United States Department of Agriculture

National Agricultural Statistics Service

Research and Applications Division

SRB Research Report Number SRB-91-04

March 1991

EVALUATION OF A MENU APPROACH TO CATI AND THE USE OF HISTORICAL DATA: 1987 DECEMBER AGRICULTURAL SURVEY

Vic Tolomeo
Terry O'Connor

EVALUATION OF A MENU APPROACH TO CATI AND THE USE OF HISTORICAL DATA: 1987 DECEMBER AGRICULTURAL SURVEY, by Vic Tolomeo and Terry O'Connor. Research and Applications Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250. March 1991. Research Report No. SRB 91-04.

ABSTRACT

A research study was conducted during the December 1987 Agricultural Survey to evaluate a design change in the Computer Assisted Telephone Interview (CATI) questionnaire instrument. The operational instrument is characterized as a step-by-step approach where the respondent is asked about each survey item of interest. The new design was directed at the crops section of the instrument and involved the use of historical or previous survey data and a menu type approach to question wording. Questions related to a specific crop were asked only if previous survey data indicated a positive response or if the respondent indicated, from a menu of items, that a particular crop had been planted during the crop year. The analysis showed that, although the menu approach resulted in a small but statistically significant reduction in the average crops section interview time, the overall interview time was not significantly different. No differences were found between estimated levels of corn and soybean planted acres. However, significant underreporting was indicated for dry bean and tobacco acreage. Also, the average number of crops reported per questionnaire was less with the menu approach, indicating underreporting of crop items.

ACKNOWLEDGEMENTS

The authors appreciate the efforts of the State Statistical Office staffs and enumerators who participated in this project. This project was initiated by the Survey Management Branch and the Survey Research Branch staff. The continued encouragement by George Hanuschak during large staff turnovers in the Nonsampling Errors Research Section is appreciated. We would also like to thank and acknowledge the contributions of Gene Danekas and the CATI section.

TABLE OF CONTENTS

SUMMARY

A split sample study was conducted in fourteen CATI states during the 1987 December Agricultural Survey to measure differences between two CATI instrument designs, the current operational design and a proposed menu design using previous survey data as crop indicators. The design changes involved only the crops section of the instrument. Other sections such as the operation description, grain stocks, and hog sections were not altered.

The operational design systematically directs the enumerator through a series of questions about specific crops of interest. Crop items vary from state to state depending on each state's particular data collection and estimation needs. The menu was designed to ask crop specific questions when previous survey information indicated that the sampled operation produced a particular crop. After the crop specific questions were asked, or if previous information were not available, a menu or listing of crops was presented. The menu used a global type question asking about crops grown. If the respondent indicated that a particular crop was grown, the enumerator keyed the corresponding code from the menu, and a series of questions related to that specific crop were asked. This process continued until the respondent indicated that no other crops were grown.

The proposed changes in the instrument design, the use of previous survey information and the menu approach, were analyzed as a combined effect. The separate effects of each could not be estimated. The analysis focused on differences between the operational and menu designs with respect to estimated crop acreage levels, average interview time, refusal rate, and the average number of crops reported per respondent. Results indicated that, although the average crops section interview time was reduced, the average overall interview time was not statistically different. No differences were found between estimated levels of corn and soybean planted acres, however, significant underestimation was indicated for dry bean and tobacco acreage. A comparison of the menu and operational designs with respect to the average number of crops reported per questionnaire shows that the menu approach resulted in an underreporting of the number of crops. Average refusal rates for the two designs were not statistically different.

Our recommendation is that the menu design, as described in this study, should not be considered as an alternative to the operational questionnaire design.

INTRODUCTION

The Agricultural Survey (AS) is a multiple frame survey of agricultural commodities, utilizing both list and area frames. The AS is conducted in December, March, June, and September for crops, grain stocks, and hogs. Surveys are conducted in January and July for cattle, sheep, and goats. List expansions of major crop items account for about 80 percent or more of the total survey acreage expansions at the national level. Throughout the crop year, approximately 40 percent of the list sample is overlapped from one survey to the next.

The purpose of this study was to evaluate a design change in the crops section of the December Agricultural Survey CATI questionnaire. The current or operational design systematically directs the enumerator through a series of questions about each state specific crop of interest. The study design involved a general menu approach modified to allow for direct questions if previous survey information during the crop year associated a particular crop with a sampled operation.

The proposed benefits of the new design were a reduction in respondent fatigue and a decrease in enumeration costs due to shortened interview time. Respondent fatigue was reduced since specific crop questions were asked only if the respondent indicated a particular crop was grown. Lower costs would result since shortening the interview time per respondent would allow more calls to be made during the telephone enumeration phase of the survey, reducing the workload of the field enumerators.

This study was conducted in response to a recommendation contained in the Enumerator Advisory Panel Report [1] which requested a menu driven CATI instrument, and comments during the NASS National Conference [2] which suggested shortening the length of the CATI instrument so as to minimize "respondent fatigue". The study design was based on the menu approach used in California during the 1985 June Acreage Survey. However, the California design did not incorporate previous survey information.

BACKGROUND

Much of the research in survey methodology has tended to promote standardization in question wording and question order [3]. An advantage of the CATI system is standardization of the interview process. Early in its evaluation stage, consistency in wording and question order, as well as the capability of structured probing, were cited as benefits of a CATI system which would result in improved data quality [5,9]. A study comparing the structured CATI system with non-CATI telephone interviewing indicated that larger interviewer error or variability was associated with the non-CATI interviews [4].

The type or form of the question can also affect the quality of the data collected. Results of studies in household and economic survey methodology indicate that a set of specific questions which cover the possible answers of the respondent is preferred to that of a single global type question. This may be particularly relevant to factual type data [3]. Open or global questions

ask respondents to provide multiple responses in their own words. Closed questions ask the respondent to select responses from a set of alternatives provided. In a study comparing telephone to personal interviews with respect to open-ended questions, it was shown that telephone respondents have a tendency to "truncate" their responses. The proportion of respondents reporting multiple answers on the telephone was less than in personal interviews [2]. Such studies and conclusions may not apply directly to NASS but they do indicate areas for concern and potential problems to guard against.

INSTRUMENT DESIGN

The current operational design can be characterized as a step-by-step approach where each respondent is asked about each particular crop of interest. For example, after the initial operation identification and acres operated questions, the general form of the crops section is:

> Next, I would like to talk about CROPS on this operation!
>
> How many acres of CORN are planted for all purposes?
>
> | | |
> |---|---|
> | <0> | NONE |
> | <1-99999> | ACRES |
> | <u> | NO ANSWER |

If the answer is "none", the screen will prompt the enumerator to ask about the next crop of interest. By indicating positive acreage or "No answer", the enumerator will then be prompted to ask about harvested acreage and production before the screen directs the enumerator to ask about the next crop.

Under the menu design, if previous information existed to identify an operation as growing specific crops, the enumerator was prompted to ask questions relating to those particular crops. Previous information was based on current crop year AS data collected during the previous March, June, or September surveys. For example, if a December sampled operator reported soybeans during the previous June survey, the screen immediately prompted the enumerator as follows:

> Next, I would like to talk about CROPS on the operation!
>
> How many acres of SOYBEANS are planted for all purposes?
>
> | | |
> |---|---|
> | <0> | None |
> | <1-99999> | Acres |
> | <u> | No answer |

After completing the information necessary for a positive response, and if additional crop indicators were not available, the next question to appear on the screen was:

What other crops do you have for harvest in 1987 on
all the land you operate?

<1>	Corn
<2>	Sorghum
<3>	Soybeans
.	.
.	.
<0>	No other crops

Note: Had the respondent indicated that soybeans were not grown this year, the question would have been changed from "What other crops" to "What crops".

This last question can be described as a global question. As the respondent indicated a particular crop, the enumerator keyed the corresponding code from a menu of state specific crop items. The screen then presented a series of questions related to the particular crop. After completion of these questions the enumerator returned to the menu. This continued until the respondent indicated that no other crops were grown.

During each cycle through the menu, the acreage was displayed beside the crop reported. Once the operator indicated that no additional crops were grown, the final review question appeared:

Let me review the acreage you've given me!

Reported acres= xxx
Total acres= xxx

<1> crop 1 xxx
<2> crop 2 _____

The final review screen displayed planted acreage beside each crop reported. Acreage for crops not reported were blank. The enumerator was able to make corrections to previously reported crops or include additional crops by keying the appropriate crop code. Questions regarding pastureland, any other land, and fall planted seedings for the 1988 crop year were asked immediately after the menu items had been completed. Although hay was a menu item, a specific question concerning hay harvested acres was again asked as a specific check. Appendix A contains the menu question for Texas.

THE SAMPLE

A split sample design involving the CATI portion of the AS list sample was used to test for differences between the operational and menu treatments. The AS list sample sizes and the proportion of the AS list sample completed by CATI are shown in Table 1. Treatment sample sizes, operational and menu, are also shown.

Table 1. Treatment sample sizes and the total AS list sample sizes by state for the December 1987 CATI Menu study.

	Sample Sizes			
State	Treatment-1 operational	Treatment-2 menu	CATI total	Total AS list
Alabama	510	451	961	1,200
California	523	488	1,011	2,216
Georgia	512	506	1,018	1,902
Indiana	789	779	1,568	2,688
Iowa	498	492	990	3,006
Louisiana	84	110	194	1,380
Minnesota	402	486	888	2,916
Mississippi	84	168	252	1,695
Nebraska	736	794	1,530	2,912
Ohio	701	669	1,370	2,328
Pennsylvania	324	307	631	1,379
Tennessee	275	223	498	1,710
Texas	485	497	982	5,155
Virginia	240	247	487	1,120
Washington	677	671	1,348	1,545
West Virginia	132	123	255	701
Total	6,972	7,011	13,983	33,853

Telephone enumerators were randomly assigned to either the operational or the menu instrument but not to both. Within a state, enumerators were first divided into two groups based on experience and abilities. Random assignments of enumerators to treatments were then made from within each group. This was done to minimize enumerator variability since each treatment was assigned similar "types" of enumerators with respect to skill levels.

RESULTS

Ideally, the menu approach should have provided estimates comparable to or as bias-free as the operational method but with a shorter average interview time. The following test comparisons were used to determine possible differences between the two treatments, menu versus operational: 1) mean interview time, 2) estimated crop totals, 3) refusal rates, and 4) average estimated number of crops reported per questionnaire. The changes in the instrument design were analyzed as a combined net effect. The two effects, the use of previous survey information and the menu approach, could not be analyzed separately.

Beginning and ending interview times were programmed into both the operational instrument and the menu instrument. Time, in minutes, was rounded to integer values. Estimated crop totals, refusal rate, and length of interview time were determined for both treatments allowing comparisons between the two. Appendix B shows the formulas used for the estimated totals, variances, and the statistical test or comparison of the two treatment totals.

Table 2 presents the differences between the menu and operational estimates for selected menu items. Differences were calculated as the menu value minus the operational value. With the exception of the cotton planted acres, differences between the menu and operational approach were negative. Statistically significant negative differences were detected for dry edible beans and tobacco planted acreage, indicating that the menu approach underestimated these items.

As shown in Table 3, differences between the menu and operational treatments for non-menu test items also resulted in negative differences, although none of the differences were statistically significant.

Table 4 presents estimated differences between the menu and operational approach related to the average interview time, average refusal rate, and the average number of crops reported per questionnaire. This last item measures the number of crops "captured" by questionnaire type and can be used to measure possible underreporting. Statistically significant differences were detected for the crops section average completion time and the number of crops reported per questionnaire. The average refusal rate and the average completion time for the entire questionnaire were not significantly different. The relatively small savings in the average crops section interview time did not have a measurable effect on the total interview length.

Table 2. Differences between estimated acreage totals for selected menu items.

Commodity	Difference	Percent[1/]	P-value
(acres)			
Corn planted	-576,235	-2.0	.277
Soybeans planted	-18,079	-0.1	.495
Hay harvested	-775,155	-5.0	.254
Dry edible beans planted	-239,050	-49.1	.003 *
Tobacco harvested	-42,334	-36.9	.008 *
Rice planted	-180,594	-15.0	.070
Cotton planted	1,184,532	15.3	.167

Table 3. Differences between estimated totals for non-menu items: hay harvested, winter wheat seedings and corn stocks.

Commodity	Difference	Percent[1/]	P-value
Cropland (acres)	-3,970,138	-2.8	.145
Winter wheat seedings (acres)	-326,252	-1.8	.399
Corn stocks (bushels)	-81,196,054	-3.2	.263

[1/] Percent = (menu - operational)/operational * 100

Table 4. Differences between estimated mean interview time, refusal rates, mean number of crops reported per questionnaire.

Average	Operational	Menu	Percent[1/]	P-value
Entire interview completion time (minutes)	7.0	7.1	1.8	.209
Crops section completion time (minutes)	3.1	3.0	-3.3	.039 *
Refusal rates (percent)	10.1	10.9	7.3	.283
Number of crops reported/questionnaire	1.29	1.23	-4.7	.004 *

[1/] Percentage = (menu - operational)/operational * 100

Finally, Table 5 shows the proportion of positive responses gathered by the operational and menu questionnaire designs for selected crops. There were statistically significant differences for corn planted acres and state specific commodities: sorghum, rice, and dry beans. The conclusion is that the menu design produces fewer positive reports of crop acreage.

Table 5. The percent of positive reports by treatment design for selected crops.

Treatment	Corn	Soybeans	Sorghum	Cotton	Tobacco	Rice	Dry Beans
			(percent)				
Operational	54.90	36.34	7.03	5.41	1.90	1.18	.86
Menu	52.66	36.39	6.07	5.40	1.63	.94	.49
% Diff.	-4.08	.15	-13.63	-.12	-14.11	-19.76	-43.20
P-value	.017 *	.479	.040 *	.495	.205	.026 *	.016 *

DISCUSSION AND RECOMMENDATIONS

A CATI menu instrument was tested for the crops section of the 1987 December Agricultural Survey. It was designed as an alternative to the current operational instrument with the aim of reducing the average interview time and the number of questions asked of the respondent. The goal was to reduce enumeration costs and respondent burden. It was proposed that, as a result of a shorter average interview time and corresponding higher CATI completion rate, costs would be reduced as fewer samples were completed by the more expensive field enumeration follow-up. Reducing the number of questions asked and length of the interview would also lessen respondent burden.

The menu design under study asked a global question modified by the use of previous survey data. If previously reported data indicated a positive response, then direct crop questions were asked. The evaluation was done by comparing estimates of totals and means derived from the menu design with that of the operational design for selected items. The analysis focused on the estimated differences resulting from a combination of two effects: 1) a change in question form and 2) the use of previous survey data. We were not able to provide separate estimates of each effect.

Acreage comparisons were made for seven menu items. The menu design resulted in lower acreage estimates for six of the seven items tested. Of these six items, statistically significant (p-value $< .05$) acreage differences were reported for two crop items, dry edible beans planted and tobacco harvested acres. The estimated difference for rice planted acres was nearly significant (p-value $=.07$). Differences for corn and soybean planted acres were small and not statistically significant. Differences for the remaining menu items tested were not statistically significant.

The analysis indicates that use of the menu design results in a lower average number of crops reported per questionnaire. The menu approach resulted in an average of 5 percent fewer crops reported per questionnaire compared to the operational design. The effects of underreporting, measured as the difference in the proportion of positive reports, were evident for corn, sorghum, rice, and dry beans.

A comparison of the refusal rates did not indicate a difference between the two designs. The menu approach did result in a small but statistically significant reduction in the average crops section interview time, however the average interview time for the complete interview was not significantly different. The conclusion is that, although the menu probably results in fewer questions asked of the respondent, significant savings in interview time were not observed.

Evaluations of the menu procedures were received from 13 of the 14 state telephone supervisors involved in the study. Eleven supervisors reported that experienced CATI enumerators preferred the menu design over the operational design. However, four supervisors reported that some enumerators tended to read the menu list of crops to the respondent. The original intent of the menu design was that enumerators could ask a general crops question and not ask questions for

each crop. One supervisor reported that some enumerators recorded the crops grown on a notepad in pencil, then used the menu to call up specific acreage and production questions.

The study findings and enumerator evaluations indicate a serious weakness in the tested menu design. As noted by Payne and supported by other studies, the inability to control enumerator effects is a major disadvantage of the open or global type question [7]. In a comparison of open versus closed questions, it was found that open questions asking for several responses were subject to interviewer differences [4]. The number of responses to an open type question is a function of the interviewers ability and tendency to probe for additional responses. Without standardized probing for each commodity of interest, the menu question "What crops do you grow?" resulted in fewer crops reported.

Comments from the enumerators suggest an alternative design which should be noted. A menu or listing of crop items could be used as a "check list" for complete reporting. The respondent would be asked about each crop of interest. If early in the interview the respondent replied with multiple answers (e.g., "I only grow corn and soybeans"), the enumerator would check these crops then verify zero reports for the other commodities on the menu. As mentioned, this type of enumerator behavior was observed in some states for the tested menu design. The CATI instrument programming would require that "yes" or "no" indicators be keyed for each listed commodity. Specific crop acreage and production questions would appear on the screen only after the menu "check list" was completed and if specific crops were reported. In this way, probing for specific crops would be standardized for all respondents. Additional probing questions could be programmed to verify the response when previous survey data indicated the presence of a particular crop but the current survey response was zero.

According to Sudman & Bradburn, "The simple reason for making each question as specific as possible is to make the task easier for the respondent, which, in turn, will result in more accurate reports" [8]. Our recommendation is that the menu approach, as described in this study, should not be considered as an alternative to the operational questionnaire design.

REFERENCES

1. Enumerator Advisory Panel Report (April 1987).

2. Groves, R.M. and R.L. Kahn (1979). Surveys by Telephone, New York: Academic Press.

3. Groves, R.M. (1989). Survey Errors and Survey Costs. New York: John Wiley & Sons.

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

5. House, C.C. (1984) "Computer Assisted Telephone Interviewing on the Cattle Multiple Frame Survey." SRS Staff Report Number 82, National Agricultural Statistics Service, U.S. Department of Agriculture.

6. National Conference Proceedings. Houston, TX. March 10-13, 1987.

7. Payne, S.L. (1951) The Art of Asking Questions, Princeton, N.J.: Princeton University Press.

8. Sudman, S. and Bradburn, N.M. (1988). Asking Questions, San Francisco: Jossey-Bass.

9. Tortora, R.D. (1985). "CATI in an Agricultural Statistical Agency." Journal of Official Statistics, Vol. 1, No. 3, 1985, pp. 301-314.

APPENDIX A: Crops menu question for Texas

>me48< What OTHER CROPS were grown for ANY PURPOSES during 1987?

		CROPLAND = XX
<1> CORN	IRR-	N/I-
<2> SORGHUM	IRR-	N/I-
<3> SOYBEANS	IRR-	N/I-
<4> COTTON	IRR-	N/I-
	PIMA-	
<5> PEANUTS	IRR-	N/I-
<6> HAY	IRR-	ALFALFA
<7> SUNFLOWERS	IRR-	N/I-
<8> RICE		
<9> POTATOES	IRR-	N/I-
<10> SWEETPOT	IRR-	N/I-
<11> OTHER CROPS		

<99> NO MORE OF THE LISTED CROPS

===>

APPENDIX B: Statistical test for the difference between two population totals.

Stratified Univariate Test:

$$H_o: T_m - T_o = 0 \quad vs \quad H_A: T_m - T_o \neq 0$$

$$if\ z > Z_{\frac{\alpha}{2}}\ then\ reject\ H_o$$

$$z = \frac{(\hat{X}_m - \hat{X}_o) - 0}{\sqrt{Var(\hat{X}_m - \hat{X}_o)}}$$

$$\hat{X} = \sum_{h=1}^{L} N_h \bar{x}_h \qquad \bar{x}_h = \frac{1}{n_h} \sum_{i=1}^{n_h} x_{hi}$$

$$Var(\hat{X}_m - \hat{X}_o) = Var(\hat{X}_m) + Var(\hat{X}_o)$$

$$Var(\hat{X}) = \sum_{h=1}^{L} N_h (N_h - n_h) \frac{s_h^2}{n_h} \qquad s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (x_{hi} - \bar{x}_h)^2$$

*U.S.G.P.O.:1991-281-097:40032/NASS