STUDIES OF RESPONSE ERRORS IN NASS SURVEYS: THE EFFECT OF USING PREVIOUS SURVEY DATA

Bradley V. Pafford, U.S.D.A., National Agricultural Statistics Service

The estimating program of the National Agricultural Statistics Service (NASS) consists of a series of monthly, quarterly, semiannual, and annual surveys of farmers and farm related businesses. The nature of these surveys is such that many questions asked of farmers overlap with a previous month's or quarter's report.

One can classify the type of questions asked for related surveys as either: (1) exact matching or, (2) related content, which is the most common. Questions that match exactly from the initial survey to the followup survey include such items as planted acreage for the June Acreage and Fall A&P surveys and harvested acreage for minor crop surveys. Most often, in NASS surveys, the same question is asked, but some changes can be expected from the initial inquiry. These surveys ask for related content.

Our survey procedures generally direct us to treat each such interview as independent of the previous one. NASS has traditionally been concerned with biasing responses by using historical data in the current interview.

The degree of adherence to this general guideline is believed to vary by State Statistical Office (SSO), survey, interview media (mail, telephone, and personal interview), interviewer, and year. Recent interest in respondent burden, concern for accurate data, and availability of high-speed computers to store and process data are a few of the reasons our SSO's have abandoned the "independence" concept and gone to using previous survey data.

Although many States have experimented with, and actually implemented the use of previous survey data in their operating programs, little research has been done to direct this effort or to provide proper guidelines. Variability in response can be affected by factors other than the use of historical data. These factors include the effect due to the enumerator, and whether the same respondent reported in the initial and in the follow up survey(s).

The California SSO undertook a research project for the 1985 Fall A&P survey to examine: (1) the effect of providing historical data in the current interview, and (2) the effects due to the interviewer, size of operation, and respondent. The historical data used were 1985 June Acreage survey planted acreages, respondents were asked to re-report these acreages in the presence or absence of previous survey data. All crops studied were planted by the June inquiry. Use of historical data, then, follows the exact question matching situation discussed previously.

Few studies on the use of historical data were found in outside journals; the primary focus is on response variability. This is trial-to-trial variability. Conceptually, this can be thought of as the proportion of individuals who reported differently on two occasions. Common terminology in the literature is reinterview research, and measurement of simple response variance (SRV).

O'Muircheartaigh (1986), working with data from the Census reinterview program, found a marked effect on the estimates of response variance; values were about twice as large for the cases where historical data were used as when these data were not used. The effect of allowing the reinterviewer to have access to the original responses was to reduce the estimate of SRV by 50 percent. Also, reinterviewers reacted differently to various respondents (for example, husbands and wives); that is, the measures of SRV varied by respondent type. O'Muircheartaigh concluded that instructions to enumerators, which direct them not to use the historical data until after the response then reconcile any differences, were not being followed. Bailar (1968) studied the length of time between a reinterview and enumerator access to the respondents' original response with the same quality control procedures for the CPS. The best procedure was the one in which the reinterview was soon after the original interview and one where the interviewers did not have access to the original responses.

In contrast to the relatively sparse literature on the use of previous survey data, there is extensive research in the area of interviewer variability. Early work was done by Mahalanobis (1946); Hansen, Hurwitz, Marks, and Maudlin (1951), and Hansen, Hurwitz and Pritzker (1964). These works and others in the fifties and sixties dealt mostly with personal interviewing. Hanson and Marks (1958) found significant interviewer effects when: (1) there was some resistance on the part of the enumerator to the question, (2) the questions were ambiguous or wordy, and (3) additional probing was done. Kish (1962) also found interviewer effects for highly ambiguous and critical attitudinal questions. Further work in the sixties and early seventies seemed to focus on developing models that examined a unified approach for response error components (Fellegi, 1964; Bailar and Dalenius, 1969; Koch, 1973)

Work on telephone interviewer variability began in the late seventies and early eighties. Groves and Kahn (1979) concluded that telephone interviewer variability was less than that for personal interviewing. However, the impact on total survey error was greater because the the average interviewer workloads were larger than for personal interviewing. The largest interviewer effects appeared for questions suggesting some sensitivity to the topic and for questions that depended on the interviewer's probing to elicit complete responses. Groves and Magilavy (1980) attempted to explain interviewer variability in terms of question types, respondent types, and interviewer types. They concluded that lower values of ρ can be achieved by close supervision, monitoring of performance, and exchange of interviewing techniques among interviewers.

The study of respondent effects are another area where there was little NASS research. Bosecker (1977) reported different levels in total acres, tract acres, and total cattle by respondent in a 1976 Oklahoma December Enumerative Survey (DES)survey. Warde (1986) reported smaller acreages and counts of hogs and pigs for spouses than for farm operators. Finally, Nealon and Dillard (1984), in a nationwide study of characteristics of farm wives and husbands found that wives had significantly more missing data and lower mean responses than husbands for an overwhelming number of characteristics.

In his CPS reinterview research, O'Muircheartaigh (1986) found the SRV lowest for self-self reports (that is, same person reports both times). Less reliable reporting was noted or a larger SRV occurred for proxy - (same)proxy respondent group. Respondent characteristics, such as age, husband/wife, or relationship to head of household also affected the consistency of reporting.

The sample units for the 1985 Fall A&P Survey were obtained from a subsample of respondents reporting durum wheat, other wheat, barley, or oats acres in the 1985 June Acreage survey. There was a total of 6,423 responses to the 1985 June Acreage survey in California, in which 2,082 units met the criteria of at least one small grain crop. These units, then, were stratified into "small," "medium," and "large" operations by collapsing the original 14 list frame strata. A random sample of 640 sample units were drawn per stratum for a total of 1,920 units from the eligible list of 2,082. The 640 samples in each stratum were then randomly assigned to a treatment by enumerator combination. Ten enumerators worked this survey and there were 4 treatments studied, or 40 treatment*enumerator cells per stratum. Dividing 40 into 640 resulted in 16 cases to complete per stratum* treatment*enumerator cell. Assignment, then, consisted of these 16 cases. Each enumerator had 12 assignments of 16 cases each to complete for a total 192 interviews.

All data for this research project were collected through CATI. A CATI instrument was designed to display the appropriate historical data treatment instrument version.

This study examines several levels of the previous data treatment and the effects of respondents, interviewers, and size of operation for the 1985 California Fall A&P survey on variables that overlap with the 1985 June Acreage survey. "Overlapping variables" refers to planted acreage questions asked in June and again in the fall.

The previous data treatment levels studied included:

(1) Control--The interview was conducted without the interviewer's and respondent's knowledge of the previous data. Historical data, if available, were not referenced in any way. The interview was treated as a "first-time" call (TREATMENT 1).

(2) Probe--Interviewers conducted the survey with previous data for that respondent on the CATI screen, but made no direct mention of these data to the respondent. This treatment was designed to reproduce the operational method in California at that time. The directions given to the interviewers were to: (1) not mention the historical figure, but to (2) re-ask the question when they thought an answer deviated too much from the orginial response. What constituted a deviation from the original response was left up to the enumerator (TREATMENT 2).

(3) Upfront--Enumerators conducted the interview with the previous responses worded into their questions. An example of this type of question was: "Our records show that xxxxx acres of CORN were reported in JUNE as planted on your operation for harvest in 1985. Is that the correct acreage?" (TREATMENT 3).

(4) Not used--Previously reported figures were used for the current Fall report when historical data were present. That is, planted acreage questions were not asked of these respondents. Instead, the CATI instrument went directly to the harvested acreage questions (TREATMENT 4).

Respondent classification was either the "same" or "different." That is, the respondent to the Fall survey was either the same person as the June respondent or a different person.

Finally, to test for interviewer effects, interviewer assignments were randomized so that each enumerator received one folder from each of the 12 treatment by stratum combinations.

Table 1 presents analysis of variance (AOV) results for the following models, respectively:

$$D_{ijkl} = \mu + S_i + T_j + ST_{ij} + I_k + SI_{ik} + TI_{jk} + \varepsilon_{l(ijk)}$$

$$Y_{ijkl} = \mu + S_i + T_j + ST_{ij} + I_k + SI_{ik} + TI_{jk} + \varepsilon_{l(ijk)}$$

$i=1,2,3 \quad j=1,2,3 \quad k=1,2,3,..,10 \quad l=1,2,..,16$

where:

D_{ijkl} = difference in the current planted acres and historic planted acreage response variable from the l-th unit stratum i , obtained by the k-th interviewer using the j-th type of historical data procedure;

Y_{ijkl} = planted acreage response variable from the l-th unit stratum i , obtained by the k-th interviewer using the j-th type of historical data procedure;

μ = the overall mean;

S_i = effect of the i-th stratum (fixed effect);

T_j = the effect of using the j-th historical data procedure (fixed effect);

ST_{ij} = the interaction effect of using the j-th historical data procedure in the i-th stratum;

I_k = the effect of the k-th interviewer (fixed effect);

SI_{ik} = the interaction effect of using the k-th interviewer in the i-th stratum;

TI_{jk} = the interaction effect of the k-th interviewer with the j-th historical data procedure;

$\varepsilon_{l(ijk)}$ = random effect within the i-th, j-th, k-th cell.

Table 1 shows the significance levels at which the null hypothesis of no difference in the effects are rejected. Nearly all tests of the overall model (53 df) were significant. That is, some of the model components did have an effect on the consistent reporting of planted acreages. The null hypotheses of no strata and historical data treatment effects were consistently rejected for nearly all the variables analyzed in table 1.

Table 1—Analysis of variance for selected crops: difference in 1985 Fall A&P and June reported planted acreage

Source 1/	DF	P-Values 2/ Difference in Fall & June Planted Acres: Small Grains F	Small Grains RST	Small Grains KWD	Oats F	Oats RST	Oats KWD	Other Wheat F	Other Wheat RST	Other Wheat KWD
Model	53	0.24	<0.01	<0.01	<0.01	<0.01	<0.01	0.99	<0.01	<0.01
S	2	.59	<.01	<.01	.99	<.01	<.01	.99	<.01	<.01
T	2	.06	<.01	<.01	.02	<.01	<.01	.70	<.01	<.01
S*T	4	.38	.17	.27	.08	.16	.21	.95	<.01	<.01
I	9	.13	.13	.27	.01	.56	.64	.92	.70	.78
S*I	18	.33	.81	.93	.62	.80	.86	.99	.41	.58
T*I	18	.49	.19	.44	.01	.31	.42	.93	.37	.54
Error	1136									
Total	1189									
Source 1/	**DF**	**Other Hay F**	**Other Hay RST**	**Other Hay KWD**	**Corn F**	**Corn RST**	**Corn KWD**	**Cotton F**	**Cotton RST**	**Cotton KWD**
Model	53	0.88	0.12	<0.01	0.08	<0.01	<0.01	<0.01	<0.01	<0.01
S	2	.27	.04	.04	.85	.01	.01	<.01	<.01	<.01
T	2	.93	<.01	.01	.04	<.01	<.01	.01	<.01	<.01
S*T	4	.24	.97	.97	.20	.62	.64	<.01	.03	.05
I	9	.82	.62	.63	.20	.55	.59	<.01	.39	.46
S*I	18	.62	.81	.82	.08	.09	.11	<.01	.35	.45
T*I	18	.54	.07	.07	.28	.22	.26	.06	.30	.40
Error	718									
Total	771									

1/ S = Strata; T = Treatment; and I = Interviewer.
2/ F = F-statistic for the difference in Fall A&P and June reported planted acreage; RST = F-Statistic based on the ranks of the absolute value of the difference in the Fall and June reports; KWD = extension of the Kruskal-Wallis ranks test for the absolute value of the difference in the Fall and June reports (see appendix A).

Table 1 also shows the results of other hypothesis tests. Focusing on interviewer effects, one can see the statistical tests were not sensitive enough to detect a consistent difference. The F-tests shown in table 1 suggest a slight effect due to the interviewer, however, when the analyses were based on ranked data, most of the significance disappeared. This distinction may be due to the sample sizes at the enumerator level. The average number of interviews completed per enumerator was 158.

The balance of the statistical tests shown in table 1 were for the interactions among historical data treatment procedures, interviewers, and strata. The p-values for these interaction tests, however, vary widely by crop and by test statistic. This same trend was not apparent for the main effects in most cases.

The cause for this can be attributed perhaps to the data, and perhaps to the test statistics used for detecting significance. That is, in the data, cell means were affected drastically by one or two observations. Cell sizes varied from around 30-36 observations for the treatment by enumerator and stratum by enumerator interaction, to 100-140 observations for the treatment by stratum interaction. The data were skewed and kurtosed (spike at zero) and severe departures from normality occurred. These affected the p-values derived from the parametric F-test procedure, and definitive conclusions from these tests could not be made. From the nonparametric testing perspective, interpretation of the interactions based on rank-transformed data were questionable.

From a NASS perspective, the major result was that the method of providing historical data did affect the level difference between the Fall and June reports. Significance is shown in all six variables analyzed. Weighted mean differences by crop and treatment are shown in table 2. There was little difference between the Fall A&P and June reports of planted acreage when interviewers directly referenced the historical (June) report (TREATMENT 3). The largest difference occurred when no historical data was presented (TREATMENT 1). That is, when farmers were given the opportunity to elicit a "same" response they were more likely to do so.

Table 2—Estimated means by treatment and selected crop.

Crop	n 2/	Mean planted acreage 1/: TREATMENT = 1 Fall A&P	TREATMENT = 1 June Ac.	TREATMENT = 1 D	TREATMENT = 2 Fall A&P	TREATMENT = 2 June Ac.	TREATMENT = 2 D	TREATMENT = 3 Fall A&P	TREATMENT = 3 June Ac.	TREATMENT = 3 D
						Acres				
Small grains	397	244	266	-22	246	254	-8	271	270	1
Oats	397	30	38	-8	39	42	-3	34	33	1
Other wheat	397	113	108	5	93	96	-3	145	145	0
Corn	257	36	38	-2	43	38	5	30	28	2
Cotton	257	45	48	-3	42	52	-10	59	61	-2
Other Hay	257	8	8	0	6	7	-1	8	6	2

1/ D = Fall A&P - June Reported Planted Acreage
2/ Average treatment response.

Table 3 gives frequencies and percentages of reported planted acreage by category and treatment for four selected crops. This table provides further detail on the effect of historical data treatment on current responses. First, the percentage of respondents failing to report planted acreage in the fall when it was reported in June (Fall=0 & June>0) and planted acreage in June when it was reported in the Fall (Fall>0 & June=0) was much greater when no historical data was referenced (TREATMENT 1), compared with when it was directly referenced (TREATMENT 3). The corresponding percentages for the TREATMENT 2 respondents were somewhere in the middle. Less of this effect was evident for the row crops than for the small grains.

Next, the percentage of respondents reporting the <u>same</u> figure in June and the fall was much greater for TREATMENT 3 than the percentage of respondents for the other two treatment groups.

These implications are paradoxical. By using the historical data directly (TREATMENT 3) one tends to get a re-report of the planted acreage; however, by not using these data respondents tended to "forget" about their crop acreage. Here "forgetting" is a general term that could include: (1) answer reported on the wrong line of June questionnaire, (2) respondent too busy at the time of interview to give a "complete" answer, (3) mix up in land operating arrangements so that the fields to be included were confusing, and (4) enough time lag between reporting in June and the fall is such that the farmer was confused with the crop year referenced. This last point may be especially true for the reporting of small grain acreage. That is, most farmers were beginning to plant or to consider planting the current year's winter grains when asked about plantings a year before. Confusion on the part of the respondent would have been natural.

Table 3—Comparison of Fall A&P and June reported planted acres by crop. number and percentage of respondents reporting by category.

CROP / CATEGORY (Planted acres)	Treatment 1 Count	Treatment 1 Percent	Treatment 2 Count	Treatment 2 Percent	Treatment 3 Count	Treatment 3 Percent
OTHER WHEAT						
Fall=0 & June>0	17	4.2	12	3.1	5	1.2
Fall>0 & June=0	13	3.2	6	1.6	4	1.0
Fall = June	315	77.0	282	73.8	380	92.0
Fall ≠ June	64	15.6	82	21.5	24	5.8
Total	409	100.0	382	100.0	413	100.0
OATS						
Fall=0 & June>0	46	11.2	35	9.2	9	2.2
Fall>0 & June=0	7	1.7	14	3.7	7	1.7
Fall = June	254	62.1	251	65.7	373	90.3
Fall ≠ June	102	24.9	82	21.5	24	5.8
Total	409	100.0	382	100.0	413	100.0
COTTON						
Fall=0 & June>0	4	1.5	4	1.6	0	0.0
Fall>0 & June=0	2	0.7	0	0.0	2	0.8
Fall = June	234	86.0	215	87.8	250	95.1
Fall ≠ June	32	11.8	26	10.6	11	4.2
Total	272	100.0	245	100.0	263	100.0
ALFALFA						
Fall=0 & June>0	9	3.3	7	2.9	2	0.8
Fall>0 & June=0	4	1.5	4	1.6	2	0.8
Fall = June	206	75.7	197	80.4	247	93.9
Fall ≠ June	53	19.5	37	15.1	12	4.6
Total	272	100.0	245	100.0	263	100.0

While tests for differences in historical data treatments are beneficial, a look at the State expansions by treatment would give an idea of the affect on the planted acreage estimates (tables 4 and 5). Table 4 presents the expanded planted acreage estimates and their precision for the small grain crops. The major findings become apparent if one observes <u>differences</u> between the Fall A&P and June reported planted acreages by treatment (last column). These differences reinforce our earlier conclusions. That is, there were large differences between the Fall A&P and June expanded totals by treatment. Providing historical data directly to the respondent (TREATMENT 3) resulted in acreages very near those given in June. Conversely, when no data was given to the respondent during the interview (TREATMENT 1), larger differences between the current and historical reports occurred. Use of historical data as a passive editing tool (TREATMENT 2) produced responses which tended be in the middle.

Table 4—Estimated planted acreage: 1985 June Acreage Survey and by historical data treatment for the Fall A&P and June Acreage surveys

Crop/ Survey	n	Treatment	Fall A&P Planted acres	Fall A&P C.V.	June Acreage Planted acres	June Acreage C.V.	Difference (000) Acres
Small grains							
June	6,422	N/A	N/A	N/A	1,211,142	4.3%	
Fall	404	1	1,072,640	14.1%	1,173,500	12.3%	-101
	377	2	1,088,683	13.8%	1,127,019	12.8%	-38
	409	3	1,214,279	16.4%	1,211,854	16.3%	+2
Other wheat							
June	6,422	N/A	N/A	N/A	547,187	6.4%	
Fall	404	1	483,858	18.7%	463,396	16.1%	+21
	377	2	402,639	16.0%	418,756	15.5%	-16
	409	3	638,053	25.2%	641,172	25.1%	-3
Oats							
June	6,422	N/A	N/A	N/A	178,563	4.6%	
Fall	404	1	151,521	10.9%	188,259	12.0%	-37
	377	2	185,347	17.0%	199,509	15.9%	-14
	409	3	166,690	14.2%	160,994	12.2%	+6
Barley							
June	6,422	N/A	N/A	N/A	383,157	6.7%	
Fall	404	1	349,269	25.6%	395,395	23.1%	-46
	377	2	426,406	30.0%	437,883	27.3%	-12
	409	3	353,102	17.1%	350,231	17.0%	+3

N/A — Not Applicable

Table 5—Estimated planted acreage: 1985 June Acreage Survey and by historical data treatment for the Fall A&P and June Acreage surveys

Crop Survey	n	Treatment	Fall A&P Planted acres	Fall A&P C.V.	June Acreage Planted acres	June Acreage C.V.	Difference (000) Acres
Corn							
June	6,422	N/A	N/A	N/A	354,970	4.4%	
Fall	269	1	218,500	17.3%	233,552	17.4%	-15
	242	2	318,686	16.9%	277,991	18.9%	+41
	261	3	194,231	22.7%	186,022	23.4%	+8
Cotton							
June	6,422	N/A	N/A	N/A	667,701	1.1%	
Fall	269	1	284,037	21.7%	301,746	21.3%	-18
	242	2	292,637	41.2%	388,362	36.0%	-96
	261	3	402,163	21.4%	419,545	21.2%	-17
Alfalfa							
June	6,422	N/A	N/A	N/A	N/C	N/C	
Fall	269	1	266,433	13.4%	286,120	14.4%	-20
	242	2	215,652	18.2%	218,042	19.2%	-2
	261	3	336,993	33.0%	329,237	33.0%	+8
Rice							
June	6,422	N/A	N/A	N/A	N/C	N/C	
Fall	269	1	37,702	37.3%	38,136	35.7%	-0.4
	242	2	12,112	59.5%	12,331	59.2%	-0.2
	261	3	25,142	65.2%	27,881	61.0%	-2.8

N/A — Not Applicable; N/C — Not Calculated

The data in table 5 for row and hay crops are less emphatic. For corn and cotton, the largest differences occur for TREATMENT 2 (+41 and -96 acres, respectively), and not for TREATMENT 1. The differences were generally smallest for TREATMENT 3. There was more variability in the statistics, due to fewer respondents planting these crops (remember, these were small grain growers in the sample), reduced sample sizes, and the possible effect of a few extreme reports.

Analyses dealing with respondent effects are limited. While complete respondent information was collected for the Fall survey, this was not the case for June. Only 214 "usable" fall respondents across all treatment groups could be linked to their respective June respondents. What one would like to show are larger mean differences and discrepant median scores for "different" compared with "same" respondents. There was some indication of this for all SMALL GRAIN acreage and LAND-IN-FARM variables when looking at mean differences. For example, the mean difference for "different" respondents was -21 acres, compared with only -11 acres for "same" respondents. This same type of discrepancy existed for land-in-farm acreage. However, this trend was not evident for OATS and OTHER WHEAT mean acreage differences nor for the median scores. There were just too few observations to make any definitive conclusions. Further examination of the respondent data in terms of their relationship to the historical data treatments or to size of operation would have been meaningless.

Discussion

The way in which historical data are used will affect a farmer's response. It generally makes the current response the same as, or very near, the historical response. "Is this a benefical or detrimental effect?" Let's discuss this issue in the perspective of the study results.

First, to understand the effect, one must first distinguish response bias and response variance. Response bias is simply the difference between the survey value and the expected value of the survey values taken over a conceptually repeatable set of identical survey conditions (that is, the "true" value). This study focuses primarily on response bias. However, in this study and in most studies of this kind, the "true" value is unknown. The results of this study can only hint at the real effect of using historical data.

Response variance, in contrast, is the "random" fluctuation about this expected value, or the variable errors in a survey. Sources of such errors would be from the interviewer who collects the data, the characteristics of the respondent and subject, and, the nature of the interview itself. Response variance, as computed in the literature (for example, O'Muircheartaigh (1986)), has not been addressed specifically in this study. However, some implications emerge from the study results.

Let's first address the issue of bias. We start with the assumption that the June planted acreage response is the "true" response. One can argue that using historical data is detrimental if it biases the current response away from this "true" value. Such situations can occur if the respondent is uninterested, or is not sufficiently motivated to think clearly about the acreage in question. Instead, the respondent says, "Yes, it is the same!" without much thought. If the original response was inaccurate, then we have done nothing more than pass on incorrect data. Many readers would argue that respondents, when given the opportunity to elicit a "same" response, will jump to the chance. Some indication of this result was found in the data. Over 90 percent of the respondents generally reported the same acreage for both surveys when historical data were directly referenced in the current interview, compared with 70-80 percent when previous survey data were not mentioned.

Conversely, using historical data is beneficial if the respondents' memory is "jogged," or their recall is improved, and they can provide a more accurate response (that is, nearer the "true" value). Some evidence for this was uncovered. That is, there was much less occurance of "forgetting" for TREATMENT 3 respondents than for respondents in the first and second treatment. 6/ Using historical data may be beneficial in this way, by helping to avoid the presence/absence of a crop from one survey to the next.

The "forgetting" effect looked to be a function of the time lag between the fall inquiry and planting/harvesting of the crop. That is, there was some evidence of memory bias. Less of a "forgetting" effect was discovered in the reporting of the row and hay crops (planted only 6-8 months earlier) than in that of small grain crops (planted the previous fall).

It would be natural to argue that the longer the time lag between initial interview and reinterview, the more important the use of historical data should be. However, creating opportunities for memory bias should be eliminated altogether. In the case of reporting small grain acreage for the Fall A&P survey, respondents were planning for and planting their current year's acreage at the time we were asking for a re-report of last year's plantings. Bailar (1968) noted increased response error as the time lag between initial and followup interview increased, and recommended that the reinterview be as close to the original interview as possible. Our findings are consistent with Bailar's. This time lag is probably far too wide to get accurate results without some method of improving recall.

The second response error issue is that of response variance. How does using historical data affect response variance? The calculation of response variance as found in the literature was not done for this report. However, it is obvious that when these data are used directly in the interview (TREATMENT 3), the variability in reporting from the June to the Fall survey decreased compared with no use of historical data (TREATMENT 1). The response variance will be smaller for TREATMENT 3 than for TREATMENT 1 respondents. Remember that 90 percent of respondents for TREATMENT 3 reported the same acreage in both time periods, compared with only 70-80 percent of TREATMENT 1 respondents.

Implications for the use of historical data are as follows. The first is the effect on total survey error (bias + sampling error + nonsampling error). It is clear that providing respondents with historical data will reduce simple response error (variability from one interview to the next), one component of nonsampling error. However, the effect on the bias term may overwhelm any gains due to reducing this response error. That is, the introduction of historical data may push the farmers response away from the "true" value. If NASS becomes interested in adding nonsampling sources of error to sampling error estimates for a total survey error approach using historical data will have some effect.

Unequivocal statements about the best approach cannot be made. It may be that some middle ground between the two extremes (no use and directly referencing) is preferable. The historical data TREATMENT 2 procedure, however, is not recommended. This procedure provided the June acreage historical data in the margins of the CATI screens for use by the interviewer. There was simply too much variability in the way in which each enumerator could use these data. Without proper controls, the number of ways in which the data are used is limited only by the number of enumerators working the survey. Indications from this project were that enumerators followed the presurvey instructions.

One approach suggested within NASS is that historical data be used as an editing tool. That is, enumerators would probe with certain defined sets of questions if the current response varied too much from the historical data.

Recommendations

Areas of NASS research and development recommended include:

(1) studying use of historical data as an editing tool, and in other data collection modes such as personal interviews and mail surveys.

(2) the development of quality control procedures to capture reinterview responses. Other quality control functions should occur simultaneously during the reinterviews to provide multipurpose uses of the reinterview data.

(3) the elimination of the TREATMENT 2 procedure as a candidate for operational use in CATI.

(6) a review of survey programs that require extensive respondent recall. Use of these data should possibly be limited.

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

Bailar, B.A., and Dalenius, Tore (1969). "Estimating the Response Variance Components of the U.S. Bureau of the Census Survey Model," Sankhya, B31, 341-360.

Bosecker, R.R. (1977). "Data Imputation Study on Oklahoma DES," Statistical Reporting Service Paper, Survey Research Section, U.S. Department of Agriculture, October.

Fellegi, I.P. (1964). "Response Variance and Its Estimation," JASA, 59, 1016-1041. Groves, R.M., and R. L. Kahn (1979). Surveys by Telephone Academic Press Inc., New York.

Groves, R.M., and L.J. Magilavy (1980). "Estimates of Interviewer Variance in Telephone Surveys," Proceeding of the American Statistical Association, Section on Survey Research Methods, 622-627.

Hansen, Morris H., William N. Hurwitz, E.S. Marks, and W.P. Maudlin (1951). "Response Errors in Surveys," JASA, 46, 147-90.

Hansen, M.H., W.N. Hurwitz, and L. Pritzker (1964). "The Estimation and Interpretation of Gross Differences and the Simple Response Variance," in C.R. Rao, ed. Contributions to Statistics Presented to Professor P.C. Mahalanobis on the Occasion of His 70th Birthday, Calcutta: Pergamon Press, Ltd., 111-136.

Hansen, M.H. and E.S. Marks (1958). "Influence of the interviewer on the accuracy of survey results," JASA, 53, 635-655.

Kish, Leslie (1962). "Studies of Interviewer Variance for Attitudinal Variables," JASA, 57, 92-155.

Koch, Gary G. (1973). "An Alternative Approach to Multivariate Response Error Models for Sample Survey Data with Applications to Estimators Involving Subclass Means," JASA, 68, 906-913.

Mahalanobis, P.C. (1946). "Recent Experiments in Statistical Sampling in the Indian Statistical Institute," JASA, 109, 325-70.

Nealon, Jack, and David Dillard (1984). "Response Comparisons Between Husbands and Wives for Farm Characteristics," Statistical Reporting Service Paper, Survey Research Secion, U.S. Department of Agriculture.

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, Reston, Va. 1985.

Warde, William D. (1986). "Examination of the Effect of the Respondent and Collectin Method on Survey Results," National Agricultural Statistics Service Paper, Survey Research Research Section.