Chapter 4.

Data Collection and Processing

INTRODUCTION

Once the planning and preparations for a census have been completed, there remains the enumeration itself, and the processing and publication of the data acquired. The data-collection phase of the 1978 Census of Agriculture involved extensive mail and telephone activities from January through November 1979. In addition, door-to-door canvass techniques were used in a sample survey of agricultural operations taken during the last 3 months of 1978.

The data-collection forms were processed and the data prepared for tabulation on a flow basis as the materials were returned to the Bureau by respondents. The processing involved (1) a clerical phase, in which the individual report forms were sorted, reviewed, edited, and the data prepared for computer processing, and (2) a computer phase, during which the Bureau's computers were used to perform certain edits of the data, impute for selected nonrespondent addresses and/or items, and tabulate the results.

DATA COLLECTION

General Information

Most of the data-collection operations associated with the 1978 Census of Agriculture were carried out by the Bureau's Jeffersonville office, with the close supervision and cooperation of the Agriculture Division. The data-collection effort comprised an initial mailout of report forms in December 1978, a subsequent mailing of reminder letters at the end of January 1979, and six followup mailings to nonrespondent addresses that began about 2 months after the initial mailout date, and were carried out at about 1-month intervals thereafter. The second, fourth, and sixth followup mailings involved sending complete sets of report forms and additional materials to nonrespondents, while the remaining three mailings were of letters requesting response, pointing out the uses of the census data, and reminding addressees of the legal requirement for response to the census.

A telephone followup operation was also used, beginning in April 1979. Selected large nonresponse cases ("must" and some "certainty" cases) were turned over to a telephone staff for enumeration, (See p. 36 for details.)

An area sample survey, designed to supplement the data collected by the mail operation, was also part of the census. Some 6,400 area segments were selected for the survey and were canvassed by the Bureau's field staff in the last 3 months of 1978. The enumerators completed an A1(A) (similar to the A1(S)) report form for each person having any agricultural operations. Respondents to the sample survey were not required to complete a mail census form, and were given identifying stickers to use on the report forms if they received materials from the initial or any followup mailing. The names and addresses of area-sample respondents were matched to the census mailing list. Those not found on the list were tabulated separately and were used to make estimates, at the State level, of the number and characteristics of farms not on the mailing list. These estimates were included in the published State totals. (The area sample survey is described in more detail below.)

Initial and Supplemental Mailouts

General information—The initial mailout for the 1978 Census of Agriculture embraced first-time mailings to any address designated to receive an A1(N) or A1(S) report form and involved over 4.2 million addresses. (The primary mailing operation for the census also included agricultural services cases, which are discussed in chapter 6.) There were also a series of supplemental mailings to cases added to the census file after the "final" address list had been compiled and mailings were underway. The components of the initial and the supplemental initial mailings were as follows:

1. The principal mailing to the addresses on the "final" census mailing address list, involving 4,240,733 report forms, released to the Postal Service in the last 2 weeks of December 1978.
2. A supplemental mailing of 132,486 report forms in early February 1979 to addresses withdrawn as problem cases from the regular address list during the final unduplication of the list. (Problem cases were identified as those without a standard name and address or ZIP code.)
3. An additional 56,414 cases mailed from February to August 1979, on a flow basis. These cases were drawn from—
   a. Special lists, especially for broilers and worms, received too late to be included on the address list.
   b. Tenant/successor adds: i.e., successors and large tenants reported on the 1978 report forms which were not included in the census address list.

c. Responses to the 78A25 "Were You Counted?" form printed in rural newspapers and farm periodicals as part of the census publicity campaign.
d. A supplement to the Hawaii mailing file, consisting of cases believed to be associated with agriculture but not included in previous mailings.

The bulk of the mailing operations were handled by Jeffersonville, although report forms for abnormal farms were sent directly from the Suitland headquarters. The mailouts were done on a flow basis with first-class postage for Alaska, Hawaii, abnormal farms, and multiunits. The remaining packages were sent by third-class bulk rate. The supplemental mailings in February 1979 were sent by third-class postage, and other miscellaneous first-time mailouts were sent first-class.

The basic characteristics of the initial December mailout for the agricultural census were as shown below.

Supplemental mailings—The supplemental list mailing packages sent to the additional special list cases were similar to the sample and nonsample packages used in the initial and February mailings, except the A7(A) first-class postage outgoing envelope was used and the A52(F) informational flyer was inserted in all packages.

Tenant/successor add cases' packages were identical to those for sample cases except that the labels carried a special list code identification number (98).

The contents of the packages for the Hawaii supplemental mailing varied considerably from that of the initial Hawaii mailout, and consisted of a BC-1266 return envelope and the report form 78-A48(S). The A48(S) form was a composite 4-page form containing a cover letter (p. 1), a short questionnaire (pp. 2-3), and excerpts from title 13, United States Code, covering authority for the conduct of the census of agriculture and the legal requirements for response.

Followup Mailings

General—The agricultural census mail file (excluding multiunits) was divided into four segments to distribute the followup and mail-receipt workloads, as follows:

Group

1 Sample forms, Alaska, Hawaii
2 Nonsample forms for geographic divisions 5 through 9 (the Southern and Western States)
3 Nonsample forms for geographic divisions 1 through 4 (New England, the Middle Atlantic, and all the North Central States)
4 Agriculture supplement (the February supplement plus special-list supplement).

All six of the followup mailings were carried out on a flow basis with groups 1-3 mailed during each of three consecutive

1978 CENSUS OF AGRICULTURE INITIAL MAILOUT

Type	Form color	Quantity[1] Initial	Quantity[1] February supplement	Form code	Other	Mail	Comments
Agriculture single units (except services)		4,232,731	132,486				
Non-sample form 78-A1 (N)	Blue	3,140,189	98,227	1N		3rd	
Sample form 78-A1 (S)							
Must (excluding Abnormals)	Green	124,908	3,871	1SI	"Must" flag present (***)	3rd	
Alaska	Yellow	1,294	60	1S2	First two digits of CFN = 94	1st	First two digits of CFN = State code
Other than Alaska	Yellow	957,740	30,111	1S2, 1S3		3rd	
Hawaii form 78-A1(H)	Blue-green	6,331	217	1H	First two digits CFN = 95	1st	First two digits of CFN = State code
Abnormal farms 78-A1(S)	"Must" green	2,269		1S1	Mail size = 14	1st	Mailed in Suitland
Multiunits		5,523 (pkgs.)				1st	
Sample forms 78-A1(S)	"Must" green	7,855		1S1	Alpha plant number and mail size = 15		
Hawaii form 78-A1(H)	Blue-green	147		1H	Alpha plant number and mail size = 15		First two digits of CFN = 95 = State code

[1] Excludes tenant/successor adds, special lists supplement, and other forms mailed March-August 1979.

weeks and group 4 a few weeks later. The closeout for each followup was usually on a Tuesday, with mailing beginning on the following Friday or Saturday. For followups that included a report form (2nd, 4th, and 6th), barcoded address labels were used, while nonbarcoded labels were used for followups involving only letters.

Reminder letter mailout—The requested response date for the 1978 Census of Agriculture was February 15, 1979. A letter was sent to all single-unit cases (excluding abnormal farms) in the initial mailouts that were still nonrespondent by the third week in January, reminding them of the upcoming due date. January 23 was the date chosen as the cutoff point for selecting nonresponse cases from the census mail file, and computer tapes listing cases still outstanding on that date were prepared for use by the label contractor to produce a set of address labels for the reminder mailing.

The mailing packages for the reminder letter mailout consisted only of the 78-A1(L2) or 78-A1(L2A) letters (the L2A was sent to nonrespondents in division 8 (the Mountain States and California)) and the outgoing envelopes; these were mechanically assembled at Jeffersonville. Mailout, on a flow basis as packages were labeled, was finished about 2 weeks after the January 23 closeout date. Closeout dates, contents of the packages, and the number of packages sent out in the "reminder" mailing were as follows:

Group	Closeout date	Form	Letter	Quantity mailed
Total . . .				3,126,717
1–3	1/23/79	78-A1 (L2) 78-A1(L2A)	78-A21 78-A21	2,762,772 276,391
4	3/20/79	78-A1 (L3)	78-A21	87,554

First followup—The first of the regular followup mailings began in late February and continued into the second week of April. Once again, the packages consisted only of letters requesting prompt response, although more specialized letters were used for nonrespondent multiunits. The characteristics of the principal followup mailing to single-unit nonrespondents were as follows:

Group	Closeout date	Letter	Outgoing envelope	Quantity mailed
Total . . .				1,842,218
1	2/20/79	78-A1(L3)	78-A21	498,264
2	2/27/79			716,987
3	3/6/79			536,501
4	4/10/79	78-A1(L4)S		90,466

The multiunit followup mailings were carried out on a flow basis during the first week of March. A total of 2,789 company packages were mailed, 2,518 to nonrespondent multiunits, and 271 to multiunits from which only a partial response had been received. The 78-A21 outgoing envelope was used for multiunit mailings as well as for single units. Two special letters were employed—the 78-A80-L1 for companies with only a single plant and the 78-A81-L1 for those with more than one plant—in the initial mailing.

A total of 671 78-A1(L3) followup letters were mailed in mid-March from the Suitland headquarters to nonrespondent abnormal farms.

Second followup—The second mail followup was the first of three that involved remailing report forms to nonrespondent cases. As was the case for the initial mailout, a private contractor printed and assembled the mailing packages for single-unit operations. The content of the individual single-unit packages was similar to that used in the initial mailout, except that new followup letters (78-A1-L4 for groups 1-3, and 78-A1-L5(S) for the group-4 addresses) were substituted for the A1(L) transmittal letters used earlier, and file copies of the report forms were included in packages for group-4 addresses only.

The quality control of the mailing packages was similar to that for the initial mailout (see chapter 3). Packages for Alaska and Hawaii, abnormal farms, multiunits, and for addresses in group 4 were mailed using first-class postage; all other packages were sent third class. As before, closeout and mailings for the followup were spread over several weeks. Immediately after each closeout date, nonrespondent addresses were selected from the mail file and a new "delinquent case" file was produced and used to print address labels. The mailout generally began within 3 days of the closeout date and continued, on a flow basis, as long as necessary. The basic characteristics of the second followup mailing, by groups, were as follows:

Group	Closeout date	Report form	Quantity mailed
Total			1,563,837
1	3/13/79	78-A1(S) 78-A1(H)	424,809
2	3/20/79	78-A1(N)	603,403
3	3/27/79	78-A1(N)	469,689
4	5/1/79	78-A1(N), 78-A1(S), and 78-A1(H)	65,936

Packages for approximately 1,900 totally or partially nonrespondent multiunit companies were mailed in the first week of April. The contents of the packages were similar to those sent in the initial mailing, except that form letters 78-A80-L2 or 78-A81-L2 were substituted for the original transmittal letters. The second followup to nonrespondent abnormal farms was carried out from Suitland in the middle of April, when 585 packages, each containing a report form (78-A1(S) "Must" (green)) and a 78-A1(L4) letter, were mailed.

Third followup—The third followup was, again, only a letter to nonrespondents. The nonrespondent address list and the mailing

labels were produced in the standard way, and mailout was completed, usually, about a week after each closeout date. Details of the mailout, by group, are given below:

Group	Closeout date	Letter	Quantity mailed
Total			1,341,741
1	4/3/79	78-A1-L5	359,473
2	4/10/79		512,924
3	4/17/79		412,957
4	5/22/79	78-A1-L6(S)	56,387

The third followup mailing to multiunits, in the first week of May, involved the use of only a single form letter, the 78-A80-L3, sent to approximately 1,600 totally or partially nonrespondent companies.

The third followup also included the establishment of a file for the telephone followup. After the closeout date for group 3 (4/17/79), approximately 43,000 addresses for large nonrespondent single-unit operations were extracted from the mail file and turned over to a telephone followup staff in Jeffersonville for enumeration. These addresses were not deleted from the mail followups until some response was obtained, so that if an operation had not been enumerated by telephone, or had not returned a completed report form by mail in time for the fourth followup cutoff date, it was included in the followup mailing. (For details of the telephone followup operation, see p. 36.)

Special April followup—By the beginning of April, the overall response rate to the census had reached 66 percent. This was considered very good, on the whole, but a significant number of individual counties had much lower rates, some as much as 10 points below the national average. In order to obtain an acceptable response rate for all counties, it was decided to mount a supplementary followup effort to nonrespondents in the 300 counties across the Nation with the lowest response rates as of the middle of April. A special followup letter was prepared, the 78-A1-L8, which used simpler language than the other transmittal letters and offered assistance in completing the census report forms. The mailing packages were assembled and labeled in Jeffersonville, and were mailed by first-class postage to 217,723 addresses during the last week of April.

Fourth followup—The fourth followup was the second that included the report forms and instructional materials. Once again, a private contractor printed all the materials and assembled the mailing packages. The packages were delivered to Jeffersonville, where they were subjected to the usual quality control procedures before being labeled for mailing. The contents of the packages were similar to the second followup, except for the use of new transmittal letters (78-A1-L6 for addresses in groups 1-3, and 78-A1-L7(S) for group 4). The mailout was as follows:

Group	Closeout date	Letter	Outgoing envelope	Quantity mailed
Total				1,053,611
1	4/24/79	78-A1(S) 78-A1(H)	78-A1(L6)	274,455
2	5/1/79	78-A1(N)		411,509
3	5/8/79	78-A1(N)		324,902
4	6/19/79	78-A1(S) 78-A1(N) 78-A1(H)	78-A1-L7(S)	42,745

Fifth followup—The fifth followup consisted of the form 78-A1-L7 letter, sent by first-class postage to approximately 855,000 nonrespondent addresses. All mailing packages were assembled and mailed from Jeffersonville. The particulars of the mailing were as follows:

Group	Closeout date	Letter	Quantity mailed
Total			854,370
1	5/22/79	78-A1-L7	222,012
2	5/29/79		338,160
3	6/5/79		257,794
4	7/10/79		36,404

The mailouts to each group were carried out in the week immediately following the closeout dates.

Sixth followup—The sixth followup included the complete package—report form, information sheet, return envelope, the 78-A2 census brochure (urging response and explaining and need for census data), and the 78-A1-L10 transmittal letter. The contractor printed all of the materials, prestuffed the mailing packages, and delivered them to Jeffersonville, where the packages were subjected to the standard quality-control procedures. The mailout was carried out on a flow basis, as follows:

Group	Closeout date	Form	Letter	Quantity mailed
Total				673,033
1	6/19/79	A1(S) A1(H)	78-A1-L10	166,836
2	6/26/79	78-A1(N)		270,192
3	7/3/79	78-A1(N)		205,931
4	8/2/79	78-A1(S), 78-A1(H) 78-A1(N)		30,074

All mailout packages were sent by first-class postage.

Tenant/successor followup mailing—The initial mailout to tenant/successor cases identified during the census processing began in February 1979. Mailings of complete packages to these cases were carried out on a flow basis by the Jeffersonville correspondence unit. Closeout for response to the initial mailings was set for June 4, at which time a list of nonrespondent cases was generated by computer and was used to produce a set of address labels. The followup package for tenant/successor nonrespondents consisted of the 78-A21 outgoing first-class envelope and the 78-A1-L6(S) followup letter. The mailing packages were prepared at Jeffersonville and were mailed to 6,922 nonrespondent addresses during the week following June 6. There was no additional followup for tenant/successor cases, although mailouts of initial census mailing packages to newly identified tenants and successors (there were only a handful of such cases) continued until the first week of August.

Hawaiian followup—The Hawaiian supplemental mailout to some 1,565 addresses not included in previous census mailings was carried out in the week of July 30. At this time, a seventh followup was made to regular Hawaiian nonrespondent cases not included in group 4. The mailing packages were the same as were used for the supplemental mailout, and were sent to the 678 nonrespondent cases in the Hawaiian file.

TELEPHONE FOLLOWUP

General Information

A telephone followup unit was established at the Bureau's Jeffersonville office to supplement the mail data-collection effort. The unit had three major functions: (1) to provide assistance to respondents calling in with questions about the census report form; (2) to verify inconsistent data reported on the forms, and to obtain missing data to resolve problem referral cases; and (3) to secure completed report forms from selected nonresponse cases. The specific cases referred to the unit included data referrals from the technical review staff, area-sample survey referrals and no-one-at-home (NOH) cases, nonrespondents in counties with low response rates, and selected nonrespondent large and multiunit operations. The unit also obtained additional information from nonmatched area-sample survey cases that might aid in matching them to the census mail list, and followed up delinquent agricultural services cases.

Telephone Followup Staff

The Bureau's Jeffersonville office included a staff and facilities to carry on various telephone operations. The telephone unit for the 1978 Census of Agriculture was established in December 1978, initially to handle incoming calls, and was expanded to carry out the telephone followup operation that was to begin in April 1979. Forty wide-area telecommunications system (WATS) lines were reserved specifically for interviews, while non-WATS lines were used for telephone-numbers research. Subject-matter specialists from the Agriculture Division and representatives of the American Telephone & Telegraph Company conducted interviewer training periodically during the period of February through April 1979.

Initially, the telephone staff consisted of only 10 supervisors, assigned in December 1978. Clerks began joining the staff in early January to handle incoming calls, and more were added as preparations for the telephone followup itself progressed. By April, the telephone staff for the agriculture census numbered 120 persons, split about equally into two shifts. The maximum staff strength of 130 was reached 3 months later (in July), but the number decreased rapidly each month thereafter, with only 11 members left by December 1979, by which time the telephone followup operation was essentially complete.

At first, the staff was divided into two shifts, each with two working units—a telephone control unit and an interview staff. However, this arrangement was soon changed to one in which a control unit was operational only during the day shift. Interviewers sometimes did telephone-numbers research on a rotational basis during each shift, but the night shift performed most of this function.

Telephone Operations

Work assignments—Two basic types of work assignments were made to the telephone unit for outgoing calls—problem referrals and nonresponse cases. Problem cases were referred to the telephone staff from the technical review and correspondence processing units beginning in early January 1979, while nonresponse cases were selected in April 1979 from the census mailing list and were, at first, limited to large farming operations. The "large" designation generally included operations believed to have had $100,000 or more in total sales for 1978 and/or had 1,000 or 5,000 acres or more of land (depending on which State was involved). In late May, the sales requirement was lowered to $40,000 in New England and $80,000 in other States, and this increased the workload of the telephone followup significantly. In September, a further increment to the workload was made when a number of nonrespondent addresses with expected sales below $80,000 in some 60 counties with response rates of less than 75 percent were turned over to the telephone unit for followup.

Subject-matter specialists from the Agriculture Division and the Jeffersonville staff were available to provide guidance and handle special problems. Members of the professional staff at Jeffersonville or from the Bureau's Suitland office supervised the telephone operation, and checked interviewers' work to verify that the data obtained were consistent and reasonable.

The control unit—Incoming cases for telephone contact were routed through the telephone control unit, which sorted the cases by State and then by group interview type (GIT), as follows:

GIT	Characteristics
1	Nonresponse; $500,000 or more in expected sales
2	Nonresponse; $100,000-$499,000 in expected sales
3	Nonresponse; less than $100,000 in expected sales, if acres exceeded 1,000
4	Nonresponse; "non-large" cases assigned by Agriculture Division

GIT	Characteristics—Con.
5	Jeffersonville "large" referral cases
6	Jeffersonville "non-large" referral cases
7	Congressional and refusal cases

All cases in GIT 1 (about 8,000 in all) were pulled from the telephone unit file and were sent to Suitland for followup by the Agriculture Division staff. The remaining cases were batched into work units of approximately 50 report forms each. (GIT 7 cases were given special handling in conformance with Agriculture Division staff instructions.) A form A82 Master Telephone Record Control Label was attached to each case or report form showing the GIT. An A410 Work Unit Control Form was then attached to each work unit and the work unit number was assigned. The work unit control number for each CFN in the unit then was posted to a master control log that included the CFN, name and address, and telephone number (if available) for each nonresponse case. This log recorded the status of each case from the time it was placed into a work unit until it left the telephone unit or was designated as "satisfied" by mail receipt or by the check-in status listing generated weekly by computer.

The telephone interviewer unit—The telephone interviewer unit performed two functions: (1) obtaining telephone numbers for nonresponse cases, and (2) carrying out telephone interviews to enumerate those cases. The interviewer unit used 40 WATS lines (with 60 telephone instruments) on a 2-shift-a-day schedule. Work units were distributed to the interview staff by WATS band (the WATS system was divided into colored "bands" designating the geographic region of the country covered by that "band") and State to ensure the most efficient use of the system. The staff was divided into four subunits, each using 15 instruments with access to the WATS lines and the Federal Telecommunications System (FTS). Initially, work units were referred to the interviewer unit for telephone numbers research, and the members of the staff checked local directories or information operators to try to obtain a telephone number. If a number was found, it was entered in the appropriate spaces on the A82 label; if none could be found, the fact was noted on the label before the work unit was returned to the control unit for assignment to individual interviewers.

Once telephone numbers research was completed, the work unit was assigned to the interviewer staff. The procedures for disposition of the various kinds of cases identified during the telephone nonrespondents followup operations were as follows:

1. Mail receipt. All cases indicated as received by mail on the weekly "alert" check-in status listing were pulled from the telephone unit file and such action was noted in the master control log.
2. In scope, completed report form. The form was completed using data obtained by telephone interview and was forwarded to the control unit, which annotated its control log accordingly and sent the form to batch for check-in.
3. Out of scope. Out-of-scope cases were noted on the control log and were referred to the coverage unit after check in.
4. Agricultural services cases. Agricultural services cases were referred to the agricultural services technical review unit.
5. Respondent claimed to have filed (Claims Filed). The interviewer attempted to obtain a CFN and the name and address on the report form supposedly completed by the respondent, and the case was referred to the microfilm mail list research team assigned to the telephone unit to verify the claim. If the claim could not be verified, the case was returned, with appropriate evidence of research, for telephone enumeration.
6. Respondent requested blank report form (Remail). If the respondent refused to give information over the telephone, but indicated the need for another form to fill out, the correct name and address were obtained and the case was referred to the correspondence unit for mailing.
7. Respondent promised to file (Will File). The date of the call and the nature of the conversation were entered on the control label and the case was placed in the suspense file in the control unit. If, after 2 weeks, the "alert" check-in status listing indicated the case was still unsatisfied, it was recycled through the telephone interviewing operation.
8. Respondent refused to give any information. The date of the refusal, the name of the person contacted, and the initials of the interviewer involved were noted on the control label. A second attempt, by a different interviewer, was made at a later date. If the subject still refused to provide the information, the details of the second call were noted on the control label and the case was checked against the microfilm mailout and current-status lists to determine if it had been otherwise resolved. If not, an analyst reviewed the case for possible special mail followup as a two-time telephone refusal. The annotated report form then was placed into the control unit's suspense file for completion by secondary sources as noted below.
9. No answer when called. After the fourth try (two attempts on each shift) the telephone number was researched again to verify that it was the correct number for the case. If it was, the control label was annotated with "N/A final try, Number Verified" and the case was placed in the work unit suspense file.
10. No listing. Cases for which no telephone numbers could be found were held in the control unit's suspense file.

Problem referral cases resolved by telephone were returned to the originating unit. Resolution often was delayed because respondents were not at home and had to be called back one or more times. Referral cases that were not resolved after several attempts at different times and days were returned to the originating unit for analyst's review.

As nonresponse cases were completed, they were sent to the check-in unit, and thereafter followed the normal processing program. Cases that had not, for whatever reason, been completed, but were known to be in scope, were usually edited by an analyst using information obtained from one or more secondary sources of data. The most important source of such data was the Agriculture Stabilization and Conservation Service (ASCS) of the USDA, which has offices in every State as well as some 2,700 county and consolidated county offices. Several

other offices of the USDA, notably the Soil Conservation Service (SCS), the Extension Service (ES), and the Farmers Home Administration (FmHA) were also major sources of data for nonrespondent farms. None of these offices or agencies of USDA has any legal obligation to provide the requested information, but since the census data are of extreme importance in formulation of USDA estimates and Government-wide agricultural policies, full cooperation was extended to the Bureau.

Results

Between January and the end of November 1979, a total of 172,217 delinquent and problem cases were referred to the telephone unit. Of this number, 29,532 were ultimately resolved by mail, 103,503 farm operators were interviewed and their report forms completed by telephone, and 32,558 more cases were completed by telephone contact with secondary sources of information. In addition, 11,447 incoming calls from respondents seeking assistance were handled (most of these occurred during the period January through May 1970).

CITRUS GROWERS

Background Information

Reports for selected citrus caretakers in Arizona, Florida, and Texas were obtained for the 1978 census by direct field enumeration. This special procedure has been used in recent censuses because of the difficulty in identifying and enumerating absentee grove owners who frequently do not have the information available that is needed to adequately complete the report form. Owners often employ "caretakers" for their groves (a citrus caretaker is an organization or individual caring for, supervising, or managing citrus groves for the owners). These caretakers are the most reliable sources of census information. Individual caretakers' operations may vary considerably: some are responsible for the entire management and care of the groves, while others perform only selected grove work; few do the harvesting.

The 1964 Census of Agriculture was the first to include a special field operation to collect data from citrus caretakers in Florida in order to improve coverage of the groves. A report form was completed for each caretaker, who was also asked to provide a list of grove owners' names and addresses and the number of acres owned by each. The names and addresses of grove owners were matched to the file of completed census report forms to eliminate possible duplication.

For the 1969 census, direct enumeration of the citrus caretakers was continued, despite the change to a mailout/mailback procedure for the general enumeration. In the 1974 census, the direct enumeration technique was extended to citrus operations in Texas, as well as in Florida. Caretakers were enumerated in May 1974 in Texas and in August and September 1974 in Florida, since these were the periods when workload in the groves was lightest and information from the bloom of 1973 would be available.

1978 Enumeration

The field enumeration of citrus caretakers for 1978 was further expanded to cover not only Florida and Texas, but Arizona as well. Staff personnel from the Bureau's Suitland headquarters carried out the enumeration in Texas in June 1978, and in Florida and Arizona in September and October 1978. For the 1969 and 1974 censuses, only the A1 agricultural questionnaire had been completed for each caretaker, but for 1978, both the 78-A1 (for citrus operations only) and the appropriate version, or versions, of the 78-A40 agricultural services report form were completed by interviewers wherever required. In order to prevent duplication of coverage, each caretaker was given a "caretaker number" and was asked to contact his or her grove owners and inform them that they should mark "citrus reported by caretaker # " on any report forms they might receive, but to be sure to supply the requested data for any other agricultural operations they might have. In order to eliminate possible duplication, the list of owners' names and addresses supplied by each caretaker was matched to the "status report list" of the regular census. Where duplicate reports were identified, the owners' citrus data were deleted from the file.

In all, 135 caretakers were enumerated in the three States covered by the special citrus enumeration effort, accounting for about 8,400 grove owners, three-quarters of whom (about 6,600) owned groves in Florida.

1978 CENSUS OF AGRICULTURE AREA SAMPLE

Background

Prior to 1969, agricultural censuses were conducted primarily through a field canvass of rural areas. In 1969, the Bureau adopted the "mailout/mailback" methodology, which, in effect, asked agricultural operators to enumerate themselves. Before each mail census, the Bureau constructed an address list of persons or businesses associated with agriculture, using as sources the administrative records of the Internal Revenue Service (IRS), the U.S. Department of Agriculture (USDA), and other Government agencies and agriculture-related associations and organizations. The source lists were combined, identifiable duplicate addresses were deleted, and the remaining names and addresses became the mailing list for the census and were sent report forms to be completed and returned.

The Bureau has routinely conducted coverage evaluations of the census of agriculture since 1945. These have indicated that neither the field enumeration nor the mail census technique has been able to attain complete coverage of agricultural operations. The percentage estimates of undercoverage in agriculture censuses over the past 25 years are as follows:

Item	1954	1959	1964	1969	1974
Number of farms . .	8.1	8.4	11.3	15.0	[1] 10.7
Land in farms	5.4	6.0	6.1	9.1	7.4
Value of products. .	(NA)	(NA)	2.9	[2] 3.3	[2] 2.9

NA Not available.

[1] The farm definition was changed in 1974. The estimated net undercoverage of number of farms by the 1959-1969 definition was 14.3 percent.

[2] Estimated value of products for missed farms only.

A primary goal in the planning for every census has been to improve coverage. When it became apparent that coverage in the 1974 census was little better than obtained for 1969, the Bureau proposed to supplement the 1978 mailing list with a direct-enumeration area-sample survey, large enough to produce reliable estimates for States, for farms not on the mailing list.

Results from the 1970 Census of Population and Housing indicated that approximately 75 percent of all households were in urban areas (primarily places with 2,500 or more population) but only 7 percent of persons classified as farmers and farm managers lived in these areas. Thus, a sample of rural areas would exclude most households while including 93 percent of farmers and farm managers. This idea was further substantiated by the 1974 Census of Agriculture, which indicated that about 80 percent of all farmers lived on the farm operated and another 9 percent lived off the farm operated but in a rural area. It was decided, therefore, to select the area sample only from rural areas, and to use a supplement to the 1978 Annual Housing Survey to provide an estimate of farm operators living in urban areas.

Sample Selection

Sample design and sampling unit—The design used for the 1978 sample survey was a stratified one-stage area-segment sample. The sample unit was a defined geographic area of land, the area segment, which could vary in geographic size and in the total number of housing units and/or farms it contained, depending on the stratum to which it was assigned. Within each stratum, the area segments were so drawn as to have approximately equal numbers of farms. A farm operator and the associated farm (or farms) would be uniquely identified within a single area segment.

Stratification and sample selection—The sampling frame for the area-sample survey in each State was the list of all enumeration districts (ED's) and block groups (equivalent of ED's in city areas where data were tabulated by block) from the 1970 Census of Population and Housing. This list contained ED identification and geographic and administrative codes from the 1970 census, together with data on population, housing units, number of farmers and farm managers, and number of farm laborers and foremen as reported in the 1970 census. These data were used to estimate the number of farms (the larger of the counts of farmers and farm managers or rural farm housing units) and to calculate farm density (the ratio of farms to housing units) in each ED. ED's then were assigned to one of six strata, based on the estimated farm density. The number of farms and housing units assigned to any area segment was also a function of farm density in the ED. The six strata are defined below:

Stratum	Estimated farm density	Segment size	
		Desired No. of farms	Maximum No. of housing units
1	.10 and above	12	120
2	.05 to .10	10	200
3	.02 to .05	5	250
4	.01 to .02	2	200
5	.005 to .01	1	200
6	Less than .005	0	150

Based on these data, a number of segments were assigned to each ED. Once stratified, and prior to sample selection, the ED's were sorted by 1978 Census of Agriculture county code and by 1970 census tract and ED. This sort gave a sequence of ED's in approximate geographic order. In theory, the sample selection was a one-stage process, but in practice, two steps were used. Desired sample size for each stratum in the State and the total number of segments in the stratum universe determined the sampling interval, which varied among States and strata.

The geographic sorting of ED's and the systematic sampling from the cumulative number of segments insured that the sample drawn for each State in each stratum was distributed fairly uniformly throughout the State.

Once the segments were selected, the actual geographic areas were identified using enumerator maps from the 1970 decennial census that showed the 1970 location of housing units. This information was used to divide each ED into the previously designated number of segments, each containing approximately equal numbers of housing units, and with recognizable physical or political boundaries (i.e., a road, river, city limit, etc.). The segments in each ED then were numbered consecutively in a serpentine fashion, beginning in the northeast corner of each ED. After this, the segments carrying the numbers previously selected for the sample were identified and enlarged maps of each segment selected were prepared for use by the enumerators.

Data Collection

Field procedures test—The Bureau expected the area-sample survey to involve 6,393 area segments encompassing an estimated 450,000 households, of which 60,000 to 70,000 were expected to include agricultural operations. While preparations for the main sample enumeration continued, a test of proposed field enumeration procedures was carried out in the spring of 1978. Twenty area segments in Colorado, Iowa, South Carolina, and Texas were selected for the test and were canvassed by personnel from the Bureau's regional offices in the first week of April. No serious problems in the procedures were noted.

Enumeration staff and training—The area-sample survey itself was supervised by the regional offices (RO's), but the canvassing of the 6,393 area segments across the country required a fairly large enumeration staff—one considerably larger than the Bureau's regular interviewer staff. While supervisors were drawn from the RO's, most enumerators were recruited specifically for this survey. Two training sessions for the approximately 230 crew leaders, a 1-day session for quality control prelist training, followed later by a 4½-day session, were carried out by the RO staffs in mid-September. The crew leaders, in their turn, trained some 2,000 enumerators during the first week of October.

Enumeration materials and procedures—Each enumerator was issued a kit that included a form 78-A10 *Enumerator's Reference Manual*, a map of the segments he or she was to canvass (each map showed all known housing units), a map of the county in which the segments were located, a form 78-A3

Record Book, and a supply of form 78-A1(A) area sample questionnaires. (The form 78-A1(A) was virtually identical in content to the 78A1(S), except that additional identification data were requested on page 1.)

Enumerators were to systematically canvass their segments, visiting every housing unit in each segment, making corrections to their segment map when necessary, and asking the head of each household, or some other responsible member of the household (if the head of the household was unavailable), a set of screening questions from the A3 *Record Book* to determine whether any agricultural operations were being carried on by any member of the household. An entry was made for each household visited, with the name and address of each head of household as well as the name and address of anyone else there who had any agricultural operations. If any agricultural operations were being carried on, the enumerator noted that in the A3 *Record Book*, and completed a form 78A1(A) questionnaire for each operation. Each operator for whom a report form was completed was given a form 78-A15 "I Have Been Counted" sticker, with instructions to apply the sticker to any census report form received as part of the regular census mailout and to return the unfilled report to the Bureau. The sticker identified area-sample cases and enabled the names and addresses of respondents to be matched to the census mail file, so that respondents to the area-sample survey were not required to complete the mailed census form.

The enumeration—The canvassing of the area segments began on October 6 and continued into December. When all segments had been completed, 600,000 households and nearly 100,000 agricultural operations had been canvassed.

Quality control coverage check—The area-sample enumeration included a quality control program aimed primarily at insuring complete coverage. Crew leaders for the enumeration staff prelisted 15 dwelling units from the first area segment in each enumerator's assignment. These 15 units were selected by picking three starting points at random throughout the segment, then listing the five consecutive housing units following each starting point. After each enumerator completed his or her first area-segment listing, the listing book was compared to these advance listings to check the enumerator's canvass.

In all, 1,682 area segments were checked in this fashion. Based on the match of prelisted and canvassed addresses, the overall "miss" rate was estimated to be 3.8 percent. Budget constraints made it impossible to revisit any of the poorly canvassed area segments, so, in effect, the purpose of the check was primarily to let enumerators know their work was being checked and enable crew leaders to identify problems that the enumerators were having in the field.

Processing

General—The processing of the report forms for the 1978 Census of Agriculture Area Sample Survey included the following operations:

1. Receipt, check-in, and filing of enumeration materials
2. Matching of area-sample report forms to the census mailing list
3. Followup and imputation for nonresponse
4. Clerical review, geocoding, and keying
5. Tabulation and publication

The initial review of the report forms from each segment was carried out in the field by supervisory personnel. The forms then were forwarded to the Bureau's Jeffersonville facility for processing. While much of this, particularly the keying, computer editing, and tabulation, was similar to that carried out for the census, some phases were significantly different. Those points at which the processing of the area-sample materials varied markedly from the regular processing cycle are discussed below.

Receipt and check-in—All the enumeration materials for each area segment—A3 record books, segment maps, completed report forms—were submitted for receipt and check-in of individual segment packages. The clerical staff at Jeffersonville opened the packages as they arrived, checked the contents against the enclosed transmittal forms, and entered the date of receipt for each segment's materials on a listing of the segments in the sample. The segment materials then were cycled through a review procedure. This included the assignment of an area-sample CFN for each report form and the computation of a check digit for each CFN. Information from the A3 record book, which clarified any part of the A1(A) report forms for an operation, was transcribed, and page 1 of the A1(A) report form, containing all of the necessary identification information, was screened to insure completeness. Incomplete A1(A)'s were referred to the technical review unit, while complete report forms were sent to the microfilm search unit.

Matching—Since one of the major objectives of the area sample operation was to estimate the number and characteristics of farms not on the mail list, each completed report form was matched to the mail list and classified as "matched" or "not matched." Respondents whose report forms were matched to the mail list were deleted from the followup mail file (processing of the area sample began too late to prevent inclusion of such cases in the initial mailout). The report forms for cases matched to the mail list then were inserted into the regular mail-list processing operation and treated as regular respondents. To ensure that all possible cases were matched, a second attempt was made to match the unmatched cases. In addition, telephone calls were made to all of the larger farms to obtain additional information that might help in the matching process. As a result of these additional checks, most of the larger farms were found to be on the mail file and thus classified as matched cases.

Clerical review, geocoding, and keying—After matching to the census mail list, each report form was assigned tabulation codes for the appropriate State and county. Non-match cases were assigned to the proper State and were given a code number for a "pseudo county" for tabulation purposes to provide State-level estimates. Since the nonmatched cases in each State were used to estimate data only at the State level, the "pseudo

county" totals would provide estimates of the number and characteristics of farms in each State that had not been represented on the census mailing list.

After geocoding, all of the report forms were clerically reviewed, and were keyed to magnetic computer tape for computer processing and tabulation.

Computer processing, tabulation, and publication—The computer records for individual operations enumerated in the area sample were subjected to essentially the same computer editing and tabulation procedures as were used for mail respondent operations. The most significant difference in procedures was the separate tabulation for "non-match" farms. These estimates and the mail-list county estimates were combined to give the State-level estimates.

For preliminary State publications, data were shown for the area sample alone and combined with estimates from the mailing list for State totals. In final State tables, totals included the area-sample data. County summary tables for each State included each actual county, plus data for the "pseudo" county enumerated in the area sample listed as "Farms not on mail list."

PRECOMPUTER PROCESSING

General Information

The precomputer processing phase of the census program had four primary objectives: (1) check-in of respondent report forms, (2) coverage check for acreage, (3) resolution of any problems with the completed forms and of correspondence from respondents, and (4) creation of data tapes (data entry) for computer processing and tabulation.

Respondents mailed their report forms to the Bureau's Jeffersonville office, where the precomputer processing of the census materials was done. Except for barcode check-in, automated sorting of the forms, and limited electronic processing done during data keying, all computer processing operations were carried out at Suitland.

The precomputer processing staff's work included the receipt and check-in of the report forms, resolution of correspondence, routing of report forms and other mail to the appropriate processing unit, screening of the report forms, resolution of problem referrals, "2+" (two or more forms received) processing, telephone followup, data keying, and, after computer editing, review of the computer changes and corrections. In all, approximately 3.98 million individual cases were processed by the Jeffersonville operation, 2.26 million of which represented in-scope farms.

Receipt and Check-in

Batch for check-in—Correspondence, report forms, and postmaster returns (PMR's) were separated upon initial receipt and batched according to the most suitable type of check-in, i.e., barcode or keying.

Receipts with barcodes were sorted by type of form and receipt (i.e., whether completed forms or PMR's), and batched into work units (WU's) of 200 report forms of one type (A1(S), A1(N), A40, etc.), or 200 PMR's, each. As each WU was assembled, a check-in status code was assigned using an A402 Check-In Work Unit Cover Sheet. These codes were as follows:

2X	Postmaster return
30	Respondent-originated correspondence
40	Form received
50	Out of scope
51	Out of scope—Requests a copy of farmer's report
7	Clerical remail
8	Computer remail
9	Out-of-scope recycle

This coding system was designed so that a higher number superseded a lower one. Many cases, especially those involving PMR's and correspondence, were checked in more than once. These cases were considered incomplete, even if something had been received from the respondent, until a completed report form was received or the case was determined to be out of scope. Two types of barcode-reading equipment were used, a laser reader and a hand-held ruby wand pen. Bulkier packages and report forms that had been removed from their mailing envelopes because of obscured labels were checked in using one of eight wands, which could accommodate packages of varying dimensions; standard mail receipts were read by the mechanical laser reader, which required envelopes of uniform size. In either case, the work unit information was keyed for the batch and the barcoded CFN's were read and stored on tape for updating on the master address file. This file was used to follow up nonrespondents at selected intervals.

Once reassembled after barcode check-in, placed in their plastic bags, and with the cover sheets attached, the individual work units were routed to the next step in the processing program. The disposition of the WU's by type, was as follows:

Type of receipt	Disposition after barcode check-in
A1(S), A1(N) and A1(H) with and without correspondence	State sort
1st time PMR's without address corrections	Central files
1st time PMR's with address corrections, and 2nd time PMR refusals	Correspondence reading
2nd and 3rd time "must" PMR's	Coverage unit
2nd and 3rd time PMR's	Correspondence typing
PMR's with area-sample sticker	Area Sample Survey clerical unit
Respondent-originated correspondence	Correspondence reading

Check-in keying—Completed report forms and PMR's that, for whatever reason, could not be checked in using the barcode readers were sorted for clerical handling into two groups, those with CFN's present, and those lacking CFN's. Materials with no CFN's were routed to the CFN researching unit, while those with CFN's were sent to the check-in keying unit. The latter were batched into work units using the same coding system

employed for barcode check-in. The CFN for each report form, piece of correspondence, and/or PMR, as well as any address changes noted, were recorded on the Bureau's data entry (key-to-disk) system, with output to computer tape. The data keying was subjected to verification and quality control procedures, and the resultant records were transmitted to Suitland via telephone datalink for updating the census master address file. Report forms rejected during keying because of faulty CFN's were removed from the work units, corrected, and recycled through keying.

After check-in keying, work units were routed to the appropriate processing unit.

State sort—After check-in, work units containing census reports were sorted by State, according to the geographic location of the operation. Sorting was done both manually and mechanically, but the majority of reports were sorted mechanically using a machine similar to the electronic check-in reader. This device contained 24 "sort" pockets and could be electronically controlled to use any pocket or group of pockets to hold selected items identified by the laser barcode reader. Sorting down to the State level required two passes: one pass sorted on geographic division code and the second by State code within division.

Reports that could not be sorted mechanically (i.e., reports without bar codes or on which the bar code could not be read) were sorted manually using the same digits described for mechanical sorting. Using a series of nine boxes, the manual sort was according to the first digit of the CFN—the division code. After the initial division sort, reports were removed from the bases and sorted on the second digit, one division at a time.

After sorting, most reports were routed to the screening or technical review unit, where they were placed in movable storage bins, by State, and held until scheduled for further processing. Reports from Alaska and Hawaii, however, were sent through final control and forwarded to Agriculture Division in Suitland for special handling.

Census File Number Research

The census file number (CFN) was the principal numeric identifier for each report form or case received and/or processed by the agricultural census operation, hence it was imperative that each case received have a CFN. Whenever a report form or piece of correspondence was received that either did not have a CFN or the CFN present was incomplete or wholly or partially obliterated, the case was referred for resolution to a special CFN research unit in Jeffersonville. This unit used 16mm microfilm reading and printing equipment and two microfilm files: (1) a complete census universe ZIP/name control file, and (2) a State/name control file for each State in which the name control (i.e., the first four characters of individual surnames (or in the case of partnership or other arrangements, the first surname), company name, association name, etc.) was used to sort and list alphabetically the complete name and address for each case originally mailed.

The ZIP/name control file was used if a ZIP code was present in the address of a referred case, otherwise the State/name control file was used. Since there were a number of names (such as Smith, Johnson, Green, etc.) that had the same name control, each such entry was reviewed and the complete addresses displayed for a comparison match to the record being researched. To be considered a match, the name on the correspondence or report had to be identical (i.e., contain the same first name, middle initial(s) (if any), and last name) to the microfilm equivalent, and the city, State, and other address elements had to be identical or very similar to that shown on the microfilm.

If a CFN was found for a case, it was transcribed to the appropriate space on the form or to the upper right-hand corner of the correspondence. A copy of the CFN entry was made from the microfilm and attached to the correspondence or report form, and the materials were referred to batching for check-in. If no CFN was located, the document was annotated "CFN NOT FOUND." If it could not be determined whether a case had been matched to the census files (e.g., because of incomplete address or name on the correspondence or report form), copies of possible CFN entries were made and the case was referred to the research unit supervisor for disposition. Area-sample materials then were referred to the area-sample processing unit, while all other materials were sent to the correspondence reading unit.

A total of 28,150 pieces of correspondence and 15,485 report forms were processed by this unit. Of these, CFN's were found for 22,064 pieces of correspondence and 11,597 report forms.

Clerical Screening and Review

General Information—The clerical screening and review unit was established in order to identify reporting errors on the 78-A1(N) and 78-A1(S) report forms that affected "keyability" and to make the necessary corrections. The unit also extracted from the regular processing cycle those forms that required special review and handling. Materials were received by the screening unit primarily from the State-sort unit. All forms with attached correspondence were screened immediately upon receipt, regardless of any State priorities that were in effect at the moment, so that congressional cases could be quickly identified and work could begin without delay on cases requiring additional respondent contact.

If the remarks or responses on any report form indicated a need for a form letter (e.g., a report form had been returned to the Bureau completely blank, or the respondent requested confirmation that his or her report form had been received), the screening clerk indicated the appropriate form letter to be used and forwarded the case to the correspondence typing subunit or, in the case of special problems, to a correspondence analyst.

Report forms without correspondence attached were clerically screened on a flow basis in State-priority order. Screening involved deleting fractions or converting them to decimal numbers, lining through extraneous material, verifying key-code assignment, and ensuring the readability of the entries to be keyed.

Screening and correction—Condition-action tables provided instructions to screening clerks for the review of the report forms. The general intent of the review was to make certain that—

1. Report forms with no agricultural operations were identified.
2. If any remarks were present that required action by the Bureau, the action was specified.
3. Remarks containing data were transcribed to appropriate data cells or were referred to technical analysts.
4. The address label contained no unprocessed changes.
5. The State code on the address label was consistent with the rest of the work unit.
6. The geographic coding on the address label was consistent with the reported State and county geographic locations.
7. Key codes were entered and/or corrected for the crop/livestock names.
8. Report forms with entries indicating total acreage was zero were referred to technical review to ascertain whether the addressee was a landlord only and, hence, out of scope.
9. All cases enumerated in the area sample were referred to the area-sample processing unit.
10. Data entries outside prescribed locations on the report form were transcribed to the proper ones.
11. Entries obscured or illegible were either deleted (if in "Quantity Harvested" or if a total for a section with detail entries) or were referred to technical analysts.
12. Alpha (i.e., "spelled out") data entries were converted to numeric entries (e.g., "five" was converted to "5," etc.).

Verification and quality control—Each work unit submitted to the screening unit had to pass the verification process. Errors detected during verification were corrected and tallied, and supervisors and clerks kept informed of the number and types of errors detected. The report forms so subjected were reviewed to make certain required referrals, data entries or transcriptions, and necessary changes in geographic area codes (GAC's) were properly made.

The verification process actually began during the training of the screening clerks. During their qualification period, the screening of the first 200 report forms by each clerk was verified on a 100-percent basis. If 4 percent or fewer of these forms contained critical screening errors, the clerk was considered qualified and subsequent work was verified on a sample basis.[1] Records of errors and corrections required were kept on each screening clerk and periodically discussed between supervisor and clerk. When an individual clerk's error rate exceeded 4 percent, additional work was verified. If a 100-percent verification of a run of 400 report forms screened by a clerk revealed an unacceptable error rate, the clerk was retrained. If, after retraining, the clerk's work still did not meet acceptable standards of error, he or she was removed.

After qualification, each clerk's work was sampled for verification at a 10-percent rate. In addition, every report form on which a geographic code change had been marked on the label had the change verified as well. Each work unit was accepted or rejected based on the number of errors found in the sample. Accepted work units were released for further processing while rejected ones had to be completely corrected. To remain on sample verification, each clerk had to have at least 8 "accept" decisions in each sequence of 10 decisions.

Correspondence

All census-related correspondence was handled by a correspondence unit established at Jeffersonville. The unit was itself divided into three subunits, dealing with (1) reading correspondence and responding to routine cases; (2) typing of address labels, letters, envelopes, etc., and handling referrals from other units; and (3) filing and followup of post-edit correspondence (PEC—i.e., replies from respondents to Bureau-originated correspondence requesting more information, etc.).

The correspondence reading subunit sorted incoming materials into those cases in which (1) the respondent claimed he or she had filed, but made no reference to a CFN; (2) the respondent requested a report form; and (3) all others. The first two groups of materials were referred to the batch unit for check-in keying in order to generate labels for mailing; all the other cases were cycled through the reading subunit, where the correspondence was read and the appropriate action decided upon.

The appropriate action generally consisted of preparing and mailing the applicable form letters and/or report forms. The function was usually performed by the typing subunit. The filing and PEC followup subunit was primarily responsible for handling replies and followup to PEC, and file maintenance.

The work of the correspondence unit was subjected to verification and quality-control measures before any materials were filed or released. The verification program in the reading unit began with an initial training period for each correspondence clerk, during which the first 100 pieces of correspondence processed were checked for errors on a 100-percent basis. Once the initial training period was completed, each clerk's work was verified on a sample basis. When an error was detected during the sample verification phase, 100-percent verification was begun once again and continued until 50 successive pieces of correspondence were found error-free, before returning to sample verification.

Verification of the typing subunit's work was designed to insure an error rate of no more than 3 percent. All of the correspondence processed by each clerk during the first 5 working days on the job was verified at a 100-percent rate. If the error rate for all the work involved was 3 percent or less, the clerk was considered qualified for sample verification. (If the error rate exceeded 3 percent, the clerk was kept on 100-percent verification for another 5 working days. If the error rate still exceeded 3 percent, the clerk was removed from the typing

[1] "Critical" errors were generally those involving failure to refer a form when it was necessary or to accurately transcribe data or complete necessary screening steps. "Noncritical" errors usually involved unnecessary actions, incorrect designation of reason for referral, etc. Noncritical errors were tallied and brought to the attention of the clerks, as their continued repetition could increase operational costs.

operation. If the rate was 3 percent or under, he or she was moved to sample verification.) Clerks qualifying for sample verification had each day's work verified at a 1-in-12 rate, beginning at a random start. If the error rate for any day's work exceeded 3 percent, the work lot was rejected and *all* the correspondence in that day's work lot were verified and errors corrected. If any clerk had more than one lot rejected in a week, the clerk was returned to 100-percent verification for requalification. If the error rate at the end of the 1-week requalification period exceeded 3 percent, the clerk was removed from the typing operation.

Some correspondence was referred to Bureau headquarters for handling. This included all congressional and potential congressional correspondence (i.e., any item that indicated the respondent was sending a copy of the letter to a Member of Congress), complex problems involving multiunits, and unusual or difficult situations that could not be resolved using routine form letters.

During later processing phases, the correspondence unit prepared letters to be sent to obtain additional information needed to edit or complete the report forms. These letters were also subject to quality control procedures. The principal form letters used by the unit in its day-to-day work are listed below:

Form No. 78-	Purpose
A101(L), A102(L), A103(L)	Reply to request for extension of time for completing report form
A104(L)	Grant time extension
A105(L)	Request completed replacement report from correspondent who claims to have filed, but whose form cannot be found
A106(L)	Request additional information (report form enclosed with items indicated)
A107(L)	Respond to request for legal authority for census; excerpts of title 13 on back page
A108(L)	After review of correspondence, advise that report form is not necessary
A109(L), A122(L)	Recommend sources of assistance in completing report form
A110(L)	Inform respondent the Bureau is unable to excuse him or her from completing report form and explain need for census
A111(L)	Notify that correct report form is enclosed and request for prompt response
A112(L)	Determine whether respondent had agricultural operations in 1978
A114(L)	After review of correspondence and report form, advise that additional information not necessary
A115(L)	Reply to refusal cases, justifying collection of census data and assuring confidentiality
A116(L)	Return report form for completion—form blank or nearly blank
A117(L)	Respond to request for payment for completing report form (no payment authorized under census law)
A118(L)	Indicate request for published data will be filled
A119(L)	Return noncensus materials included with report form
A120(L)	Advise additional materials sent to respondent, as requested
A121(L)	Acknowledge receipt of report form
A123(L)	Request census file number for response concerning respondent's report form
A124(L)	Request census file number
A125(L)	Explain need for census
A126(L)	Brief description of sources used in compilation of the census mailing list.
A127(L)	Furnish additional report form when original not received
A128(L)	Original addressee deceased, request executor provide information
A129(L)	Acknowledge receipt of report form
A130(L)	Original addressee deceased, request respondent provide additional information
A135(L)	Post-edit correspondence for deceased operator/addressee cases. Successor's information needed.
A136(L)	Reply to respondent who no longer had agricultural operations

Postmaster returns (PMR's)—The correspondence unit was also responsible for handling PMR's and Postal Service notifications of address corrections. Materials were referred from check-in to the correspondence unit on a flow basis, presorted as follows: (1) A1 first-time PMR's with address changes or "deceased" indicated[2]; (2) A40 first-time PMR's with address change or "deceased" indicated[2]; (3) follow-up letter PMR's with address changes or "deceased" indicated; (4) A1 second-time refusal PMR's (non-must); (5) A1 second- and third-time PMR's (non-"must")[3]; and (6) Postal Service notifications of address corrections.

The typing subunit prepared new mailing labels for remail to all cases where address changes or "deceased" were indicated,

[2]All first-time PMR's without address changes or "deceased" indicated were sent to central files after check-in. Mailing labels for these cases were generated by computer as a result of the check-in actions assigned, and affixed to prestuffed packages for remail. All second and third-time A40 PMR's were referred to the agricultural services processing unit after check-in.

[3]Second- and third-time "must" level PMR's were referred to the coverage processing unit after check-in.

and to other cases designated for remailing by an analyst. Complete packages were assembled, containing the appropriate report form, information sheet, return envelope, brochure, and transmittal letter for all cases to be remailed. New labels were affixed and the packages were mailed on a flow basis. The typing subunit prepared all materials that had name and/or address corrections, and referred them to data keying in order to update the mail file.

All second- and third-time PMR's were sorted to identify cases with name/address changes, large cases (i.e., those with 1974 farm sales of $40,000 or more, or 1974 acreages of 500 or more), and special-list cases. Cases requiring name or address corrections were processed and remailed as described above. Large and special list cases were sent to the final control unit for referral to Suitland, as were second-time refusal PMR's. The remaining PMR's were sent to central files.

The volume of first-time PMR's processed in the 1978 census was similar to that experienced in 1974, that is, about 355,000. However, there were 220,000 second-time returns, a considerable increase over the rate experienced for 1974. This reflected the inclusion for 1978 of several large lists that contained many out-of-date addresses.

Technical Review

General procedures—The technical review unit consisted of a staff of technical analysts and clerks, who reviewed cases referred to their unit and made corrections or transcriptions as necessary to facilitate data-keying of the census report forms. In addition, cases rejected by the computer edit program were pulled from the computer processing cycle and the data items "flagged" by the computer were reviewed and corrected as necessary.

Most of the workload for the technical review unit came from the clerical screening referrals. The technical review staff edited each report form, resolved any problems if possible, and routed on to the analysts those forms that contained problems it could not resolve.

Computer-edit rejection review—The technical review unit also received records rejected by the computer format and edit programs because either the entire report or certain data items within were unacceptable and received disposition lists and/or batch edit listings that noted the reason(s) for failure. These reasons included the following:

The census file number check digit failed.

The county code was outside the acceptable range for its State.

The form code (1H, 1N, 1S1, 1S2, or 1S3) differed from others in the batch.

An invalid State code was detected.

A nonnumeric character (such as an asterisk or a slash) was detected in the address label code.

An invalid area sample characteristics code was detected.

A data item was identified, but no data responses were included.

Nonnumeric data were entered in a data field.

In addition to these specific reasons for rejection, 10 or more item-rejection messages for a given report form would result in the rejection of the form. The particular reasons for rejecting an individual item included—

Miskeying of legitimate data, or extraneous data-keying error.

Invalid key code for the State in which the operation was located.

Data outside acceptable ranges.

Wrong date listed.

Write-in entry out of its section.

After reports were corrected, they were recycled through data keying.

Edit rejects—Detailed comparisons of crop, livestock, sales, etc., entries were made to determine if the values were consistent. Also, checks were made to determine if respondent entries were within the limits established by subject-matter specialists for each data item.

Rejected reports were referred to the review unit from the document-control area. Clerks matched the census file numbers on the batch edit listings to the report forms, reviewed the latter and took the necessary corrective action. The corrected batch edit sheets were routed to the batch unit for data keying and were re-edited by the computer correction program. If the number of changes to a case was above prescribed limits, the report was rekeyed.

Correspondence referral—When a problem case referred for technical review required followup correspondence, usually because a critical item or section was not reported, the clerk involved was to indicate on a form A404 that an A106(L) letter was necessary, staple the A404 to the report form, and refer it to the correspondence typing subunit. The correspondence unit normally held the report form for one month after the necessary documents had been mailed to the respondent concerned, and then sent the report form and the response, if any, back to the technical review unit. Clerks in the review unit then transcribed the data from the correspondence response to the report form, and returned the latter to the processing cycle. In critical cases, when a response was not received, telephone calls were made to resolve the problem.

Coverage unit—The coverage unit consisted of a staff of technical analysts and statistical clerks, who reviewed the "must" cases, multiunits, and abnormal farms cases, and prepared these reports for data keying. The workload for the unit came primarily from the check-in unit in the form of pre-identified "must" and multiunit reports. A portion of the workload also came from the screening unit where large acreage and/or high value cases were identified. All reports entering the coverage unit were screened and made keyable. Forms showing 10,000 acres or more on a place, and/or showing $1 million or more in sales of agricultural products, were identified and referred to the Agriculture Division in Suitland for review. Multiunits were screened and the coverage unit prepared any followup mailings required.

Data Keying

The data-keying system—Data keying for the 1978 Census of Agriculture was done using an electronic key-to-disk-to-tape system to prepare the census data for computer processing. The data-keying unit at Jeffersonville employed 8 to 10 "keying systems," each consisting of 16 to 20 individual keying stations. Each station was made up of a keyboard and a cathode-ray tube (CRT) viewing screen that enabled the operator and supervisors to monitor and edit the keyed data. Each of the keying systems was linked to a computer disk drive serving one to four computer disks. Each of these disks had a capacity of approximately 2.5 million characters. Programs and control instructions occupied about 20 percent of the total capacity of each disk, while the remainder was available for data from the census report forms. Since each report form required (on the average) about 250 characters, a disk, if used to its maximum capacity, might hold data for as many as 8,000 returns. In practice, the capacity of each disk assigned to a keying system was divided among the individual stations comprising that system, and a portion of each disk's capacity was needed to insure there was no overlap of entries from the different stations.

The verification process enabled rekeyed data to be recalled for visual comparison to those already on the disk and, in cases of conflicts between data items, the appropriate report form would be consulted. When necessary, the original keying was corrected. Inasmuch as the verification procedures required that *every* difference in keying be checked, the need for quality control of the verification process was eliminated.

After verification and correction, the data were moved automatically from the disk to a magnetic pooler tape containing data for only one State, and then were transmitted to the Suitland computer facilities by telephone datalink. As soon as the data were "read" at Suitland, the Jeffersonville pooler tape was erased and reused. To safeguard against the loss of data through accident or technical breakdown, the contents of each disk were transcribed to "systems-save" magnetic tapes every 2 hours during the keying operation. These "save" tapes were held for approximately a week, or until the data they contained were accounted for as having been correctly read at Suitland.

Data-keying operations—After screening and technical review, the report forms were batched by State into work units of 100 forms. Separate batches were maintained for A1(N), A1(S), and "must" cases. A form A405 Data Keying Work Unit Cover Sheet was attached and the work unit was sent on to the data-keying unit where clerks, following detailed instructions, keyed the entries on the work-unit cover sheet and each report form in the unit. For each report, data were keyed from the address label and from each item to which there had been a response.

As the codes and responses were keyed, certain checks were performed electronically in a series of data-entry edits that were changed according to the type of form being keyed. The purpose of these "input edits" was to (1) insure that check digits and entries in State, item, and sub-item fields were valid; (2) insure that all identification, item code, and data-field entries (except the form code from the label) were numeric; (3) perform sequence and valid-code checks by section for each report form; (4) insure data were keyed by item code; and (5) compare consecutive item codes for duplicates.

Rejected records were referred to agriculture subject-matter analysts for review. (The balance of the editing of the records was done during computer processing.)

After data keying and verification, the report forms were placed in a holding area until data were satisfactorily processed through the computer. Once that was accomplished, the processed forms were moved to central files for boxing and storage.

Verification—As with the other major clerical operations in the census processing program, data-keying was subjected to a process of verification to insure that keying was complete and accurate. In the verification phase of the operation, attention was centered upon the key operators, each of whose work was reviewed for errors. An error was defined as (1) a keystroke error in keying an item code or data entry, (2) an omission, or (3) duplication of an item code or datum. The verification process was carried out in three stages, during which the key operators progressed through three periods—training, qualification, and process control. During the training period, each operator's work was verified on a 100-percent basis. Each operator was to become familiar with the key structures of the census report forms and was to key at least three complete work units. Operators with a cumulative error rate for these three work units of 3.5 percent or less advanced to the qualification period stage, while those failing to do so were retrained.

During the qualification period, key operators' work was verified on a 10-percent basis; the specific records verified were selected at random from each work unit. To move on to process-control verification, an operator had to have a sequence of four successive "accept" decisions within a maximum of eight decisions. Operators failing to achieve this were allowed a second chance to qualify for process-control verification, but a second failure to do so meant the operator would not be retained.

Successful completion of the qualification stage meant operators were moved into process-control verification. During this stage a 4-percent sample of report forms (excluding "green must" forms—see below) was randomly selected and verified from each work unit keyed. Operators had to have a minimum of 7 "accept" decisions in each sequence of 10, or else had to requalify for process control.

At every verification point, errors detected were corrected before the data were transmitted to Suitland. "Must" cases (form 78-A1(S) "green") were assigned only to keyers qualified for process control and were verified and corrected 100 percent. To maintain their standing, these "must" case keyers had to maintain an error rate of 2.5 percent or less. The verification plan was designed to allow an estimated outgoing error rate of no more than 2.5 percent for all records keyed. These goals were slightly improved upon in practice.

COMPUTER PROCESSING

General Information

The use of computers to process census data has increased progressively from census to census, since the first automatic

processing equipment was introduced in 1890. While a considerable amount of manual sorting, reviewing, and checking was done for the 1978 census, the greater part of the actual assembly, editing, and tabulation of the data was by computer at the Bureau's Suitland headquarters. The computer processing phase of the agricultural census started as soon as report forms began to arrive in Jeffersonville and were processed through the clerical screening and keying operations. The first farm records were processed by the computers in February 1979, and the operation continued until the final tabulations were completed in July 1981. Data were processed on a flow basis as records were received, although there were occasional interruptions in the access to the computer facilities because of water damage to the computers during a mishap in August 1979. Nevertheless, approximately 3.2 million individual agricultural census records were edited; some 2.26 million of these met the Bureau's definition of a farm and were incorporated into the agricultural census file.

The computer processing operation can be divided into three major phases: (1) formatting, (2) edit and failed-edit correction, and (3) tabulation of the data. These phases are described below.

Formatting

The first step in the computer processing was the formatting of the data into binary records that could be manipulated electronically. A computer record was established for each census report form; each consisting of a section of variable-length data segments and one fixed-length segment. The fixed-length section contained the report form's identifying information, such as State and county codes; serial number; farm definition, farm criteria, and SIC codes; and all the other data necessary to create a complete, consistent, and individually identifiable data record. The variable-length segments of each record contained a computer "word," or record segment, for each item reported, imputed, or changed in the record. Each data item was identified by the item key code associated with it on the report form. (For example, item 67 on the A1 form represented the acres of corn-for-grain harvested.) Data items for which nothing was reported or imputed contained no information and were omitted from the detail data record.

The following major operations were performed during the computer format run:

1. Data for crop production were converted into standard units of measure for those crops showing more than one such unit on the report form.
2. "Landlord only" and other types of out-of-scope records identified during the format run were separated from the general data file.
3. Invalid codes were identified and classified by type, and appropriate action was taken, as follows:

   a. **Invalid State, county, and form codes:** These records were printed out and dropped from the format run. The sample and nonsample report forms involved were corrected and then rekeyed.

   b. **Rejected item codes.** These were codes that were either not assigned anywhere on that particular report form or were valid crop item codes that were invalid for a specific State (e.g., codes for sugarcane in Maine). For listing-identification purposes, the offending item code, the item code immediately preceding it, and the two item codes immediately following, together with all the associated data, were printed out. Valid item codes that appeared out of sequence, including duplications, were handled in the same way. All invalid or otherwise offending codes and their data were omitted from the formatted record and were printed out for review. Corrections then were made and were carried to the formatted record in a correction match program.

   c. **Maximum acceptable rejects exceeded.** The number of errors in any given record was limited. Once the total number of errors exceeded 9, the record in question was pulled from the formatting cycle, displayed in its entirety, and reviewed. Corrections were made, as necessary, to the appropriate report form or forms, which were then returned to the data-keying subunit and recycled through the processing operation.

Computer Editing and Failed-Edit Correction

Computer editing—Computer editing is the mechanized process of validating, cross-checking, and refining reported data. The computer processing programs for the 1978 census included an editing program that tested key ratios within the data for reasonableness and consistency. The ratios were tested by matching them to tolerance limits based on experience in previous censuses and surveys, after which the computer corrected errors by rounding the individual data items, substituting the sum of the detail items for a reported total, or imputing on the basis of one of several ratios that included the challenged component.

The computer programs written to perform these tasks were necessarily long and complex. The individual tests and checks comprised several thousand steps in total, although generally only a relatively small fraction of these were involved in the editing of the data from any one report form.

Computer edit specifications were transmitted from the subject-matter specialists to the computer programmers by means of a decision logic table (DLT), i.e., a tabular display of all the elements of an edit problem, from conception to solution, with flowcharts and texts attached when additional information was needed. About 3,000 pages of DLT's and related materials were needed for the computer edit of the standard agriculture census report forms, including several rounds of revisions carried out to improve precision and consistency in the edit. (An effort on a somewhat smaller scale was necessary to edit the various A40 forms, which were processed separately.)

The actual computer editing was done by State. Batches for editing, each consisting of formatted records sorted by State, county, and CFN, were assembled by setting cutoff dates; records received during a specified period (2 weeks early in the

census, but as much as 8 weeks later) became part of a single batch, which was then edited. The last batch was assembled and processed after the analysts' review of "must" cases.

The computer edit of the standard report forms—

1. Supplied missing entries.
2. Reconciled acres reported for individual crops with acres reported as total cropland.
3. Imputed production for crops when the reported production was outside acceptable limits.
4. Edited to assure consistency between and within the different sections of the report forms.
5. Calculated and checked values for products sold, using average prices in each State for each production item, and substituted these values for reported values if the latter were outside acceptable limits.

The computer edit also determined whether each record met the criteria for a farm or was out of scope, and coded (classified) the farm records according to acreage, tenure, value of agricultural products sold, and type of organization. Records that did not meet the minimum criteria for a farm were deleted from the data file and were transferred to an out-of-scope file. A list of these out-of-scope addresses was sent to Jeffersonville, where the clerical staff reviewed the related report forms to insure that they had been accurately keyed and correctly classified.

The minimum criterion for meeting the Bureau's definition of a farm was annual sales of agricultural products of $1,000 or more. The computer edit identified and retained, as farm records, data for those places that did not have, but normally would have, a total value of annual sales of agricultural products of $1,000 or more. Places not meeting the $1,000 definition were tested against a set of criteria designed to identify potential farms and farms that would normally meet the sales minimum but which did not because of extenuating circumstances, such as drought or crop failure. A set of 46 criteria codes was established for these farms, each code indicating the broad type of product (cash grains, vegetables, livestock, pastureland, etc.) involved, with a minimum quantity or acreage specified.

In addition to determining whether records were in scope, the computer edit program also converted nonsample records to sample records if they met the certainty criteria. Certainty criteria varied by State, with sales of from $40,000 to $200,000 or a minimum indicated acreage of from 1,000 to 5,000 acres. Institutional operations and other special cases were also included in the certainty group even if they did not meet the minimum criteria, as were all farm addresses in counties with fewer than 100 farms in 1974. When data from a nonsample record met the certainty criteria for the sample, the additional detailed information was obtained by correspondence or was imputed on the basis of responses from farms of similar size in the same geographic area. Any such conversions based on reported sales or acreage were coded as "certainty" cases. Occasionally, a sample record was converted to a nonsample record. This usually was done if the farm in question had originally been sent both a sample and a nonsample form, and the sample name and address were retained with the nonsample data. A record of the changes made for these individual farm records was printed out periodically during the computer batch edit phase of the operation and was sent to Jeffersonville for review.

Failed-edit correction—Once the computer edit was completed, the high-speed printer was used to produce a failed-edit listing that included a printout for each report form that had one or more items flagged by the edit program. This listing displayed the items from each form that (1) had failed the edit, (2) had not failed the edit but had been changed by the edit, and/or (3) had a referral flag. The printout for a given farm record occasionally ran to two or more pages, but no page contained items for more than one farm record.

The failed-edit listings were sent to Jeffersonville where they were separated, sorted by State, placed in portfolios in lots of 500 consecutively numbered records, and matched to the report form file. The listing sheets and the corresponding report forms then were reviewed clerically. A set of procedures was provided for referral of records to agricultural analysts when this was necessary. Disposition codes were assigned to the individual records, indicating the general action to be taken for each, as follows:

Code	Action
1	Make corrections—re-edit record
2	Make corrections—bypass specified sections of the edit
3	Make corrections—bypass the edit except coding (edit section 51), SIC coding (54), and summing (75)
4	No corrections—change failed-edit flag to passed edit
5	Delete record from file
6	Change RD (referral disposition) code to 3—make corrections and re-edit record
7	Convert record from sample to nonsample, make corrections and re-edit record

An item-locator code was assigned to each location within every farm data record where an edit failure occurred. These codes were used when corrections were inserted into the farm data file, and a file of corrections or changes, called the change index, was compiled. Every time an item was changed during processing, this was noted in the index; ultimately, the item, the value of the item before the alteration, and the value as changed were listed on a microfilmed "universe of changes" file. (This "universe" was used as a review tool during the analytical review of the tables.)

If no corrections were required, a disposition code of 4 was assigned. In cases that required numerous corrections, the form was corrected, rekeyed, and then recycled through the computer processing operation. For most cases, corrections were marked on the failed-edit listing sheets and the sheets were batched for keying (data to be keyed were underlined). The corrections were keyed to tape, verified 100 percent, and trans-

mitted by datalink to Suitland for a computer match to the data file. The corrected files were re-edited to insure that the corrections had been made properly and to determine whether further corrections were necessary.

Final data merge—After the computer edit and the failed-edit corrections were completed, the corrected files for each State were merged into a single data file in sequential order by State, county, and identification number. The file was then unduplicated by matching CFN's and all duplicate records were displayed for review. Unless circumstances dictated otherwise (e.g., the duplicate CFN's, through some error, represented entirely different operations), the first of the duplicate records displayed was retained while the rest were deleted. The merge program also tallied farms by size, total value of products sold, and type. These tallies were used to help impute data for nonrespondents. (Imputation is discussed below.) In addition, the merge program identified other problem records and displayed them for further review and possible correction before tabulation.

Imputation for nonresponse—Data for nonrespondents were imputed after the data files were corrected, merged, and unduplicated. Data for a respondent within the same size group and county were duplicated to represent each nonrespondent. All this meant was that the data for the farm selected for duplication was counted twice, to approximate the contribution to the county totals that would have been made by the nonrespondent operation.

Approximately 12 percent of all names and addresses on the mailing list did not respond to the census. Adjustments to compensate for this were made to the data at the county level, using a three-step procedure. First, a stratified sample of nonrespondents designed to provide State estimates was selected and each sample nonrespondent was mailed a short report form. Those not responding to the mailed form were contacted by telephone to determine if they operated a farm. The sample nonrespondents were classified on the basis of this survey as either "farm" or "nonfarm," and results of this classification procedure were used to estimate the number of nonrespondent census farms in each State. (About 60 percent of the nonrespondent sample units were classified as nonfarm.) A synthetic estimator was developed to estimate the number of nonrespondents by size strata for each county of a State. Finally, a sample of respondents was selected to represent the missing nonrespondent farm operators in the census on the basis of their expected total value of sales as recorded on the census mail list. Farms with expected values of sales of up to $40,000 were candidates for duplication in the census data file, since most nonrespondent farms were in this range of value of sales. Any farm with an expected value of sales in excess of $40,000 was a certainty case and was subject to a 100-percent followup. In the rare instance where a response from an operation of this latter size could not be obtained from the operator, administrative records were used to estimate totals rather than impute the data using the usual techniques.

Stratification and sample weighting—The use of sampling from the mail list introduced into the census data several elements that could cause substantial variation and a potential bias. First among these was the fact that half of the addresses on the mail list from which the sample was selected did not represent farms and were not identifiable as such at the time the sample was selected. Hence, both farm and nonfarm addresses were included in the sample. Data actually tabulated came from only part of the sample—those names and addresses that represented farms. Second, stratification by size of operation of the addresses in the sample was based on information from several sources of variable quality. Further, the response rate for addresses in the sample may have been different than for nonsample addresses.

In order to improve the precision of the estimates from the sample, post-stratification was used to produce adjusted estimates. Basically, this consisted of classifying all farms into relatively homogeneous strata and weighting sample farms within each stratum by the ratio of total farms to sample farms.

Farms meeting certainty size criteria during sample selection, and those identified during processing as meeting similar criteria, were assigned to a certainty stratum. All other farms were assigned to 64 strata. Farms with sales of less than $2,500 were classified into eight size-of-farm groups (less than 10 acres, 10 to 49 acres, 50 to 69 acres, 70 to 99 acres, 100 to 199 acres, 200 to 259 acres, 260 to 499 acres, and 500 acres or more) within each of two value-of-sales groups (less than $1,500 and $1,500 to $2,499) and within each of two type-of-farm groups (crop or general farms and livestock or poultry farms). Farms with sales of $2,500 or more were classified into four size-of-farm groups (less than 50 acres, 50 to 99 acres, 100 to 259 acres, and 260 acres or more) within each of the four value-of-sales groups ($2,500 to $4,999, $5,000 to $9,999, $10,000 to $19,999, and $20,000 or more) and within each of two type-of-farm groups (crop or general farms and livestock or poultry farms).

Each stratum was examined and collapsed into another stratum if (1) the stratum contained less than 20 sample farms; or (2) the calculated weight for the stratum was greater than 10 in counties sampled at a rate of 1 in 5 or greater than 4 in counties sampled at a rate of 1 in 2.

The post-stratification provided weights to be assigned to farms in each of the final collapsed strata such that the total of the weights for sample farms in the stratum would be equal to the total number of farms in the stratum.

Estimates were prepared for items in sections 22 through 27 of the report form by multiplying the data for each item for each farm in the sample by the weight assigned to the farm. The weight for a certainty farm was 1.

Tabulating and Reviewing Data

General information—After the records had been edited, corrected, and merged, the data were ready for tabulation. The individual records were tabulated by computer into detailed data matrices, each containing over 12,000 different items that were designed to provide the basic data input for most of the data tables drawn from the 1978 Census of Agriculture. Analytical tabulations were prepared, using these matrices, in a

detailed format by county, to aid in locating problems in the preliminary data. After the tabulations were reviewed and the farm records had been corrected, a special tabulation run was made of corrected records both before and after correction. The data from the uncorrected records were subtracted from those for the corrected records, and the resultant net totals were merged into the data matrices. The corrected matrices were used as the data source for the preliminary reports, the major results reports, and for the review copies of the tables for volume 1. State-level cross-tabulations were not run until the final volume 1 corrections were completed; a separate computer pass of the data records was required for these cross-tabulations.

Analytical tabulations—All of the items reported on individual report forms were tabulated for each county and State, for all farms, and for farms with sales of $2,500 or more. Comparable historical data drawn from the 1974 census final reports were included in the tables for use in reviewing the accuracy and completeness of the 1978 data. The analytical tables served as the basic documents for review by Agriculture Division staff. A substantial amount of related check data, most of it from USDA estimates, was also used in the review. Detailed criticisms of questionable data were prepared and were transmitted, together with suggested remedial action, to the Jeffersonville staff.

Prior to these transmissions, representatives from USDA's Economic Statistics Service's (ESS) State offices also reviewed the analytical tables and criticism sheets prepared by Agriculture Division staff. The ESS reviewers indicated which criticisms they considered unnecessary, offered additional or alternative solutions to identified problems, and added comments and/or criticisms on problems or potential problems not identified by the Agriculture Division review.

Criticisms arising from the review of these materials were acted upon by the Jeffersonville staff; the actions taken included—

1. Review criticisms and suggested actions made by the Agriculture Division staff and by ESS State representatives.
2. Verify the validity of data questioned in the criticisms submitted, or make necessary corrections to the data.
3. Obtain reports from farm operators for places that had not been included in the tabulations.
4. Correct data-keying, reporting, and processing errors.
5. Identify and eliminate duplicate records not previously detected by matching CFN's.
6. Assign correct State and county code numbers for large operations to ensure that these operations were tabulated in the proper State and county.

County data corrections—When the review of the analytical tables was completed, corrections were made to individual farm records in the same way as they had been after the initial computer edit. These corrections were reviewed by the Agriculture Division staff for accuracy and to ensure that the data criticisms were satisfied. The preliminary reports then were tabulated and reviewed. If any additional corrections were necessary, the data were changed by computer, or hand corrections were made to the tabulation printouts. The data file was corrected as often as necessary to ensure its accuracy.

Tabulations for counties, States, divisions, regions, and the United States—County and State tables were drawn from the matrices and State cross-tabulations were prepared directly from the data file. Data for divisions, regions, and the United States were obtained by summing data from the State matrices. The historical data for the 1978 tables were taken from the 1974 computer matrices.

Final disclosure analysis—The Bureau of the Census is prohibited by law from publishing data that could be used to identify individual respondents to any of its censuses or surveys. To insure that confidentiality is maintained, all data tables are completely reviewed (disclosure analysis) before they are released for publication. While part of the analysis of the 1978 data was done by computer, the computer programs were incapable of completing the entire analysis and much of it had to be done by statisticians. Essentially, this involved the identification and suppression of figures that (1) would result in direct disclosures, or (2) could be used to reveal information about individual operations by derivation (e.g., adding or subtracting a published subtotal from a published total would expose individual data).

At the county level, for a county with fewer than 10 farms, no data were released at all because of the possibility of disclosure of individual information.

The established guidelines usually set a lower limit on the number of farms that must report data for an item before those data would be released for publication. If more than the minimum required number of farms reported an item, the data could be published, unless comparison of different tables could result in disclosing that one or two farms had a very large percentage of the total. Exceptions in the application of these rules were made, but generally only for the very large specialty operations —e.g., poultry, feedlots, greenhouses, etc.—any of which might easily be identified as a specific farm, but whose absence from the counts would grossly distort the data. Publication of the number of farms reporting an item was not in itself considered a disclosure; only related information about the item was suppressed.

Several of the tables contained the same information arranged according to a different classification, so that when it was necessary to suppress a figure in one table, it would also be necessary to check other tables and suppress it in them as well. Similarly, when it was found necessary to suppress an item in one of the county tables, all of the tables for the county had to be reviewed and the item in question had to be suppressed in all of them.