Chapter 9.

Coverage Evaluation and Research

COVERAGE EVALUATION PROGRAM

Purpose

The Bureau of the Census routinely evaluates various phases of all of its major censuses and makes available to data users information on the limitations of the statistical data. A coverage evaluation has been carried out for each census of agriculture since 1945, and the results have been published for every census since 1950. The essential methodology has remained relatively unchanged, although techniques have been refined and sample design has been improved.

The coverage evaluation program for the 1978 Census of Agriculture, like its predecessors, was designed primarily to provide—

1. National and regional measures of the accuracy of census farm counts and of selected data items, such as land in farms and value of farm sales, to aid users in the utilization and interpretation of the data.
2. Estimates indicating the characteristics of missed farms.
3. Estimates of the accuracy of the census area sample, and identification of potential problem areas.
4. Information on factors associated with census error.
5. Identification of problem areas to improve coverage in future censuses.

General Procedures

The general procedures for the 1978 coverage evaluation were as follows:

1. An area probability sample of farm operators was obtained from a post-enumeration survey (PES) of the census of agriculture area sample (CAAS) and a sample of farms identified in the 1978 Annual Housing Survey (AHS). These served as a representative basis for measuring the census universe. (A sample was employed in the evaluation because more intensive enumeration and processing techniques, which would yield higher quality results, could be used than were possible in a nationwide census.)
2. The farms in the base sample were matched to the census mailing lists and census reports to establish the relationship between the base sample units and the census.
3. Followup operations were carried out to check and clarify differences between base sample farms and census responses and to establish "true" values.
4. The results were processed, tabulated, analyzed, and published.

The various phases of the evaluation are described in more detail below.

Sample Survey Designs and Methodologies

The 1978 Annual Housing Survey (AHS)—The census of agriculture evaluation sample from the AHS was used primarily to obtain information on the number and characteristics of farms operated by individuals living in urban areas (i.e., places with populations of 2,500 or more, and thus not covered in the census of agriculture area sample (CAAS)), as well as to provide certain measures of error for misclassified farms on the mail list.

The regular AHS sample for 1978 consisted of delineated sample areas spread geographically across the 50 States. Sampling rates differed in rural and urban areas; for the former, the sample rate was approximately 1 in 650, while the rate in the latter was 1 in 1,300. The sample thus selected included about 72,000 housing units. Field interviewers visited each housing unit between October 1978 and January 1979 and interviewed the occupants or, if the unit was vacant, informed persons (i.e., landlords, rental agents, neighbors, etc.). A series of screening questions were attached to the standard AHS questionnaire for purposes of identifying households with agricultural operations. If a household had agricultural operations, the address was included in the census evaluation sample.

The CAAS Post-Enumeration Survey (PES)—The PES was an evaluation study carried out in December 1978, following the completion of the data-collection phase of the CAAS. The PES sample consisted of a 1-in-30 subsample of the approximately 6,400 area segments used in the CAAS, selected systematically across the 48 contiguous States and including 212 segments in all, each containing about 75 households.

An intensive field enumeration of the PES segments was carried out by a field staff under the supervision of the Bureau's regional offices. All members of each household in each segment were listed in the PES listing books (only the "head of the

household" had been listed in the CAAS itself), and the most knowledgeable member of the household was interviewed for the survey. In each case, the interviewee was asked, for each person in the household, the following five screening questions:

1. Did (this person) have any cattle, hogs, sheep, poultry, or horses for sale or use during 1978?
2. Did (this person) have any other livestock or animal specialties during 1978?
3. In 1978, did (this person) raise or sell any crops, such as corn or hay?
4. Did (this person) raise FOR SALE any vegetables, berries, nursery or greenhouse products?
5. Did (this person) have 20 or more fruit or nut trees?

Callbacks were made in the event that no one was home when an interviewer first contacted a household. When several return visits were made without result, the case was referred to the telephone unit.

Matching, Mailing, and Processing Operations

The principal processing operations for the coverage evaluation were as follows:

1. Receipt of PES and AHS responses.
2. Clerical match of PES and AHS sample cases, on last-name basis, to 1978 census mail list, and classification of sample cases as matches, possible matches, and nonmatches.
3. Review after matching operations.
4. Mailout of report form A90, "Evaluation of the 1978 Census of Agriculture," to all nonmatch and possible match cases, with followup of nonrespondents.
5. Attempt to match returned A90 report forms to the census mailing list.
6. Telephone followup of nonrespondents to resolve matching problems and differences.
7. Preparation of data for keying.
8. Computer edit and edit review.
9. Data tabulation and publication.

Bureau headquarters received the PES and AHS data early in 1979, and the match to the census mail file was carried out from January through April. When a positive match was made, no further search was done, but possible matches and nonmatches were subject to verification and were held in the active matching file until the entire census mail list had been checked.

The first matching operation was completed in April 1979, and the mailing list for the A90 evaluation questionnaires was prepared. The A90 requested basic data on acreage and ownership, crops harvested, livestock and poultry, location, operator characteristics (residence, race, age, etc.), and census status (i.e., whether or not a 1978 census report form had been received). In June 1979, A90 questionnaires were mailed to approximately 4,300 possible matches and nonmatches from the PES and AHS address lists. Three followup mailings to nonrespondents were made, at approximately 3-week intervals, beginning in early July 1979. By the end of September, approximately 60 percent of the report forms had been received. A telephone followup of the remaining nonrespondent cases was begun in October and was closed out 2 months later.

The second matching operation took place in February and March 1980. This was a further attempt to locate evaluation-sample farms on the census mailing list, using the additional information available from completed A90 report forms. When a match occurred, a computer printout of the census data record for that case was obtained from the census data file. The data from the census and the corresponding A90 report form were compared, particularly with respect to acreage reported and farm classification. Individual operations then were assigned a code identifying them as included, overcounted, or missed in the census. Each of these major categories had subdivisions within them relating to acreage, part of the sample, or part of the census involved.

Tabulation

Preparation of most of the coverage check data for keying, and the keying itself, were completed in February 1981. The computer program for the consistency edit was completed early in 1981, and computer edit and tabulation of the data were finished in June.

Publication

The results of the coverage evaluation program were published in early 1982, in the *1978 Census of Agriculture*, Volume 5, *Special Reports*, Part 3, *Coverage Evaluation.* The publication included tables showing estimates for the number of farms, land in farms, and total value of products sold. Since the sample was too small to provide reliable county- or State-level data, estimates were published only for regions and the United States.

PROCESSING EVALUATION SAMPLE

The purpose of the processing evaluation sample of the 1978 Census of Agriculture was to investigate the effect of the data processing operations on census data. The sample of addresses used for this evaluation study was selected from the final census mailing list immediately after completion of the second matching and unduplication operation in the fall of 1978. In order to facilitate detailed evaluation of the data, separate samples were drawn at the county, State, and national levels. Each of these samples, in turn, was divided into three subsamples. The individual records in each sample were identified during processing by a special processing-sample code printed on the mailing label.

The county-level sample consisted of all addresses on the 1978 census mailing list for six counties in six different States; the counties selected and the number of addresses from each in the sample were as follows:

County	No. of addresses
Total	16,332
Bingham Co., Idaho	2,548
Jones Co., Miss.	2,142
Madison Co., Ill.	3,295
Robeson Co., N.C.	4,431
Roosevelt Co., N.Mex.	1,583
Wayne Co., N.Y.	2,333

These counties were chosen for the sample because they were (1) geographically dispersed, (2) a reasonable cross-section of farming in the United States, (3) among those counties for which individual data records from the 1974 census had been retained, and (4) large enough to provide useful data, but small enough to be manageable.

The State-level sample was drawn from the entire census address list for Kansas. The list was stratified, based on estimated economic size codes, to permit more detailed study of individual records as follows:

Stratum	Total value of products (TVP) sold	Sampling interval	No. of addresses
Total . .			8,171
1	$200,000 or more, abnormal farms, and multiunits	4	4,737
2	$80,000–$199,999	9	977
3	$20,000–$79,999	32	973
4	Less than $20,000	53	1,484

Each stratum was systematically sampled.

The national sample was selected in much the same way, using the same stratum definitions but different random starts and intervals. The characteristics of the national sample were as follows:

Stratum	Interval	No. of addresses
Total		6,136
1	70	1,217
2	200	1,207
3	500	1,468
4	1,200	2,244

If an address selected for the national sample was found to also be on either the State or a county sample, the address was dropped from the national sample and was not replaced.

The sample cases were mailed the appropriate census report forms and were followed up in exactly the same manner as the regular census cases. Upon receipt in Jeffersonville, each report form was keyed and submitted to the processing operation. Sample cases were identifiable by a code number on their mailing labels, and this code was entered into the computer tape record of each case with the other data. Thereafter, each time a sample case was processed through the computer edit, a copy of the record was made. Since a number of cases were recycled through the edit because of errors or omissions, there were several copies of their records, each showing any changes made as a result of the processing. Once all the sample records had been processed successfully, they were combined into one file and sorted by CFN. Thereafter, separate files were produced for the Nation, Kansas, and the six sample counties, for use as in-house resource material for Bureau planning and program design.

SPECIAL COVERAGE STUDIES

Introduction

The AHS and PES samples, used principally for overall coverage evaluation, also provided a good deal of information useful in examining the specific characteristics of the agricultural operations missed, as well as of coverage obtained in the census of agriculture area sample. In late 1980, the Bureau undertook three special, small-scale studies of the data files obtained through the AHS and PES to investigate cases of (1) farms missed in the census due to misclassification as out-of-scope, (2) farms missed in the census of agriculture area sample (CAAS), and (3) farms overcounted in the CAAS.

General Procedures

The misclassification study—The misclassification study used the AHS sample as the basis for an investigation of the misclassification of agricultural operations as out-of-scope by the census. The contents of the evaluation folder maintained for each operation was carefully reviewed, including the coverage evaluation report form (A90) itself, telephone followup materials, materials from the mail list search, and the AHS supplementary questionnaire. If new information was found for any case, an additional search of the mail list was made. Analysts conducted a final review to determine the correct classification of each case and reasons for the misclassifications.

The results of the study, when weighted to provide estimates for the entire census, indicated the net "missed farms" total was 42,688, or about 1.9 percent of all farms.

Farms missed in the CAAS—This study was an attempt to determine why farms were missed in the CAAS. The 21 missed-farm cases from the PES were analyzed. When weighted, these cases represented 22,320 missed farms in the CAAS. Each case was classified with respect to the coverage by various Bureau mail-list and enumeration efforts.

The characteristics of the sample were as follows:

No. of cases	Weight No. of farms	Coverage classification
21	22,320	Total
3	4,320	Not listed in CAAS; non-match to A4 (Farm and Ranch Identification Survey)
1	690	Not listed in CAAS; match to A4
15	16,140	Listed but no agricultural operations in CAAS; non-match to census mail list; non-match to A4
2	1,170	Listed but no agricultural operations in CAAS; non-match to census mail list; match to A4 out-of-scope

A detailed review of each of the 21 sample cases was completed for this study. The CAAS and PES listing books were reviewed to verify match or nonmatch status, any notes in the books were reviewed, CAAS and PES maps were matched to determine if the household locations were the same, and all other information in the evaluation folder for each case was rechecked.

In general, the conclusions reached by the study, although based upon a relatively small number of cases, were that (1) problems often arose when the respondent for a household in the CAAS was not the operator of the farm, (2) enumerators needed further instruction about obtaining the full name of the head of the household, (3) the importance of enumerators' exploring every road or driveway to cover isolated households should be further emphasized, (4) only the latest maps should be used in a door-to-door canvass, and (5) small farm operations are those that are missed most often, frequently because the operators do not consider themselves to be running farms or ranches and are unfamiliar with the Bureau's definition of a farm.

The CAAS overcount study—Estimates made using PES data indicated that approximately 7,000 operations were overcounted in the CAAS. Overcount occurred when a CAAS farm should have been matched to the census mailing list but was incorrectly classified as a nonmatch; thus the same farm was included in both the CAAS and the regular census. The overcount study involved a careful review and analysis of the five overcount cases from the PES. When weighted, these cases represent 7,700 overcounted farms in the CAAS. All materials in their respective folders were examined and all microfilm search records were rechecked. Farm data from the CAAS and the census were also compared to determine if there was duplication.

The study of these five cases could lead to only the most tentative conclusions. Nevertheless, analysis indicated that there were three recurrent problems with these specific cases: (1) misspelled names in the CAAS and/or the census mail list, (2) different addresses for the same operation, and (3) alternative names for the sample operations.