Appendix A.
Census of Agriculture Methodology

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 NASS farm definition, that is, an operation that produces, or would normally produce and sell, $1,000 or more of agricultural products per year. 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 are 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 if they meet the NASS farm definition. For the 2007 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 2007 CML started in 2004. Between 2004 and 2007, NASS conducted a series of Agricultural Identification Surveys (AIS) on approximately 1.7 million records, which included nonrespondents from the 2002 census and newly added records from outside list sources. The AIS report form collected information that was used to determine if an operation met the NASS farm definition. If the definition was met, the operation was added to the NASS list and subsequently to the CML. Addressees that were nonrespondents 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 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.

The official CML was established on September 1, 2007. The list contained 3,194,373 records. There were 2,198,410 records that were thought to meet the NASS farm definition and 995,963 potential farm records, which included AIS nonrespondents, other records added to the CML by the NASS field offices, and late adds to the CML that were not included in any previous AIS or State screening survey.

Not on the Mail List

To account for farming operations not on the CML, NASS used its area frame. The NASS area frame covers all land in the U.S. and includes all farms. The land in the U.S. is stratified by characteristics of the land. Segments of approximately equal size are

delineated within each strata and designated on aerial photographs. A probability sample of segments is drawn within each strata for the NASS annual area frame survey, known as the June Agricultural Survey (JAS). The JAS sample of segments is allocated to strata to provide accurate measures of acres planted to widely grown crops and inventories of hogs and cattle. Sampled segments in the June Survey are personally enumerated. Each operation identified within a segment boundary is known as a tract.

The 2007 JAS sample was allocated to strata so that it would provide additional measures of small and minority owned farms. The 2007 JAS consisted of 10,912 regular sampled segments, supplemented with 3,692 Agricultural Coverage Evaluation Survey (ACES) segments – segments selected to provide measures of small and minority owned farms. These additional ACES segments targeted farming demographics that typically had lower coverage rates on the list.

The information from each tract (operation) within a segment is matched against operations on the NASS list to determine the amount of undercoverage that exists for a wide range of farming sectors and farmer demographics. The names and addresses collected in the 2007 JAS and 2007 ACES were matched to the CML and checked for duplication. Farms from the June 2007 survey that did not match were determined to be Not on the Mail List (NML) and sent a report form of a different color to be easily identified. Data from the NML operations provided a measure of the undercoverage of the CML operations. Instructions on the census report form guided the respondent to complete the CML form and mail back both CML and NML forms together if duplicate forms were received. Those who returned a CML census form and an NML census form had been erroneously classified as NML and were removed from the NML.

The percentage of farms not represented on the CML varied considerably by State. In general, farms not on the mail list tended to be small in acreage, production, and sales of agricultural products. Farm operations were missed for various reasons, including the possibility that the operation started after the mail list was developed, the operation was so small that it did not appear in any agriculture-related source lists, or the operation was erroneously classified as a nonfarm prior to mailout.

The NML consisted of 12,821 tracts. The CML was used with the NML in multiple frame estimation to represent all farming operations across all States, with the exception of Alaska. It is financially and logistically unfeasible to maintain an area frame in Alaska due to its vast land mass and relatively sparse agriculture.

DATA COLLECTION

Method of Enumeration

Mailout and mailback was the primary data collection method. It was supplemented with Electronic Data Reporting (EDR) on the Internet and non-response follow-ups by telephone and personal enumeration. The enumeration methods used in the 2007 census were similar to those used in the 2002 census.

Report Forms

A master report form was developed that included all data items to be collected in the census. From the master, two types of report forms were developed to be used in the 2007 census - a regionalized report form with 7 versions and a national report form. Each of the 24-page regionalized report forms (07-A0201, 07-A0202, 07-A0203, 07-A0204, 07-A0205, 07-A0206, 07-A0207) were designed to facilitate reporting crops most commonly grown within the report form region. The 12-page national report form (07-A0100) was designed for operations throughout the country with few commodities. The national report form collected the same information as the regional form, but it was formatted to fit on fewer pages. All of the forms allowed respondents to write in specific commodities that were not identified on their form. The national form was mailed to approximately 528,000 addresses on the CML (about 20 percent) and the regional form was mailed to 2.67 million addresses on the CML (about 80 percent).

Report Form Mailings and Respondent Follow-up

The initial mailout took place at the end of December 2007. Approximately 3.2 million packets

were mailed. Each packet contained a cover letter, instruction sheet, a labeled report form, and a return envelope. Mailout packet preparation, initial mailout, and two follow-up mailings to nonrespondents were handled by the Census Bureau's National Processing Center (NPC) in Jeffersonville, IN. The first follow-up was mailed during the last two weeks of February 2008 to approximately 1.3 million nonrespondents. The second follow-up was mailed the beginning of April 2008 to approximately 1.0 million nonrespondents. Additionally, NPC received, checked-in, scanned, and keyed (from image) returned report forms. NASS statisticians on site at NPC provided technical guidance and monitored NPC processing activities.

Select groups of census records were identified to receive special handling procedures. Report forms were labeled at NPC and shipped to the field offices for enumeration. These respondents were excluded from the initial and both follow-up mailings, and were referred to as "must" operations. Each "must" operation was enumerated by telephone or face-to-face. If a record was determined to be no longer in operation, their non-farm status was verified and documented. The field offices were responsible for enumerating or resolving all non-response "must" records in their State. Computer Assisted Telephone Interview (CATI) calling for nonrespondent "must" records was conducted between March 2008 and June 2008. Once enumerated, the report forms were either sent to NPC for check-in and data capture or the data were keyed directly from the form at the field office. The 169,000 "must" records fell into one of five groups.

The first "must" group consisted of 46,000 records "tagged" by the NASS field offices for personal enumeration rather than mailout and mailback enumeration. The second "must" group consisted of 4,000 "specialized" records including such operations as grazing associations, governmental units, research farms, college farms, etc.. The third "must" group was characterized by location. All 3,000 records in Alaska and Rhode Island were identified as "must" records because census statistics for these two States were based on responses to the CML because nonresponse was not permitted. The last two groups consisted of a total of 116,000 records expected to have either a large number of acres in farm land or a large value of sales. Threshold levels were identified for each State.

Advanced Follow-up was conducted between February 2008 and April 2008. It focused on three groups of nonrespondents that included: respondents least likely to respond because they were nonrespondents to the 1997 and 2002 Censuses of Agriculture, even though they may have responded to other NASS surveys; respondents viewed as easy and quick interviews based on expected sales of zero, including respondents who received Conservation Reserve Program (CRP) payments and respondents to the AIS with expected future sales; and new records whose farm status was uncertain due to unsuccessful earlier screening attempts. The field offices conducted CATI and field enumeration for operations in their State. This phase was followed by Low-Response County Follow-up to attempt to reach a minimum response rate of at least 75 percent in all counties. It was conducted by the field offices using CATI between March 2008 and June 2008.

DATA COLLECTION OUTREACH AND PROMOTIONAL EFFORTS

NASS engaged in an unprecedented level of public outreach for the 2007 Census of Agriculture, seeking to increase the level of awareness and response among U.S. agricultural producers and, in particular, minority and small farm operators. This was accomplished through an integrated marketing communications program that focused on four primary areas: partnership building, public relations, paid media, and the Internet. External support was provided by a private agricultural marketing communications agency.

The unifying force behind the 2007 marketing campaign was the theme "Your Voice, Your Future, Your Responsibility." This was accompanied by supporting messages and artwork that created a consistent look and feel for all census communications.

Partnership

At the national level, NASS officials met with leaders from dozens of key agricultural organizations and other USDA agencies, successfully securing

their support in promoting the census among their constituencies through publications, special mailings, speeches, and other communications. In addition, NASS made special efforts to reach out to minority and limited-resource farmers and ranchers by partnering with a number of community-based organizations. The national-level outreach was mirrored by field offices at the State and local levels. Among the features of these collective efforts was the production of State-specific radio public service announcements (PSAs) featuring State secretaries and commissioners of agriculture, as well as a national radio PSA featuring the U.S. Secretary 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: 2007** (Table A in the Alaska publication), 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 the agriculture media at the national level and equipped NASS's 46 field offices with communications tools that enabled them to deliver the right message to producers in their States. From customizable press releases, to radio public service announcements, to a video news release, to newsletter articles and letters to the editor, the public relations strategy was designed to ensure NASS fields offices could easily and effectively deliver the census message to local media. As a result, in the print media alone, the public relations efforts generated 27 million media impressions.

Paid Media

Because there were certain constituencies that were difficult to reach through partnership or public relations, NASS also employed a paid media strategy that was narrowly targeted to reach previously under-represented populations. NASS purchased limited print and radio advertising in areas where there were high concentrations of minority farmers and where 2002 census response rates were low.

Internet

For the 2007 census, NASS created a dedicated website, www.agcensus.usda.gov. This became a repository for all types of census information, including basic background materials, previous years' census data, sample report forms, and news releases and other publicity materials. The website also enabled individuals to submit their contact information to ensure that they were on the mailing list to receive a census form. NASS also enhanced its online presence by purchasing banner ads and pay-per-click advertisements on key agricultural websites as well as major search engines.

REPORT FORM PROCESSING

Data Capture

All report forms returned to NPC were immediately checked in, using bar codes printed on the mailing label, and removed from follow up 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 format program. The program verified that record identifiers were valid and checked the basic integrity of the data fields. Rejected records were referred to analysts for correction. Accepted records were sent to a 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 NPC.

All 2007 census records were passed through a complex computer edit. The edit determined whether a reporting operation met the minimum criteria to be counted as a qualifying farm (in-scope). Operations failing to meet the minimum criteria (out-of-scope) were referred to analysts for verification. The edit examined each in-scope record for reasonableness and completeness and determined whether to accept the recorded value for each data item or take corrective action. 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 edit determined a replacement value. Strategies for determining replacement values are discussed in the next section.

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 defined the criteria for acceptable data. Problems that could not be resolved within the edit were referred to an analyst for intervention. Analysts in the NASS field offices used additional information sources, examined the scanned image, and determined an appropriate action. Field office analysts used an interactive version of the edit program to submit corrected data and immediately re-edit the record to ensure a satisfactory solution.

Imputing for Missing Data

Missing data occurred whenever a respondent failed to report in a cell that should have a positive value or when the edit determined a value was not reasonable and should be changed. The edit performed a sequence of steps that determined the best value to impute for the missing item. If an item could not be calculated directly from other data reported on the current form, the edit checked for previously reported data. Acreage, production, and inventory items may have been reported on a recent NASS crop or livestock survey. Operator characteristics, such as race and gender, were brought forward from the previous census if the operator had not changed in five years. Administrative data from the Farm Service Agency was used for a few items, such as Conservation Reserve Program acreage. When these deterministic sources failed to produce a solution, the edit invoked an automated imputation system which searched for a reporting farm of similar type, size, and location to provide a value for the missing data item. If the imputation algorithm failed to provide a solution, the record was referred to an analyst for resolution.

The guiding principal for imputation was to find a close match to the farm with the missing item. The census imputation algorithm relied on pre-established donor pools, one for each State. A donor pool included a collection of completed reports that had successfully navigated the edit. Each pool was further divided into groups of similar type and size, referred to as profiles. When the edit determined the need to impute an item, it went to the appropriate profile and searched for the best fit. Best fit was determined by calculating "distance" between the incomplete report and each candidate donor using a set of match variables. Match variables were specific to each section of the report form and included the latitude and longitude of the principal county of operation. The distance was the sum of the squared differences between the reported values of the match variables. The donor with the smallest distance was considered the "nearest neighbor" and became the source for the imputation action. The value returned may have been a direct copy of the donor's value. In many cases, a relationship between two related variables on the donor record was applied to a reported value on the incomplete record. Using crop production as an example, the donor's production was divided by its harvested acres (yield) and multiplied by the recipient's harvested acres to obtain imputed production.

The imputation process was imbedded in the edit. When the edit determined an item required imputation, the edit program launched the algorithm, waited for a value to be returned, validated that the returned value was satisfactory, and resumed editing. Since imputation was conducted independently for each occurrence, reports requiring multiple imputations drew from multiple donors.

Initial donor pools were established before the first batch edits were run. These donor pools were "seeded" with 2002 census data that were "mapped" to look like 2007 data and passed through the 2007 edit to ensure they were consistent using the 2007 data relationships. In addition, data from the 2005 Census Content Test were similarly mapped and edited. As 2007 data were successfully processed, new records systematically replaced the older records in the donor pool. The older records disappeared entirely from the donor pool after the first few batch edits.

The donor pool for each State was refreshed weekly during the first couple of months of editing. As the flow of new data slowed, the donor pools were refreshed biweekly. During the early stages of editing, records that needed to impute production for field crops or hay were set aside. When the donor pool no longer contained old data, these records were brought back and passed through the edit, ensuring 2007 yields were imputed.

In some cases, nearest-neighbor imputation was not possible. The requirement of a positive imputed value could have ruled out all available donors, resulting in an imputation failure. An imputation failure could have occurred if there were no donors 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.

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.

WHOLE FARM NONRESPONSE ESTIMATION

Whole farm nonresponse adjustments were necessary because some farm operators did not respond to the census, despite numerous attempts to contact them. Statistical estimation procedures were used to account for these CML nonrespondents. The objectives of the nonresponse adjustments included estimating the number of in-scope records (farms) included in the total number of nonrespondents of a similar size and type by increasing the weights of reporting farms of that size and type. This procedure

was intended to account for those farms that failed to return a report form. These procedures were applied in all States, except Alaska and Rhode Island where staff were required to submit data for every record on the CML due to the low level of farming operations in these States. Large or unique farms (Must records) for which a report was required (and thus given a nonresponse weight of one) were exempt from this weighting procedure. These farms received intensive follow-ups. Data were imputed for the record if all followup contacts failed (rather than using the nonresponse weighting procedure).

After census data collection was completed, all CML records in a State were put into mutually exclusive weighting groups based on a list of farm characteristics known at the time of mail-out and the census response status of the record. Data mining techniques systematically checked selected variables, identifying those groups with differences in response rates that were statistically significant. The algorithm would take one characteristic, divide all names into two groups, and check for statistical significance between the response rates of the two groups. If a significant difference was found, these groups became permanent and the next characteristic would be examined within those two groups. If the response rate between two groups was not statistically significant, the groups were rejoined and the next characteristic was tested. This stepwise process continued until all characteristics were checked and no further statistical significance could be found. Since the "path" taken by the algorithm was driven by an individual State's response pattern, the final breakout of weighting groups was customized for the State.

Within each weighting group, the percent of responding in-scope farms was computed. This rate was applied to the count of nonresponding farms to estimate the number of in-scope nonrespondents. The weights of the responding in-scope farms in each weighting group were scaled to account for nonresponding farms in that group.

This procedure was applied to all of the weighting groups except the one that consisted primarily of records who were included on the CML but had not responded to data collection efforts either during CML development activities or during the census data collection phase. The estimate of in-scope records (farms) within this group was not reliable. To get a more reliable estimate, NASS conducted a nonresponse follow-up activity. After scheduled census data collection efforts were completed, a target sample of 5,000 records was selected from across all States. These 5,000 records were personally interviewed by NASS staff to determine if they were indeed in-scope records (farm) or out-of-scope records (nonfarm). Each record fell into one of these two categories. The percent of in-scope records was used to form the weight for this group.

When NASS summarizes the census of agriculture, it assigns the data from an individual report to the "principal" county. The principal county is the one county in which the majority of agricultural products are produced from a respondent. This is a question on the census report form and is therefore determined by the respondent. 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.

The percent of the total that came from the whole farm nonresponse estimate is shown for selected census data items in Tables A and C. The estimates provided in Tables A and C do not reflect the effect of item nonresponse on individual census data items. The effect of this item nonresponse is discussed in the section on "Item Nonresponse" in "Nonmeasured Census Error."

COVERAGE ADJUSTMENT

Although much effort was expended making the CML as complete as possible, the census did not count all U.S. farms. NASS's goal was to produce agricultural census totals for publication that were fully adjusted for list undercoverage at the county level. NASS used its area frame with the CML in a dual-frame estimation procedure to measure the number of farms in the population and key characteristics of those farms. Area frame segments were enumerated using field enumerators (as

described in the first section of this appendix) who personally visited the tract operators within a segment. Because field enumeration is significantly more expensive than other modes of data collection, NASS's area frame sample allocation is only designed to generate reliable estimates at the State, regional, and U.S. level. Therefore, in order to produce estimates that represented all farms at the county level, NASS used an allocation process known as "calibration" to distribute the dual-frame estimates across counties.

Once all CML and NML data were collected, NASS analysts went through an extensive process to generate adjusted estimates. The weights of the CML respondents had been previously adjusted to account for all of the CML nonrespondents, referred to as list plus nonresponse (CML+NR). Simultaneously, NASS summarized the NML tract records to generate state-level NML survey estimates. These two pieces were then combined in a dual-frame estimation procedure to form State estimates of totals that represented all farms. These estimates are annotated as [(CML+NR) +NML]. The state-level totals for these variables were summed to yield national totals.

The whole farm nonresponse and list undercoverage record weighting processes were initially applied at the State level to produce adjusted estimates of farm numbers and land in farms for 65 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; total number of farms and land in farms (2); 4 sales categories for each of 10 major commodities (40); and 7 farm type groups. The national-level adjusted estimates were smoothed across States to get initial State farm operation coverage targets because state-level farm-count estimates based on this two-piece formula sometimes had unacceptably high state-level standard errors and apparent biases. This often occurs when estimating a rare item, such as female farm operators, using a general purpose survey.

The smoothing process examined the proportion of the total JAS estimate attributable to the NML, for each of the 65 variables in each State and the U.S. Since the CML was built using standard national methods, the NML percentages were expected to be uniform across States. The smoothed NML value for each of the 65 variables in a given State was calculated as the product of the state-level NML value and the weighted average of the ratios of the NML for a given variable in the State to the overall NML in the State and the NML for the given variable in the U.S. to the overall NML in the U.S. The weighting factor was chosen to minimize the mean square error under a random effects model with the control that the sum of the State smoothed NML values was equal to the total NML estimate for each of the 65 variables. This methodology effectively draws the state-level NML undercoverage proportions of the JAS toward the national estimate of undercoverage with the most extreme values adjusted the most. The smoothed NML values for each variable were added to the (CML + NR) totals to form calibration targets for each variable. Subject-matter experts in headquarters reviewed all targets.

However, these State estimates were general purpose in that they did not provide any control over expected levels of commodity production of the 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 were added to the calibration algorithm, known as commodity coverage targets. These targets were commodity totals from administrative sources or from NASS surveys of non-farm 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.

The calibration algorithm addressed farm operation undercoverage and commodity coverage concurrently. The algorithm was controlled by the 65 State farm operation coverage targets and the State commodity coverage targets. In order 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 sometimes made it possible for the calibration algorithm to produce a set of satisfactory, adjusted weights that it would not have otherwise.

Ranges for the list 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 [(CML+NR)+NML] plus or minus one-half of one estimated standard error of NML 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.

Adjusted 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-adjusted weights and added a second stage weight to simultaneously satisfy all farm operation coverage and commodity coverage calibration targets. 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 and the calibration algorithm was rerun. Additionally, the CML was assumed to be complete for very large and unique farms with their weight being controlled to 1 during the calibration adjustment process.

Weight computations in the nonresponse and final coverage calibration algorithms were performed to several decimals. Thus, the fully-adjusted weights were non-integer numbers. To insure 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 insured 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 adjustments are displayed in Tables A and C. Some estimated coverage adjustments could be negative. The use of commodity targets in calibration indirectly exposed some duplication on the census list or over adjustment by the nonresponse algorithm resulting in negative coverage adjustments.

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.

Based on agency standards, data cells were determined to be sensitive to a disclosure of information if they violated either of two criteria. First, 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. Second, a 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 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.

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 complement was chosen.

MEASURES OF CENSUS QUALITY

An important objective of the 2007 Census of Agriculture was to provide data with a high level of quality. However, every census or survey has the potential for error in its processes. These errors impact the quality of the data estimates. When feasible, measurements of those errors are provided with individual data items or used to make adjustments to the census or survey estimates. In conducting the 2007 Census of Agriculture, efforts were initiated to measure error associated with the adjustment for farm operations that were not respondents to the request to CML records, the coverage adjustment for farms not on the CML using the NML and calibration, and the integerization process. Other errors present in the census of agriculture include respondent or enumerator error, error in classification of farm operations, other types of processing errors, error associated with imputation for item nonresponse, and matching error associated with dual-frame estimation. These latter errors were not measured in the census of agriculture process. Information relating to these errors is provided in the sections that follow.

The 2007 Census of Agriculture process measured the error introduced by the nonresponse algorithm, the coverage algorithm, and integerization. The root mean squared error (RMSE) 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 all possible outcomes of the census collection, including variants as to who was on the census list, who returned a census form, and which weights were chosen to be rounded up. The RMSE was used rather than the standard error because it could capture additional error arising from integerization and the potential for bias in the calibration targets. The RMSE is the square root of the sum of the weighted differences between the final recorded value and its expected value squared divided by the number of reports.

Table B presents the fully adjusted total with the root mean squared error for selected items. The relative root mean squared error is obtained by dividing the root mean squared error by the value of the estimate and then multiplying by 100. The table also includes the percent contribution to the mean squared error (the square of the root mean squared error) from nonresponse adjustment and from coverage adjustment.

NONMEASURED CENSUS ERROR

As noted in the previous section, sampling errors can be introduced from the nonresponse and coverage adjustment procedures. This error is measureable. However, nonsampling errors are imbedded in the census process which 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 dual frame 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.

Census Response Rate

The response rate is an 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. The response rate for the 2007 Census of Agriculture is 85.2 percent as compared with a response rate of 88.0 for the 2002 Census of Agriculture and 86.2 percent for the 1997 Census of Agriculture. There was no effort to measure nonresponse bias for the census. However, the census will be used to measure nonresponse bias in NASS surveys.

The response rate for the 2007 Census of Agriculture was calculated as the ratio of the total respondents after data collection was completed to the number of CML records after those that were undeliverable-as-addressed were removed. The total respondents consisted of three groups – those respondents not eligible for the nonresponse survey, those in the universe for the nonresponse survey but who responded prior to the selection of the nonresponse survey sample, and an estimate of the potential respondents in the nonresponse survey sample universe from the response rate to the nonresponse survey. Additional details of the nonresponse study are found in the section on "Whole Farm Nonresponse Estimation."

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 and 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 nonsampling 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.

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) where as 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 Area 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.

Classification Error

Classification error results when a response to the census is misclassified – either as a farm operation if it does not meet the definition or not as a farm operation when it meets the definition. The definition of a farm operation in the 2007 Census of Agriculture is an operation that has $1,000 in agricultural sales or the potential for $1,000 in agricultural sales. A Classification Error Study (CES) has historically been conducted after the census of agriculture. The objectives of a CES are to examine the procedures used to determine farm status (in-scope or out-of-scope) to see if they are producing accurate decisions, document the sources of errors resulting in overcounts and undercounts, and recommend strategies to eliminate them from future censuses. Classification error is a component of census coverage error in addition to coverage error resulted from list incompleteness or duplication. Historically, measures have indicated that the error is small. There has not been any attempt to incorporate this error measure in the coverage adjustment procedure for the 2007 Census of Agriculture.

Prior to 1997 a list based re-interview sample of census respondents was used to measure classification error in the census – specifically the number of farms incorrectly classified as non-farms (undercount) and the number of duplicate farms (overcount). Additionally, an area frame survey was used separately to measure the largest component of census coverage error – incompleteness of the census list. Following the 1997 census, NASS conducted the CES for the 11 western States of Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. The 1997 CES used information from the June Area Survey (JAS) enumeration in lieu of re-interviews; estimates were based on the JAS. The 1997 CES

results indicated a net undercount of 27,971 farms (non-farms incorrectly classified as farms minus duplicate farms and farms incorrectly classified as non-farms) in the eleven States. While the standard error of this estimate is not available to determine statistical significance, even if statistically significant, it represents a relatively small portion of the overall undercount.

Following the 2002 census, the CES similarly used an area-based approach that was conducted in all States. The 2002 CES matched census records to JAS records to identify the differences in farm status of an operation. The JAS area frame-based survey data were assumed to be truth and the estimates of misclassification (records which were incorrectly classified as farms or non-farms and duplicates) were based on this assumption. The 2002 CES results indicated a net overcount of 51,345 farms at the US level, with a standard error of 6,456. In this case, substantial resources were expended to estimate something relatively small. Estimates of overcount and undercount were computed but were not used to adjust totals. Results of the 2002 CES were documented in an internal NASS research report titled "Results from the 2002 Classification Error Study" dated April 2007.

For the 2007 Census of Agriculture, a classification error research study (CES) was conducted in five States -- Arizona, Georgia, Minnesota, New York, and Washington. Estimates of net error were not generated, as the CES was quality research and limited to the five States. Review of the 2002 CES indicated the assumption that the JAS was the truth was inappropriate and re-interviews were reinstated. The 2007 CES used data from the 2007 JAS and the 2007 census to examine farms incorrectly classified as nonfarms, nonfarms incorrectly classified as farms, and to examine records with significant discrepancies in reporting of land between the JAS and census reports. The overall objectives of the 2007 CES were to identify legitimate changes in operations and determine the source of potential errors in the data.

Records in the 2007 JAS were matched to the 2007 census using probabilistic record linkage. From the set of matched records, three groups of interest were identified: 1) in-scope JAS records that were out-of-scope on the census, 2) census in-scope and JAS non-agricultural records, and 3) in-scope census and JAS records with acreage differences of more than 25 percent. Farms whose farm status was in disagreement were interviewed to determine which source was correct; a reason for the change of status on the census was recorded. For records with a discrepancy between the data reported on the 2007 JAS and the 2007 census forms, respondents were re-contacted and asked to verify their data and resolve the difference.

Results of the 2007 CES showed that true changes in size of operations between the JAS and census were rare. Most discrepancies in farm status were the result of errors in reporting with respondents indicating most often that the census data rather than the JAS data were correct, challenging the previous assumption that the JAS data was the truth. Results of the 2007 CES will be used as input for redesign efforts for the JAS operational procedures and the 2012 census report form and instructions.

Table A. **Summary of State Nonresponse and Coverage Adjustments: 2007**

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

Item	Total	Percent from nonresponse adjustment	Percent from coverage adjustment
Farms ..... number	31,169	17.4	8.0
Land in farms ..... acres	43,666,403	11.9	1.9
Farms by size:			
1 to 9 acres ..... farms	920	13.4	19.0
acres	4,485	12.2	21.1
10 to 49 acres ..... farms	3,898	13.2	18.4
acres	103,542	13.6	17.3
50 to 69 acres ..... farms	920	12.7	16.1
acres	53,465	12.6	16.0
70 to 99 acres ..... farms	1,492	14.3	13.0
acres	120,817	14.3	13.0
100 to 139 acres ..... farms	1,283	15.1	12.4
acres	147,938	15.1	12.3
140 to 179 acres ..... farms	2,214	14.5	13.1
acres	349,640	14.6	13.1
180 to 219 acres ..... farms	882	15.9	12.0
acres	174,908	15.9	12.0
220 to 259 acres ..... farms	929	17.4	10.4
acres	221,051	17.4	10.5
260 to 499 acres ..... farms	4,063	19.7	9.0
acres	1,501,281	19.9	8.9
500 to 999 acres ..... farms	4,714	22.3	5.9
acres	3,386,253	22.5	5.7
1,000 to 1,999 acres ..... farms	4,362	23.5	0.6
acres	6,124,293	23.5	0.1
2,000 acres or more ..... farms	5,492	13.7	-1.2
acres	31,478,730	8.0	1.1
Market value of agricultural products sold $1,000	6,570,450	11.3	4.7
Farms by value of sales:			
Less than $1,000 ..... farms	6,981	13.2	18.0
$1,000	465	11.0	24.5
$1,000 to $2,499 ..... farms	1,296	12.4	23.9
$1,000	2,174	12.1	24.2
$2,500 to $4,999 ..... farms	1,077	15.8	8.2
$1,000	3,860	16.0	7.9
$5,000 to $9,999 ..... farms	1,433	13.7	12.0
$1,000	10,349	14.1	12.3
$10,000 to $19,999 ..... farms	1,809	17.5	5.0
$1,000	26,034	17.3	5.2
$20,000 to $24,999 ..... farms	706	17.4	4.4
1,000	15,727	17.5	4.4
$25,000 to $39,999 ..... farms	1,576	18.1	5.6
$1,000	50,300	18.1	5.6
$40,000 to $49,999 ..... farms	939	20.3	5.6
$1,000	41,910	20.3	5.7
$50,000 to $99,999 ..... farms	3,409	20.7	6.3
$1,000	247,296	20.4	6.4
$100,000 to $249,999 ..... farms	5,511	23.2	2.2
$1,000	919,156	23.2	3.1
$250,000 to $499,999 ..... farms	3,588	23.8	0.9
$1,000	1,267,732	23.4	2.0
$500,000 to $999,999 ..... farms	1,687	11.3	-5.9
$1,000	1,177,490	10.3	-4.5
$1,000,000 or more ..... farms	1,157	2.2	11.4
$1,000	2,807,957	1.3	10.0
Farms by type of organization:			
Family or individual ..... farms	26,633	17.9	8.3
acres	29,348,956	14.8	1.0
Partnership ..... farms	2,658	15.3	7.2
acres	7,247,814	7.7	5.8
Corporation:			
Family held ..... farms	1,299	14.2	4.5
acres	3,804,217	7.1	1.6
Other than family held ..... farms	122	12.3	3.3
acres	180,565	5.5	-0.1
Other - cooperative, estate or trust, institutional, etc. ..... farms	457	8.8	8.3
acres	3,084,851	0.8	2.0

Item	Total	Percent from nonresponse adjustment	Percent from coverage adjustment
Tenure:			
Full owners ..... farms	15,626	15.3	11.6
acres	11,646,998	9.8	3.2
Part owners ..... farms	12,754	19.6	3.3
acres	28,889,546	12.6	1.3
Tenants ..... farms	2,789	19.0	9.6
acres	3,129,859	12.9	2.7
Principal operator characteristics by-			
Sex of operator:			
Male ..... farms	28,775	17.5	7.7
acres	41,768,446	11.9	2.0
Female ..... farms	2,394	15.4	11.6
acres	1,897,957	10.9	0.3
Primary occupation:			
Farming ..... farms	18,775	18.7	4.6
Other ..... farms	12,394	15.3	13.2
Spanish, Hispanic, or Latino origin (see text) ..... farms	116	13.8	16.4
acres	87,645	13.9	2.7
Race:			
American Indian or Alaska Native ..... farms	748	15.1	13.6
acres	3,357,973	8.8	-2.1
Asian ..... farms	19	21.1	15.8
acres	6,481	12.3	1.2
Black or African American ..... farms	10	20.0	20.0
acres	(D)	(D)	(D)
Native Hawaiian or Other Pacific Islander ..... farms	1	0.0	0.0
acres	(D)	(D)	(D)
White ..... farms	30,299	17.4	7.9
acres	40,034,883	12.2	2.3
More than one race reported ..... farms	92	15.2	7.6
acres	252,116	8.9	-5.4
Reporting primary occupation as farming by age group:			
Under 25 years ..... farms	129	15.5	17.1
25 to 34 years ..... farms	1,227	18.7	8.4
35 to 44 years ..... farms	2,228	18.0	5.0
45 to 54 years ..... farms	5,002	18.5	4.3
55 to 64 years ..... farms	4,716	19.6	3.0
65 years and over ..... farms	5,473	18.6	4.9
Reporting primary occupation as other than farming by age group:			
Under 25 years ..... farms	113	13.3	24.8
25 to 34 years ..... farms	886	15.5	18.6
35 to 44 years ..... farms	1,817	15.9	13.0
45 to 54 years ..... farms	3,698	15.4	13.6
55 to 64 years ..... farms	3,119	14.7	13.3
65 years and over ..... farms	2,761	15.6	10.3
All operators by age group [1]:			
Under 25 years ..... farms	755	17.1	13.2
25 to 34 years ..... farms	3,884	17.1	9.8
35 to 44 years ..... farms	6,877	16.3	8.9
45 to 54 years ..... farms	13,026	16.6	8.5
55 to 64 years ..... farms	10,924	17.5	7.0
65 to 74 years ..... farms	6,409	17.8	6.6
75 years and over ..... farms	3,989	17.5	6.1

[1] Data were collected for a maximum of three operators per farm.

Table B. **Reliability Estimates of State Totals: 2007**

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

Item		Total	Root mean squared error (RMSE)	Relative RMSE (percent)	Nonresponse contribution to MSE (percent)	Coverage adjustment contribution to MSE (percent)
Farms	number	31,169	214	0.7	4.1	95.9
Land in farms	acres	43,666,403	379,116	0.9	7.3	92.7
Farms by size:						
1 to 9 acres	farms	920	36	3.9	20.2	79.8
	acres	4,485	202	4.5	20.4	79.6
10 to 49 acres	farms	3,898	85	2.2	12.1	87.9
	acres	103,542	2,304	2.2	13.9	86.1
50 to 69 acres	farms	920	34	3.7	20.0	80.0
	acres	53,465	1,959	3.7	19.8	80.2
70 to 99 acres	farms	1,492	43	2.9	19.6	80.4
	acres	120,817	3,469	2.9	19.6	80.4
100 to 139 acres	farms	1,283	39	3.1	20.8	79.2
	acres	147,938	4,565	3.1	20.8	79.2
140 to 179 acres	farms	2,214	53	2.4	19.9	80.1
	acres	349,640	8,365	2.4	19.9	80.1
180 to 219 acres	farms	882	33	3.7	23.4	76.6
	acres	174,908	6,472	3.7	23.5	76.5
220 to 259 acres	farms	929	33	3.6	23.8	76.2
	acres	221,051	7,896	3.6	23.7	76.3
260 to 499 acres	farms	4,063	72	1.8	21.3	78.7
	acres	1,501,281	27,022	1.8	21.9	78.1
500 to 999 acres	farms	4,714	78	1.7	22.1	77.9
	acres	3,386,253	57,129	1.7	22.7	77.3
1,000 to 1,999 acres	farms	4,362	77	1.8	19.2	80.8
	acres	6,124,293	108,056	1.8	19.3	80.7
2,000 acres or more	farms	5,492	69	1.3	16.6	83.4
	acres	31,478,730	324,857	1.0	11.4	88.6
Market value of agricultural products sold	$1,000	6,570,450	67,004	1.0	5.4	94.6
Farms by value of sales:						
Less than $1,000	farms	6,981	177	2.5	3.1	96.9
	$1,000	465	24	5.2	13.2	86.8
$1,000 to $2,499	farms	1,296	99	7.6	2.0	98.0
	$1,000	2,174	167	7.7	2.2	97.8
$2,500 to $4,999	farms	1,077	57	5.3	7.0	93.0
	$1,000	3,860	203	5.3	7.4	92.6
$5,000 to $9,999	farms	1,433	73	5.1	5.8	94.2
	$1,000	10,349	525	5.1	6.1	93.9
$10,000 to $19,999	farms	1,809	55	3.0	12.8	87.2
	$1,000	26,034	799	3.1	13.2	86.8
$20,000 to $24,999	farms	706	31	4.4	18.3	81.7
	1,000	15,727	686	4.4	18.5	81.5
$25,000 to $39,999	farms	1,576	50	3.2	16.9	83.1
	$1,000	50,300	1,612	3.2	17.1	82.9
$40,000 to $49,999	farms	939	36	3.8	21.8	78.2
	$1,000	41,910	1,592	3.8	21.9	78.1
$50,000 to $99,999	farms	3,409	78	2.3	14.8	85.2
	$1,000	247,296	5,649	2.3	15.7	84.3
$100,000 to $249,999	farms	5,511	121	2.2	8.7	91.3
	$1,000	919,156	19,813	2.2	10.0	90.0
$250,000 to $499,999	farms	3,588	76	2.1	14.0	86.0
	$1,000	1,267,732	27,237	2.1	14.3	85.7
$500,000 to $999,999	farms	1,687	26	1.5	31.7	68.3
	$1,000	1,177,490	18,460	1.6	29.5	70.5
$1,000,000 or more	farms	1,157	22	1.9	8.1	91.9
	$1,000	2,807,957	46,397	1.7	4.4	95.6
Farms by type of organization:						
Family or individual	farms	26,633	196	0.7	6.2	93.8
	acres	29,348,956	295,304	1.0	10.1	89.9
Partnership	farms	2,658	50	1.9	22.8	77.2
	acres	7,247,814	158,131	2.2	8.6	91.4
Corporation:						
Family held	farms	1,299	32	2.5	24.7	75.3
	acres	3,804,217	76,433	2.0	17.9	82.1
Other than family held	farms	122	9	7.6	21.1	78.9
	acres	180,565	10,918	6.0	24.0	76.0
Other - cooperative, estate or trust, institutional, etc.	farms	457	19	4.2	21.7	78.3
	acres	3,084,851	47,045	1.5	5.5	94.5
Tenure:						
Full owners	farms	15,626	152	1.0	10.3	89.7
	acres	11,646,998	130,000	1.1	14.5	85.5
Part owners	farms	12,754	137	1.1	9.8	90.2
	acres	28,889,546	312,450	1.1	9.6	90.4
Tenants	farms	2,789	58	2.1	23.1	76.9
	acres	3,129,859	92,388	3.0	12.8	87.2
Principal operator characteristics by-						
Sex of operator:						
Male	farms	28,775	209	0.7	4.5	95.5
	acres	41,768,446	372,591	0.9	7.4	92.6
Female	farms	2,394	69	2.9	9.5	90.5
	acres	1,897,957	51,835	2.7	20.7	79.3
Primary occupation:						
Farming	farms	18,775	168	0.9	7.3	92.7
Other	farms	12,394	141	1.1	10.4	89.6
Spanish, Hispanic, or Latino origin (see text)	farms	116	12	10.0	27.7	72.3
	acres	87,645	15,748	18.0	22.9	77.1

See footnote(s) at end of table.

--continued

Table B. **Reliability Estimates of State Totals: 2007** - Con.

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

Item		Total	Root mean squared error (RMSE)	Relative RMSE (percent)	Nonresponse contribution to MSE (percent)	Coverage adjustment contribution to MSE (percent)
Principal operator characteristics by- Con.						
Race:						
American Indian or Alaska Native	farms	748	75	10.0	2.6	97.4
	acres	3,357,973	140,276	4.2	11.0	89.0
Asian	farms	19	4	18.4	33.0	67.0
	acres	6,481	3,039	46.9	35.5	64.5
Black or African American	farms	10	2	21.5	50.2	49.8
	acres	(D)	(D)	(D)	(D)	(D)
Native Hawaiian or Other Pacific Islander	farms	1	0	0.0	-	-
	acres	(D)	(D)	(D)	(D)	(D)
White	farms	30,299	221	0.7	3.9	96.1
	acres	40,034,883	361,060	0.9	6.7	93.3
More than one race reported	farms	92	9	10.3	26.5	73.5
	acres	252,116	14,455	5.7	22.3	77.7
Reporting primary occupation as farming by age group:						
Under 25 years	farms	129	18	14.1	12.2	87.8
25 to 34 years	farms	1,227	52	4.2	12.6	87.4
35 to 44 years	farms	2,228	49	2.2	23.8	76.2
45 to 54 years	farms	5,002	77	1.5	19.0	81.0
55 to 64 years	farms	4,716	73	1.6	21.0	79.0
65 years and over	farms	5,473	81	1.5	20.2	79.8
Reporting primary occupation as other than farming by age group:						
Under 25 years	farms	113	16	14.2	13.7	86.3
25 to 34 years	farms	886	45	5.0	12.5	87.5
35 to 44 years	farms	1,817	48	2.6	21.3	78.7
45 to 54 years	farms	3,698	69	1.9	18.8	81.2
55 to 64 years	farms	3,119	64	2.1	18.4	81.6
65 years and over	farms	2,761	60	2.2	17.8	82.2
All operators by age group [1]:						
Under 25 years	farms	755	40	5.3	15.2	84.8
25 to 34 years	farms	3,884	110	2.8	9.9	90.1
35 to 44 years	farms	6,877	105	1.5	19.9	80.1
45 to 54 years	farms	13,026	147	1.1	17.1	82.9
55 to 64 years	farms	10,924	131	1.2	19.2	80.8
65 to 74 years	farms	6,409	96	1.5	20.8	79.2
75 years and over	farms	3,989	73	1.8	20.8	79.2
Net cash farm income of operations (see text):						
Farms with gains of [2] -						
Less than $1,000	farms	846	34	4.0	19.0	81.0
	$1,000	400	18	4.6	21.1	78.9
$1,000 to $4,999	farms	2,485	56	2.2	17.8	82.2
	$1,000	7,118	172	2.4	18.6	81.4
$5,000 to $9,999	farms	1,822	46	2.6	19.6	80.4
	$1,000	13,452	350	2.6	19.9	80.1
$10,000 to $24,999	farms	3,529	69	2.0	18.9	81.1
	$1,000	59,604	1,203	2.0	19.6	80.4
$25,000 to $49,999	farms	3,423	68	2.0	20.9	79.1
	$1,000	124,836	2,521	2.0	21.2	78.8
$50,000 or more	farms	9,878	120	1.2	10.6	89.4
	$1,000	2,211,942	25,834	1.2	9.3	90.7
Farms with losses of -						
Less than $1,000	farms	949	37	3.9	19.7	80.3
	$1,000	460	20	4.4	21.8	78.2
$1,000 to $4,999	farms	2,878	72	2.5	13.9	86.1
	$1,000	8,015	212	2.6	14.8	85.2
$5,000 to $9,999	farms	1,806	52	2.9	18.2	81.8
	$1,000	12,920	379	2.9	18.6	81.4
$10,000 to $24,999	farms	1,888	50	2.6	21.9	78.1
	$1,000	29,587	791	2.7	22.6	77.4
$25,000 to $49,999	farms	810	29	3.6	27.0	73.0
	$1,000	28,021	1,031	3.7	27.7	72.3
$50,000 or more	farms	855	28	3.2	27.1	72.9
	$1,000	120,353	4,514	3.8	13.2	86.8

[1] Data were collected for a maximum of three operators per farm.
[2] Farms with zero net cash income are included as farms with gains of less than $1,000.

Table C. **Summary of Nonresponse and Coverage Adjustments by County: 2007**

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

Geographic area	All farms			Land in farms			Sales		
	Total (number)	Nonresponse adjustment (percent)	Coverage adjustment (percent)	Total (acres)	Nonresponse adjustment (percent)	Coverage adjustment (percent)	Total ($1,000)	Nonresponse adjustment (percent)	Coverage adjustment (percent)
STATE TOTAL									
South Dakota	31,169	17.4	8.0	43,666,403	11.9	1.9	6,570,450	11.3	4.7
COUNTIES									
Aurora	379	17.9	7.9	364,612	14.3	15.4	102,716	10.6	11.6
Beadle	750	17.3	11.9	769,855	13.8	6.5	195,425	9.5	7.2
Bennett	265	16.6	12.1	753,263	8.6	6.9	38,128	11.7	10.3
Bon Homme	563	19.0	9.4	308,583	20.1	8.7	109,173	15.1	16.2
Brookings	986	16.5	11.7	462,579	16.7	6.7	186,725	11.8	4.8
Brown	1,036	17.7	4.5	1,085,020	13.5	-1.0	248,765	10.7	-1.0
Brule	370	19.5	4.6	518,462	17.7	-0.7	99,712	15.3	0.4
Buffalo	86	17.4	8.1	312,068	15.6	-1.7	25,045	15.8	2.1
Butte	584	15.1	7.2	1,140,405	5.2	-1.2	55,443	8.1	-1.3
Campbell	318	17.3	3.8	400,871	15.2	-1.5	49,391	12.5	3.2
Charles Mix	693	18.9	5.3	660,519	18.3	-1.9	176,225	13.9	-0.2
Clark	577	18.0	8.5	508,768	14.0	6.1	146,496	10.5	3.4
Clay	484	17.4	20.7	266,697	16.5	20.1	80,821	15.2	17.5
Codington	663	16.6	9.0	367,107	14.0	4.2	107,791	9.8	4.4
Corson	392	17.1	4.8	1,283,038	11.8	-2.7	65,475	9.3	1.7
Custer	359	14.8	10.0	601,129	7.9	-5.4	14,377	7.8	-4.0
Davison	406	17.0	9.9	279,524	16.5	10.5	78,142	14.8	7.5
Day	675	17.9	7.4	567,218	16.3	2.1	97,814	14.2	-0.1
Deuel	583	18.4	6.0	317,164	18.4	0.0	105,092	13.6	0.5
Dewey	410	16.1	4.6	1,449,585	11.2	-3.0	49,090	9.6	2.9
Douglas	363	20.1	4.7	225,166	25.4	-4.0	107,066	15.3	-1.8
Edmunds	425	18.8	3.1	656,678	13.1	1.9	162,523	8.8	5.5
Fall River	330	14.8	8.2	949,697	8.8	-5.8	96,927	3.0	-2.2
Faulk	294	19.0	3.4	614,607	13.6	3.5	109,658	9.3	5.4
Grant	555	18.4	9.0	363,689	18.1	3.2	133,526	11.5	1.8
Gregory	511	20.2	4.5	654,445	17.5	4.4	73,425	17.0	9.0
Haakon	284	13.0	12.3	1,151,144	6.1	8.2	53,037	6.6	11.0
Hamlin	449	17.6	8.2	309,740	13.9	4.4	110,977	12.1	1.0
Hand	484	17.6	7.6	898,741	12.0	7.3	163,949	9.2	8.3
Hanson	308	18.2	8.1	219,023	17.2	3.9	67,319	11.5	3.6
Harding	252	10.7	10.3	1,596,101	3.5	5.3	41,251	4.2	6.5
Hughes	305	16.1	6.9	411,199	15.8	-5.4	60,772	14.2	-4.7
Hutchinson	723	19.5	6.9	509,775	18.9	8.5	192,352	13.3	9.3
Hyde	181	18.2	2.2	480,989	12.0	-5.9	47,169	11.7	-6.0
Jackson	297	18.9	-3.7	1,184,156	10.7	-6.3	36,882	13.6	-4.2
Jerauld	239	18.4	0.4	328,624	15.1	-6.1	68,736	11.0	-5.1
Jones	163	19.0	-3.1	519,314	10.5	-6.1	28,811	10.7	-4.5
Kingsbury	551	18.9	7.4	477,481	16.9	1.6	172,487	12.0	4.4
Lake	514	17.5	10.5	314,946	17.0	7.0	131,847	12.9	6.2
Lawrence	301	15.0	10.0	133,503	18.7	-7.8	11,620	14.9	0.2
Lincoln	855	17.5	13.1	332,762	17.3	14.3	158,109	12.7	16.5
Lyman	443	15.1	12.9	976,457	9.2	12.7	84,445	8.4	15.1
McCook	545	18.3	10.8	363,408	16.8	6.2	129,206	12.5	6.7
McPherson	398	18.8	0.8	518,187	16.9	-5.9	84,858	10.7	-4.1
Marshall	523	18.0	5.9	534,178	15.6	0.3	161,290	7.2	1.1
Meade	879	15.0	9.9	2,208,880	9.0	5.5	78,408	9.1	8.5
Mellette	216	19.0	-5.6	729,778	16.8	-10.0	53,268	9.5	-3.8
Miner	356	17.4	8.7	300,076	16.3	6.9	74,972	13.6	8.6
Minnehaha	1,194	16.7	14.4	421,416	17.5	7.8	190,342	13.4	7.5
Moody	556	16.9	9.5	293,395	14.2	7.1	158,114	8.4	7.0
Pennington	655	13.1	13.6	1,185,055	8.3	4.2	56,038	8.2	4.4
Perkins	432	12.7	7.2	1,829,157	6.1	2.8	59,485	7.5	3.5
Potter	238	16.8	-0.8	516,683	9.8	-3.9	90,376	11.2	-3.9
Roberts	887	19.7	7.2	592,889	19.7	0.6	135,347	17.6	0.1
Sanborn	354	17.8	8.8	318,254	16.2	4.8	63,587	12.4	7.2
Shannon	250	15.2	20.0	1,333,708	9.0	1.5	19,803	12.7	10.3
Spink	624	16.5	6.1	907,643	12.4	5.4	229,139	9.4	5.2
Stanley	165	14.5	1.8	921,110	4.6	0.5	35,208	6.5	6.4
Sully	195	17.9	-3.6	608,976	9.5	2.4	116,459	7.0	2.9
Todd	258	19.4	0.0	869,445	9.6	-4.8	35,806	11.6	-1.0
Tripp	624	19.2	2.4	1,014,336	14.1	2.2	136,677	10.6	4.4
Turner	722	19.9	6.0	371,436	17.6	1.2	173,442	11.9	4.6
Union	521	18.6	10.4	278,916	18.2	5.3	128,205	13.6	2.8
Walworth	279	19.0	-1.8	444,128	15.0	-11.1	54,612	17.4	-14.2
Yankton	658	16.4	17.9	322,242	13.7	25.5	123,591	10.7	25.1
Ziebach	234	13.2	1.7	1,058,403	7.9	-3.4	37,481	8.3	6.7

Table D. **American Indian or Alaska Native Operators: 2007**

[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 [1]	Other [2]		Total	Individually reported [1]	Other [2]
STATE TOTAL				**COUNTIES** - Con.			
South Dakota	1,291	1,291	-	Hughes	4	4	-
				Hutchinson	1	1	-
COUNTIES				Hyde	-	-	-
				Jackson	79	79	-
Aurora	-	-	-	Jerauld	-	-	-
Beadle	4	4	-	Jones	-	-	-
Bennett	74	74	-	Kingsbury	1	1	-
Bon Homme	2	2	-	Lake	5	5	-
Brookings	4	4	-	Lawrence	3	3	-
Brown	12	12	-	Lincoln	1	1	-
Brule	5	5	-				
Buffalo	19	19	-	Lyman	24	24	-
Butte	10	10	-	McCook	-	-	-
Campbell	-	-	-	McPherson	-	-	-
				Marshall	6	6	-
Charles Mix	11	11	-	Meade	12	12	-
Clark	-	-	-	Mellette	39	39	-
Clay	3	3	-	Miner	1	1	-
Codington	9	9	-	Minnehaha	6	6	-
Corson	79	79	-	Moody	4	4	-
Custer	5	5	-	Pennington	38	38	-
Davison	1	1	-				
Day	11	11	-	Perkins	11	11	-
Deuel	3	3	-	Potter	-	-	-
Dewey	237	237	-	Roberts	16	16	-
				Sanborn	-	-	-
Douglas	-	-	-	Shannon	237	237	-
Edmunds	4	4	-	Spink	-	-	-
Fall River	8	8	-	Stanley	5	5	-
Faulk	-	-	-	Sully	-	-	-
Grant	4	4	-	Todd	128	128	-
Gregory	16	16	-	Tripp	20	20	-
Haakon	4	4	-				
Hamlin	1	1	-	Turner	7	7	-
Hand	-	-	-	Union	-	-	-
Hanson	1	1	-	Walworth	4	4	-
				Yankton	-	-	-
Harding	4	4	-	Ziebach	108	108	-

[1] Data were collected for a maximum of three operators per farm.
[2] Data represent American Indian or Alaska Native farm or ranch operators on reservations who did not report individually. Data obtained from reservation officials.