

EVALUATION OF THE EFFECTIVENESS OF DATA COLLECTION PROCEDURES FOR THE 1982 CENSUS OF AGRICULTURE

Donna R. Ruggles, Jane Y. Dea, Flora K. Kwok, and Cindy A. Carman, U.S. Bureau of the Census

1. INTRODUCTION

The 1982 Census of Agriculture was the 22nd nationwide census of agriculture taken in the United States. The census generally has been taken at 5-year intervals and collects data on land in farms, land use, agricultural production and sales, inventories and sales of livestock and poultry, as well as other agricultural data. These data are used by farmers, government agencies, and private organizations for making decisions, benchmarking surveys, and market research.

This census was the fourth consecutive census to be conducted using mail-out/mail-back self enumeration. This procedure, first implemented in the 1969 census, compares favorably with collection of data by personal interview as previously used and reduces the cost of data collection. The mail list for the 1982 enumeration was compiled from the previous census list, Internal Revenue Service farm tax returns, agricultural administrative records, and agricultural trade association lists.[1] Census report forms were mailed initially in late December 1982 requesting return by mid-February. Six follow-up reminders were mailed approximately at 3-week intervals. Nonrespondents of large size (expected sales of \$100,000 or more) were contacted by telephone, some with computer assistance between the fourth and tenth months of data collection.[2] Total data collection efforts achieved a census response rate of 86 percent.

Evaluation of the effectiveness of these census data collection procedures has included: 1) a mail variation test using an alternate sequence of mailed reminders, 2) a study examining characteristics of response rates, and 3) a sample survey of nonrespondents providing state-level estimates of the number of mail list nonrespondents that qualify as farms. This paper presents a description of the methodologies and results obtained from these studies.

2. DATA COLLECTION

The initial census mail follow-up was a postcard reminder sent after the mid-February due date to all nonrespondent addresses. The second and sixth follow-up mailings were census report forms with instructional materials. The remaining follow-up mailings were letters requesting response pointing out the uses of census data and reminding addressees of their legal requirement to respond to the census. In order to accommodate the large bulk of mailings from the 3.6 million record file, the file was divided into approximately eight equal segments of state groups. Just prior to the time of a follow-up mailing, name and address labels for nonrespondents were made by segment. Thus, each follow-up mailing was carried out on a flow basis within a 2-week period of the initial mail date. Table 1 provides detailed information on the follow-up mailings.

By the beginning of April, the overall response

rate to the census had reached 65 percent. This was considered adequate at the time, but a number of individual counties had much lower response rates. In order to encourage response from areas with lower response rates, it was decided to initiate a supplementary mail follow-up effort to 309,000 nonrespondents in selected states or counties with response rates lower than the national level. The selected states were South Carolina, Georgia, Alabama, Mississippi, Louisiana, Montana, New Mexico, Arizona, Utah, and Nevada; selected counties were located in North Carolina, Arkansas, Oklahoma, and Texas. A special follow-up letter was prepared which used simpler language than the other follow-up letters and offered assistance in completing the census report forms.

At the time of the second mail follow-up, a file of names and addresses of nonrespondents was established for later use in telephone follow-up. This included 417,000 potentially "large" farm operators. The "large" designation generally included operations believed to have had \$100,000 or more in total sales for 1982. The names on this telephone file were retained on the file used for mail follow-up until either a mail or telephone response was obtained. In addition, because some county response rates were still below an acceptable level (75 percent) in May, a supplementary telephone follow-up was initiated to enumerate a sample of the nonrespondents from the 252 counties with the lowest response rates. This effort was conducted by state between May and August and was considered necessary to obtain acceptable county-level response rates for published data.

TABLE 1. Follow-up Mailings for the 1982 Census of Agriculture

Follow-up	Initial Mail Date	Type of Follow-up	Response Rate at Mailout	Number Mailed (000)
Initial Mailout	Dec. 1982	Report Form and Letter	--	3,600
First	02/22/83	Reminder Card	48.4	1,900
Second	03/15/83	Report Form and Letter	57.3	1,600
Third	04/13/83	Letter	70.2	1,071
Fourth	05/12/83	Letter	75.6	890
Fifth	05/25/83	Letter	77.2	790
Sixth	06/21/83	Report Form and Letter	80.4	708

3. RESPONSE RATES AS A MEASURE OF DATA COLLECTION EFFECTIVENESS

The broad universe covered by the farm definition complicates collection of the desired data. The initial report form and accompanying letter, and the mail follow-ups may not effectively communicate to all recipients the necessity for their response, whether or not they perceive that their activities are agricultural. Thus, a number of farm and nonfarm operations are not reported. Telephone interviewers may be more effective in obtaining information leading to identification of farm or nonfarm status from nonrespondents who don't perceive that their activities are agricultural. This procedure, however, was primarily used in 1982 to obtain information from nonrespondents who were thought to have either a large or a unique farm operation.

The final 1982 Census of Agriculture data were based on 3.1 million responses from a mail list of 3.6 million names and addresses. Of these respondents, 67.4 percent were agricultural operations. There are several procedural factors that might have affected the response rate for the census of agriculture. These include differences in the response rate by type and by frequency of follow-up. Another factor thought to affect response is the length of the form. Response rates over time as well as by census mail list classification of size (measure of size derived from indicators present in mail list source records[1]) were examined to gain some insight regarding the optimum frequency of follow-up mailings.

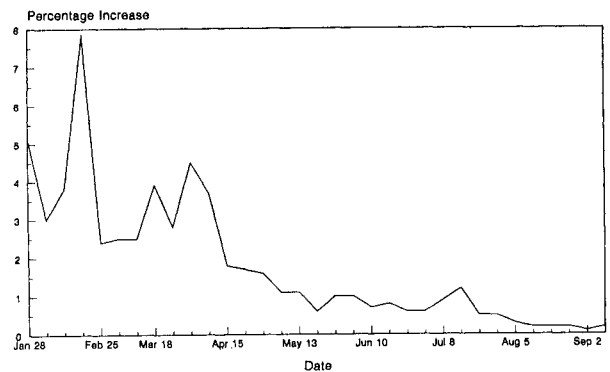
Weekly response rates for the period January 21 through September 9, 1983, were calculated as the total number of returns divided by the total number of report forms mailed out. Returns consisted of all report forms mailed back (whether completed or not), all correspondence from the report form recipients, and undeliverable report forms returned by the post office. All potential farms to whom report forms were mailed initially were classified into 16 categories based on their expected 1982 sales. These categories then were aggregated into five groups for this study. The expected sales of these groups were: A -- at least \$100,000, B-- \$10,000 to \$99,999, C -- \$1,000 to \$9,999, D -- less than \$1,000, and E -- unknown.

About 25 percent of the mail list addresses were mailed long report forms. The long report forms contained all of the questions that were on the short report forms as well as some additional questions. The recipients of long report forms came from two groups. The "certainty" group consisted of recipients whose size and source code indicated either "large" (expected sales of \$100,000 or more) or unique type of farm operation. Approximately 328,000 or 9 percent of the mail list were certainty cases. The other recipients of long forms were sampled from the remaining addresses on the mail list. These 573,000 recipients of long forms were referred to as the "noncertainty sample." The remaining mail list addresses, 2.8 million, were mailed short forms.

A cumulative national response rate of 46 percent was achieved by the February 15 due

date. The cumulative national response rate increased at the highest rate between January and the middle of April, tapered off until mid-July, and leveled off from then until the end of the data collection period. Each of the six mail follow-ups to nonrespondents was effective in increasing the response rate. Of these mail follow-ups only the second and the sixth contained a report form with the letter. The largest weekly increase occurred three weeks after the second follow-up with a 12.8 percent increase between the second and third follow-ups (Graph 1). Three weeks after the sixth followup, the response rate was higher than in the ten preceding weeks with a 2.7 percent increase. This implies that a report form may be more effective than a letter in eliciting response.

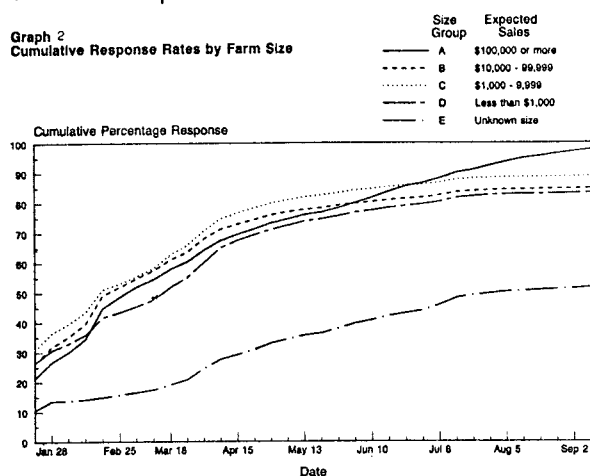
Graph 1
Weekly Increase in Cumulative National Response Rate



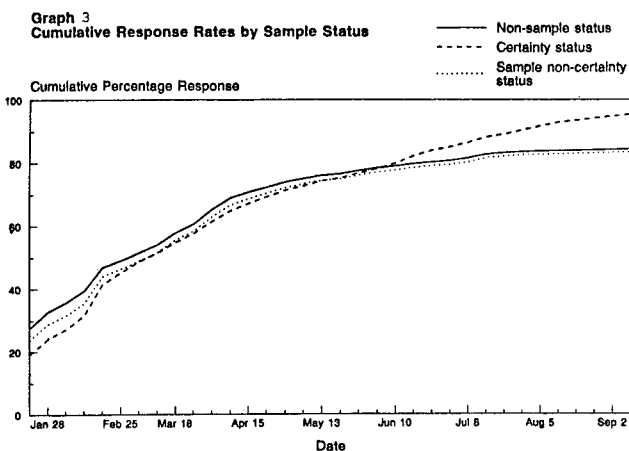
Mail Date	Follow-up	Increase in Response Rate (percent)	Cumulative Response Rate (percent)
Dec. 1982	Initial Mailout	31.4	31.4
02/22/83	First Follow-up	17.0	48.4
03/15/83	Second Follow-up	8.9	57.3
04/13/83	Third Follow-up	12.8	70.1
05/12/83	Fourth Follow-up	5.5	75.6
05/25/83	Fifth Follow-up	2.6	78.2
06/21/83	Sixth Follow-up	2.1	80.3
07/15/83		2.7	83.0
09/09/83		2.2	85.2

The response rates for all of the potential farm groups, except for those of unknown size, followed a similar pattern of increase between January and the middle of June (Graph 2). The mail follow-ups had about the same effect on each size group. The telephone follow-up to all nonrespondents in size group A (beginning in mid-May) caused the cumulative response rate for that group to increase at a faster rate than the cumulative response rates for the other groups. By the end of June, group A had the highest cumulative response rate among all groups and the response rate of A continued to increase until data collection ceased, achieving a final response rate of 97.7 percent. This relatively high response rate for group A illustrates the effectiveness of the telephone follow-up operation.

The cumulative response rates for groups B, C, and D had similar patterns of increase between the middle of June and the end of the tabulation period. The final response rates of these groups differed by only 3 percentage points. The response rate for group E was substantially lower than for those addresses for which an expected size classification was possible. The proportion of actual farm operations among respondents in this group was considerably lower than among the other groups. The low response rate for this group may reflect inadequate instructions as to who should complete the census form.



The length of the report form did not seem to have much impact on the response rates (Graph 3). However, in June telephone follow-up began for the certainty portion of the group receiving the long form increasing the response rate for that portion of the long form group. Response rates for the sample portion of the group receiving the long form and the nonsample group receiving the short form were very close. This indicates that the additional respondent burden associated with the longer report form did not have an adverse effect on response.



On the basis of this evaluation of response rates, several conclusions can be reached. The mail follow-ups in which a report form was sent along with a letter appear to be more effective in increasing the response rates than the mail follow-ups in which only a letter was sent. This idea was explicitly tested in the mail variation

test. However, the telephone follow-up to all nonrespondents whose expected sales were \$100,000 or greater was the most effective procedure for increasing the response rate for that group. The length of the report form did not appear to affect response rate to the census. The response of census recipients grouped by expected sales did not appear to be correlated with response rate.

4. MAIL VARIATION TEST

The objective of the mail variation test was to determine if there was a statistical difference in mail response between a report form follow-up and a letter follow-up. In order to test this hypothesis, a test group and a control group were selected. The procedure used for the initial mailout and first follow-up for the test group and control group was identical to that used for the census. For the second and third follow-up mailings, the test group received first a letter and then a report form, reversing the order of follow-up used in the census and for the control group.

Cost considerations limited the sample selection to 13 states and a sample size of 100,000. The states from which the sample was drawn were chosen because they were representative of two very different areas in terms of farm size. Seven of the states (Virginia, North Carolina, South Carolina, Georgia, Kentucky, Tennessee, and Alabama) were from the South where there are more small farms (in both size and product value), and for which the state response rates have been the lowest in past censuses. The remaining six states (Ohio, Indiana, Illinois, Iowa, Nebraska, and Kansas) were selected from the Midwest where the response rates have been among the highest for past censuses. An additional factor in the design of the experiment was the length of the report form -- long and short. The systematic sampling procedure selected approximately 5,000 mail list addresses per report form and group from each of the 13 states.

Up until the first time point (March 19), the test group and control group had been treated identically by receiving the initial report form and a postcard reminder. There was no significant difference in response at this point. March 19 was during the mailout of the second follow-up in which the control group received a report form and the test group received a letter. A second time point (April 23) was chosen at the end of the mailout of the third follow-up. By this time both the test group and control group had been mailed a letter and a report form, but in different order. The final test point (May 21) occurred during the fourth follow-up mailing.

A multivariate analysis of covariance model was used with the cumulative response rate at time one as the covariate and the cumulative response rate at time two and three as the dependent variables. Three factors: group (control versus test), report form (long versus short), region (South versus Midwest), and one interaction term (form by region) were represented in the model. The analysis showed that there was a group difference (no significant form or interaction difference) and a region difference. Also, the

covariate had a mean difference between regions and was a significant term in the model. Graph 4 visually suggests the results of the analysis.

The analysis of covariance model with the same terms as above also was used to analyze the data. In this model, time two (April 23) was used as the covariate and time three (May 21) was used as the dependent variable. The results were the same as above--a group and region difference and no significant form or interaction difference.

The results of the mailout variation test clearly indicate that report form follow-ups are more effective than letter follow-ups. This conclusion will be important in planning the program of mail follow-up for the 1987 Census of Agriculture. A key step toward improving this program will be to conduct a cost analysis of the more costly report form follow-ups versus the less costly letter follow-ups. An important factor in evaluating the cost-effectiveness of the report form follow-up will be the potential increase in data quality obtained from earlier response.

5. SURVEY OF NONRESPONDENTS

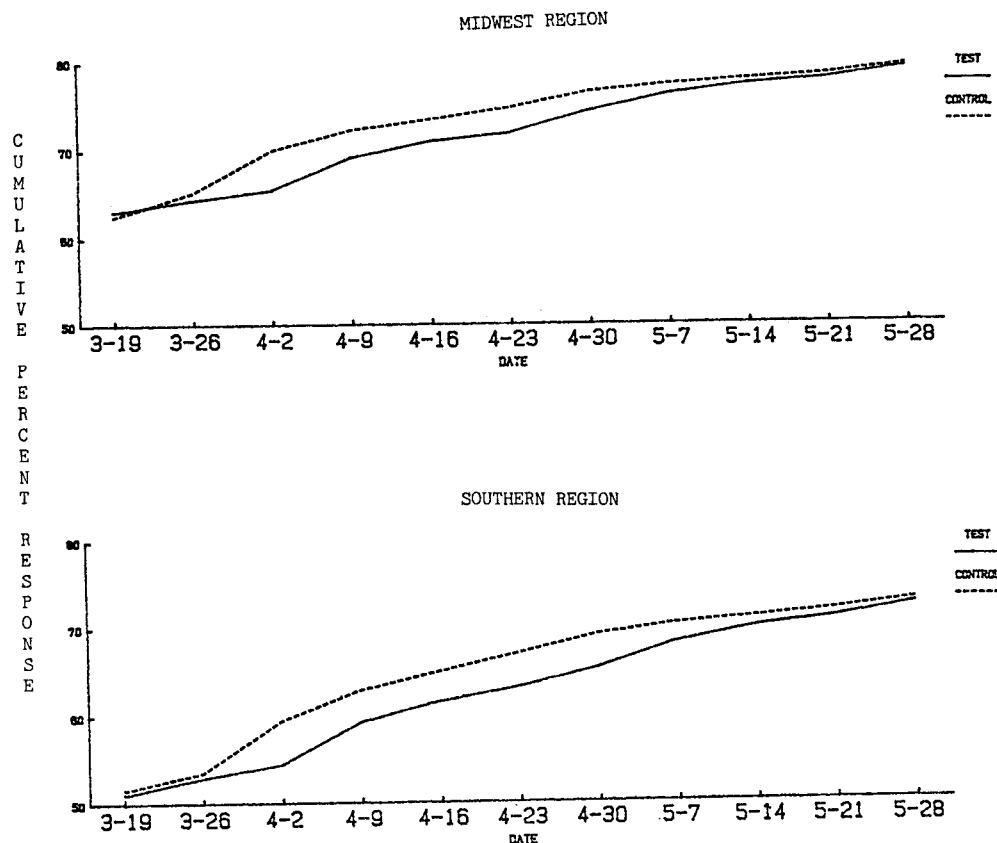
The total data collection effort, consisting of both mail and telephone follow-up, achieved a response rate of 86 percent. Continuing these efforts after the scheduled period was thought to result only in a marginal increase in the response rate. As previously indicated, only a portion of the final census nonrespondents are

actual farm operators. In order to publish data for the entire farm universe, collected data are weighted to account for nonresponding farm operators. A survey of census nonrespondents was designed to provide state-level estimates of these proportions for use in adjusting census enumerated data to represent census nonrespondent farms. The nonrespondent survey also provided some information on characteristics of nonrespondents.

The nonresponse sample was a single stage, stratified, systematic sample with selection rates varying by stratum and by state. Nonrespondents within a state were divided into six strata. Strata were based on mail size classification, administrative record source, and special handling codes. Approximately 13,000 names and addresses were selected from the April universe of census nonrespondents. Sample names and addresses were selected with equal probability from each strata. The variable selection rates used for each state were designed to estimate the number of nonrespondent census farms in each state with a relative error of about 6 percent.

The report form for the nonrespondent survey was designed to differentiate between addressees or census recipients that qualified as census farms and those that did not. The scope of this report form was similar to the census so that there would be adequate information to classify an operation. However, it was less lengthy and detailed than the census in order not to discourage response.

Graph 4
Cumulative Response by Region



The nonrespondent sample was split into four geographic groups of states according to the census processing schedule. Report forms were mailed by group at approximately 4-week intervals beginning the end of April. The report form was followed two weeks later by a second request form to survey nonrespondents. Telephone follow-up was used heavily to obtain survey response. Since data collection for the census was current with the nonrespondent survey, if a sampled nonrespondent ultimately responded to the census that nonrespondent was dropped from the sample. During the collection and processing of the nonrespondent survey, census forms were completed by 18 percent of the original sample of nonrespondents.

The nonresponse survey achieved a final response rate of 64 percent by October 1983. The remaining 36 percent either refused to provide information or could not be interviewed by mail or telephone. The relatively high proportion of nonclassified records presents a potential for biased estimation. On the basis of a small study of nonclassified records from the previous agricultural census, an assumption was made that the nonclassified records had a high degree of similarity to the classified records.

Of the 64 percent who responded to the survey and provided adequate information for classification, 42 percent represented farm operations (see Table 2). This was a substantially lower percentage of farm operations than on the census (67 percent). This difference indicates that nonfarm units choose not to respond to a greater degree than farm units. It also reflects the effect of the 100 percent telephone follow-up of "large" units which usually had a higher percentage of farm operations. This yield, however, varied considerably from state to state with a low of 17.9 percent nonrespondent farm operators in North Carolina to a high of 59.2 percent in Wisconsin.

Table 2. Nonresponse Survey Summary

	Number
Original Sample	13,489
Late Census Report Returned	<u>2,453</u>
Adjusted Sample	11,036
Classified	<u>7,057</u>
. Farm	2,941
. Nonfarm	4,116
Nonclassified	<u>3,979</u>

6. RESEARCH TO IMPROVE THE CENSUS OF AGRICULTURE RESPONSE RATE

Achieving a high response rate for the census of agriculture is a key component in improving the quality of the statistics resulting from the census. Because the 1982 census final response rate was 86 percent and the 1978 final response rate was 88 percent, a research project examining

factors affecting response to the agriculture census was begun. Thus far, the research has identified initiatives described below as potential ways in which the response rate might be increased.

6.1 Review of Focus of Census Mailings

The census mailings primarily have been directed to an audience of farm operators. However, a large number of census recipients are not or do not think of themselves as farm operators. Responses need to be encouraged from all census recipients. The focus and content of the census report form, material accompanying the form, and mail follow-ups will be reviewed to identify ways to encourage all census recipients to respond, whether they consider themselves farm operators or not. Testing of different approaches to use on the report form and on the mailings will be done in conjunction with planning for the 1987 census.

6.2 Study of Operator Characteristics By Geographic Area

Demographic and Agricultural characteristics of farm operators and operations, respectively, will be studied for selected geographic areas to determine if varying census procedures by area would be productive. Response rates differ by state and by counties within a state. Characteristics of farm operators, reported type of occupation, tenure, age, race, operations total value of product sold, and SIC code will be studied in sampled low and high response counties. If response rate is correlated with differences in the studied characteristics, publicity and procedures for follow-up can be developed to complement the specific characteristics of the low response areas.

6.3 Study of Respondent Characteristics By Mail List Record Source

In addition, characteristics of mail list sources will be studied to determine whether different procedures for follow-up by group would improve response. The yield of farm addresses and the response rate of various mail list sources differs considerably. Response characteristics of census recipients from various mail list sources will be examined to determine what differences are significant. Differential procedures will be proposed and tested for source groups.

6.4 Study of Reasons for Nonresponse

A study to examine specific reasons for nonresponse to the agriculture census is being designed. Such a study is done best in conjunction with the census in order to obtain responses that are not affected by a time lapse. A study is currently being planned in conjunction with the 1987 Census of Agriculture. In addition to examining specific reasons for agricultural census nonresponse, it will be designed to provide current estimates of the response rate and to indicate when alternative census procedures should be used. A side benefit of the study will be to provide additional census publicity for the sampled units.

The objective of the response research is to increase the response rate of all census

recipients -- both farm and nonfarm. This will reduce the proportion of the mail list for which data imputation is needed, and consequently improve the reliability of the census estimates. A higher response rate on the census will produce better quality agricultural data.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the assistance of Margaret Bruce, Eileen Gray, and Kathleen Campbell for their invaluable manuscript word processing, and a number of colleagues for their constructive comments.

REFERENCES

- [1] Dea, Jane Y., Tommy W. Gaulden, D. Dean Prochaska, "Record Linkage For the 1982 Census of Agriculture Mail List Development Using Multiple Sources," 1984 Proceedings on the Section on Survey Methods, American Statistical Association.
- [2] Ferrari, Pamela, Richard R. Storm, Francis Tolson, "Computer Assisted Telephone Interviewing During the 1982 Census of Agriculture," 1984 Proceedings on the Section on Survey Methods, American Statistical Association.