

---

## Appendix A.

# Census of Agriculture Methodology

---

The purpose of a census is to enumerate all objects with a defined characteristic. For the census of agriculture, that goal is to account for “any place from which \$1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year.” To do this, NASS creates a Census Mail List (CML) of agricultural operations that potentially meet the farm definition, collects agricultural information from those operations, reviews the data, corrects or completes the requested information, and combines the data to provide information on the characteristics of farm operations and farm operators at the national, State, and county levels. In this appendix, these census processes are described.

### THE CENSUS POPULATION

#### The Census Mail List

The National Agricultural Statistics Service (NASS) maintains a list of farmers and ranchers from which the Census Mail List (CML) is compiled. The goal is to build as complete a list as possible of agricultural places that meet the farm definition. The CML compilation begins with the list used to define sampling populations for NASS surveys conducted for the agricultural estimates program. Each record on the list includes name, address, and telephone number plus additional information that is used to efficiently administer the census of agriculture and agricultural estimates programs.

NASS builds and improves the list on an ongoing basis by obtaining outside source lists. Sources include State and federal government lists, producer association lists, seed grower lists, pesticide applicator lists, veterinarian lists, marketing association lists, and a variety of other agriculture-related lists. NASS also obtains special commodity lists to address specific list deficiencies. These outside source lists are matched to the NASS list

using record linkage programs. Most names on newly acquired lists are already on the NASS list. Records not on the NASS list are treated as potential farms until NASS can confirm their existence as a qualifying farm. Staff in NASS field offices routinely contact these potential farms to determine whether they meet the farm definition. For the 2012 Census of Agriculture, NASS made a concerted effort to work with Community-Based Organizations not only to improve list coverage for minorities but also to increase census awareness and participation.

List building activities for developing the 2012 CML started in 2009 by updating list information from respondents to the 2007 Census of Agriculture. Between 2010 and 2012, NASS conducted a series of National Agricultural Classification Surveys (NACS) on approximately 1.7 million records, which included nonrespondents from the 2007 census and newly added records from outside list sources. The NACS report forms collected information that was used to determine whether an operation met the farm definition. If the definition was met, the operation was added to the NASS list and subsequently to the CML. Addressees that were nonrespondents to a NACS were also added to the CML and identified with a special status code.

Measures were taken to improve name and address quality. Additional record linkage programs were run to detect and remove duplicate records both within each State and across States. List addresses were processed through the United States Postal Service’s National Change of Address Registry and the Locatable Address Conversion System to ensure they were correct and complete. Records on the list with missing or invalid phone numbers were matched against a nationally available telephone database to obtain as many phone numbers as possible. To reduce costs, operations with characteristics that indicated they were unlikely to be farms, according to the farm definition, were

removed from the list.

The official CML for the 2012 Census of Agriculture was established on September 1, 2012. The list contained 3,009,641 records. There were 2,387,326 records that were thought to meet the NASS farm definition and 622,315 potential farm records, which included NACS nonrespondents, other records added to the CML by the NASS field offices after the record linkage process, and late adds to the CML that were not included in any previous NACS or State screening survey.

### **Not on the Mail List (NML)**

Extensive efforts are directed toward developing a CML that includes all farms in the U.S. However, some farms are not on the list, and some agricultural operations on the list are not farms. NASS uses its June Agricultural Survey (JAS) to quantify the number and types of farms not on the CML. The tracts in the JAS that are not on the CML are said to be in the Not on the Mail List (NML) domain. If a tract in the NML domain is determined to be a farm during the census, it is an NML farm. The NML farms are used to estimate the undercoverage associated with the census.

The NASS area frame, which is used for the JAS, covers all land in the U.S. and includes all farms. The land in the U.S. is stratified by characteristics of the land. A probability sample of segments is drawn within each stratum for the JAS. Segments of approximately equal size are delineated within each stratum and designated on aerial photographs. The JAS sample of segments is allocated to strata to provide accurate measures of acres planted to widely grown crops, farm numbers, and inventories of cattle. Sampled segments in the JAS are personally enumerated. Each operation identified within a segment boundary is known as a tract.

The 2012 JAS sample was increased to improve the farm counts for operations that produced specialty commodities or had socially disadvantaged or minority operators. The total sample consisted of 14,376 segments of which 3,291 were additional segments added to facilitate the use of the JAS as an Agricultural Coverage Evaluation Survey (ACES). The additional segments were added based upon

multivariate sample allocations to target specific items at the U.S. level. The 2012 JAS consisted of sample segments from all States, with the exception of Alaska where NASS does not maintain an area frame.

During the JAS prescreening operation, each tract is identified as either agricultural or non-agricultural. Each JAS agricultural tract is identified as a farm or non-farm in June based on the farm definition. Non-agricultural tracts are further classified into categories; with farm potential, with unknown farm potential, or with no farm potential. The names and addresses collected in the 2012 JAS were matched to the CML. Those from the JAS 2012 survey that did not match were determined to be in the NML domain and sent a yellow census report form so that they could be differentiated from the green report form sent to those addressees on the CML. Instructions on the census report form directed any respondent who received duplicate forms to complete the CML form and to mail all duplicate forms back together. Those who returned a CML and an NML form had been misclassified as NML and were removed from the NML domain.

The initial NML mailout consisted of 36,021 records. An additional 403 June area tracts linked to Census records that were Undeliverable as Addressed (UAA) were later added to the NML domain. A total of 36,424 NML records were summarized of which 5,565 records were truly NML and in-scope.

The farm/nonfarm status of each NML domain operation was determined based on the reported data in the census form. An operation in the NML domain that was determined to be a farm is referred to as an NML farm. Characteristics of NML farms and their operators provided a measure of the undercoverage of farms on the CML. The percentage of farms not represented on the CML varied considerably by State. In general, NML farms tended to be small in acreage, production, and sales of agricultural products. Farm operations were missing from the CML for various reasons, including the possibility that the operation started after development of the CML, the operation was so small that it did not appear in any agriculture-related source list, or the operation was misclassified as a nonfarm prior to census mailout. The CML was used with the NML in

a capture-recapture framework to represent all farming operations across all States in the JAS sample.

## **DATA COLLECTION OUTREACH AND PROMOTIONAL EFFORTS**

NASS planned and executed a multi-phase strategic communications campaign for the 2012 Census of Agriculture, to increase the level of awareness and response among all U.S. agricultural producers.

- Phase 1 ran from October 2011 – July 2012. It raised awareness about the census and list building, encouraged producers to sign up in response to NASS mailings and at community, association, and other stakeholder meetings where NASS partners reached out.
- Phase 2 ran from July 2012 – December 2012. It notified farm operators and agricultural organizations that the census would be mailed in December, and encouraged communications regarding the census.
- Phase 3 ran from December 2012 – July 2013. It focused on census data collection with messaging urging response, reminding operators that it's-not-too-late-to-respond, and thank-you messaging.
- Phase 4 began in February 2014. It communicated information about the data release plan, which has four phases:
  - Phase A (November 2012 – December 2013) focused on thanking farmers for their participation in the census and partners for their leadership.
  - Phase B (January 2014 – February 2014) drew attention to the preliminary census release.
  - Phase C (February 2014 through May 2014) focused on the final census release.
  - Phase D (ongoing) continues to focus on the census findings as they are released.

As part of the plan, NASS targeted selective communications and outreach efforts on beginning and minority farm operators. All of these efforts were accomplished through an integrated communications program that focused on four primary areas: partnership building, local-level outreach, public relations, and paid media. External support was provided by a private agricultural communications agency.

The unifying force behind the 2012 communications campaign was the theme “There’s Strength in Numbers.” This was accompanied by supporting messages and artwork that created a consistent look and feel for all census communications. All messages and materials served the purpose of inspiring action: *Grow Your Farm Future - Shape Your Farm Programs - Boost Your Rural Services - Fill out your Census of Agriculture - Do your part to be counted - There’s strength in numbers.*

## **Partnership and Local-Level Outreach**

At the national level, NASS officials met with leaders from dozens of key agricultural organizations, State departments of agriculture, and other USDA agencies, to successfully secure their support in promoting the census among their constituencies. Stakeholders partnered with NASS to promote the 2012 Census of Agriculture through publications, special mailings, speeches, social media, websites, and other communications. In addition, through grassroots-level outreach and efforts, NASS partnered with a number of community-based organizations to reach minority and limited-resource farmers and ranchers. All national-level outreach was encouraged and mirrored at the regional, State, and local levels. Among the highlights of these partnership efforts was the production of more than 40 television and radio public service announcements (PSAs) featuring the U.S. Secretary of Agriculture, State secretaries, directors, and commissioners of agriculture and leaders from community-based organizations. The PSAs, available in both English and Spanish, encouraged farmers and ranchers to respond to the 2012 Census of Agriculture.

## **Coverage of American Indian and Alaska Native Farm Operators**

To maximize coverage of American Indian and Alaska Native farm operators, special procedures were followed in the census. A concerted effort was made to get individual reports from every American Indian and Alaska Native farm operator in the country. If this was not possible within some reservations, a single reservation-level census report was obtained from knowledgeable reservation officials. These reports covered agricultural activity

on the entire reservation. NASS reviewed these data and removed duplication with any data reported by American Indian or Alaska Native farm operators who responded on an individual census report form. Additionally NASS obtained, from knowledgeable reservation officials, the count of American Indian and Alaska Native farm operators (on reservations) who were not counted through individual census report forms, but whose agricultural activity was included in the reservation-level report form.

This information is summarized in Table D, **American Indian and Alaska Native Operators: 2012**, providing the number of farm operators (for up to three operators per farm) reported as American Indian or Alaska Native in the race category, either as a single race or in combination with other races, on the individual census report forms, plus the total number of American Indian or Alaska Native operators farming on reservations as reported by reservation officials. The count from the individual report forms is summarized in the “Individually reported” column. It includes operators on or off reservations. The “Other” column provides counts of operators on reservations as reported by a reservation or tribal official. The “Total” column is simply a sum of the “Individually reported” and the “Other” columns. Tables in other parts of the publication count the reservation-level reports as single farms.

## **Public Relations**

In the public relations arena, NASS and the contractor worked with internal and external stakeholders to equip them with communications tools and resources to deliver the census communications message to their audiences. NASS utilized its Intranet to deliver materials to the 12 regional and 46 field offices and created a “Partner to Promote the Census” portal on the census website to deliver public relations materials and tools to external stakeholders. The materials included, but were not limited to: customizable news releases, feature stories, newsletter articles, blogs; drop-in advertisements; website buttons and banners; PowerPoint templates; brochures; and more. In addition, at the national level NASS issued a dozen news releases citing department and agency spokespeople and published timely and relevant

pieces to the USDA blog highlighting the census. These public relations efforts at the national, State, and local levels helped ensure that NASS’s message about the census was continually in the media, including print and online publications, a variety of social media, radio, and some television programs. Media outlets included both those specializing in agriculture and more general outlets.

## **Paid Media**

For the 2012 Census of Agriculture, NASS placed special emphasis on reaching new and beginning farmers, while continuing efforts to improve its reach within previously under-represented populations. Even with increasingly limited budgets and resources, NASS was able to apply a portion of funds towards paid media. Strategically, NASS purchased limited print and online advertising in areas where there was the potential for high concentrations of under-represented populations and new and beginning farmers and ranchers.

## **DATA COLLECTION**

### **Method of Enumeration**

Data collection was accomplished primarily by mailout/mailback, but supplemented with Electronic Data Reporting (EDR) on the Internet, and personal enumeration for special classes of records in the census operations. Personal enumeration (interviewing) involved the use of both Computer-Assisted Telephone Interviewing (CATI) and Computer-Assisted Personal Interviewing (CAPI). Enumerators at the NASS National Operations Center in St. Louis, MO conducted CATI data collection. In addition, enumerators under contract with NASS through the National Association of State Departments of Agriculture (NASDA) conducted phone and personal interviews with respondents. For the 2012 Census of Agriculture, NASS implemented a pre-notification strategy in an effort to increase awareness, improve overall responses, and encourage respondents to report early to avoid continued correspondence. All records in the initial mailout received either a postcard or pre-recorded voice message announcing the census mail packets were coming.

## Report Forms

There were seven regionalized versions of the report forms used for the 2012 Census of Agriculture. The report form versions were designed to facilitate reporting crops most commonly grown within each report form region. Additionally, an American Indian report form was developed to facilitate reporting for operations on reservations in Arizona, New Mexico, and Utah. The regional report form numbers are: 12-A101, 12-A102, 12-A103, 12-A104, 12-A105, 12-A106 and 12-A107 (HI). The American Indian report form is 12-A200. All of the forms allowed respondents to write in specific commodities that were not listed on their form.

## Report Form Mailings

Pre-notification by postcard or pre-recorded message began December 10, 2012. Approximately 3.0 million mail packets were mailed in December 2012. Each packet contained a cover letter, instruction sheet, a labeled report form, and a return envelope. The Census Bureau's National Processing Center (NPC) in Jeffersonville, IN was contracted to perform mail packet preparation, initial mailout, and two follow-up mailings to nonrespondents.

The initial mailout was followed by a thank-you reminder postcard that was delivered in January 2013 to all operations that received mail packets. First follow-up mail packets were mailed in mid-February 2013 to approximately 1.0 million nonrespondents. Second follow-up mail packets were mailed in mid-March 2013 to approximately 750,000 nonrespondents.

## Personal Follow-up

Operating concurrently with NPC's mail data collection efforts, NASS telephone call centers targeted selected groups of census nonrespondents for telephone enumeration. NASS field offices targeted selected groups of census records for in-person enumeration. These efforts were referred to as:

- Suspicious Out of Scope Follow-up
- Criteria Record Follow-up
- Must Case Follow-up
- American Indian and Alaska Native Farm

- Operator Follow-up
- Low Response County Follow-up
- Last Call Nonresponse Follow-up
- Not on Mail List (NML) Follow-up

**Suspicious Out-of-Scope Follow-up.** The Suspicious Out-of-Scope Follow-up was a phone follow-up that began in February 2013 and was conducted through May 2013. It included records that mailed their form back with a response that they were no longer farming. These operations had reported agricultural information in another survey during 2012. The operations were re-contacted with a CATI instrument to either verify the respondent was not farming or complete a census report form.

**Criteria Record Follow-up.** Nonrespondents and refusals to the National Agricultural Classification Surveys received unique coding on the CML and are referred to collectively as Criteria Records for follow-up data collection. These Criteria Records typically had a lower probability of meeting the farm definition and were less likely to respond. It was critical to identify those records in this group that represented farms to provide coverage of the small farm population. Small farms make up a significant portion of the overall U.S. farm population.

For the 2012 Census of Agriculture, 276,043 Criteria Records were included in the Census Mail List (CML). A sample of 23,739 Criteria Records was selected for targeted data collection efforts. The sampled records were first contacted by telephone using the census CATI instrument beginning in February 2013 after the initial mail returns were processed. Certified mail to 18,831 respondents was used for those who could not be contacted by telephone. Data collection resulted in 10,887 returns from both telephone and certified mail. The in-scope rate from the returns was applied to the remaining criteria records during replication, which is described in the next sub-section.

**Must Case Follow-up.** Must cases were known large operations, the absence of which could have significantly affected the accuracy of census results. For the 2012 Census of Agriculture, 118,533 records were categorized as Must cases. Each active Must operation was accounted for by mail receipt, phone interview, or personal enumeration; if an operation was no longer in operation, its nonfarm status was

documented. CATI calling of nonrespondent Must cases was undertaken by call centers from March 2013 through May 2013, after the initial and first follow-up mailing. Following the CATI calling, the remaining nonresponse Must cases were assigned to field offices for personal enumeration. Because of the potential importance of Must cases, they were all accounted for and therefore not eligible for nonresponse weighting adjustment.

**American Indian and Alaska Native Farm Operator Follow-up.** The American Indian report form (12-A200) was mailed to all operations in Arizona, New Mexico and Utah thought to have an American Indian or Alaska Native operator. It was included in the initial mailout, but due to poor mail response a personal enumeration data collection strategy was utilized with no additional mail follow-up. A concerted effort was made to get individual reports from every American Indian and Alaska Native farm operator in the country. If this was not possible within a reservation, a single reservation-level census report was obtained from knowledgeable reservation officials. These reports covered agricultural activity on the entire reservation. The NASS reviewed these data and removed any duplicate data reported by American Indian or Alaska Native farm operators from that reservation who responded on an individual census report form. Additionally NASS obtained, from knowledgeable reservation officials, the count of American Indian and Alaska Native farm operators (on the reservations) who were not counted through individual census report forms, but whose agricultural activity was included in the reservation-level report form.

**Low Response County Follow-up.** The Low Response County (LRC) follow-up activity was used to increase the response rate in all counties to at least 75 percent. CATI was used for this follow-up activity. NASS utilized an adaptive design technique to identify particular records for telephone contact, in an effort to increase coverage on minority operations and operations known to produce specialty commodities. In early April 2013, NASS identified nonresponse cases in counties with a response rate of less than 75 percent. Nonresponse records in these counties were then prioritized so that minority operations and specialty commodity producers were the primary records delivered to

phone enumerators. Nonrespondent telephone contact information was transmitted electronically to NASS call centers and incorporated into their CATI instrument. CATI follow-up activities began in mid-April 2013 and continued through mid-June 2012. Automated procedures were employed biweekly to ensure that the record selection procedures were targeting counties that would meet the goals of increasing minority operation coverage and to monitor the number of respondents needed to reach the 75 percent county response rate. When the required number of completions was achieved for a given county, LRC activity was suspended in that county.

**Last Call Nonresponse Follow-up.** The Last Call Nonresponse Follow-up activity was utilized to increase the national response rate to 80 percent. All remaining nonresponse records with an expected value of sales greater than \$50,000 in counties that had not achieved a 75-percent response rate were eligible for this phone follow-up activity. CATI was used for this activity and began in mid-July 2013 and lasted until August 1, 2013. Automated procedures were employed to monitor the number of respondents needed and completed. When a 75 percent response rate was achieved for a given county, follow-up in that county was suspended. NASS achieved its goal of an 80-percent national response rate utilizing Last Call Nonresponse Follow-up.

**Not on the Mail List (NML) Follow-up.** To account for farming operations not on the CML, NASS used its 2012 JAS supplemented sample from the NASS area frame. The NASS area frame covers all land in the U.S. with the exception of Alaska and includes all farms. As previously described, the NASS conducted a record linkage operation between the CML records and the records from the 2012 JAS. Those 2012 JAS records that did not match records on the CML were designated as “Not on the Mail List (NML)” records. These records were mailed a yellow census form so that it could be differentiated from the green forms mailed to CML records. The NML records were mailed at the same time as the census mailing and received the same follow-up procedures as the census mailing through the first follow-up in mid-February 2013. Beginning in March 2013, CATI was used for nonresponse follow-up for NML nonrespondents.

## Replication

Replication is utilized to improve efficiency and reduce respondent burden. To adjust for nonresponse associated with criteria records in the 2007 Census of Agriculture, NASS replicated a set of respondents determined to be in-scope from the last mailing of the Agricultural Identification Survey (AIS), conducted in December 2006. The replicated records represented operations that were relatively small in size and homogeneous in nature. Replicated records were assumed to be in-scope, based on their AIS reported data.

For the 2012 Census of Agriculture, a first mailing was sent to the criteria records, a subpopulation consisting of all of the approximately 74,000 respondents to the 2011 NACS mailing. This included pre-notification using a pre-recorded message, the first mailing, and the thank-you reminder post card. No further follow-up efforts were conducted on this subpopulation. As in 2007, the agricultural operations in this subpopulation were relatively small in size and homogeneous in nature. The responses from the criteria records were used to estimate the in-scope rate for the 20,168 nonrespondents from this subpopulation.

Records were selected randomly for replication or coding as out-of-scope based on the estimated in-scope rate. The use of the in-scope rate after one mailing is supported by analysis of 2007 census data, which indicated the early in-scope rate was a reasonable proxy for the in-scope rate for the subpopulation of criteria records that did not respond to the NACS immediately preceding the census mailing. Of the 20,168 NACS records with no response, 16,762 records were selected to be in-scope.

Data relationships between the 2012 responses and their respective NACS data were applied to the NACS data for the nonrespondents selected to be in-scope to derive values to seed replication. Then replication was conducted through imputation.

Criteria records with no response to the December 2011 NACS were excluded in the capture-recapture adjustments for coverage, response, or correct classification. The in-scope records were each given an initial weight of one. However, for calibration, the

replicated in-scope records were eligible for a coverage adjustment.

## REPORT FORM PROCESSING

### Data Capture

The Census Bureau's National Processing Center (NPC) in Jeffersonville, IN was contracted to process returned mail packets. NASS staff on site at the NPC provided technical guidance and monitored NPC processing activities. All report forms returned to the NPC were immediately checked in, using bar codes printed on the mailing label, and removed from follow-up report form mailings. All forms with any data were scanned and an image was made of each page of a report form. Optical Mark Recognition (OMR) was used to capture categorical responses and to identify the other answer zones in which some type of mark was present.

Data entry operators keyed data from the scanned images using OMR results that highlighted the areas of the report forms with respondent entries. The keyer evaluated the contents and captured pertinent responses. Ten percent of the captured data were keyed a second time for quality control. If differences existed between the first keyed value and the second, an adjudicator handled resolution. The decision of the adjudicator was used to grade the performance of the keyers, who were required to maintain a certain accuracy level.

The images and the captured data were transferred to NASS's centralized network and became available to field offices and headquarters on a flow basis. The images were available for use in all stages of review. Images were computer generated for reports obtained from the telephone interviews and the Internet.

### Editing Data

Captured data were processed through a computer formatting program, which verified that records were valid – that the record identification number was on the list of census records, that the reported counties of operation and production were valid, and other related criteria. Rejected records were referred to

analysts for correction. Accepted records were sent to a complex computer batch edit process. Each execution of the computer edit in batch mode consisted of records from only one State and flowed as the data were received from the NPC, the NASS Electronic Data Reporting (EDR) web utility, or the Computer-Assisted Telephone Interview (CATI) applications.

The computer edit determined whether a reporting operation met the qualifying criteria to be counted as a farm (in-scope). The edit examined each in-scope record for reasonableness and completeness and determined whether to accept the recorded value for each data item or to take corrective action. Such corrective actions included removing erroneously reported values, replacing an unreasonable value with one consistent with other reported data, or providing a value for an overlooked item. To the extent possible, the computer edit determined a replacement value. Strategies for determining replacement values are discussed in the next section. Operations failing to meet the qualifying criteria were categorized as out-of-scope for the census; that is, they were classified as being a nonfarm. Out-of-scope records that NASS had reason to believe might be in-scope (indications of recent and/or significant agricultural activity reported on NASS surveys, for example) were referred to analysts for verification.

The edit systematically checked reported data section-by-section with the overall objective of achieving an internally consistent and complete report. NASS subject-matter experts had previously defined the criteria for acceptable data. Problems that could not be resolved within the edit were referred to an analyst for intervention. Prior to the census mailout, NASS established a group of 90 analysts in a Census Editing Unit in the National Operations Center in St. Louis, MO who examined the scanned images, consulted additional sources of information, and determined an appropriate action. Field office analysts also participated using an interactive version of the edit program to submit corrected data and immediately re-edit the record to ensure a satisfactory solution.

## Imputing Data

The edit determined the best value to impute for reported responses that were deemed unreasonable

and for required responses that were absent. If an item could not be calculated directly from other current responses, the edit determined whether acreage, production or inventory items had been reported for that farm on a recent NASS crop or livestock survey. For operators who had not changed in five years, demographic variables such as race and sex were taken from the previous census. Administrative data from the Farm Service Agency were used for a few items, such as Conservation Reserve Program acreage. When deterministic edit logic and previously-reported data sources proved inadequate, data from a reporting farm of similar type, size, and location (a donor farm) were considered. In cases where automated imputation was unable to provide a consistent report, the record was referred to an analyst for resolution.

Separate system processes were established to efficiently provide data from a similar farm to the edit when donor imputation was required. The farm characteristics used to define similarity between a recipient record and its donor record were determined dynamically by the edit logic. Euclidean distance was used for similarity computations, with each contributing similarity characteristic scaled appropriately. The most similar farm based on this criterion (the “nearest neighbor”) was identified and returned to the edit for use as a donor. The calculated distance between the centroids of the principal counties of production of the donor and recipient was always included as one of the measures of similarity.

To provide donors to the automated edit, a pool of successfully edited records was maintained for each section of the report form. These donor pools began with 2007 census data, reconfigured to emulate 2012 data and then edited using 2012 logic. Data from the 2010 Census Content Test were similarly remapped and edited before being added to the original donor pools. As 2012 records were successfully processed, they were added to the donor pools, which maintained the most recent data for each farm. Donor pools were updated approximately every other week, as determined by edit processing schedules. After several updates, all initial data records were dropped, leaving only 2012 records in the donor pools. After each update, donor pool records were grouped into strata containing farms in the same state of similar type and size, using a data-



driven algorithm to define strata. Certain American Indian farms were treated as a separate group, effectively having their own donor pool.

In response to each donor request issued by the edit, a dedicated system process would search the appropriate stratum and respond with the most similar donor, while giving preference to more recent donors. In relatively rare instances where it was unable to provide a donor, the donor selection process issued an appropriate failure message to the edit. Imputation failures occurred for several different reasons. The requirement that an imputed value be positive could have ruled out all available donors, as could have the necessity for the donor record to satisfy a particular constraint – say, that the donor record has cattle, but no milk cows. In general, an imputation failure occurred if there was no satisfactory donor in the same profile as the report being edited. Records with imputation failures were either held until more records were available in the donor pool or referred to an analyst. In addition, when such a failure occurred in finding a donor for expenditure data, a program provided values from a table of donor pool averages in lieu of values from an individual donor, wherever possible. This ‘failover’ utility was new for the 2012 census imputation process, and significantly reduced the number of imputation failures among the expenditure and labor variables. During the early stages of editing, records requiring imputation for production (and hence yields) of field crops or hay, land values, or certain expenditure variables were set aside or “parked.” These records were edited when the donor pools contained only 2012 records, ensuring that 2012 data were used in imputations for these variables.

After receiving a donor's data, the edit substituted the values into the edited record. In many cases, the donor record's data value was scaled using another data field specified in the edit logic. In such cases, the size of the auxiliary field's value in the edited record, relative to its value in the donor record, was used to inflate or reduce the donor record's value for the imputed field. The imputed data were then validated by the same edit logic to which reported data were subject. Since imputation was conducted independently for each occurrence, reports requiring multiple imputations may have drawn from multiple donors.

## Data Analysis

The complex edit ensured the full internal consistency of the record. Successfully completing the edit did not provide insight as to whether the report was reasonable compared to other reports in the county. Analysts were provided an additional set of tools, in the form of listings and graphs, to review record-level data across farms. These examinations revealed extreme outliers, large and small, or unique data distribution patterns that were possibly a result of reporting, recording, or handling errors. Potential problems were researched and, when necessary, corrections were made and the record interactively edited again.

When NASS summarizes the census of agriculture, it assigns the data from an individual report to the “principal” county. The principal county is based on the operator's response to a census question and is the one county in which the majority of agricultural products are produced. Because some large operations have significant production in multiple counties, some reports were broken up into multiple source counties, to more accurately allocate the data. Similarly, large farms operating in more than one State were treated as distinct, state-specific operations. A separate report form was completed for each county or State and a separate record was added.

## ACCOUNTING FOR UNDERCOVERAGE, NONRESPONSE, AND MISCLASSIFICATION

Although much effort was expended making the CML as complete as possible, the CML did not include all U.S. farms, resulting in list undercoverage. Some farm operators who were on the CML did not respond to the census, despite numerous attempts to contact them. In addition, although each operation was classified as a farm or a nonfarm based on the responses to the census report form, some were misclassified; that is, some nonfarms were classified as farms and some farms were classified as nonfarms. NASS's goal was to produce agricultural census totals for publication that were fully adjusted for list undercoverage, nonresponse and misclassification at the county level.

In the 2007 Census of Agriculture, adjustments for undercoverage and nonresponse were estimated independently. In 2007, as in earlier censuses, the NASS area frame was used to adjust for undercoverage. This process assumed that the area frame provided complete coverage and that all operations were correctly classified as farm/nonfarm. To determine the extent of undercoverage in 2007, the CML records were matched to the area-frame tracts designated as agricultural, non-agricultural with potential, or non-agricultural with potential unknown in June. The area-frame tracts that did not match a CML record were designated as being in the Not on the Mail List (NML) domain. In 2007, tracts that were determined to be non-agricultural without potential during the pre-screening phase of the June Agricultural Survey (JAS) were not considered in the NML domain construction. The NML domain tracts were sent a census form and, if a tract was associated with a farm, then that farm contributed to the correction for undercoverage.

To adjust for nonresponse in 2007, each responding CML record was given a probability of being a farm using a classification tree. The inverse of this probability became the nonresponse weight for that record. For undercoverage, the adjustment provided State-level values. A State-level estimate was based on the weighted sum of the responders with an adjustment for the non-responders within that State plus the State-level undercoverage adjustment. Because State-level farm count estimates based on this two-step process sometimes had high standard errors and apparent biases, the national-level adjusted estimates were smoothed across States, producing initial State-level farm operation coverage targets.

Research following the 2007 Census of Agriculture led to the realization that some area-frame operations were misclassified as farm/nonfarm, which was in conflict with the previous assumption that the JAS farm classification was the accurate classification. Further, because nonresponse could only occur if the operation was on the CML, undercoverage and nonresponse were dependent. Thus in 2012, NASS used capture-recapture methodology to adjust for undercoverage, nonresponse, and misclassification. To implement capture-recapture methods, two independent surveys were required. The 2012 Census of Agriculture (based on the CML) and the

2012 JAS (based on the area frame) were those two surveys. Historically, NASS has been careful to maintain the independence of these two surveys.

A second assumption was that the proportion of JAS farms with a given set of characteristics captured by the census was equal to the proportion of U.S. farms with those same characteristics captured by the census.

For a farm to be identified as a farm, and thus captured by the census, it must be on the CML, respond to the census report form and, based on the census response, be classified as a farm; that is, the capture probability  $\pi_C$  is of interest:

$$\pi_C = \pi(\text{CML, Responded, Farm on Census} | \text{Farm})$$

Two types of classification error can occur. First, a farm can be misclassified as a nonfarm. This type of misclassification is accounted for in determining the probability of capture  $\pi_C$ . The second type of classification error results when a response to the census is classified as a farm operation when it does not meet the definition of a farm. That is, some farms on the CML may be misclassified from their census report response and may be nonfarms. To account for the misclassification of nonfarms as farms, the probability of a farm on the census being classified correctly must be estimated; that is,

$$\pi_{CCFC} = \pi(\text{Farm} | \text{Farm on Census})$$

where *CCFC* represents Correct Census Farm Classification. To adjust for undercoverage, nonresponse, and misclassification, each CML record classified as a farm based on its response to the census report form was given a weight of the ratio of the estimated probability of correct classification of a farm on the census and the estimated probability of capture ( $\hat{\pi}_{CCFC} / \hat{\pi}_C$  where the hat symbol ( $\hat{\cdot}$ ) denotes an estimate). To estimate the number of farms with a given set of characteristics, the weights of CML records responding as farms on the census and having that set of characteristics were summed. This estimator is

referred to as the capture-recapture estimator (*CR*):

$$CR = \sum_{i \in F} \frac{\hat{\pi}_{CCFC,i}}{\hat{\pi}_{C,i}}$$

where *F* is the set of all CML records classified as farms based on their responses to the census questionnaire.

To estimate the capture and correct census farm classification probabilities, a matched dataset consisting of JAS records and census records was created. Records in the 2012 JAS sample were matched to the 2012 census using probabilistic record linkage. The CML records that matched with JAS tracts represent the Census sample. Note: The Census Sample is a subset of the CML records and includes only those records matching a JAS tract. Both agricultural and non-agricultural tracts were included in the matched dataset. (This differs from the 2007 processes, which considered only the agricultural tracts and non-agricultural tracts with potential or with potential unknown. It also included CML records that responded to the census as a farm or nonfarm and CML records that did not respond to the census.)

## Resolving Farm Status

The farm status based on census responses to either the CML or NML census data collection and the JAS agreed in most cases; these records are referred to as having resolved farm status. However, in other cases, a record was identified as a farm (nonfarm) on the JAS and as a nonfarm (farm) by the census through either the CML or the NML. Such records are said to have conflicting or unresolved farm status. An operation identified as a farm is referred to as in-scope; one identified as a nonfarm is referred to as out-of-scope. From the set of matched records, three groups with conflicting farm status were identified: 1) in-scope JAS records that were out-of-scope on the census and 2) census in-scope and JAS out-of-scope records, and 3) in-scope JAS records that did not have a census response. The records with conflicting farm status were sent to regional field offices for review. In each case, efforts were made to determine whether (1) the status had changed between June and December when the

census was conducted, (2) the JAS farm status was correct, (3) the census farm status was correct, (4) the records were incorrectly matched, or (5) the farm status could not be resolved. Not all of the records with conflicting farm status could be resolved. In 2012, 11.6 percent of the records in the Census Sample had unresolved farm status. Of these, 18.9 percent were from nonresponse to the census report form.

The probability an operation is a farm was estimated for the records with unresolved farm status. Using the 2012 matched dataset, a logistic model of the probability an operation is a farm based on the records with resolved farm status was developed; that is, the operations where the farm (or nonfarm) status agreed between the JAS and the census were used to develop a missing data model, which was then used to resolve farm status. The final missing data model was used to impute the probability that each of the agricultural operations with unresolved farm status is a farm. For the resolved farms and nonfarms, the probability of the operation being a farm was 1 and 0, respectively. Five-fold cross-validation was used to develop and to compare competing models. The accuracy of the model was thereby not overstated due to fitting and evaluating the model on the same set of data. To ensure that each of the cross-validation samples covered the U.S., the five cross-validation samples of JAS segments were drawn within State-stratum combinations. Characteristics of the JAS tracts were considered as potential covariates in the model. Because limited information is available for JAS nonfarm tracts, county-level socio-demographic variables from the most recent U.S. population census were also considered. The sample weight associated with each JAS tract was multiplied by the probability of being a farm. This adjusted weight was used in all subsequent modeling.

## Capture Probabilities

Recall that, for a farm to be identified as a farm, and thus captured, by the census, it must be on the CML, respond to the census report form and, based on the census response, be classified as a farm. These adjustments are dependent so that the probability of capture  $\pi_C$  may be written as

$$\pi_C = \pi(\text{CML, Responded, Farm on Census}|\text{Farm}) = \pi(\text{CML}|\text{Farm})\pi(\text{Responded}|\text{CML, Farm})\pi(\text{Farm on Census}|\text{CML, Responded, Farm})$$

The probability of capturing a farm depends on the characteristics of the farm. Using five-fold cross-validation, three logistic models were developed based on the matched dataset. The first model estimated the probability of a farm being on the CML. The second model estimated the probability that a farm on the CML responded to the census report form. The final model estimated the probability that a farm that was on the CML and responded to the census was identified as a farm based on its response. The probability that a farm is captured by the census of agriculture is then the product of the three conditional probabilities that a farm is on the CML, responds, and is identified as a farm.

Note 1: Responses were required for Must cases. These operations were only included in modeling the probability of a farm being on the CML. Consequently, the weight associated with a Must record was the reciprocal of the probability of a farm being on the CML.

Note 2: Two sets of models were created. One set estimated the probability of capture for Texas farms. The other set provided estimated capture probabilities for farms in the remaining States, except for Alaska.

Note 3: Because Alaska is not included in the JAS and thus has no area frame, the Alaskan agricultural operations were not included in the capture-recapture process. No adjustments were made for undercoverage or misclassification. To account for nonresponse, the CML records were divided into three groups: (1) the Must records, (2) the Criteria Records, and (3) the remaining CML records. The must records received a weight of one, thereby receiving no adjustment for nonresponse. The probability of response for each of the other two groups was the proportion of responders within the group. Each record within the group was then given a weight equal to the reciprocal of the probability of response.

## Misclassification

An operation is misclassified if (1) it meets the definition of a farm, but is classified as a nonfarm on the census or (2) it does not meet the definition of a farm, but is classified as a farm on the census. The first type of misclassification is accounted for when modeling the probability of capture. An adjustment is still needed for the misclassification of nonfarms as farms. As with farm status and capture, the probability of this misclassification depends on an operation's characteristics. Thus, a final logistic model was developed. Given that an operation was classified as a farm on the CML, the probability of its being a farm was modeled based on its characteristics. Five-fold cross-validation was used to ensure that the model was not over-fitted.

## CALIBRATION

Each operation identified as being in-scope on the CML was given a weight equal to the probability of misclassification divided by the probability of capture. This weight accounted for undercoverage, nonresponse, and both types of misclassification.

The record weighting processes were initially applied at the State level to produce adjusted estimates of farm numbers and land in farms for 63 different categories of 8 characteristics of the farm operation or the farm operator -- value of agricultural sales (8); age (2); female; race (4); Hispanic origin of principal farm operator ; 4 sales categories for each of 10 major commodities (40); and farm type groups (7). The State-level number of farms and land in farms were two additional adjusted estimates, resulting in 65 categories. To reduce the intercensal variation at the State level, the State targets were smoothed by averaging the 2012 estimates from capture-recapture and the published 2007 state estimates with the restrictions that the smoothed targets were within one standard error of the capture-recapture estimates. The smoothed State targets were rescaled so that they summed to the national capture-recapture estimates.

These State estimates were general purpose in that they did not provide any control over expected levels of commodity production of the individual farm operation. As a result of this limitation, the

procedures could have over-adjusted or under-adjusted for commodity production. To address this, a second set of variables, known as commodity targets, was added to the calibration algorithm. These targets were commodity totals from administrative sources or from NASS surveys of nonfarm populations (e.g. USDA Farm Service Agency program data, Agricultural Marketing Service market orders, livestock slaughter data, cotton ginning data). The introduction of these commodity coverage targets strengthened the overall adjustment procedure by ensuring that major commodity totals remained within reasonable bounds of established benchmarks. Commodity coverage targets with acceptable ranges were established by subject-matter experts for each State, with New England treated as a State.

Each State was calibrated separately. The calibration algorithm addressed commodity coverage. The algorithm was controlled by the 65 State farm operation coverage targets and the State commodity coverage targets. To ensure that the calibration process converged with so many constraints, it was desirable to provide some tolerance ranges for each target. Although full calibration to a single point estimate would assure that the weighted total among census respondents equaled its target for each calibration variable in either set, it was not always possible to calibrate to such a large number of target values while ensuring that farm weights were within a reasonable range and not less than one. Because of this and because calibration targets are estimates themselves subject to uncertainty, NASS allowed some tolerance in the determination of the adjusted weights. Rather than forcing the total for each calibration variable computed using the adjusted weights to equal a specific amount, NASS allowed the estimated total to fall within a tolerance range. This tolerance strategy made it possible for the calibration algorithm to produce a set of satisfactory, adjusted weights.

Ranges for the farm operation coverage targets were determined differently from the commodity targets. The State target for number of farms had no tolerance range. The tolerance range for the 64 other State farm operation coverage targets was the estimated smoothed State total for the variable plus or minus one-half of the standard error of the capture-recapture estimate. This choice limited the

cumulative deviation from the estimated total for a variable when State totals were summed to a U.S. level total. The commodity target tolerance ranges were determined by subject-matter experts, based on the amount of confidence in the source, and usually were less than plus or minus two percent of the target. Ranges were not necessarily symmetric around the target value.

Census data collection was assumed to be complete for very large and unique farms with their weight being controlled to 1 during the calibration adjustment process. For all other farms, adjustment weights were obtained using truncated linear calibration which forced the final census record weights to fall in the interval [1,6]. Adjustments began with the nonresponse and misclassification adjusted weights. Through calibration, a second stage weight that simultaneously satisfied all farm operation coverage and commodity coverage calibration targets was obtained. Calibration was seldom able to adjust weights so that all State targets were met. Within the calibration process, the highest priority for meeting a target was given to the number of farms, total land in farms, and top cash-receipt commodities accounting for 80 percent of the State's production. All remaining targets associated with commodities and characteristics of farms and farm operators had equal priority. If a value within the tolerance range of any variable could not be achieved in a given State, the variable was removed as a target in that State and the calibration algorithm was rerun.

Weight computations in the final algorithms were performed to several decimals. Thus, the fully-adjusted weights were non-integer numbers. To ensure that all subdomains for which NASS publishes summed to their grand total, fully-adjusted weights were integerized. This eliminated the need for rounding individual cell values and ensured that marginal totals always added correctly to the grand total. As an example of how the integerization process worked, assume there were five census records in a county with final noninteger coverage weights of 2.2, for a total of 11. The integerization process randomly selected four of these records and rounded their final weight down to 2.0 and rounded the fifth record up to 3.0, for a total of 11.

The proportions of selected census data items that

are due to coverage, response, and classification adjustments are displayed in Tables A and C.

## **DISCLOSURE REVIEW**

After tabulation and review of the aggregates, a comprehensive disclosure review was conducted. NASS is obligated to withhold, under Title 7, U.S. Code, any total that would reveal an individual's information or allow it to be closely estimated by the public. Cell suppression was used to protect the cells that were determined to be sensitive to a disclosure of information. Farm counts are not considered sensitive and are not subject to disclosure controls.

Based on agency standards, data cells were determined to be sensitive to a disclosure of information if they violated either of two criteria rules. The threshold rule was violated if the data cell contained less than three operations. For example, if only one farmer produced turkeys in a county, NASS could not publish the county total for turkey inventory without disclosing that individual's information. The dominance rule was violated if the distribution of the data within the cell allowed a data user to estimate any respondent's data too closely. For example, if there are many farmers producing turkeys in a county and some of them were large enough to dominate the cell total, NASS could not publish the county total for turkey inventory without risking disclosing an individual respondent's data. In both of these situations, the data were suppressed and a "(D)" was placed in the cell in the census publication table. These data cells were referred to as primary suppressions.

Since most items were summed to marginal totals, primary suppressions within these summation relationships were protected by ensuring that there were additional suppressions within the linear relationship that provided adequate protection for the primary. A detailed computer routine selected additional data cells for suppression to ensure all primary suppressions were properly protected in all linear relationships in all tables. These data cells were referred to as complementary suppressions. These cells were not themselves sensitive to a disclosure of information but were suppressed to protect other primary suppressions. A "(D)" was also placed in the cell of the census publication table to

indicate a complementary suppression. A data user could not determine whether a cell with a (D) represented a primary or a complementary suppression.

Field office analysts reviewed all complementary suppressions to ensure no cells had been withheld that were vital to the data users. In instances where complimentary suppressions were deemed critically important to a State or county, analysts requested an override and a different complementary cell was chosen.

## **CENSUS QUALITY**

The purpose of the census of agriculture is to account for "any place from which \$1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the census year." To accomplish this, NASS develops a CML that contains identifying information for operations that have an indication of meeting the census definition, develops procedures to collect agricultural information from those records, establishes criteria for analyst review of the data, creates computer routines to correct or complete the requested information, and provides census estimates of the characteristics of farms and farm operators with associated measures of uncertainty.

It is not likely that either the CML includes all operations that meet the definition of a farm or that all those that do meet the definition of a farm respond to the census inquiry. The goal is to publish data with a high level of quality. There are many ways to measure the quality of a census.

One of the first indicators used is a measure of the response to the census data collection as it has generally been thought that a high response rate indicates more complete coverage of the population of interest. This is a valid assumption if the enumeration list, the CML here, has complete coverage of the population of interest. In the case of the census of agriculture, the definition requiring advance knowledge of sales makes achieving a high level of coverage difficult. To ensure that the census of agriculture is as complete as possible, records are included that might not meet the census definition of a farm – in fact, almost 50 percent more records than the anticipated number of qualifying farm operations

were included in the 2012 CML. A second indicator of quality then is the coverage of the farm population by the CML. Other indicators of quality relate to the accuracy and completeness of the data, and the validity of the procedures used in processing the data.

In some cases, NASS was able to produce measures of quality – such as the response rate to the data collection, the coverage of the census mail list, and the variability of the final adjusted estimates. In other cases, measures were not produced but descriptions of procedures that NASS used to reduce errors from the procedures were subsequently provided.

### Census Response Rate

The response rate is one indicator of the quality of a data collection. It is generally assumed that if a response rate is close to a full participation level of 100 percent, the potential for nonresponse bias is small, although this has been questioned recently in the literature. Because the CML contains both farm and nonfarm records, the response rate is an indicator of replying to the census data collection effort, but does not reflect whether those responding met the farm definition. The response rate for the 2012 Census of Agriculture CML is 80.1 percent as compared with a response rate of 85.2 percent for the 2007 Census of Agriculture and 88.0 for the 2002 Census of Agriculture.

The 2012 Census of Agriculture response rate used the fourth response rate formula from the American Association of Public Opinion Research Response Rate Standard Definitions manual:

$$RR4 = \frac{C_{adj}}{C_{adj} + R + NC + O + Replicated + e(U)} \quad (100)$$

where

- $C_{adj}$  = number of fully and partially completed records, excluding replicated records
- $R$  = number of explicit refusals
- $NC$  = number of non-contacted operations
- $O$  = number of other types of nonrespondents
- $Replicated$  = number of replicated records
- $U$  = number of operations of unknown eligibility

$e(U)$  = estimated number of operations of unknown eligibility assumed to be eligible

Records were classified into the above variables based on the combination of their active status (AS) codes, in-scope status, and replication status. Active status refers to the eligibility status of records for selection on the CML. All replicated records were considered to be a form of nonresponse and were classified into other nonrespondents; in-scope status was considered immaterial.

Certain active status classifications indicated records of unknown agricultural status. These classifications included records to be removed from the CML but had data from outside sources indicating agricultural activity, new records from outside data sources, nonrespondents and refusals to the NACS, records for regional office handling only, and records with Farm Service Agency or Conservation Reserve Program data on operations that are not owned by the principal operator. These records were stratified (grouped) based on their probabilities of being in-scope had they responded. The estimated number of in-scope nonrespondents was calculated for the  $h$ th stratum (group) by the following formula:

$$e(U_h) = \left( \frac{C_{in-scope,h}}{C_h} \right) U_h$$

where

- $e(U_h)$  = estimated number of operations of unknown eligibility assumed to be eligible in the  $h$ th group
- $C_{in-scope,h}$  = the number of completed and in-scope census records in the  $h$ th group
- $C_h$  = the number of completed census records in the  $h$ th group
- $U_h$  = number of operations of unknown eligibility in the  $h$ th group

### Census Coverage

As a side-product of the statistical adjustment used to account for undercoverage, nonresponse of farms on the CML, and misclassification of responses to the census, the proportion of the adjustments due to each of those factors can be derived. The percentages of final census estimates due to adjustments for

undercoverage, nonresponse, and misclassification as well as the total percent adjustment for selected items are displayed in Tables A and C.

## MEASURED ERRORS IN THE CENSUS PROCESS

Although the census of agriculture does not inherently rely on a sample, it uses statistical procedures in compiling the CML, in its data collection procedures, in data editing and processing, and in compiling the final data. Additionally, it uses statistical procedures to both measure errors in the various processes and in making adjustments for those errors in the final data. One example is the statistical process used to account for undercoverage, nonresponse of farms on the CML, and misclassification of responses to the census. The basis of the undercoverage adjustment is the capture-recapture procedure that uses the area sample enumeration from the June Agricultural Survey. The largest contribution to error in the census estimates is due to the adjustments for nonresponse, undercoverage, misclassification, calibration and integerization.

### Variability in Census Estimates due to Statistical Adjustment

In conducting the 2012 Census of Agriculture, efforts were initiated to measure error associated with the adjustments for farm operations that were not on the CML, for farm operations that were on the CML but did not respond to the census report form, for farms and nonfarms that were misclassified as nonfarms and farms, respectively, for calibration, and for integerization. These error measurements were developed from the standard error of the estimates at the national, State, and county levels and were expressed as coefficients of variation (CVs) at the national and State levels and as generalized coefficients of variation (GCVs) at the county levels.

The standard error of an estimate is an estimate of the standard deviation of the sampling distribution of the estimator. Because Texas and Alaska were modeled separately from the other States, the variances of a national-level data item for these two States were computed separately and added to the

variance of that data item for the rest of the U.S. The standard error was then the square root of the total variance. In each case, standard errors were computed using the group jackknife approach. To conduct the jackknifing,  $k$  mutually exclusive and exhaustive groups of JAS segments were formed. The groups were selected using a stratified random design so that each group reflected the survey design, including State and agricultural strata within a State. In turn, each group,  $j = 1, 2, \dots, k$ , was deleted and the capture-recapture estimate  $CR_i^{(j)}$  was computed for each data item  $i$  at the specified geographical level, such as nation, State, or county, using the remaining  $(k - 1)$  groups. Estimates of the variance and standard error associated with the capture-recapture estimate  $CR_i$  are then, respectively,

$$\sigma_i^2 = \frac{k-1}{k} \sum_{j=1}^k (CR_i^{(j)} - CR_i)^2; \quad SE(CR_i) = \sqrt{\sigma_i^2}$$

Increasing  $k$  improves the estimate of the variance but, as  $k$  increases, the observations become too sparse to reflect the survey design and to provide country-wide coverage. Based on 2007 data,  $k = 10$  was determined to be the largest number of groups that could be formed and still have each group provide adequate coverage within all States and agricultural strata. Thus, 10 jackknife groups were used to provide standard errors for 2012 State and national estimates. To capture the additional variability from calibration and integerization, the standard errors were computed using the calibrated, integerized capture-recapture estimates from the jackknife groups. For the estimate of the number of farms with a given set of characteristics, only the CML records with those characteristics were used to obtain the overall estimate as well as the estimates from each jackknife group.

When the constraints of the calibration process produced an artificially small standard error, the more conservative capture-recapture standard error was used. Note that the jackknife groups must only be constructed once, and different subsets of the records were used to compute estimates and standard errors for the data items.

The CV is a measure of the relative amount of error



associated with the sample estimate:

$$CV = \frac{SE(CR_i)}{CR_i} 100\%$$

where  $SE(CR_i)$  is the standard error of the capture-recapture estimate for data item  $i$ . This relative measure allows the reliability of a range of estimates to be compared. For example, the standard error is often larger for large population estimates than for small population estimates, but the large population estimates may have a smaller CV, indicating a more reliable estimate. For county-level estimates, a generalized coefficient of variation (GCVs) was determined for each estimate within a State. A generalized variance function relates a function of the variance of an estimator to a function of the estimator. Within a State, the standard error of an estimate for a data item was often found to be linearly related to the estimate of that item with an intercept of zero. Based on this modeled relationship, the GCV is the slope of the line relating the standard error to the estimate, multiplied times 100 to represent the GCV as a percentage.

The standard error is the product of the CV (or GCV for county estimates) and the estimate divided by 100. As an example, if the GCV for a State is 25 percent and a county's estimate is 4, then the standard error is  $25(4)/100 = 1$ . The standard error of an estimated data item from the census provides a measure of the error variation in the value of that estimated data item based on the possible outcomes of the census collection, including variants as to who was on the CML, who returned a census form, who was misclassified either as a farm or as a nonfarm, and the uncertainty associated with calibration and integerization. With 95 percent confidence, an estimate is within two standard errors of the true value being estimated. For this example, with 95 percent confidence, the estimate of 4 is within  $2(1) = 2$  of the true county value.

Table B presents the fully adjusted estimates with the coefficient of variation for selected items.

## NONMEASURED ERRORS IN THE CENSUS PROCESS

As noted in the previous section, sampling errors can

be introduced from the coverage, nonresponse and misclassification adjustment procedures. This error is measurable. However, nonsampling errors are imbedded in the census process that cannot be directly measured as part of the design of the census but must be contained to ensure an accurate count. Extensive efforts were made to compile a complete and accurate mail list for the census, to elicit response to the census, to design an understandable report form with clear instructions, to minimize processing errors through the use of quality control measures, to reduce matching error associated with the capture-recapture estimation process, and to minimize error associated with identification of a respondent as a farm operation (referred to as classification error). The weight adjustment and tabulation processes recognize the presence of nonsampling errors; however, it is assumed that these errors are small and that, in total, the net effect is zero. In other words, the positive errors cancel the negative errors.

## Respondent and Enumerator Error

Incorrect or incomplete responses to the census report form or to the questions posed by an enumerator can introduce error into the census data. Steps were taken in the design and execution of the census of agriculture to reduce errors from respondent reporting. Poor instructions and ambiguous definitions lead to misreporting. Respondents may not remember accurately, may give rounded numbers, or may record an item in the wrong cell. To reduce reporting and recording errors, the report form was tested prior to the census using industry accepted cognitive testing procedures. Detailed instructions for completing the report form were provided to each respondent. Questions were phrased as clearly as possible based on previous tests of the report form. Computer-assisted telephone interviewing software included immediate integrity checks of recorded responses so suspect data could be verified or corrected. In addition, each respondent's answers were checked for completeness and consistency by the complex edit and imputation system.

## Processing Error

Processing of each census report form was another potential source of nonsampling error. All mail

returns that included multiple reports, respondent remarks, or that were marked out of business and report forms with no reported data were sent to an analyst for verification and appropriate action. Integrity checks were performed by the imaging system and data transfer functions. Standard quality control procedures were in place that required that randomly selected batches of data keyed from image be re-entered by a different operator to verify the work and evaluate key entry operators. All systems and programs were thoroughly tested before going on-line and were monitored throughout the processing period.

Developing accurate processing methods is complicated by the complex structure of agriculture. Among the complexities are the many places to be included, the variety of arrangements under which farms are operated, the continuing changes in the relationship of operators to the farm operated, the expiration of leases and the initiation or renewal of leases, the problem of obtaining a complete list of agriculture operations, the difficulty of contacting and identifying some types of contractor/contractee relationships, the operator's absence from the farm during the data collection period, and the operator's opinion that part or all of the operation does not qualify and should not be included in the census. During data collection and processing of the census, all operations underwent a number of quality control checks to ensure results were as accurate as possible.

### **Item Nonresponse**

All item nonresponse actions provide another opportunity to introduce measurement errors. Regardless of whether it was previously reported data, administrative data, the nearest neighbor algorithm, or manually imputed by an analyst, some risk exists that the imputed value does not equal the actual value. Previously reported and administrative data were used only when they related to the census reference period. A new nearest neighbor was randomly selected for each incident to eliminate the chance of a consistent bias.

### **Record Matching Error**

The process of building and expanding the CML involves finding new list sources and checking for

names not on the list. An automated processing system compared each new name to the existing CML names and "linked" like records for the purpose of preventing duplication. New names with strong links to a CML name were discarded and those with no links were added as potential farms. Names with weak links, possible matches, were reviewed by staff to determine whether the new name should be added. Despite this thorough review, some new names may have been erroneously added or deleted. Additions could contribute to duplication (overcoverage) whereas deletions could contribute to undercoverage. As a result, some names received more than one report form, and some farm operators did not receive a report form. Respondents were instructed to complete one form and return all forms so the duplication could be removed.

Another chance for error came when comparing June Agricultural Survey tract operator names to the CML. Area operators whose names were not found on the CML were part of the measure of list incompleteness, or NML. Mistakes in determining overlap status resulted in overcounts (including a tract whose operator was on the CML) or undercounts (excluding a tract whose operator was not on the CML). All tracts determined to not be on the list were triple checked to eliminate, or at least minimize, any error. NML tract operators were mailed a report form printed in a different color. In order to attempt to identify duplication, all respondents who received multiple report forms were instructed to complete the CML version and return all forms so duplication could be removed.

Records in the 2012 JAS were matched to the 2012 census using probabilistic record linkage. The records of operations with unresolved farm status were reviewed by the field offices. If farm status could not be resolved, the probability of an operation being a farm was imputed using a missing data model. The uncertainty associated with this estimate, with the exception of model uncertainty, was accounted for, but errors not found through this process were not.

### **Model Uncertainty Error**

Five logistic models were developed in the process of adjusting the farm numbers for undercoverage,

nonresponse, and misclassification. One model estimated the probability of an agricultural operation with unresolved farm status being a farm. The remaining four models estimated the probability of coverage, response, and correct classification of farms and of nonfarms. Each model was fit independently by two people. For some models, both statisticians obtained the same model. Although the

covariates in the two selected models differed some for the other logistic models, the estimated probabilities were similar, but not identical. The reported standard errors account for the variability in the parameter estimates of the selected models, but not for the additional variation due to model uncertainty. They also do not account for any bias associated with a model.

**Table A. Summary of State Coverage, Nonresponse, and Misclassification Adjustments: 2012**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
Farms .....number	30,961	2,813	37.6	7.9	19.5	10.2
Land in farms .....acres	39,262,613	1,282,136	23.9	2.7	17.5	3.6
<b>Farms by size:</b>						
1 to 9 acres .....farms	689	222	48.5	27.3	5.5	15.7
.....acres	2,143	714	49.8	25.8	6.8	17.2
10 to 49 acres .....farms	2,711	625	57.3	16.5	21.4	19.4
.....acres	77,476	16,256	56.1	16.2	20.8	19.1
50 to 69 acres .....farms	1,063	166	51.9	13.2	19.4	19.3
.....acres	61,655	9,616	51.9	13.1	19.4	19.3
70 to 99 acres .....farms	1,455	200	47.2	12.6	17.3	17.3
.....acres	118,037	16,335	47.2	12.6	17.2	17.3
100 to 139 acres .....farms	1,407	197	47.1	12.9	17.9	16.3
.....acres	163,414	22,818	47.0	12.8	17.9	16.3
140 to 179 acres .....farms	2,632	370	44.6	11.7	17.8	15.1
.....acres	414,560	57,996	44.5	11.7	17.7	15.1
180 to 219 acres .....farms	958	94	43.2	10.7	17.8	14.7
.....acres	189,716	18,602	43.2	10.7	17.8	14.7
220 to 259 acres .....farms	862	78	41.1	10.1	17.2	13.8
.....acres	204,469	18,513	41.1	10.1	17.2	13.8
260 to 499 acres .....farms	4,086	379	36.7	8.5	16.5	11.7
.....acres	1,504,518	136,031	36.4	8.3	16.7	11.5
500 to 999 acres .....farms	4,237	420	38.9	4.6	25.9	8.5
.....acres	3,017,670	296,340	38.4	4.4	25.8	8.2
1,000 to 1,999 acres .....farms	4,424	431	30.8	2.4	24.0	4.4
.....acres	6,333,780	623,311	30.6	2.4	24.0	4.2
2,000 acres or more .....farms	6,437	151	21.6	1.4	18.2	1.9
.....acres	27,175,175	692,905	19.0	1.8	15.6	1.6
<b>Irrigated land use:</b>						
Harvested cropland .....farms	729	98	24.4	4.7	14.3	5.5
.....acres	217,057	11,693	16.7	1.6	12.3	2.8
Pastureland and other land .....farms	33	5	45.5	9.4	22.5	13.5
.....acres	1,350	162	50.5	9.9	21.6	19.1
Market value of agricultural products sold .....\$1,000	10,950,680	336,530	18.7	1.6	15.3	1.8
<b>Farms by value of sales:</b>						
Less than \$1,000 .....farms	9,669	1,276	44.2	13.4	13.3	17.5
.....\$1,000	299	71	60.1	20.3	19.4	20.5
\$1,000 to \$2,499 .....farms	778	111	50.8	13.7	21.1	15.9
.....\$1,000	1,284	182	50.0	13.8	20.4	15.8
\$2,500 to \$4,999 .....farms	976	162	54.0	17.1	20.3	16.6
.....\$1,000	3,639	598	54.3	16.9	20.5	16.8
\$5,000 to \$9,999 .....farms	1,275	217	53.4	16.8	20.4	16.2
.....\$1,000	9,087	1,597	53.5	16.7	20.5	16.2
\$10,000 to \$19,999 .....farms	1,306	226	45.9	9.7	24.1	12.0
.....\$1,000	18,856	3,268	46.2	9.6	24.5	12.1
\$20,000 to \$24,999 .....farms	510	69	44.3	9.4	23.7	11.2
.....1,000	11,387	1,542	44.4	9.4	23.7	11.3
\$25,000 to \$39,999 .....farms	1,170	128	46.2	7.3	25.8	13.1
.....\$1,000	37,667	4,078	46.3	7.3	26.1	12.9
\$40,000 to \$49,999 .....farms	647	65	47.8	6.9	28.8	12.0
.....\$1,000	28,839	2,874	47.7	6.9	28.8	12.0
\$50,000 to \$99,999 .....farms	2,074	210	41.0	5.4	27.1	8.6
.....\$1,000	151,312	15,061	41.5	5.3	27.6	8.6
\$100,000 to \$249,999 .....farms	3,304	221	35.7	3.0	28.5	4.2
.....\$1,000	577,981	37,120	38.0	3.1	30.5	4.4
\$250,000 to \$499,999 .....farms	3,156	178	34.4	2.4	28.5	3.5
.....\$1,000	1,194,099	64,129	36.7	2.5	30.4	3.7
\$500,000 to \$999,999 .....farms	2,913	791	21.9	1.1	19.2	1.7
.....\$1,000	2,163,140	512,495	24.4	1.2	21.3	1.8
\$1,000,000 or more .....farms	3,183	34	10.3	0.9	8.7	0.7
.....\$1,000	6,753,090	128,843	11.0	1.4	9.0	0.7
<b>Net cash farm income of operations (see text):</b>						
<b>Farms with gains of -</b>						
Less than \$1,000 .....farms	687	93	46.1	13.6	15.7	16.8
.....\$1,000	349	46	45.6	13.0	16.3	16.3
\$1,000 to \$4,999 .....farms	2,432	290	44.2	12.4	15.6	16.2
.....\$1,000	6,949	828	44.0	12.2	15.6	16.1
\$5,000 to \$9,999 .....farms	1,916	242	40.2	10.0	16.0	14.1
.....\$1,000	14,133	1,855	40.2	9.8	16.3	14.1
\$10,000 to \$24,999 .....farms	3,487	355	41.2	8.4	19.4	13.4
.....\$1,000	57,943	5,816	41.3	8.3	19.6	13.3
\$25,000 to \$49,999 .....farms	2,554	262	39.5	6.6	21.2	11.6
.....\$1,000	92,843	9,518	39.3	6.6	21.3	11.5
\$50,000 or more .....farms	11,806	885	25.0	2.2	19.6	3.3
.....\$1,000	4,647,092	268,191	16.0	1.3	13.2	1.6
<b>Farms with losses of -</b>						
Less than \$1,000 .....farms	687	102	49.2	14.2	17.3	17.7
.....1,000	317	46	49.3	14.0	17.3	18.1
\$1,000 to \$4,999 .....farms	1,921	302	50.9	15.3	18.9	16.6
.....1,000	5,522	915	51.9	15.7	19.4	16.8
\$5,000 to \$9,999 .....farms	1,624	253	51.1	15.6	18.6	16.9
.....1,000	11,848	1,862	51.4	15.6	18.7	17.1
\$10,000 to \$24,999 .....farms	2,118	380	53.5	16.3	21.3	15.9
.....1,000	33,034	5,964	53.5	16.2	21.6	15.8
\$25,000 to \$49,999 .....farms	714	90	52.9	12.0	28.1	12.9
.....1,000	24,459	3,069	53.3	12.0	28.3	12.9
\$50,000 or more .....farms	1,015	81	40.5	5.9	26.7	7.9
.....1,000	188,911	11,265	31.0	4.7	21.4	4.9
<b>Farms by legal status for tax purposes:</b>						
Family or individual .....farms	27,578	2,582	38.6	8.0	20.2	10.4
.....acres	31,181,498	1,153,895	25.4	2.6	19.0	3.8
Partnership .....farms	2,124	135	29.3	6.4	14.5	8.5
.....acres	5,470,405	330,341	20.3	3.6	13.2	3.5

See footnote(s) at end of table.

--continued

**Table A. Summary of State Coverage, Nonresponse, and Misclassification Adjustments: 2012 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
<b>Farms by legal status for tax purposes: - Con.</b>						
<b>Corporation:</b>						
Family held ..... farms	662	29	22.8	4.1	13.9	4.8
..... acres	1,504,821	34,902	17.8	1.9	13.8	2.1
Other than family held ..... farms	45	7	26.7	8.3	9.8	8.5
..... acres	73,509	7,643	15.9	3.6	9.3	3.1
Other - cooperative, estate or trust, institutional, etc. .... farms	552	67	35.7	11.6	10.4	13.6
..... acres	1,032,380	50,314	5.5	2.9	1.6	1.0
<b>Tenure:</b>						
Full owners ..... farms	15,967	1,926	41.7	11.3	16.0	14.5
..... acres	7,832,469	468,374	27.3	5.2	14.1	8.0
Part owners ..... farms	12,173	625	30.8	3.2	22.8	4.8
..... acres	28,298,194	709,725	22.6	1.9	18.3	2.4
Tenants ..... farms	2,821	308	43.2	7.8	27.2	8.2
..... acres	3,131,950	157,711	27.1	3.5	20.6	3.0
<b>Principal operator characteristics by-</b>						
<b>Sex of operator:</b>						
Male ..... farms	27,728	2,391	37.1	7.4	20.2	9.5
..... acres	37,674,703	1,180,093	23.9	2.5	18.0	3.4
Female ..... farms	3,233	462	41.3	12.6	13.0	15.7
..... acres	1,587,910	129,443	24.2	6.9	9.8	7.4
<b>Primary occupation:</b>						
Farming ..... farms	17,509	1,088	32.3	4.7	21.1	6.4
Other ..... farms	13,452	1,764	44.4	11.8	17.9	14.8
<b>Spanish, Hispanic, or Latino origin (see text) ..... farms</b>						
..... acres	122	72	63.9	20.1	24.1	19.7
..... acres	138,097	25,955	45.5	10.1	27.5	8.0
<b>Race:</b>						
American Indian or Alaska Native ..... farms	348	106	62.4	20.0	20.0	22.3
..... acres	392,080	39,686	27.4	5.2	16.3	5.9
Asian ..... farms	16	(H)	62.5	21.5	13.7	27.3
..... acres	(D)	(D)	(D)	(D)	(D)	(D)
Black or African American ..... farms	3	2	33.3	12.3	15.3	5.7
..... acres	2,230	825	15.9	3.9	10.1	1.9
Native Hawaiian or Other Pacific Islander ..... farms	4	(H)	75.0	1.7	45.8	27.5
..... acres	(D)	(D)	(D)	(D)	(D)	(D)
White ..... farms	30,539	2,664	37.3	7.8	19.4	10.0
..... acres	38,831,224	1,250,547	23.8	2.7	17.6	3.6
More than one race reported ..... farms	51	25	41.2	18.8	15.9	6.5
..... acres	33,855	9,221	32.8	10.1	17.6	5.1
<b>Reporting primary occupation as farming by age group:</b>						
Under 25 years ..... farms	241	33	69.3	22.6	32.6	14.1
25 to 34 years ..... farms	1,468	144	43.8	7.1	30.4	6.3
35 to 44 years ..... farms	2,115	144	35.1	3.9	26.0	5.2
45 to 54 years ..... farms	4,004	239	31.3	3.3	24.3	3.8
55 to 64 years ..... farms	4,951	250	27.5	3.9	18.5	5.1
65 years and over ..... farms	4,730	348	31.3	5.9	14.8	10.7
<b>Reporting primary occupation as other than farming by age group:</b>						
Under 25 years ..... farms	89	30	71.9	26.5	26.8	18.5
25 to 34 years ..... farms	964	268	59.4	19.7	23.1	16.6
35 to 44 years ..... farms	1,457	241	52.3	11.2	26.5	14.5
45 to 54 years ..... farms	2,599	388	47.9	11.8	24.5	11.6
55 to 64 years ..... farms	3,675	383	41.0	11.1	16.6	13.3
65 years and over ..... farms	4,668	510	39.2	10.5	10.2	18.5
<b>All operators by age group <sup>2</sup>:</b>						
Under 25 years ..... farms	746	100	55.9	15.0	29.2	11.7
25 to 34 years ..... farms	4,010	599	47.0	11.5	25.3	10.3
35 to 44 years ..... farms	5,648	551	41.4	7.7	24.4	9.2
45 to 54 years ..... farms	10,120	939	38.2	7.3	23.0	7.8
55 to 64 years ..... farms	12,132	889	33.8	7.2	17.7	8.9
65 to 74 years ..... farms	7,069	622	35.0	8.4	13.0	13.5
75 years and over ..... farms	4,652	422	35.5	7.5	13.3	14.7
<b>Livestock and poultry:</b>						
Cattle and calves inventory ..... farms	9,868	893	39.1	7.3	24.5	7.3
..... number	1,809,613	68,144	35.6	3.9	27.1	4.6
Beef cows inventory ..... farms	8,964	808	37.8	7.0	24.0	6.9
..... number	881,682	55,041	31.2	3.4	23.8	4.1
Milk cows inventory ..... farms	243	24	40.7	4.1	31.3	5.4
..... number	17,876	1,358	13.1	0.5	11.6	1.1
Hog and pigs inventory ..... farms	218	28	44.5	16.3	19.8	8.4
..... number	133,653	3,802	1.4	0.3	1.0	0.1
Layers inventory ..... farms	840	125	52.7	14.7	25.3	12.8
..... number	92,754	4,112	12.4	4.6	5.1	2.8
Broilers sold ..... farms	87	16	46.0	17.9	17.0	11.1
..... number	60,366	1,917	6.0	3.0	1.9	1.1
Aquaculture sold ..... farms	7	2	14.3	7.1	4.2	3.0
..... \$1,000	738	68	(Z)	(Z)	(Z)	(Z)
<b>Selected crops harvested:</b>						
Corn for grain ..... farms	6,651	306	29.1	2.5	23.5	3.1
..... acres	3,465,997	45,644	26.0	2.4	21.6	2.0
Wheat, winter ..... farms	1,769	60	24.8	1.9	20.2	2.6
..... acres	729,244	11,206	21.4	1.7	18.1	1.6
Wheat, durum ..... farms	2,177	318	17.4	1.4	13.8	2.3
..... acres	1,329,835	174,639	7.0	0.5	5.9	0.6

See footnote(s) at end of table.

--continued

**Table A. Summary of State Coverage, Nonresponse, and Misclassification Adjustments: 2012 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
Selected crops harvested: - Con.						
Wheat, spring ..... farms	8,692	337	28.3	2.1	22.8	3.4
..... acres	5,708,405	322,408	22.8	1.5	19.4	1.9
Soybeans for beans ..... farms	7,223	349	28.0	2.2	22.6	3.2
..... acres	4,729,137	79,577	24.0	1.8	20.3	1.9
Sorghum for grain ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Rice ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Cotton ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Peanuts ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Barley ..... farms	3,148	108	28.7	2.1	23.5	3.2
..... acres	1,006,554	66,017	25.0	1.8	21.0	2.1
Oats ..... farms	1,162	67	33.8	2.9	26.1	4.9
..... acres	109,519	4,004	33.1	2.5	26.6	4.0
Forage - land used for all hay and all haylage, grass silage, and greenchop (see text) ..... farms	10,141	832	35.6	5.2	22.2	8.2
..... acres	2,172,738	122,597	28.0	2.8	20.2	5.0
Land in vegetables (see text) ..... farms	262	112	21.0	7.3	6.8	6.9
..... acres	88,741	7,461	4.6	1.2	2.4	1.0
Potatoes ..... farms	189	77	14.8	4.8	5.2	4.8
..... acres	85,844	7,510	4.6	1.1	2.6	1.0
Tomatoes in the open ..... farms	59	(H)	33.9	14.6	6.2	13.1
..... acres	15	(H)	17.6	8.2	3.4	6.0
Sweet corn ..... farms	69	25	34.8	11.0	13.5	10.3
..... acres	129	42	18.2	5.7	7.1	5.4
Lettuce ..... farms	1	(H)	(Z)	(Z)	(Z)	(Z)
..... acres	(D)	(D)	(D)	(D)	(D)	(D)
Land in orchards ..... farms	54	15	40.7	19.9	5.5	15.4
..... acres	145	38	48.5	23.6	6.8	18.1
Apples ..... farms	22	8	40.9	17.2	6.8	16.9
..... acres	21	8	41.4	18.9	5.7	16.8
Grapes ..... farms	31	9	38.7	20.4	4.1	14.3
..... acres	63	19	44.2	24.7	4.7	14.8
Oranges ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Almonds ..... farms	-	-	-	-	-	-
..... acres	-	-	-	-	-	-
Land in berries ..... farms	57	12	40.4	15.6	10.6	14.2
..... acres	64	12	37.9	15.8	8.9	13.2

<sup>1</sup> Farms with total production expenses equal to market value of agricultural products sold, government payments, and farm-related income are included as farms with gains of less than \$1,000.

<sup>2</sup> Data were collected for a maximum of three operators per farm.

**Table B. Reliability Estimates of State Totals: 2012**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Coefficient of variation (percent)	Item	Total	Coefficient of variation (percent)
Farms .....number	30,961	9.1	Farms by legal status for tax purposes: - Con.		
Land in farms .....acres	39,262,613	3.3	Partnership .....farms	2,124	6.3
			.....acres	5,470,405	6.0
Farms by size:			Corporation:		
1 to 9 acres .....farms	689	32.2	Family held .....farms	662	4.3
.....acres	2,143	33.3	.....acres	1,504,821	2.3
10 to 49 acres .....farms	2,711	23.1	Other than family held .....farms	45	16.3
.....acres	77,476	21.0	.....acres	73,509	10.4
50 to 69 acres .....farms	1,063	15.6	Other - cooperative, estate or		
.....acres	61,655	15.6	trust, institutional, etc. ....farms	552	12.2
70 to 99 acres .....farms	1,455	13.8	.....acres	1,032,380	4.9
.....acres	118,037	13.8	Tenure:		
100 to 139 acres .....farms	1,407	14.0	Full owners .....farms	15,967	12.1
.....acres	163,414	14.0	.....acres	7,832,469	6.0
140 to 179 acres .....farms	2,632	14.0	Part owners .....farms	12,173	5.1
.....acres	414,560	14.0	.....acres	28,298,194	2.5
180 to 219 acres .....farms	958	9.8	Tenants .....farms	2,821	10.9
.....acres	189,716	9.8	.....acres	3,131,950	5.0
220 to 259 acres .....farms	862	9.0	Principal operator characteristics by-		
.....acres	204,469	9.1	Sex of operator:		
260 to 499 acres .....farms	4,086	9.3	Male .....farms	27,728	8.6
.....acres	1,504,518	9.0	.....acres	37,674,703	3.1
500 to 999 acres .....farms	4,237	9.9	Female .....farms	3,233	14.3
.....acres	3,017,670	9.8	.....acres	1,587,910	8.2
1,000 to 1,999 acres .....farms	4,424	9.7	Primary occupation:		
.....acres	6,333,780	9.8	Farming .....farms	17,509	6.2
2,000 acres or more .....farms	6,437	2.3	Other .....farms	13,452	13.1
.....acres	27,175,175	2.5	.....acres		
Irrigated land use:			Spanish, Hispanic, or		
Harvested cropland .....farms	729	13.4	Latino origin (see text) .....farms	122	59.2
.....acres	217,057	5.4	.....acres	138,097	18.8
Pastureland and other land .....farms	33	15.7	Race:		
.....acres	1,350	12.0	American Indian or		
Market value of agricultural			Alaska Native .....farms	348	30.5
products sold .....\$1,000	10,950,680	3.1	.....acres	392,080	10.1
Farms by value of sales:			Asian .....farms	16	(H)
Less than \$1,000 .....farms	9,669	13.2	.....acres	(D)	(D)
.....\$1,000	299	23.6	Black or African American .....farms	3	77.5
\$1,000 to \$2,499 .....farms	778	14.2	.....acres	2,230	37.0
.....\$1,000	1,284	14.1	Native Hawaiian or		
\$2,500 to \$4,999 .....farms	976	16.6	Other Pacific Islander .....farms	4	(H)
.....\$1,000	3,639	16.4	.....acres	(D)	(D)
\$5,000 to \$9,999 .....farms	1,275	17.0	White .....farms	30,539	8.7
.....\$1,000	9,087	17.6	.....acres	38,831,224	3.2
\$10,000 to \$19,999 .....farms	1,306	17.3	More than one race reported .....farms	51	49.4
.....\$1,000	18,856	17.3	.....acres	33,855	27.2
\$20,000 to \$24,999 .....farms	510	13.6	Reporting primary occupation as		
.....\$1,000	11,387	13.5	farming by age group:		
\$25,000 to \$39,999 .....farms	1,170	11.0	Under 25 years .....farms	241	13.8
.....\$1,000	37,667	10.8	25 to 34 years .....farms	1,468	9.8
\$40,000 to \$49,999 .....farms	647	10.0	35 to 44 years .....farms	2,115	6.8
.....\$1,000	28,839	10.0	45 to 54 years .....farms	4,004	6.0
\$50,000 to \$99,999 .....farms	2,074	10.1	55 to 64 years .....farms	4,951	5.1
.....\$1,000	151,312	10.0	65 years and over .....farms	4,730	7.4
\$100,000 to \$249,999 .....farms	3,304	6.7	Reporting primary occupation as		
.....\$1,000	577,981	6.4	other than farming by age group:		
\$250,000 to \$499,999 .....farms	3,156	5.6	Under 25 years .....farms	89	33.7
.....\$1,000	1,194,099	5.4	25 to 34 years .....farms	964	27.8
\$500,000 to \$999,999 .....farms	2,913	27.2	35 to 44 years .....farms	1,457	16.6
.....\$1,000	2,163,140	23.7	45 to 54 years .....farms	2,599	14.9
\$1,000,000 or more .....farms	3,183	1.1	55 to 64 years .....farms	3,675	10.4
.....\$1,000	6,753,090	1.9	65 years and over .....farms	4,668	10.9
Net cash farm income of operations (see text):			All operators by age group <sup>2</sup> :		
Farms with gains of <sup>1</sup> -			Under 25 years .....farms	746	13.4
Less than \$1,000 .....farms	687	13.5	25 to 34 years .....farms	4,010	14.9
.....\$1,000	349	13.2	35 to 44 years .....farms	5,648	9.8
\$1,000 to \$4,999 .....farms	2,432	11.9	45 to 54 years .....farms	10,120	9.3
.....\$1,000	6,949	11.9	55 to 64 years .....farms	12,132	7.3
\$5,000 to \$9,999 .....farms	1,916	12.6	65 to 74 years .....farms	7,069	8.8
.....\$1,000	14,133	13.1	75 years and over .....farms	4,652	9.1
\$10,000 to \$24,999 .....farms	3,487	10.2	Livestock and poultry:		
.....\$1,000	57,943	10.0	Cattle and calves inventory .....farms	9,868	9.1
\$25,000 to \$49,999 .....farms	2,554	10.3	.....number	1,809,613	3.8
.....\$1,000	92,843	10.3	Beef cows inventory .....farms	8,964	9.0
\$50,000 or more .....farms	11,806	7.5	.....number	881,682	6.2
.....\$1,000	4,647,092	5.8	Milk cows inventory .....farms	243	9.9
Farms with losses of -			.....number	17,876	7.6
Less than \$1,000 .....farms	687	14.8	Hog and pigs inventory .....farms	218	12.8
.....1,000	317	14.6	.....number	133,653	2.8
\$1,000 to \$4,999 .....farms	1,921	15.7	Layers inventory .....farms	840	14.9
.....1,000	5,522	16.6	.....number	92,754	4.4
\$5,000 to \$9,999 .....farms	1,624	15.6	Broilers sold .....farms	87	18.2
.....1,000	11,848	15.7	.....number	60,366	3.2
\$10,000 to \$24,999 .....farms	2,118	18.0	Aquaculture sold .....farms	7	27.1
.....1,000	33,034	18.1	.....\$1,000	738	9.2
\$25,000 to \$49,999 .....farms	714	12.6	Selected crops harvested:		
.....1,000	24,459	12.5	Corn for grain .....farms	6,651	4.6
\$50,000 or more .....farms	1,015	8.0	.....acres	3,465,997	1.3
.....1,000	188,911	6.0			
Farms by legal status for tax purposes:					
Family or individual .....farms	27,578	9.4			
.....acres	31,181,498	3.7			

See footnote(s) at end of table.

--continued

**Table B. Reliability Estimates of State Totals: 2012 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Item	Total	Coefficient of variation (percent)	Item	Total	Coefficient of variation (percent)
Selected crops harvested: - Con.			Selected crops harvested: - Con.		
Wheat, winter ..... farms	1,769	3.4	Land in vegetables (see text) .....farms	262	42.6
..... acres	729,244	1.5	..... acres	88,741	8.4
Wheat, durum ..... farms	2,177	14.6	Potatoes .....farms	189	40.9
..... acres	1,329,835	13.1	..... acres	85,844	8.7
Wheat, spring ..... farms	8,692	3.9	Tomatoes in the open .....farms	59	(H)
..... acres	5,708,405	5.6	..... acres	15	(H)
Soybeans for beans ..... farms	7,223	4.8	Sweet corn .....farms	69	36.8
..... acres	4,729,137	1.7	..... acres	129	32.7
Sorghum for grain ..... farms	-	-	Lettuce .....farms	1	(H)
..... acres	-	-	..... acres	(D)	(D)
Rice ..... farms	-	-	Land in orchards .....farms	54	28.3
..... acres	-	-	..... acres	145	26.0
Cotton ..... farms	-	-	Apples .....farms	22	34.1
..... acres	-	-	..... acres	21	35.7
Peanuts ..... farms	-	-	Grapes .....farms	31	27.8
..... acres	-	-	..... acres	63	30.3
Barley ..... farms	3,148	3.4	Oranges .....farms	-	-
..... acres	1,006,554	6.6	..... acres	-	-
Oats ..... farms	1,162	5.8	Almonds .....farms	-	-
..... acres	109,519	3.7	..... acres	-	-
Forage - land used for all hay and all haylage, grass silage, and greenchop (see text) ..... farms	10,141	8.2	Land in berries .....farms	57	21.9
..... acres	2,172,738	5.6	..... acres	64	19.6

<sup>1</sup> Farms with production expenses equal to market value of agricultural products sold, government payments, and farm-related income are included as farms with gains of less than \$1,000.

<sup>2</sup> Data were collected for a maximum of three operators per farm.



**Table C. Summary of Coverage, Nonresponse, and Misclassification Adjustments by County: 2012**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	Total (number)	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
<b>ALL FARMS</b>						
<b>State Total</b>						
North Dakota .....	30,961	2,813	37.6	7.9	19.5	10.2
<b>Counties</b>						
Adams .....	392	36	35.6	6.9	17.7	11.0
Barnes .....	855	67	35.0	7.1	17.8	10.1
Benson .....	563	39	38.7	8.1	20.4	10.3
Billings .....	197	20	34.0	6.5	21.4	6.1
Bottineau .....	863	76	36.4	7.6	18.1	10.7
Bowman .....	348	30	35.5	8.0	19.1	8.5
Burke .....	488	43	33.6	7.1	16.0	10.5
Burleigh .....	1,014	150	49.6	11.9	22.3	15.4
Cass .....	968	94	35.2	7.1	19.3	8.8
Cavalier .....	667	42	35.3	5.6	21.0	8.8
Dickey .....	543	45	37.4	8.4	19.6	9.4
Divide .....	452	38	26.8	4.7	14.0	8.1
Dunn .....	628	80	39.6	9.5	20.5	9.6
Eddy .....	331	29	38.2	8.6	17.5	12.1
Emmons .....	609	46	41.2	6.8	23.7	10.7
Foster .....	310	22	34.1	7.7	18.5	8.0
Golden Valley .....	251	23	34.0	6.7	19.9	7.4
Grand Forks .....	970	75	35.1	8.3	15.9	10.9
Grant .....	508	50	37.8	9.1	19.5	9.1
Griggs .....	456	37	35.8	7.2	18.1	10.5
Hettinger .....	494	39	35.1	7.8	17.1	10.2
Kidder .....	559	50	38.5	7.1	20.6	10.8
LaMoure .....	642	49	34.4	7.2	17.7	9.4
Logan .....	379	28	39.3	7.1	22.0	10.1
McHenry .....	911	88	40.8	7.8	21.8	11.2
McIntosh .....	471	40	43.0	9.9	22.8	10.4
McKenzie .....	574	67	38.0	6.6	21.8	9.6
McLean .....	868	70	36.4	8.1	18.7	9.6
Mercer .....	422	45	45.6	8.7	25.6	11.3
Morton .....	887	114	42.3	9.7	22.8	9.7
Mountrail .....	670	62	33.1	5.8	18.4	8.8
Nelson .....	603	48	36.1	8.1	15.1	12.9
Oliver .....	290	31	42.4	8.6	23.8	10.0
Pembina .....	584	54	34.4	7.5	16.7	10.1
Pierce .....	521	40	38.6	7.9	20.0	10.6
Ramsey .....	573	45	37.3	7.8	18.0	11.4
Ransom .....	548	51	37.1	7.8	18.2	11.1
Renville .....	304	18	33.1	4.5	22.4	6.2
Richland .....	854	66	32.4	6.6	17.7	8.2
Rolette .....	649	89	47.5	11.7	20.8	15.0
Sargent .....	537	39	33.8	6.1	18.7	9.1
Sheridan .....	370	31	39.1	7.5	20.9	10.6
Sioux .....	176	42	41.4	11.7	18.9	10.8
Slope .....	221	19	31.0	6.3	17.2	7.4
Stark .....	837	123	43.2	9.2	21.9	12.1
Steele .....	355	26	32.5	6.8	16.5	9.1
Stutsman .....	1,028	89	38.4	9.6	19.3	9.5
Towner .....	529	40	38.5	8.0	20.5	10.0
Trails .....	468	38	31.9	6.4	16.8	8.7
Walsh .....	962	84	36.3	7.9	17.4	11.0
Ward .....	961	91	40.6	8.5	22.8	9.4
Wells .....	543	43	35.1	7.5	17.8	9.9
Williams .....	758	81	35.3	7.7	18.0	9.6
<b>LAND IN FARMS</b>						
<b>State Total</b>						
North Dakota .....	39,262,613	1,282,136	23.9	2.7	17.5	3.6
<b>Counties</b>						
Adams .....	601,272	42,300	21.9	3.1	14.8	4.0
Barnes .....	936,780	21,996	22.3	2.6	16.4	3.4
Benson .....	802,059	67,611	30.9	3.4	22.5	5.0
Billings .....	722,275	26,628	11.4	2.3	7.8	1.2
Bottineau .....	899,431	28,315	21.7	2.4	15.8	3.5
Bowman .....	730,327	38,353	17.1	2.8	11.9	2.4
Burke .....	595,094	22,872	21.0	2.8	14.4	3.9
Burleigh .....	951,021	64,095	32.5	4.4	21.3	6.7
Cass .....	1,107,398	23,731	22.1	1.9	17.7	2.5
Cavalier .....	940,331	21,431	25.2	1.9	20.7	2.7
Dickey .....	633,295	74,930	27.0	3.3	19.8	3.9
Divide .....	564,975	54,280	11.0	1.1	8.1	1.9
Dunn .....	1,031,359	59,476	22.3	2.4	16.7	3.2
Eddy .....	395,811	27,330	36.4	4.0	25.8	6.7
Emmons .....	743,925	28,421	31.8	3.6	23.3	5.0
Foster .....	373,970	17,767	22.0	2.0	17.6	2.5
Golden Valley .....	562,453	43,090	17.5	1.3	14.7	1.5
Grand Forks .....	816,478	38,399	21.8	2.6	15.6	3.7
Grant .....	1,049,893	34,254	24.2	2.8	18.4	3.1
Griggs .....	445,480	15,287	25.8	3.1	17.9	4.7
Hettinger .....	715,958	23,991	23.8	2.7	17.3	3.7
Kidder .....	780,319	33,078	28.5	2.8	20.5	5.2
LaMoure .....	725,867	18,181	25.4	2.7	18.9	3.8

--continued

**Table C. Summary of Coverage, Nonresponse, and Misclassification Adjustments by County: 2012 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	Total (number)	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
<b>LAND IN FARMS - Con.</b>						
<b>Counties - Con.</b>						
Logan.....	571,599	21,882	35.8	3.4	27.4	5.0
McHenry.....	1,061,267	46,921	31.2	2.9	23.0	5.2
McIntosh.....	589,716	33,792	36.9	3.9	27.5	5.6
McKenzie.....	1,064,191	45,131	18.6	2.3	13.3	3.0
McLean.....	1,112,659	31,778	24.1	2.8	17.9	3.5
Mercer.....	503,164	24,751	31.6	3.4	23.1	5.1
Morton.....	1,220,004	52,722	21.6	2.0	16.8	2.8
Mountrail.....	963,672	62,100	14.9	1.7	10.8	2.5
Nelson.....	560,478	49,417	29.4	4.3	17.9	7.2
Oliver.....	394,528	24,193	30.8	3.5	22.4	4.8
Pembina.....	692,140	22,481	17.3	2.1	12.1	3.0
Pierce.....	598,228	33,957	35.8	3.7	25.9	6.2
Ramsey.....	698,412	31,498	27.7	3.0	19.7	5.0
Ransom.....	501,632	40,655	20.5	3.5	13.6	3.4
Renville.....	500,082	15,142	18.6	1.1	15.9	1.5
Richland.....	868,923	39,695	20.4	2.0	16.0	2.4
Rolette.....	534,301	29,433	31.9	4.0	20.9	6.9
Sargente.....	512,830	22,890	20.9	2.2	15.5	3.3
Sheridan.....	513,511	28,482	26.6	3.5	18.3	4.7
Sioux.....	573,113	30,104	20.9	3.6	14.6	2.6
Slope.....	674,345	30,567	10.0	2.8	6.2	1.1
Stark.....	829,547	43,526	23.1	2.4	17.0	3.7
Steele.....	425,834	21,498	19.4	1.9	15.3	2.2
Stutsman.....	1,302,623	34,203	28.8	3.7	20.3	4.8
Towner.....	645,445	19,248	31.5	3.2	23.8	4.5
Traill.....	547,732	18,678	17.8	1.7	13.9	2.2
Walsh.....	802,372	23,249	20.4	2.4	14.0	4.0
Ward.....	1,073,283	43,840	23.7	1.9	18.8	3.0
Wells.....	738,102	17,698	23.7	2.7	17.4	3.6
Williams.....	1,063,109	46,647	19.1	2.0	14.4	2.8
<b>SALES</b>						
<b>State Total</b>						
North Dakota.....	10,950,680	336,530	18.7	1.6	15.3	1.8
<b>Counties</b>						
Adams.....	107,537	8,354	18.7	2.4	14.3	2.1
Barnes.....	376,350	13,932	19.4	1.7	16.0	1.7
Benson.....	240,629	10,262	24.5	2.6	19.6	2.2
Billings.....	36,667	1,504	28.6	2.7	22.8	3.2
Bottineau.....	254,032	13,388	16.0	1.2	13.5	1.3
Bowman.....	111,417	11,668	12.6	2.8	8.5	1.3
Burke.....	105,142	3,274	19.1	1.7	15.4	2.0
Burleigh.....	179,612	30,675	29.3	3.9	20.8	4.5
Cass.....	567,108	32,758	16.0	1.3	13.6	1.2
Cavalier.....	334,532	11,122	21.9	1.4	19.0	1.4
Dickey.....	266,829	47,454	21.0	2.1	17.4	1.6
Divide.....	93,788	11,794	7.6	0.5	6.3	0.7
Dunn.....	124,788	6,627	20.1	2.0	15.7	2.4
Eddy.....	105,105	4,509	31.6	2.3	26.5	2.8
Emmons.....	171,284	12,913	25.1	2.7	20.0	2.4
Foster.....	169,410	12,614	17.0	2.0	13.8	1.3
Golden Valley.....	59,569	3,387	23.7	1.9	19.6	2.3
Grand Forks.....	428,762	33,987	13.1	1.1	10.8	1.2
Grant.....	157,090	5,257	21.9	2.1	17.5	2.3
Griggs.....	127,828	4,844	20.7	1.7	17.2	1.8
Hettinger.....	159,032	3,662	16.0	1.3	13.5	1.2
Kidder.....	148,317	12,111	21.6	1.8	16.9	2.9
LaMoure.....	292,431	16,784	20.9	1.8	17.3	1.8
Logan.....	172,099	4,332	27.7	2.6	22.1	2.9
McHenry.....	198,216	13,136	22.7	2.0	18.2	2.5
McIntosh.....	139,771	10,318	33.2	2.7	27.4	3.1
McKenzie.....	114,448	5,066	15.5	1.5	11.1	2.9
McLean.....	293,406	13,522	18.7	2.0	15.1	1.5
Mercer.....	76,705	3,238	28.4	2.7	22.1	3.5
Morton.....	225,239	10,252	14.2	1.2	11.7	1.3
Mountrail.....	154,881	17,930	12.1	1.0	9.8	1.3
Nelson.....	145,760	7,558	20.7	2.0	17.0	1.8
Oliver.....	85,495	5,507	27.6	3.4	21.1	3.1
Pembina.....	405,976	10,476	11.5	1.0	9.1	1.4
Pierce.....	142,050	7,616	33.3	2.5	27.8	3.0
Ramsey.....	236,079	3,176	22.5	1.7	19.2	1.6
Ransom.....	179,265	24,259	13.4	1.0	11.1	1.3
Renville.....	156,440	7,170	14.6	0.7	12.9	1.0
Richland.....	535,658	34,125	14.9	1.3	12.3	1.3
Rolette.....	107,929	9,708	25.9	2.3	20.5	3.1
Sargente.....	243,428	22,566	13.1	1.1	10.9	1.1
Sheridan.....	107,779	5,091	20.8	2.0	16.7	2.1
Sioux.....	62,730	1,867	16.0	3.6	10.7	1.8
Slope.....	67,557	5,981	14.5	2.0	11.2	1.3
Stark.....	152,583	14,330	15.7	1.6	12.4	1.6
Steele.....	210,571	12,596	18.0	1.5	15.2	1.3
Stutsman.....	464,568	17,570	22.9	2.2	18.6	2.1
Towner.....	196,814	4,020	25.1	1.6	21.7	1.8
Traill.....	308,827	14,525	11.7	1.0	9.4	1.4
Walsh.....	424,033	31,018	13.8	1.1	10.8	1.9
Ward.....	274,458	16,405	18.3	1.3	15.5	1.6

--continued

**Table C. Summary of Coverage, Nonresponse, and Misclassification Adjustments by County: 2012 (continued)**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	Total (number)	Standard error	Adjustment as percent of total	Percent of total adjustment from coverage	Percent of total adjustment from nonresponse	Percent of total adjustment from misclassification
<b>SALES - Con.</b>						
<b>Counties - Con.</b>						
Wells .....	271,957	12,081	18.0	1.4	15.3	1.2
Williams .....	178,703	11,190	13.3	0.9	11.1	1.4

**Table D. American Indian or Alaska Native Operators: 2012**

[For meaning of abbreviations and symbols, see introductory text.]

Geographic area	American Indian or Alaska Native farm operators			Geographic area	American Indian or Alaska Native farm operators		
	Total	Individually reported <sup>1</sup>	Other <sup>2</sup>		Total	Individually reported <sup>1</sup>	Other <sup>2</sup>
<b>State Total</b>				<b>Counties - Con.</b>			
North Dakota .....	652	652	-	McHenry .....	3	3	-
<b>Counties</b>				McKenzie .....	30	30	-
Adams .....	7	7	-	McLean .....	22	22	-
Barnes .....	3	3	-	Mercer .....	2	2	-
Benson .....	7	7	-	Morton .....	6	6	-
Bottineau .....	10	10	-	Mountrail .....	32	32	-
Burke .....	1	1	-	Nelson .....	2	2	-
Burleigh .....	12	12	-	Pembina .....	5	5	-
Cass .....	7	7	-	Pierce .....	4	4	-
Divide .....	4	4	-	Richland .....	2	2	-
Dunn .....	39	39	-	Rolette .....	323	323	-
Eddy .....	2	2	-	Sargent .....	3	3	-
Golden Valley .....	2	2	-	Sioux .....	68	68	-
Grand Forks .....	5	5	-	Steele .....	5	5	-
Grant .....	12	12	-	Stutsman .....	2	2	-
Hettinger .....	4	4	-	Towner .....	4	4	-
Kidder .....	4	4	-	Ward .....	4	4	-
				Wells .....	1	1	-
				Williams .....	15	15	-

<sup>1</sup> Data were collected for a maximum of three operators per farm.

<sup>2</sup> Data represent American Indian or Alaska Native farm or ranch operators on reservations who did not report individually. Data obtained by reservation officials.