviews.

Respondent combinations were also grouped into categories. The first category is 'operator-operator' and includes all the interview-reinterview combinations where an operator (which includes partner, hired manager, and individual operator) responded both times. This category can be used to measure the response bias resulting from interviewing the operator. The second category is 'other-operator', which includes all cases where the operator was the reinterview respondent but not the original respondent. This category can be used to measure bias in reporting by individuals other than the operator. The last category is 'other combinations', which contains all remaining respondent combinations. This category is difficult to interpret since it contains several effects which are not easily separated.

In the operational survey, enumerators are instructed to contact the operator whenever possible because the operator is assumed to be the most knowledgeable person about the operation. Similarly, the reinterview enumerators were also instructed to reinterview the operator, if possible, regardless of who the original respondent was. A large proportion of the original CATI interviews and the reinterviews involved the operator. Of the original CATI respondents in the reinterview sample, 840 (89%) were the operator. Operators accounted for 87% of the reinterview respondents. Table 4 presents the breakdown of respondent combinations and respondent categories. As shown, the operator-operator category comprised 80.7% of the total respondent combinations. Most of these were individual operators (71.4% of the total). The 'other-operator' category accounted for only 6.4% of the cases, and the 'other combinations' category accounted for the remaining 12.9%.

Table 4. Frequency distribution of original and reinterview respondent combinations for the 3 states combined in the March 1988 reinterview study.

Respondent:

CATI interview-reinterview	Frequency		Percent	
Operator-Operator	763		80.7	
Ind. operator-ind. operator		675		71.4
Partner-same partner		59		6.3
Manager-manager		5		0.5
Partner-different partner		24		2.5
Other-Operator	61		6.4	
Spouse-ind. operator		36		3.8
Other-ind. operator		20		2.1
Spouse-partner		3		0.3
Other-partner		2		0.2
Other Combinations	122		12.9	
Spouse-spouse		28		3.0
Other-same other		6		0.6
Ind. operator-spouse		48		5.2
Ind. operator-other		21		2.2
Partner-spouse		5		0.5
Partner-other		3		0.3
Spouse-other		3		0.3
Other-spouse		6		0.6
Any other combination		2		0.2
Total		946		100.0

Reasons for Differences

Reasons for differences between the two interview responses were provided by the reinterview respondents and recorded on the reconciliation forms by the enumerators. These written explanations were later coded for tabulation purposes.

The reasons for differences between the CATI response and the reconciled value have been grouped into 3 categories. The frequency distribution for these categories is shown in Table 5. The first category is 'estimating/rounding reasons', which includes cases where the respondent said they were just estimating the answer or they used their records for one of the two interviews. At the 3-state level, estimating reasons accounted for 40% of the differences for corn stocks and 34% of the differences for soybean stocks. The second category is 'definitional reasons'. These reasons cover a wide range of problems associated with what exactly should be included and excluded in the reporting of stocks. Some examples are including off-farm stocks,

forgetting to report stocks in one bin, not reporting as of the March 1 reference date, and confusion about whether or not the operation is actually in business (according to the NASS definition). Definitional reasons accounted for 36% of the differences in corn stocks and 46% in soybean stocks. The final category, 'other reasons', includes the reasons that were not attributable to either definitional or estimating problems. 'Other reasons' for differences range from problems associated with telephone interviewing to cases where no explanation was provided by the respondent (the enumerator was not permitted to record a reason unless it was provided by the respondent). The other reason category accounted for 24% of the differences in corn stocks and 20% of the differences in soybean stocks. Appendix B lists all of the reasons given for differences between the CATI and the reconciled responses for corn stocks (Appendix B, Table 1) and soybean stocks (Appendix B, Table 2) at the state level.

Table 5. Frequency distribution of reason categories for corn stocks and soybean stocks for the 3 states combined in the March 1988 reinterview study.

Reason Category	Corn Stocks		Soybean Stocks	
	Freq.	%	Freq.	%
Estimating/Rounding	106	40	22	34
Definitional	95	36	30	46
Other	65	24	13	20
Total	266	100	65	100

Bias and the Effect of Respondent and Reason for Difference

Next we look at the actual contribution to the bias by respondent combination and reason category. Tables 6 and 7 show these bias estimates for corn stocks and soybean stocks, respectively.

Looking at the row totals for corn stocks in Table 6, it can be seen that definitional reasons accounted for about 61% of the bias, other reasons accounted for about 39%, and estimating reasons virtually nothing. Although 40% of the differences were attributable to estimating reasons (as shown in Table 5), they contributed very little to the overall bias. For soybean stocks, estimating reasons were again a small contributor (although in the opposite direction) to the overall bias. Definitional reasons were the largest contributors (75%), with other reasons contributing 28%.

The operator-operator respondent category accounted for 72% and 80% of the bias for corn stocks and soybean stocks, respectively. However, as

shown in Table 4, this category also accounts for 80% of all respondent combinations. Tables 6 and 7 show that definitional type reasons accounted for most of the bias within the operator-operator category for both corn and soybean stocks. The other reason category accounted for most of the bias in the other-operator and other combination categories for corn stocks. There are too few observations for soybean stocks in the last two respondent categories to determine which reason category contributed most to the bias.

For those operations with a bias in corn stocks reporting, 236 (89%) of them were in the operator-operator category. The expanded average underreporting per respondent was -301,000 bushels (-71,030,000 / 236). This was much less than the expanded average amount of underreporting for the other two respondent categories, which was -991,000 bushels for other-operator and -803,000 bushels for other combinations.

Table 6. Bias estimates for corn stocks by respondent and reason categories. Data are from the March 1988 reinterview study.

	Respondent category				
Reason for difference	Operator-operator (1000 bu)	Other-operator (1000 bu)	Other combinations (1000 bu)	Total (1000 bu)	% of total
Estimating	3,063	-3,845	997	215	<1%
Definitional	-61,508	-973	2,441	-60,040	61%
Other	-12,585	-11,030	-14,680	-38,295	39%
Total	-71,030 (n=236)	-15,848 (n=16)	-11,242 (n=14)	-98,120 (n=266)1/	
% of Total	72%	16%	12%		100%

1/ includes those observations with a bias (positive or negative).

9

Table 7.   Bias estimates for soybean stocks by respondent and reason categories.   Data are from the March 1988 reinterview study.

| Reason for difference | Respondent category | | | Total (1000 bu) | % of total |
	Operator-operator (1000 bu)	Other-operator (1000 bu)	Other combinations (1000 bu)		
Estimating	361	-26	0	335	-3%
Definitional	-6,662	0	-1,256	-7,918	75%
Other	-2,132	-27	-780	-2,939	28%
Total	-8,433	-53	-2,036	-10,522	
	(n=59)	(n=2)	(n=4)	(n=65)1/	
% of Total	80%	1%	19%		100%

1/   includes those observations with a bias (positive or negative).

Chi-square tests for checking the relationships between respondent type, reason category, and magnitude of bias were conducted.   Table 8 shows the frequency distribution of the reason category crosstabulated with the respondent category.   No significant relationship between reason and respondent categories was detected (p-value =.56).

Table 8.   Two-way frequency table of respondent category by reason category for corn stocks in the 3 states combined. Data are from the March 1988 reinterview study. 1/

| Reason for difference | Respondent category | | |
	Operator-operator	Other-operator	Other combinations
Estimating	93 (41%)	5 (31%)	3 (21%)
Definitional	82 (35%)	6 (38%)	7 (50%)
Other	56 (24%)	5 (31%)	4 (29%)
Total	236 (100%)	16 (100%)	14 (100%)

Chi-Square   p-value = 0.56

1/   Includes those observations with a bias (positive or negative).

10

To look at the relationship between the level of the bias and the respondent category or reason category, a relative bias was calculated for each observation with a corn stocks bias. Two categories of bias were then generated - one for those observations with a relative bias less than 25% in magnitude (positive or negative) and the other for those with a bias greater than 25% in magnitude.

Table 9 shows the frequency distribution and significance level for the relative bias crosstabulated with respondent category. The results show that the three categories of respondent combinations did not have the same distribution of relative biases (p-value = .07). There were fewer large biases when the respondent was the operator in both interviews. This is not surprising since one would expect larger differences when individuals less knowledgeable about an operation were involved in either the original interview or reinterview and the same person did not respond both times.

Table 10 shows the frequency distribution for the relative bias crosstabulated with reason category. There was a very strong relationship between the size of the bias and the reason that it occurred (p-value < .01). Of all the estimating reasons, 83% of them resulted in biases that were less than 25% in magnitude, with the remaining 17% resulting in biases larger than 25%. Almost the opposite was true for the definitional and other reasons. In these two categories, 73% and 72% of the reasons for a difference resulted in biases larger than 25% in magnitude. This shows that estimating reasons (more frequent in occurrence) accounted for relatively few of the large biases.

As mentioned earlier, the average bias or underreporting for respondents was smaller in the operator-operator category when compared to the other two respondent categories. This is because estimating reasons account for proportionately more of the observed differences in the operator-operator category, yet their contribution to the overall bias is small.

11

Table 9. Two-way frequency table of respondent category by relative bias for corn stocks in the 3 states combined. Data are from the March 1988 reinterview study. 1/

| Relative bias 2/ | Respondent category | | |
	Operator-operator	Other-operator	Other combinations
Between -25% and +25%	123 (52%)	5 (31%)	4 (29%)
Greater than +25% or less than -25%	113 (48%)	11 (69%)	10 (71%)
Total	236 (100%)	16 (100%)	14 (100%)

Chi-Square p-value= 0.07

1/ Includes those observations with a bias (positive or negative).
2/ Relative bias =
    100 * (CATl value - reconciled value) / reconciled value.

Table 10. Two-way frequency table of reason category by relative bias for corn stocks in the 3 states combined. Data are from the March 1988 reinterview study. 1/

| Relative bias 2/ | Reason for difference | | |
	Estimating	Definitional	Other
Between -25% and +25%	88 (83%)	26 (27%)	18 (28%)
Greater than +25% or less than -25%	18 (17%)	69 (73%)	47 (72%)
Total	106 (100%)	95 (100%)	65 (100%)

Chi-Square p-value < .01

1/ Includes those observations with a bias (positive or negative).
2/ Relative bias =
    100 * (CATI value - reconciled value) / reconciled value.

Bias by Size of Operation

The bias in corn stocks was calculated based on size of operation. Tables 11 and 12 show the results using land in farms and cropland acres for the size classification variable. Total expanded corn bias and the average reported bias for each size category are presented. In both tables, it can be seen that as farm size increases, the average amount of corn stocks underreported also increases. The smallest farms actually had a positive average unexpanded bias, yet the overall expanded total for this category was negative.

Table 11. Corn stocks bias estimates by size of operation based on land in farms for the 3 states combined. Data are from the March 1988 reinterview study.

Land in farms classification (acres)	Corn stocks bias (expanded) (1000 bu)	Response (n)	Average bias per response (unexpanded)
0-99	-19,258	158	282
100-249	-5,475	180	-217
250-499	-29,417	203	-823
500-999	-21,387	192	-948
1000+	-19,060	135	-1708
missing 1/	-3,523	10	-
Total	-98,120	878	-780

1/ Operations where a final "true" value for land in farms could not be determined were put in this category.

Table 12. Corn stocks bias estimates by size of operation based on cropland acres for the 3 states combined. Data are from the March 1988 reinterview study.

Cropland acres classification	Corn stocks bias (expanded) (1000 bu)	Response (n)	Average bias per response (unexpanded)
0-99	-19,464	219	196
100-249	-10,772	215	-302
250-499	-23,051	214	-642
500-999	-25,271	153	-1068
1000+	-19,483	66	-4126
missing 1/	-79	11	-
Total	-98,120	878	-678

1/ Operations where a final "true" value for cropland acres could not be determined were put in this category.

13

Specific Reasons for Differences and the Resulting Bias

Table 13 shows some of the specific reasons that contributed most to the definitional bias in corn stocks. The largest definitional contributor was the reason "didn't report someone else's grain". It accounted for 36% of the total definitional bias, and occurred 15 times. Several other reasons are presented, including one where the bias was in the positive direction. However, this positive bias was not large enough to counterbalance the amount of underreporting or negative differences that occurred.

Table 13. The definitional reasons for differences that contributed the most to the bias for corn stocks for all 3 states combined. Data are from the March 1988 reinterview study.

Total Definitional Bias = -60,060 (1000 bu)
Total number of observations contributing to the bias = 95

Reason	Frequency	Contribution to bias (1000 bu)	% of total definitional bias
Didn't report someone else's grain.	15	-21,758	36.2%
Has moved, now has only storage.	1	-8,020	13.4
Forgot to include a bin, silo, or other structure.	18	-7,448	12.4
Didn't include bins on rented land.	4	-7,406	12.3
Didn't include government stored corn.	3	-6,084	10.1
In process of selling operation, but still has stocks.	1	-5,010	8.3
Included grain or capacity on another operation or in town.	7	+11,250	-18.7

DISCUSSION

Results for the March 1988 reinterview study were similar to those for the December 1987 study. Significant underreporting of corn and soybean stocks was found in both studies. Comparable results were found for the distribution of respondent combinations, bias estimates by respondent category and reason category, and the relative magnitude of observed differences with respect to both the respondent category and reason category. The results cannot be compared directly because the surveys were done at different periods and involved different states.

14

The studies indicate that the biases calculated are influenced by two factors, type of respondent and type of reason for the bias. Large relative biases are shown to be associated with respondents other than the operator and with definitional type reasons. We have already noted that operators accounted for nearly 90% of the original CATI respondents subsampled for reinterview. This indicates that CATI is generally successful in contacting operators. Enumerator training should continue to emphasize the importance of contacting the operator. Efforts to reduce the bias should focus on specific reasons for the differences. Many of the reasons for errors that involve excluding or including stocks incorrectly can most likely be reduced through questionnaire design changes and enumerator training. As noted in Pafford (1989), this is particularly true of the definitional type reasons.

The problem of significant underreporting of on-farm grain stocks has been observed for at least 20 years in the Agency. This study was limited in that it only included subsamples from CATI interviews. This does not, however, imply that non-CATI telephone interviews would be better or worse for on-farm grain stocks. Only future studies that include subsamples from both CATI and non-CATI interviews could provide any insight into that comparison. Research by House (1984) on the use of CATI for a cattle survey indicates that CATI interviews are at least as accurate as non-CATI interviews.

Finally, it was noted that a single observed difference can have a significant impact on the total bias estimated. Appendix C, Figure 1, shows the distribution of soybean stocks biases, excluding zeroes. The total number of observations with a bias was 65. Three extreme negative values can be read from the frequency chart. The bias estimate was recalculated excluding these three observations. Although the negative bias decreased from -10.6% to -6.7%, it was still statistically significant.

The December 1988 and December 1989 reinterview studies were conducted in all six of the states that participated in the December 1987 or March 1988 reinterview surveys. Future reports will be able to compare the bias levels over time. This will be important in determining whether the bias is constant or needs to be measured on a regular basis. Reasons for the biases which are common to each survey and the magnitude of the biases attributable to these reasons will be identified. Future recommendations for design changes or additional training will depend upon the frequency of occurrence of particular reasons and the magnitude of the resulting biases. In the March 1988 reinterview study, for example, 39% of the corn stocks differences that were categorized as definitional were attributable to three specific reasons. They were failure to report someone else's grain, failure to include a bin or other grain storage structure, and failure to include bins on rented land (see Table 13). These three reasons accounted for 61% of the total definitional bias and 37% of the total bias. Similar analysis and comparison of the later reinterview studies may enable us to identify specific reasons contributing most to the biases.

REFERENCES

House, Carol (1984) "COMPUTER ASSISTED TELEPHONE INTERVIEWING ON THE CATTLE MULTIPLE FRAME SURVEY," SFSRB Staff Report Number 82, Statistical Reporting Service, U.S. Department of Agriculture.

Pafford, Brad (1989) "USE OF REINTERVIEW TECHNIQUES FOR QUALITY ASSURANCE: The Measurement of Response Bias in the Collection of December 1987 Quarterly Grain Stocks Data Using CATI," SRB Research Report Number SRB-89-08, National Agricultural Statistics Service, U.S. Department of Agriculture.

AGRICULTURAL SURVEY
MARCH 1, 1988
REINTERVIEW FORM

Form Approved
O.M.B. Number 0535-0213
Approval Expires 12/31/89

Pennsylvania

PART C

999
1

Strata	ID	Tract	Subtr.

Dear Reporter:

Information requested in this survey is used to improve the quality of our agricultural statistics. It is strictly confidential, and your response is voluntary.

Respectfully,

Richard D. Allen
Chairperson
Agricultural Statistics Board

SECTION 1 - IDENTIFICATION

1. Please verify name and address of this operation.
   Is it correct?   ☐ YES   ☐ NO   (Make corrections on lablel)

   `001`

2. On land operated by the farm, ranch or individual(s) listed on the label:

   a. Will crops be grown or hay cut at any time during 1988? .... ☐ YES ☐ NO

   b. Will grains or oilseeds be stored at any time during 1988, or do you have storage facilities used for storing grain? .... ☐ YES ☐ NO

   ➤ If **NO** to all questions, **GO TO SECTION 7.**

   c. Were any hogs on this operation from December 1, 1987 through February 29, 1988?..... ☐ YES ☐ NO

   `002`

3. Does this operation do business under any name other than as shown on label?

   ☐ **NO**   ☐ **YES** - Enter name: _____
   (Do you want this name to appear on the label?) ☐ YES ☐ NO

   `003`

4. Are the day-to-day decisions for this farming (or ranching) operation made by:

   ☐   An Individual Operator?

   ☐   Partners? Enter number of partners, including self.........

   ☐   A Hired Manager?

   `004`

   4a. Are the decisions still made by the same person(s) making them on June 1, 1987?

   ☐ YES   ☐ NO   Would you please explain what changed?

   _____
   _____

   `005`

Continue On Next Page

SECTION 2 - ACRES OPERATED

1. How many **total acres of land** were in this operation on **March 1**? . . . .

900

    Include: Farmstead, all cropland, woodland, pastureland, wasteland,
               government program land, all land owned, rented or managed.

    Exclude: Land rented to others and all grazing land used on an
               AUM *(fee per head)* basis

2. Of the total acres in this operation, how many acres would be considered
cropland *(include land in hay and cropland in government programs)*? . . . .

802

SECTION 3 - CROPS

How to complete this section:

--- Report for all the land you operate, including land rented from others
--- Include cover crops planted on government program land.

Please report for crops **PLANTED** and **TO BE PLANTED** for the
1988 CROP YEAR.

ACRES PLANTED AND TO BE PLANTED THIS SPRING 1988
530

1. **CORN** for all purposes (exclude popcorn and sweet corn) . . . .

600

2. **SOYBEANS** for all purposes . . . .

Crops: Incomp.	1 - Has 2 - Unk 3 - No	138
Comp.	4	

SECTION 4 - GRAINS AND SOYBEANS IN STORAGE

Please account for whole grains and soybeans on hand or stored March 1 on the total acres
operated, whether for feed, seed, or sale. They may have belonged to you or someone else,
or been stored under a government program *(loan, farmer owned reserved, or CCC).*

	NO	YES		1987 and earlier crop years
1. On March 1, was any **WHOLE GRAIN CORN** on hand or stored on the total acres operated? . . .	☐	☐	How many bu?	121
2. Were any **SOYBEANS** on hand or stored on these acres?	☐	☐	How many bu?	125
3. What about **WHEAT**? . . . . *(winter, durum, and spring)*	☐	☐	How many bu?	126

GRAIN STORAGE CAPACITY

4. On March 1, what was the **TOTAL STORAGE CAPACITY** of all the bins,
cribs, sheds, and other structures normally used to store whole grains
or oilseeds on the total acres operated? . . . . bushels

808

Stocks: Incomp.	1 - Has 2 - Unk 3 - No	141
Comp.	4	

18

SECTION 5 - HOGS AND PIGS

1. On March 1, were any **HOGS** or **PIGS**, regardless of ownership, on the total acres operated?

`006`

☐ **YES**   ☐ **NO** ——▶ **1a.** Were any **HOGS** or **PIGS** on the total acres operated, at any time, during the period of December 1, 1987 through February 29, 1988?

☐ **YES** - *GO TO Section 6.*   ☐ **NO** - *GO TO Section 6.*

2. Of the **HOGS and PIGS** for **BREEDING** on hand March 1, how many were:

a. **Sows, gilts and young gilts** bred and to be bred . . . . `301` ★

b. How many were **boars and young males** for breeding . . . . `302` ★

c. How many were **sows and boars** no longer used for breeding . . . . `303` ★

3. How many **HOGS and PIGS FOR MARKET AND HOME USE** were on hand March 1? *(Exclude breeding hogs reported in Item 2.)* . . . . `007` ★

4. **TOTAL NUMBER OF HOGS** and **PIGS** on hand March 1. (Add ★ *Items 2a through 3)* . . . . `300`

Hogs: Incomp.	1 - Has 2 - Unk 3 - No	499
Comp.	4	

SECTION 6 - PARTNER'S NAMES

1. Did you check partners in Section 1, Item 4, on Face Page?

☐ **NO** - *GO TO Section 8.*

☐ **YES** - *Continue.*

2. Please identify the other person(s) in this partnership in boxes below, then go to Section 8.
(make necessary corrections if names have already been entered)

Name _____
 (First)   (Middle)   (Last)
Address _____
 (Rt. or St.)
City _____ State _____
Zip Code _____ Phone _____
 Did this person operate land individually
 in this State on June I, 1987? ☐ **YES** ☐ **NO**

Name _____
 (First)   (Middle)   (Last)
Address _____
 (Rt. or St.)
City _____ State _____
Zip Code _____ Phone _____
 Did this person operate land individually
 in this State on June I, 1987? ☐ **YES** ☐ **NO**

Continue On Next Page

19

SECTION 7 - CHANGE IN OPERATOR

Has this operation (name on label) been sold, or turned over to someone else?

☐ **NO** - *GO TO Section 8.* ☐ **YES** - *Please identify the new operator(s).*

| 011 |

Name _____

Address _____ Phone _____

City _____ State _____ Zip _____

Did this person operate land individually in this State on June I, 1987? ☐ **YES** ☐ **NO**

| 012 |

SECTION 8 - CONCLUSION

| 013 |

1. Do you make any day-to-day decisions for another farm or ranch?

☐ **NO** ☐ **YES** - *List other operation(s)* _____

_____

2. Thank you for your help in completing this questionnaire. Now I would
   like to compare these responses with those from the original telephone
   interview.

| GO TO THE RECONCILIATION FORM. |

Reported by _____ Date _____

Telephone(Area Code) _____ (Number) _____

Respondent		Response Code		Sup /Enum	Eval	Date
1-Op	101	3-Int	910	098	100	095
2-Sp		8-IR				
3-Oth		9-Inac				

AGRICULTURAL SURVEY - MARCH 1988

RECONCILIATION FORM

THIS FORM IS NOT TO BE OPENED UNTIL AFTER THE REINTERVIEW
RESPONSES HAVE BEEN OBTAINED. In order to obtain measures of
quality of our data we must maintain independence between the
initial and reinterview surveys. Viewing the initial response
before the reinterview may hurt this relationship.

Strata	ID	Tract	Subtr	County

LABEL

+--+
| |
| |
| |
| |
| |
+--+

Reported by:_____ Date:_____

Telephone : _____

Please check one:
| Completed--with |
| differences   ___ |
| Completed--with |
| no differences ___ |
| Refusal   ___ |
| Inaccessible   ___ |

Sup./Enum. code :_____

PA

Initial Respondent: Date of Initial Interview: Mar.

QUESTION	ORIGINAL (1)	REINTERVIEW (2)	What is correct? (3)		Source of error? R1-orig.resp. R2-reint.resp E1-orig.enum. E2-reint.enum (4)
			O	R	
SECTION 1 - IDENTIFICATION					
1. Label Correct? (YES or NO)					
Corrections to Label:					
2. Crops OR Grains OR Hogs on Farm? (YES or NO)					
3. Business under another name? (Y or N)					
name:					
4. Day-to-day decisions made by: I=ind. oper., P=partnrs, H=hired mgr.					

EXPLANATION			
(Explain as fully as possible why the original and reinterview information differed)			
Section Number	Item		Reason for difference

PA

Initial Respondent: Date of Initial Interview: Mar.

QUESTION	ORIGINAL (1)	REINTERVIEW (2)	What is correct? (3)		Source of error? R1-orig.resp. R2-reint.resp E1-orig.enum. E2-reint.enum (4)
			O	R	
SECTION 2 – ACRES OPERATED					
1. TOTAL ACRES of Land					
2. CROPLAND ACRES					
SECTION 3 – CROPS					
1. CORN planted and to be planted					
2. SOYBEANS planted and to be planted					

EXPLANATION
(Explain as fully as possible why the original and reinterview information differed)

Section Number	Item	Reason for difference

PA

Initial Respondent: Date of Initial Interview: Mar.

QUESTION	ORIGINAL (1)	REINTERVIEW (2)	What is correct? (3) O	R	Source of error? R1-orig.resp. R2-reint.resp E1-orig.enum. E2-reint.enum (4)
SECTION 4 - GRAINS AND SOYBEANS IN STORAGE					
1. CORN in STORAGE					
2. SOYBEANS in STORAGE					
3. ALL WHEAT in STORAGE					
4. GRAIN STORAGE CAPACITY					

EXPLANATION
(Explain as fully as possible why the original and reinterview information differed)

Section Number	Item	Reason for difference

PA

Initial Respondent: Date of Initial Interview: Mar.

QUESTION	ORIGINAL (1)	REINTERVIEW (2)	What is correct? (3)		Source of error? $R1$-orig.resp. $R2$-reint.resp $E1$-orig.enum. $E2$-reint.enum (4)
			O	R	
SECTION 5 – HOG AND PIG INVENTORY					
1. HOGS or PIGS? (YES or NO)					
1a. HOGS or PIGS between Dec 1 and Feb 29? (YES or NO)					
2a. SOWS, GILTS BRED AND TO BE BRED					
b. BOARS AND YOUNG MALES FOR BREEDING					
c. SOWS AND BOARS NO LONGER USED FOR BREEDING					
3. HOGS & PIGS FOR MARKET AND HOME USE					
4. TOTAL HOGS & PIGS					

EXPLANATION
(Explain as fully as possible why the original and reinterview information differed)

Section Number	Item	Reason for difference

PA

Initial Respondent: Date of Initial Interview: Mar.

QUESTION	ORIGINAL (1)	REINTERVIEW (2)	What is correct? (3) O	R	Source of error? R1-orig.resp. R2-reint.resp E1-orig.enum. E2-reint.enum (4)
SECTION 6 - PARTNER'S NAMES					
Name 1:					
Name 2:					
Name 3:					
Name 4:					
SECTION 7 - CHANGE IN OPERATOR					
Operation name on label sold, or turned over? (YES or NO)					
Name:					
Operate Indiv. on June 1? (YES or NO)					
SECTION 8 - CONCLUSION					
1. Day-to-day decisions for another farm or ranch (YES or NO)					
Name:					

EXPLANATION
(Explain as fully as possible why the original and reinterview information differed)

Section Number	Item	Reason for difference

PA

EXPLANATION		

(Explain as fully as possible why the original and reinterview information differed)		
Section Number	Item	Reason for difference

PA

APPENDIX B: REASONS FOR DIFFERENCES BY STATE

Table 1. Reasons for differences in corn stocks reporting between
 CATI and the reinterview, by state, for the March 1988 reinterview study.

Reason	Iowa Freq.	Iowa %	Nebraska Freq.	Nebraska %	Pennsylvania Freq.	Pennsylvania %
ESTIMATING/ROUNDING REASONS	48	45%	25	32%	33	40%
Figure was estimated	32		18		24	
Used records or actually counted	1		–		1	
Both estimated, either could be right	5		3		8	
Rounding	2		2		–	
Respondent reported CCC amount, but thinks there is more	1		–		–	
Final correct value was refigured	–		1		–	
Checked records for a final correct value	1		–		–	
Final correct value was between the original and reinterview responses	3		–		–	
Final correct value was average of the original and reinterview responses	3		1		–	

Table 1. (cont.) Reasons for differences in corn stocks reporting between CATI and the reinterview, by state, for the March 1988 reinterview study.

Reason	Iowa Freq.	%	Nebraska Freq.	%	Pennsylvania Freq.	%
DEFINITIONAL REASONS	39	37%	35	45%	21	26%
Did not report as of March 1	-		3		-	
Didn't report "on total acres operated"	-		2		-	
R didn't think it was enough to report	-		1		-	
Enumerator asked question wrong	1		-		-	
Reported 1987 amount, not 1988	1		-		-	
Enumerator thinks respondent misunderstood question	-		2		-	
Confusion over what to include or exclude	4		-		1	
Forgot to include a bin, silo, or other structure	7		6		5	
Gave capacity, not grain stored	2		2		-	
Didn't include bins on rented land	2		1		1	
Didn't report someone elses grain	6		6		3	
Didn't include government stored corn	2		1		-	
Included grain or capacity on another operation or in town	1		4		2	
Gave milo stored, not just corn	-		1		-	
Sold some, but is still on operation	-		1		-	
Didn't include high moisture storage	-		1		-	
Reported sealed bushels, not actual	1		1		-	
Didn't report stored & sealed grain	1		1		-	
Didn't report ear corn amount or storage	1		-		4	
Didn't report corn for feed	2		-		-	
Farm rented out, but does have storage	2		-		-	
Reported beans with corn, on operation and in town	1		-		-	
Included grains that were not whole	-		-		3	
Didn't report seed grain	1		-		-	
Respondent told enumerator not really a farm, but has hay	-		1		-	
In process of selling operation, but still has	-		1		-	
Does have stocks (no crops or hay)	-		-		1	
Has moved, now has only storage	1		-		-	
Has hay & storage, so still in business	1		-		-	
Most land is sons (respondent), but operator still has some	1		-		-	
"Retired" but has stocks (forgot to report)	1		-		-	
R1 said turned over, R2 said no	-		-		1	

29

Table 1. (cont.) Reasons for differences in corn stocks reporting between CATI and the reinterview, by state, for the March 1988 reinterview study.

Reason	Iowa Freq.	%	Nebraska Freq.	%	Pennsylvania Freq.	%
OTHER REASONS	19	18%	18	23%	28	34%
No explanation	1		1		7	
Respondent said they thought they had reported the correct value the first time	4		4		1	
Respondent said "don't know where that answer came from"	-		-		2	
Resp said was not asked on phone interview	-		-		3	
Too rushed on phone to figure accurately	-		2		-	
Misunderstanding between enumerator and respondent	2		5		1	
"Gave wrong answer" or "added wrong"	6		1		4	
Forgot to report	1		1		-	
Phoned too late (or too early) to give good answer	-		-		1	
Respondent had difficulty hearing on phone	1		1		-	
Spouse didn't know	-		-		1	
Enumerator switched corn and soybean storage amounts	-		1		-	
Forgot son had removed his grain	-		1		-	
Possible problem in converting from tons to bushels	-		-		8	
Respondent didn't remember any phone interview	-		1		-	
Respondent said enumerator didn't ask about grain storage	1		-		-	
Original respondent was office	1		-		-	
Respondent gave no reasons, but enumerator thinks reinterview answers are better	1		-		-	
Operator and wife (first respondent) retarded, 2nd respondent is their guardian	1		-		-	
Total	106	100	78	100	82	100

Table 2. Reasons for differences in soybean stocks reporting between CATI and the reinterview, by state, for the March 1988 reinterview study.

Reason	Iowa Freq.	Iowa %	Nebraska Freq.	Nebraska %	Pennsylvania Freq.	Pennsylvania %
ESTIMATING/ROUNDING REASONS	15	33%	5	38%	2	29%
Figure was estimated	10		2		1	
Used records or actually counted	1		1		-	
Both estimated, either could be right	2		1		1	
Rounding	1		-		-	
Final correct value was best estimate	-		1		-	
Final correct value was average of the original and reinterview responses	1		-		-	
DEFINITIONAL REASONS	22	49%	4	31%	4	57%
Did not report as of March 1	4		1		-	
Reported 1987 amount, not 1988	1		-		-	
Confusion over what to include or exclude	2		-		1	
Forgot to include a bin, silo, or other structure	2		1		-	
Didn't include bins on rented land	2		-		-	
Didn't report someone elses grain	4		2		-	
Included grain or capacity on another operation or in town	2		-		2	
Didn't report stored & sealed grain	1		-		-	
Reported beans with corn, on operation and in town	1		-		-	
Didn't report seed grain	-		-		1	
Has moved, now has only storage	1		-		-	
Most land is sons (respondent), but operator still has some	1		-		-	
Tenant operates land (all rented out)	1		-		-	

Table 2. (cont.) Reasons for differences in soybean stocks reporting between CATI and the reinterview, by state, for the March 1988 reinterview study.

Reason	Iowa Freq.	%	Nebraska Freq.	%	Pennsylvania Freq.	%
OTHER REASONS	8	18%	4	31%	1	14%
No explanation	1		–		–	
Respondent said they thought they had reported the correct value the first time	1		–		–	
Resp said was not asked on phone interview	–		3		–	
"Gave wrong answer" or "added wrong"	3		–		1	
Forgot to report	1		–		–	
Enumerator switched corn and soybean storage amounts	–		1		–	
CATI off 2 decimal places	1		–		–	
Operator and wife (first respondent) retarded, 2nd respondent is their guardian	1		–		–	
Total	45	100%	13	100%	7	100%

APPENDIX C: FREQUENCY CHART OF SOYBEAN BIASES

Figure 1. Frequency chart for the unexpanded soybean stock biases (does not include observations with no bias). Data are from the March 1988 reinterview study.

```
Soybean stocks                             Number of      Percent
  bias interval                            observations   of total
                                |
(-15,000,  -14,000]             | *                  1        1.54
(-14,000,  -13,000]             |                    0        0.00
(-13,000,  -12,000]             |                    0        0.00
(-12,000,  -11,000]             |                    0        0.00
(-11,000,  -10,000]             | *                  1        1.54
(-10,000,   -9,000]             |                    0        0.00
(  -9,000,  -8,000]             |                    0        0.00
(  -8,000,  -7,000]             | *                  1        1.54
(  -7,000,  -6,000]             |                    0        0.00
(  -6,000,  -5,000]             | *                  1        1.54
(  -5,000,  -4,000]             | ***                3        4.62
(  -4,000,  -3,000]             | **                 2        3.08
(  -3,000,  -2,000]             | ****               4        6.15
(  -2,000,  -1,000]             | *************      14       21.54
(  -1,000,       0)             | ***************    16       24.62
(       0,   1,000]             | *************      14       21.54
(   1,000,   2,000]             | ****               4        6.15
(   2,000,   3,000]             | ***                3        4.62
(   3,000,   4,000]             | *                  1        1.54
                                 ----+---+---+---+--
                                    4   8  12  16
                                 Frequency
```

33