United States Department of Agriculture

National Agricultural Statistics Service

Research and Development Division
Washington DC 20250

RDD Research Report
Number RDD-10-03

December 2010

Impact of the Screening Procedures of the June Area Survey on the Number of Farms Estimates

Denise A. Abreu
Jaki S. McCarthy
Leslie A. Colburn

EXECUTIVE SUMMARY

The National Agricultural Statistics Service uses its annual June Area Survey (JAS) as the vehicle to generate annual estimates of farm numbers. A farm is defined as a place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the year, and the computation includes any government agricultural payments received. Every five years, the annual numbers of farms estimates are compared to ones obtained from the quinquennial Census of Agriculture (conducted for all years ending in 2 and 7). The annual numbers have been declining steadily between censuses, especially between the 2002 and 2007 Censuses. Furthermore, they have been considerably lower than farm numbers from the census, and the difference cannot simply be attributed to sampling error. This trend has prompted concerns leading to a sequence of actions intended to address and resolve the current issues with farm number estimation. Additionally, results from a 2007 qualitative study revealed that agricultural operations were being incorrectly classified as non-agricultural during the screening procedures of the June Area Survey.

In an attempt to get a better understanding of misclassification on the JAS, a post-June Survey intensive screening called the Farm Numbers Research Project (FNRP) was undertaken in the fall of 2009. The study's main objective was to determine the extent of misclassification resulting from the operational screening procedures for the June survey and its immediate impact on number of farms estimation. The study focused primarily on the newly rotated-in segments which comprised 20 percent of the 2009 JAS sample. It targeted three types of tracts in these segments: 1) non-agricultural tracts; 2) agricultural tracts that refused to participate in the JAS and were estimated; and 3) agricultural tracts that were inaccessible in June and were estimated. Respondents completed a short and simple questionnaire.

The primary purpose of the survey was to verify farm/non-farm status, utilizing a survey instrument that contained questions similar to those on the JAS screening form. The study introduced the concepts of subsampling and subtracts, which are not routinely part of the JAS. Also, it was believed that farming operations might be missed during the operational screening process because enumerators are instructed to visually inspect residential areas of one-half acre or less per parcel for any agricultural activity. For these small parcels, if there is no clear indication of agricultural activity, the tract is identified as non-agricultural and no JAS questionnaire is completed. In order to be able to assess the potential impact of this operational process for the FNRP study, enumerators were instructed to ignore the ½-acre rule.

The evaluation of inaccessible and refusal tracts was essential since the farm status of these tracts from the JAS is based solely on observed and/or estimated data. Based on the results of the study, the operational JAS slightly underestimated (by 5,210) the number of farms from these tracts. Even with the small net change in number of farms resulting from re-screening these estimated agricultural tracts, the effort was worthwhile in confirming that there was indeed not a huge problem with them. Also, the additional time spent re-screening them proved to be very beneficial in converting refusals and contacting the inaccessible ones.

Of more significance from the FNRP results, though, was the finding that the operational

screening efforts have resulted in misclassification of a substantial number of non-agricultural tracts. Tracts identified as non-agricultural in June expanded to 576,613 additional farming operations as a direct result of the FNRP re-screening effort. An even more surprising finding was that the vast majority of these farms were found in June non-agricultural tracts identified as having no potential for agriculture.

The results also showed that the JAS is probably not missing a high number of farming operations as a result of the ½-acre rule, since a minute number of farming operations in the FNRP sample had ½ acre or less inside the June tract. The subsampling strategy employed showed that over half of the farming operations found had only one subtract in the sampled June tract – indicating a proper parceling of operating arrangements using the operational procedures.

Another supposition of this study was that the JAS was missing farms in agri-urban or commercial strata. Although the JAS is indeed missing some farms in this type of stratum, these do not constitute a large percentage of the total farms missed or misclassified. However, the findings were consistent with expectations that mostly smaller operations are being missed or misclassified, since additional ones identified through the re-screening are primarily small, with less than 24 acres in the entire operation and a value of production or sales less than $10,000.

The overall results of this study point directly to potential improvements in the screening questionnaire and to enhanced enumerator training guidelines. The results are very encouraging, suggesting that with more time and resources a better quality screening of the non-agricultural tracts can be implemented, resulting in the proper identification of more farming operations.

RECOMMENDATIONS

1. ***Determine the agricultural/non-agricultural status of each tract by asking the screening questions rather than by observation.*** It is important that all tract operators be questioned whether or not there is any agricultural potential in the tract.

2. ***Add a question/column to the screening questionnaire similar to Question 1 of the Land Utilization Survey: Did you own, rent, or operate land with potential for agricultural production?*** Results of the FNRP showed that close to 63.3 percent of agricultural operations answered 'yes' to this question, indicating that potentially more non-agricultural tracts could be identified as agricultural during the screening. If a question is not added, emphasize to enumerators and office staff that this should not be determined by enumerator observation (if at all possible) since land operated outside the segment cannot be observed.

3. ***Retain the "½-acre rule."*** It appears that relatively few farm operations will be missed, compared to the gain in resource savings, by continuing to employ the "½-acre rule."

4. ***Provide better training for enumerators and field office (FO) survey administrators concerning screening procedures.*** This should be a standard part of the mid-year training school. Additionally, emphasize that field enumerators and FO staff should utilize all tools available (including asking questions of the respondent's neighbors) to gain valuable information about agricultural producers in the segment. This way, enumerators will essentially ask for information about every person in the segment; thereby accounting for every farm in the segment. Training should emphasize proper drawing of boundaries and proper allocation of ownership/operator definition and identification. FO staff noted that when they conducted the follow-up interviews, it was obvious that several mistakes were made in June in identifying tract operators and/or tract boundaries. They found that the majority of wooded areas were not being allocated to the proper tract and/or owner/operator. Strengthen the section in the Survey Administration Manual which refers to screening and include more examples. This recommendation is already being adopted by the Survey Administration Branch.

5. ***Extend the screening period to start several weeks earlier than the usual start date to facilitate more intensive screening*** FNRP results showed that this may reduce the number of refusal and inaccessible tracts for which data are estimated. Ensure that all FOs prescreen new segments and, for June 2010, potentially prescreen "unknown" operators from older segments. This recommendation was implemented beginning in June 2010.

6. ***Obtain better quality information on the non-agricultural tracts in order to facilitate more accurate screening.*** Results of the post-screening showed that the information obtained from the First American Spatial Solutions was useful to the states in accomplishing this. Additionally, states made use of Google maps as an additional

source of name and address information for non-agricultural tracts. FOs were able to categorize numerous office hold cases more accurately and many were classified as farms. The methods the FOs used in the FNRP should be reviewed to see if those used by individual states could be used more broadly.

7. ***Conduct a yearly un-duplication review of all the June Area Survey (JAS) tracts (agricultural and non-agricultural).*** There were several FNRP subtracts which were linked to other JAS operations located in different segments or tracts. Conducting an un-duplication review could help identify the actual farm status of these non-agricultural tract operators. The name and address information on the non-agricultural tracts could be used by record linkage techniques to attempt to match them to existing agricultural tracts. If these match to other tracts on the JAS, their status could be updated from non-agricultural to agricultural.

8. ***Evaluate the results from this study for their potential impact on the Census of Agriculture number of farms estimates.***

9. ***Use FNRP to enhance regression models being developed by the NASS/NISS[1] team to adjust for misclassification in the JAS during non-census years.*** Currently, the model uses 2007 probabilities of misclassification with a verification of the status of the non-agricultural tracts using Census of Agriculture information.

[1] NASS has a two year collaborative research program with the National Institute of Statistical Sciences (NISS) called the Cross-Sector Research in Residence Program. This program is composed of three academic-government teams focusing on important NASS research issues. One of the teams was entrusted to work on potential improvements to the methodology and design of the June Area Survey.

Impact of the Screening Procedures of the June Area Survey on the Number of Farms Estimates

Denise A. Abreu[2]
Jaki S. McCarthy
Leslie A. Colburn

Abstract

The National Agricultural Statistics Service uses its annual June Area Survey (JAS) as the vehicle to generate annual estimates of farm numbers. These estimates are compared to ones obtained from the quinquennial Census of Agriculture (conducted for all years ending in 2 and 7). The annual estimate of the number of farms from the JAS has been declining steadily between censuses, especially between the 2002 and 2007 Censuses. Furthermore, these have been considerably lower than farm numbers from the census and the difference cannot simply be attributed to sampling error. Some insight into this issue was provided when the 2007 Classification Error Survey, a census follow-up qualitative study, revealed that agricultural operations were being incorrectly classified as non-agricultural during the screening procedures of the JAS.

Therefore, to determine the extent of misclassification resulting from the operational screening procedures for the June survey and its immediate impact on number of farms estimation, a post-June Survey intensive screening called the Farm Numbers Research Project was undertaken in the fall of 2009. The study focused primarily on the newly rotated-in segments which comprised 20 percent of the 2009 JAS sample. It targeted non-agricultural, refusal and inaccessible tracts in these newly rotated-in segments. The primary purpose of the survey was to verify farm/non-farm status, utilizing a survey instrument that contained questions similar to those on the JAS screening form. The study introduced the concepts of subsampling and subtracts, which are not routinely part of the JAS. Also, enumerators were instructed to screen residential areas of one-half acre or less per parcel, yet another deviation from standard JAS procedures. The results of this effort are presented in this report.

KEY WORDS: area frame, re-screening, non-agricultural tracts, misclassification

[2] Denise A. Abreu and Jaki S. McCarthy are statisticians with the National Agricultural Statistics Service (NASS) - Research & Development Division, located at Room 305, 3251 Old Lee Highway, Fairfax, VA 22030. Leslie A. Colburn is a NASS statistician with the agency's Statistics Division, located at 1400 Independence Avenue SW, Washington, DC 20250.

I. INTRODUCTION

On an annual basis, the National Agricultural Statistics Service (NASS) produces numerous estimates for a number of commodities of agricultural importance such as corn, cotton and soybeans, as well as for total number of farms in the United States (U.S.). A farm is defined as a place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold, during the year. The calculation also includes any government agricultural payments. NASS uses its annual June Area Survey (JAS) as the vehicle to generate annual estimates of number of farms. These estimates are compared to ones obtained from the quinquennial Census of Agriculture (conducted for years ending in 2 and 7). The annual estimates of the number of farms from the JAS have been declining steadily between censuses (especially between the 2002 and 2007 Censuses). Furthermore, these have been considerably lower than farm numbers from the census, and the difference cannot simply be attributed to sampling error. Additionally, results from a qualitative study, known as the Classification Error Survey (CES), revealed that agricultural operations were being incorrectly classified as non-agricultural during the screening procedures of the JAS. These developments have prompted concerns leading to a series of actions intended to address and resolve the current issues with farm number estimation.

To address these issues, an aggressive agenda to identify and implement ways to improve the number of farm indications from the JAS was developed that focused on explaining reasons for discrepancies in farm status between the annual JAS and the Census of Agriculture. The developed plans included 1) measures to improve the area sample for farm numbers, 2) the implementation of an intensive post-June survey re-screening and follow-up for bias reduction, and 3) the evaluation of capture-recapture estimation methodology to improve farm number estimation.

This report presents the results of component 2 above -- an intensive post-June survey re-screening effort intended to assess issues with screening procedures and to study misclassification in the JAS.

II. BACKGROUND

NASS conducts an annual area frame based survey which collects information about U.S. crops, livestock, grain storage capacity, and type and size of farms. Since the distribution of crops and livestock can vary considerably across a state, the precision of the survey indications is substantially improved by dividing the land in the state into homogeneous groups or strata and optimally allocating the total sample to the strata. The basic stratification employed by NASS involves: (1) dividing the land into land-use strata such as intensively cultivated land, urban areas and range land, and (2) further dividing each land-use stratum into substrata by grouping areas that are agriculturally similar. The JAS uses a sample of designated land areas (segments) which field enumerators visit to collect data on all agricultural activity occurring therein. A typical segment is about one square mile (640 acres). Each segment is outlined on an aerial photo which is provided to the appropriate field enumerator.

Through field enumeration, a segment is divided into tracts of land, each representing a unique

land operating arrangement. An area screening form is completed for all sample segments. It inventories all tracts within the segment and contains screening questions that determine whether or not each tract has agricultural activity. In this way, all land inside the segment is screened for agricultural activity and the screening applies to all land in the identified operating arrangement (both inside and outside the segment). Those operations (tracts) that qualify as agricultural are interviewed using the area version questionnaire, which collects detailed agricultural information specifically about the operator's land, again both inside and outside the segment.

The area frame is a theoretically complete sampling frame with every acre of land having a known chance of selection. As such, it can be used to estimate the number of farms and land in farms independently of the list frame, as well as to measure incompleteness in the list. The area frame uses a replicated sample design. A sample rotation scheme is used to reduce respondent burden caused by repeated interviewing and to avoid the expense of selecting a completely new area sample each year. Sample rotation is accomplished each year by replacing segments from specified replicates in each land-use stratum with newly selected segments. Approximately 20 percent of the replicates in each land-use stratum are replaced annually.

In addition to the JAS and the annual list-based surveys, NASS conducts a Census of Agriculture every five years (for years ending in 2 and 7). The Census of Agriculture is a complete count of U.S. farms and ranches and the people who operate them. The census collects data on land use and ownership, operator characteristics, production practices, income and expenditures, and many other characteristics. The outcome, when compared to earlier censuses, helps to measure trends and new developments in the agricultural sector of our nation's economy. The information is used only for statistical purposes and data are published only in tabulated totals. Census forms are sent to all known and potential agricultural operations in the U.S. The census provides the most uniform, comprehensive agricultural data for every county in the nation. It employs a dual frame: an independent list frame and the area frame from the JAS. The area frame is used as a measure of incompleteness of the list frame.

Historically after each census, an evaluation has been conducted to measure misclassification of farms on the census mail list. This evaluation involves either recontacting a sample of census respondents or overlap matching the census mail list to the area frame. For the 1997 and 2002 Censuses of Agriculture, classification errors were measured by comparing an operation's status on the census to its status on the area frame based JAS. In cases where there were discrepancies between the two, the JAS was assumed to be correct, and the operations were counted as misclassified on the census. For 2007, the primary focus of the misclassification evaluation was to identify reasons for discrepancies between the JAS and the census. The 2007 Classification Error Survey (CES) was a qualitative examination of why classification and reporting errors occur. The 2007 CES results showed that most of the discrepancies were actual errors which occurred in the JAS, *not* the census. The U.S. Bureau of the Census has also found errors in screening area frame samples in its household surveys (Manheimer and Hyman, 1949; Eckman, 2009). These results also suggested that screening methods for the JAS should be reviewed. There were numerous cases of operations incorrectly classified as non-agricultural based upon non-agricultural land inside the segment, without recognizing the associated agricultural land outside the segment.

III. MOTIVATION AND PROJECT DEVELOPMENT

Results from the 2007 CES indicated that misclassification was not limited to the census, but that some agricultural operations were incorrectly classified as non-agricultural during the screening procedures of the 2007 JAS (Abreu et. al., 2009). However, the CES was a very small study and produced no estimates of either JAS or census misclassification. In an attempt to get a better understanding of misclassification on the JAS, a post-June Survey intensive re-screening called the Farm Numbers Research Project (FNRP) was undertaken in the fall of 2009.

The general idea for and framework of this project was formed out of a cross-Divisional meeting in December 2008, called to identify ways in which NASS might address the increasing spread between the number of farms indications from the JAS and the Census of Agriculture. In January 2009, the meeting notes/draft proposal document from the December meeting was widely circulated in Headquarters for comment. This document was refined based on the resulting feedback, and the final version (included in Appendix A) was presented and approved for implementation at the March 2009 Program Planning Council[3] (PPC) meeting.

In anticipation of PPC approval, a cross-Divisional team, later to be called the Farm Numbers Research Project (FNRP) Team, was formed in January to flesh out the details of this very complex and extensive project. Unlike most research projects, there was to be no small scale pre-testing for this project. A basic premise of the project development was that it would be conducted as an operational activity in all states following the 2009 JAS. The concern being addressed was considered serious enough that senior management felt that the Farm Numbers Agricultural Statistics Board would need results from this project for every state for use in December 2009, when setting the 2009 number of farms estimates. In trying to address this need, the Team met weekly, with growing membership (as additional operational units' participation was required), from January until the project was ultimately launched in August.

FNRP's main objective was to determine the extent of misclassification resulting from the operational screening procedures for the June survey. The study focused primarily on the 20 percent newly rotated-in segments of the 2009 JAS. Addressing these segments specifically would help provide better long term estimates of farm numbers, due to their longevity in future samples, in addition to laying the ground work for better screening procedures for future segments.

The project targeted three types of tracts in the newly rotated-in segments:

1. Non-agricultural tracts,
2. Agricultural tracts that were estimated, because their operators refused to participate in the JAS, and
3. Agricultural tracts that were estimated, because their operators were inaccessible in June.

These tracts were to be re-contacted via mail, phone or through personal interviews as efficiently as possible, commensurate with achieving high quality results. Tracts which were entirely composed of Public, Industrial, or Grazing Association (PIGA) land were excluded, as PIGA

[3] The Program Planning Council is responsible for short-term and medium-term planning, implementing the strategic vision, monitoring programs, and determining future actions at NASS.

land is used on a fee-per-head or Animal Unit Month (AUM) basis, and is not included in the annual land in farms estimates.

IV. SAMPLE SELECTION

There were 2,465 segments in the 2009 JAS that were eligible for the study. However, only 2,209 segments contained tracts in any of the three targeted categories described above. From those segments, there were 10,204 total tracts identified for this project, of which 8,552 were identified as non-agricultural during the 2009 JAS screening. An additional 1,652 were agricultural tracts whose operators refused or were inaccessible during the 2009 interview period.

The sample included only operations from the newly rotated-in segments which comprised 20 percent of the 2009 JAS sample. It is important to note that not every state had 20 percent of their JAS sample rotated in for 2009. For the three states that received an entirely new frame in 2009 -- Minnesota, South Dakota and Washington -- only selected replicates (approximately 20 percent) were contacted. For New Hampshire, Rhode Island, and West Virginia (states that did not have any new segments in 2009) the rotation group targeted to leave the sample last was selected. Florida, Utah, and Wyoming had very few new segments in 2009, so for these states, segments in prior years were selected in addition to the few newly rotated-in segments.

In most states, sample sizes were relatively small (see Appendix B for sample sizes by state). The sample also included specific segments from the Agricultural Coverage Evaluation Survey (ACES), which were selected only in Arkansas, Connecticut, Maine, Massachusetts, Michigan, Montana, Ohio, South Carolina, and Vermont to aid with farm number estimates for the operational June survey, *not* for FNRP. These supplemental ACES segments were allocated to less cultivated strata where small farms are prevalent and more likely to be missed.

V. DATA COLLECTION

In conducting this study, all operators in the sampled June tracts needed to be identified, including all individuals in residential or other non-agricultural tracts. One major issue with non-agricultural tracts has been the lack of good name and address information. The rigorous re-screening undertaken for this post-survey effort focused heavily on obtaining mailable names and addresses (and phone numbers, if available) for all the selected places of interest in the targeted tracts. A place of interest is defined as any individual residing within the sampled tract or any land area that could have been drawn off as a separate tract within the sampled tract, regardless of size.

JAS segments were overlapped with a real estate parcel dataset from First American Spatial Solutions[4] (FASS) to get name and address information for as many tract operators within the

[4] First American Spatial Solutions (FASS) is a provider of property location information and geospatial services. They are now a division of Core Logic and aggregate up-to-date tax and municipal databases from around the country which provide geospatial analytics and solutions targeted around these areas. The ParcelPoint boundary database, which is a digitized and accurate spatial database of actual parcel boundaries of 125 million parcels in the U.S, was obtained from FASS. In addition, this database provided tax identification numbers, names and addresses.

scope of the study as possible. FASS data provided field offices (FOs) with an additional source of information to identify owners of the land within the segment. The FASS database contained name, parcel address, owner address if different from parcel, telephone number, and tax identification number. Maps were produced showing all parcels within each JAS segment. One map was created for each segment, and each FASS parcel on the map contained a number referring to an accompanying report delivering the information described above. See Appendix C for examples of maps of JAS segments displaying FASS parcel numbers. FO staff received JAS segments with FASS information and used it to build a mailing list of potential names and addresses for subtracts for which this information was not otherwise available. This process proved especially beneficial in densely populated areas, since it provided a listing of land owners, which helped enumerators with pre-screening. In addition to the use of FASS data, FO staff also looked to obtain operators' information through Google maps and any other sources identified locally.

Operators of the selected places of interest in the study tracts were asked to participate in this effort, referred to in the field as the Land Utilization Survey. The primary purpose of the survey was to verify farm/non-farm status, and respondents to it completed a short and simple questionnaire that contained questions similar to those on the JAS screening form. The questionnaire did not have the exhaustive list of crops and livestock items that are in the June questionnaire; however, it collected enough information to calculate points for tracts which did not report sufficient sales. The calculated points allowed for a definitive farm/nonfarm determination for these tracts. See Appendix D for a copy of the survey instrument used for this study.

The current JAS screening procedures implement a ½-acre rule which instructs enumerators to visually inspect residential areas of ½ acre or less per parcel for any agricultural activity (Bosecker et. al., 1988). If there is no clear indication of agricultural activity, the tract is identified as non-agricultural and no JAS questionnaire is completed. This is done to minimize interviewing costs in enumerating every unit in a segment, which is especially important for densely populated areas of land where the cost of enumeration can be very high. For the FNRP, enumerators were instructed to ignore the ½-acre rule and attempt to go to every place of interest (i.e., operator) in a specific June tract. As a result of this, the concepts of subsampling and subtracts, which are not routinely part of the JAS, were introduced. It is important to note that current NASS procedures define each tract as a unique land operating arrangement. However, for densely populated June-defined tracts it is possible that multiple operations (places of interest) may be present. For example, a residential development with multiple houses may be designated as one tract and labeled "houses." For the FNRP, houses within the selected tract had to be screened individually.

To get the best possible results from the FNRP study, commensurate with containing data collection costs at an acceptable level, a special subsampling scheme was developed and the concept of subtracts was introduced. For any sampled June tract containing seven or fewer places of interest, all of them were interviewed. For tracts with 8-20 places of interest present, enumerators were instructed to randomly interview 1/2 of them. For tracts with 20 or more places of interest present in the target tract, enumerators were instructed to interview 1/6 of them. The selected units in the target tract were determined through rolling a die for each place of

interest. No more than 20 places of interest were interviewed in any given tract. Additionally, no fewer than three places of interest were interviewed in each tract, unless there were fewer than three total places of interest in the tract.

The project allowed substantial time and flexibility in scheduling and coordinating data collection efforts. FOs could collect their data in any manner they deemed appropriate beginning in late August and continuing through early November. For the most part, FOs had an initial mail-out in late September with additional mailings to follow. They employed several different data collection strategies; however, much of the phone/field follow-up occurred in October. Editing and summarization of the results occurred in November.

VI. RESULTS AND CONCLUSIONS

There were 10,204 tracts within the scope of FNRP's selected segments. The subsampling scheme employed for FNRP identified 6,987 additional places of interest eligible to be interviewed. The resulting total of 17,191 places of interest, or subtracts, represented a 68.5 percent increase, indicating that there were potentially additional farming operations within a NASS tract. Overall, the study resulted in 12,847 completed subtracts, a 74.7 percent completion rate.

Because the JAS is a probability-based sample, each tract has an inclusion probability π_i and an expansion factor $e_i = 1/\pi_i$. Within each tract identified as a farm, a proportion of a farm is observed. This proportion, the tract-to-farm ratio, is t_i = tract acres / farm acres. Both of these are used in calculating the current JAS estimate for the number of farms, which is defined as follows,

$$\sum_{i=1}^{l} \sum_{j=1}^{s_i} \sum_{k=1}^{n_{ij}} e_{ijk} a_{ijk}$$

where

i indexes stratum
j indexes substratum
k indexes segment
l = Number of land-use strata
s_i = Number of substrata in stratum *i*
n_{ij} = Number of segments in substratum *j* within stratum *i*
e_{ijk} = Expansion factor or the inverse of the probability of the selection for each segment in substratum *j* in land-use stratum *i*

$$a_{ijk} = \sum_{m=1}^{x_{ijk}} t_{ijkm}$$

m indexes tract

x_{ijk} = Number of tracts identified as farms in the given segment

t_{ijkm} = Tract-to-farm ratio of the tract = $\frac{\text{tract acres for the } m^{\text{th}} \text{ tract}}{\text{farm acres for the } m^{\text{th}} \text{ tract}}$

The results in Table 1 show the total number of subtracts (or places of interest in the study), and the resulting number of farms and expanded number of farms by tract type (agricultural or non-agricultural tract). Henceforth, any mention of subtract or place of interest will refer to FNRP, and tract will refer to the operational JAS.

Table 1: Results by Type of Agricultural Tract

Type of Agricultural Tract	FNRP Total Subtracts	Number of FNRP Farms	Percent of Subtracts that were Farms	Number of FRNP Non-farms	Percent of Subtracts that were Non-farms	Net Expanded Number of Farms
Estimated agricultural tracts	1,712	1,503	87.8%	209	12.2%	5,210
Non-agricultural tracts w/ potential	487	95	19.5%	392	80.5%	38,346
Non-agricultural tracts w/ unknown potential	364	56	15.4%	308	84.6%	37,479
Non-agricultural tracts w/out potential	14,628	905	6.2%	13,723	93.8%	500,338
Totals	**17,191**	**2,559**	**14.9%**	**14,632**	**85.1%**	**581,373**

Of the 17,191 subtracts in the study, 14.9 percent (or 2,559 subtracts) were identified as farm operations. More than half of these farms (1,503) were from agricultural tracts which were estimated during the JAS summary. These expanded to 256,613 farms, but since most of these agricultural tracts were counted as farms during the June Survey, the net increase in farms from the FNRP re-screening of them was only 5,210.

The remaining 1,056 farms, which expanded to 576,163, were from tracts identified as non-agricultural during the JAS. The coefficient of variation (CV) associated with this latter indication is 10.9. It is important to note that as a result of the subsampling scheme used for the FNRP, JAS tract expansion factors (e_i) had to be adjusted by subsampling rates. This adjustment resulted in over 150 FNRP subtracts identified as farms receiving expansion factors exceeding 1,000. One very important note is that most of the farms added in the FNRP (86.1% of them) were farm operations identified in tracts that had been classified as non-agricultural without any potential for agriculture.

Of the 2,559 subtracts qualifying as farm operations, 122 (4.8 percent) were from ACES segments. These ACES subtracts, which only occurred in nine states, expanded to 77,289 farms (9.3 percent of all expanded farms). Since the ACES subtracts contributed a proportionately small percentage of the farms identified by the study, all results presented in this report will include the ACES subtracts.

Re-screening of Estimated Agricultural Tracts (Refusals & Inaccessibles)

The survey period of the JAS (including summarization of the data) lasts one month. FOs have two weeks prior to the start of data collection to prescreen new and residential segments to identify tract operators and agricultural activity. Due to this limited timeframe, enumerators do not devote as much time as is probably needed to identify tract operators of non-agricultural tracts. The study focused on two types of tracts: 1) tracts classified as agricultural which refused or were inaccessible during the JAS data collection period and thus estimated, and 2) tracts identified as *not* having any agricultural activity. The farm status of inaccessible and refusal tracts from the JAS is based solely on observed and/or estimated data. It was important to evaluate these tracts further to determine whether they had been correctly classified. Based on the results of the FNRP, the number of farms from these tracts was underestimated by 5,210.

The FNRP sample initially consisted of 1,652 estimated and inaccessible tracts which comprised 16.2 percent of the total FNRP sample. Through the re-screening and further follow-up, the study identified 60 additional places of interest (i.e., FNRP subtracts), resulting in three percent more subtracts in FNRP than in the original JAS. Table 2 below shows the total number of estimated and inaccessible subtracts in the study, their JAS farm status, their FNRP farm status, and the expanded number of farms from both the FNRP and the JAS.

Table 2: Refusal and Inaccessible Tracts – JAS vs FNRP Farm Status

	JAS Total Tracts	FNRP Total Subtracts	FNRP Non-Farm	FNRP Farm	FNRP Expanded Farms	JAS Expanded Farms (n=1,652)	FNRP Net Expanded Farms
JAS Non-Farm	115	121	84	37	13,032	0	13,032
JAS Farm	1,537	1,591	125	1,466	243,581	251,403	(7,822)
Totals	**1,652**	**1,712**	**209**	**1,503**	**256,613**	**251,403**	**5,210**

The study results confirmed the farm status of 92.1 percent $\left(\frac{1,466}{1,591}\right)$ of the tracts identified as farms during the JAS. The study also found, though, that 30.6 percent $\left(\frac{37}{121}\right)$ of the tracts identified as non-farms during the JAS should actually have been classified as farms. The expanded number of farms for the 1,503 FNRP farms from refusal and inaccessible estimated tracts is 256,613. By comparison, the June data for the 1,652 refusal and inaccessible tracts in the FNRP sample expanded to 251,403 farms. Therefore, the study results indicate that there should be an upward correction of about 2.1 percent or 5,210 farms $\left(\frac{256,613-251,403}{251,403}\right)$ to the total number of farms from the JAS from estimated tracts, indicating slightly more farms from the FNRP re-screening even for these.

The farm status for over 80 percent of the tracts for which the JAS and FNRP farm status disagreed (37 and 125 in Table 2) was actually determined through data collection for both the JAS and the FNRP, rather than held in the office or estimated. Most of these farms were borderline ones with low values of sales. Also, many of them did not report sufficient sales to

qualify them directly as a farm. Therefore, in order to have their final farm/non-farm status determined, $1,000 worth of points were calculated if they had enough agricultural inventory, both in the operational JAS and the FNRP. Additionally, most of them had a single place of interest in the tract, simplifying the screening process.

It is important to examine how the farm/nonfarm status of these tracts was determined through the FNRP re-screening. Toward that end, Table 3 shows the final status of each of the JAS refusal and inaccessible tracts by their original refusal/inaccessible status on the JAS.

Table 3: Final FNRP Response Status of JAS Refusal and Inaccessible Tracts in the Study

	Total Sub-tracts	Completed FNRP Re-screenings			FNRP Estimated Refusals			FNRP Estimated Inaccessibles			FNRP Estimated Office Hold		
		Sub-tracts	Farms	Exp. Farms	Sub-tracts	Farms	Exp. Farms	Sub-tracts	Farms	Exp. Farms	Sub-tracts	Farms	Exp. Farms
JAS refusals	728	228	213	33,670	124	122	18,719	55	50	5,272	321	318	42,241
JAS Inacc.	984	544	413	93,155	78	69	12,653	158	129	27,537	204	189	23,366
Totals	1,712	772	626	126,825	202	191	31,372	213	179	32,809	525	507	65,607

Of the 728 JAS refusal tracts, 31.3 percent or 228 were completed via telephone, mail, or face-to-face interviewing. Although the study managed to convert a number of these refusals, a fairly high percentage of them were also nonresponse in the FNRP. About 17.0 percent of the JAS refusals remained refusals and 7.5 percent became inaccessible, 124 and 55 respectively. However, JAS refusals continue to be an issue of concern since 44.1 percent (321 reports) of all subtracts were held in the office and not part of any other data collection effort.

The results were more encouraging for the original 984 JAS inaccessible tracts. The FNRP re-screening was completed for over half of these tracts via one of the three major modes of data collection employed. About 7.9 percent (78 subtracts) of the JAS inaccessible tracts were refusals for the FNRP re-screening, and only 16.1 percent (158 subtracts) remained as inaccessible. Even though 20.7 percent were held in the office and not part of any data collection effort, the overall results indicate that it might be possible to obtain responses for these tracts if more data collection time were available.

Table 4: Results of JAS Estimated FNRP Subtracts by FNRP Mode of Interview

FNRP Mode of Interview	Total Subtracts	Total FNRP Farms from Estimated JAS Tracts	Total FNRP Farms		FNRP Expanded Farms from JAS Estimated Tracts
			from JAS Refusal Tracts	from JAS Inaccessible Tracts	
Mail	267	239	102	137	39,724
Telephone	410	378	170	208	70,555
Face-to-Face	291	211	43	168	55,104
CATI	108	89	40	49	16,695
Web	0	0	0	0	0
Fax	1	1	1	0	8
Other[5]	110	78	29	49	8,919
Office Held[6]	525	507	318	189	65,607
Total	**1,712**	**1,503**	**703**	**800**	**256,613**

For a majority of the JAS refusal and inaccessible tracts, FOs were able to get responses via mail or telephone. However, a substantial number of the inaccessible tracts were ultimately re-screened via face-to-face interviewing (See Table 4).

Table 5: Results for JAS Estimated Tracts in FNRP by Number of Places of Interest

Number of Places of Interest	Number of JAS Estimated Tracts	FNRP Total Subtracts	FNRP Farms	FNRP Expanded Farms
1	1,623	1,623	1,466	249,132
2-7	28	78	30	7,469
8-20	1	11	7	12
20 or more	0	0	0	0
Totals	**1,652**	**1,712**	**1,503**	**256,613**

As indicated in Table 5, subtracts and subsampling were not that much of a concern for refusal and inaccessible tracts, since most of the farms (97.5 percent or $\left(\frac{1,466}{1,503}\right)$) were found in JAS tracts with only one FNRP subtract. The subsampling did, however, identify about 7,500 more farming operations among JAS tracts having two to seven FNRP subtracts.

[5] These subtracts were linked to other JAS operations located in different segments or tracts and their JAS information was copied over.

[6] Office held cases were June refusal or inaccessible tracts that the FO chose not to recontact.

Re-screening of Non-Agricultural Tracts

In addition to the refusal and inaccessible tracts discussed in the previous sections of this report, the FNRP study focused on tracts which were identified as non-agricultural during the JAS screening procedures. In the June survey, each tract is screened to determine whether it is part of an agricultural operation. Even if there is no agriculture in the tract, it should be classified as agricultural if it is part of an operation with agricultural production elsewhere. Additionally, a non-agricultural tract is classified into one of the following three categories: has potential for agriculture, does not have potential for agriculture, or potential for agriculture is unknown. The rigorous re-screening undertaken for this post-survey effort focused heavily on obtaining mailable names and addresses (and phone numbers, if available) for all tracts in the study segments.

The overall evaluation of the non-agricultural tracts confirmed that the operational screening efforts have resulted in misclassification of a substantial number of these tracts. The study found 1,056 JAS tracts that had been classified in June as non-agricultural, that were actually part of farming operations. These misclassified, non-agricultural tracts expanded to 576,161 farms. An even more surprising finding was that 85.7 percent (905 out of 1,056) of these farms were found in non-agricultural tracts identified as having no potential for agriculture (see Table 6). Recall that the agricultural potential of a non-agricultural tract is normally determined by the enumerators during the screening procedures of the JAS. The results of this study point directly to potential improvements in the screening questionnaire and to enhanced enumerator training guidelines. Overall the results are very encouraging, suggesting that with more time and resources a better quality screening of the non-agricultural tracts could be implemented, which would result in the proper identification of more farming operations.

Table 6: Results of JAS Non-Agricultural Tracts in FNRP by FNRP Response Code

FNRP Response Code	FNRP Total Subtracts	FNRP Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
Completed	12,075	983	90	36,884	49	34,636	844	483,703
Refusal - Estimated	190	16	3	679	0	0	13	616
Inaccessible Estimated	1,713	42	1	385	6	2,815	35	15,190
Office Held Estimated	173	15	1	398	1	28	13	829
Known Zero[7]	1,328	0	0	0	0	0	0	0
Totals	**15,479**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

[7] Known zero records refer to non-agricultural tracts such as schools, cemeteries, railroads, etc. where there is clearly no possibility of the land being part of an agricultural operation.

Table 6 presents results by response category for the FRNP study. The overall completion rate for the study was 74.7 percent, with the completion rate for non-agricultural tracts slightly higher at 78.0 percent (12,075/15,479). Only 5.5 percent (58/1,056) of the farms from non-agricultural tracts were refusals and inaccessibles, which had to be estimated. Also, a small number (15) of these subtracts were linked to other JAS records. These were held in the office and their data were estimated.

The concepts of subtracts and subsampling played a much more important role for the study of non-agricultural tracts. Recall that for a specific tract, all places of interest were interviewed if there were seven or fewer units in the June tract. For tracts with 8-20 places of interest present, enumerators were instructed to randomly interview 1/2 of them. For tracts with 20 or more places of interest present, enumerators were instructed to interview 1/6 of them.

Table 7: Results for JAS Non-Agricultural Tracts in FNRP by Number of Places of Interest

Number of Places of Interest	FNRP Total Subtracts	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
1	7,102	646	82	33,629	36	15,737	528	157,872
2-7	2,457	174	9	4,302	12	2,872	153	37,780
8-20	1,778	88	4	415	3	1,164	81	61,823
20 or more	4,142	148	0	0	5	17,706	143	242,862
Totals	**15,479**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

Table 7 shows the results by the number of places of interest in each type of non-agricultural tract. It is evident that most of the unexpanded farms (77.7 percent or 820/1,056) were found in low density tracts (7 or fewer places of interest). Furthermore, 646 of these farms were found in tracts with only *one* place of interest. This indicates that better in-depth screening of these tracts and more effort in identifying the operator of the tract would yield the identification of more farming operations. A thorough screening of these 646 tracts would have added an estimated 207,238 more farming operations. On the other hand, farms found in high density tracts cannot be ignored as they also constitute a significant percentage of the misclassified farms.

Table 8: Acreage Distribution by Subtract of JAS Non-Agricultural Tracts in the FNRP

Subtract Acreage Distribution	FNRP Total Subtracts	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
0.5 Acre or less	2,551	34	0	0	0	0	34	2,941
0.6 -0.9 Acre	338	9	0	0	0	0	9	1,587
1-24.9 Acres	9,714	628	39	12,976	28	17,762	561	338,055
25-99.9 Acres	1,893	255	37	15,591	17	15,630	201	114,654
100+ Acres	983	130	19	9,779	11	4,087	100	43,101
Totals	**15,479**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

One of the issues to be addressed through this intensive re-screening was a concern over the potential impact of the ½-acre rule. It was believed that farming operations might be missed during the screening as a result of this field enumeration efficiency. In the current operational procedures, enumerators are instructed to visually inspect residential areas of ½ acre or less per parcel for any agricultural activity. For these small parcels, if there is no clear indication of agricultural activity, the tract is identified as non-agricultural and no JAS questionnaire is completed.

Table 8 shows the distribution of subtract acreage by type of non-agricultural tract. The results showed that there were only 34 farms identified from non-agricultural tracts which had ½ acre or less. These expanded to only 2,941 farms, indicating that the JAS is probably not missing a high number of farming operations as a result of the ½-acre rule.

Table 9: Farm Acreage Distribution of JAS Non-Agricultural Tracts in the FNRP

Farm Acreage Distribution	FNRP Total Farms	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
0.5 Acre or less	1	1	0	0	0	0	1	26
0.6-0.9 Acre	3	3	0	0	0	0	3	1,136
1-24.9 Acres	382	301	19	10,796	14	16,741	268	307,203
25-99.9 Acres	419	265	33	16,254	15	14,605	217	118,058
100+ Acres	1,754	486	43	11,296	27	6,133	416	73,915
Totals	**2,559**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

As indicated in Table 9, the results of this study showed that most of the operations being misclassified or missed in the non-agricultural tracts are small farms with acreage between 1 and 24.9 acres (301 farms). These accounted for 334,740 expanded farms. Furthermore, for 28.9 percent of the 1,056 FNRP farms from non-agricultural tracts, the entire operation fell inside the subtract. These operations accounted for 58.3 percent of all expanded farms.

Table 10: Results for JAS Non-Agricultural Tracts in FNRP by FNRP Mode of Interview

FNRP Mode of Interview	FNRP Total Subtracts	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
Mail	1,650	160	19	6,399	4	2,384	137	46,473
Telephone	1,962	253	23	8,032	15	5,359	215	67,384
Face-to-face	8,882	519	42	16,757	29	27,550	448	334,297
CATI	569	84	8	5,929	4	840	72	45,624
Web	13	2	0	0	0	0	2	532
Fax	0	0	0	0	0	0	0	0
Other	902	23	2	831	3	1,318	18	5,199
Known Zero	1,501	15	1	398	1	28	13	829
Totals	**15,479**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

The results for non-agricultural tracts by mode of interview, as displayed in Table 10, indicate that FOs were able to get a reasonable number of responses via mail and telephone. However, for logistical purposes, enumeration for the vast majority of the tracts with eight or more places of interest was conducted via face to face interviews, as FOs had been instructed. Only 23 percent of all subtracts with 8 or more dwellings (5,931) were conducted by means of other modes of data collection and not face to face interviews. Overall, the vast majority of the farms (86.1 percent or 500,338) were found in tracts that had been identified as having no potential for agriculture.

Table 11: Distribution by Land-Use Strata of Non-Agricultural JAS Tracts in the FNRP

Land-Use Strata	FNRP Total Subtracts	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
50% + cultivated	4,868	392	22	3,883	18	4,647	352	80,934
15-49% cultivated	5,575	365	43	16,372	14	3,427	308	177,340
<15% cultivated	671	287	29	18,091	23	29,404	235	237,648
Agri-urban/ Commercial	4,359	12	1	0	1	0	10	4,416
Non-ags	6	0	0	0	0	0	0	0
Totals	**15,479**	**1,056**	**95**	**38,345**	**56**	**37,479**	**905**	**500,338**

Another supposition of this study was that the JAS was missing farms in agri-urban or commercial strata. However, the results in Table 11 show that the 697 total subtracts screened in these strata only identified 12 additional (4,416 expanded) farms. Although the JAS is missing farms in this type of stratum, these do not constitute a large percentage of the farms misclassified. The study did find a significant number of additional farms in the less than 15% cultivated stratum, indicating that current screening procedures more effectively identify agricultural operations in highly cultivated areas than they do in areas with less agriculture.

Table 12: Distribution by Sales of Agricultural Products of Non-Agricultural JAS Tracts in the FNRP

Sales of Agricultural Products	FNRP Total Number Farms	FNRP Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
Less than $10,000	898	630	60	28,290	33	31,041	537	389,701
$10,000-$24,999	258	158	17	5,510	12	5,215	129	70,331
$25,000-$49,999	168	78	7	4,101	5	694	66	27,217
$50,000-$99,999	173	53	2	332	2	246	49	5,852
$100,000-$499,999	547	81	2	15	2	259	77	6,536
$500,000 +	515	56	7	98	2	25	47	702
Totals	**2,559**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

Table 12 presents the results of the non-agricultural JAS tracts in the FNRP by value of sales (including government agricultural payments). These results solidify the hypothesis that small farms are missed or misclassified more often, since most of the farms from non-agricultural tracts (630/1,056) had less than $10,000 in sales of agricultural products. These 630 farms represented 77.9 percent (389,700/500,338) of all farms from JAS non-agricultural tracts identified in the study. However, there were 47 non-agricultural tracts with no potential that reported over $500,000 in sales in FNRP. Interestingly, four of these reported over $5,000,000 in sales. These operations should certainly be on the NASS list frame; however, an extensive search using their name and address information revealed that three out of the four were not, indicating that larger farms are being missed as well.

Table 13: Distribution by Sales Class of Non-Agricultural JAS Tracts in the FNRP

Type of Farm	FNRP Total Farms	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
			Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
Point Farms	667	487	53	27,808	27	29,164	407	328,631
Value of Sales Farms	1,892	569	42	10,538	29	8,315	498	171,707
Totals	**2,559**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

As mentioned previously, a farm is defined as a place from which $1,000 or more of agricultural products were produced and sold, or normally would have been sold during the year, and the computation includes any government agricultural payments received. To ensure a definitive determination of farm/non-farm status for each subtract in the survey, the survey instrument for the FNRP collected enough information to calculate dollar "points" for subtracts which did not report sufficient sales to qualify them directly as a farm. Any land operating arrangement with more than $1,000 worth of points based on agricultural inventory qualifies as a farm on the "normally would have been sold" sales criterion of the farm definition, even if it had no actual agricultural sales. Farming operations qualifying by this criterion are referred to as "point farms." Table 13 shows that the vast majority of the expanded farms identified were from point farms.

Table 14: Number of Farms Identified by Screening Question in the FNRP (Land Utilization) Instrument by Type of Agricultural Tract

Screening Questions * Multiple answers were allowed	FNRP Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
		Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
Question 1: Any Land w/ Ag Potential?	915	83	33,304	48	34,416	784	434,004
Question 2: Any Cropland?	4	1	531	0	0	3	3,043
Question 3: Any Government Agricultural Payments?	8	0	0	1	220	7	1,827
Question 4: Any Hay or Pasture?	26	2	1,395	0	0	24	15,599
Question 5: Raise Livestock?	13	3	1,321	0	0	10	6,981
Question 6: Any Equine?	9	1	333	0	0	8	20,670
Question 7: Any Poultry or Birds?	0	0	0	0	0	0	0
Question 8: Any Bee Colonies?	1	0	0	0	0	1	290
Question 9: Any Aquaculture?	0	0	0	0	0	0	0
Question 10: Any Agricultural Sales?	1	0	0	0	0	1	9

The screening portion of the survey instrument for the FNRP study included ten screening questions designed to screen in as many respondents as possible to complete the four page questionnaire, which ultimately determines farm/nonfarm status. The results by screening question are presented in Table 14. From the table, it is clear that the most important questions that could be asked are Any Potential for Agriculture (Question 1), Any Hay or Pasture (Question 4), Any Livestock (Question 5), Any Government Agricultural Payments (Question 3), and Any Equine (Question 6), in that order. As shown in the table, there were 784 subtracts found in tracts which were classified as non-agricultural with NO potential for agriculture in June that answered that they had land with agricultural potential. These results suggest that it is important to ASK about land with agricultural potential, not just make a determination based on observation of the tract. Previous research by Manheimer and Hyman (1949) and Eckman (2009) also indicate that casual observation of units by enumerators is not always adequate and has led to misjudging and misclassifying units during screening procedures. This is especially true in the JAS screening, since observation can only indicate what's inside the segment, and farm status is determined by all land under that land operating arrangement, both inside and outside the segment.

For FNRP, it is important to note that for face to face, telephone and CATI interviews, interviewers were instructed to complete the questionnaire as soon as they obtained a 'yes' to any one of the screening questions. Thus, questions were *not* to be asked every time from every

respondent. An analysis of the 160 FNRP reports from June non-agricultural tract mail respondents which were subsequently identified as farms showed most of the respondents (94.3 percent) answered the potential for agriculture screening question (Question 1) most often, skipped all other screening questions, and then proceeded to complete the rest of the questionnaire.

Table 15: Distribution by Age of Frame of Non-Agricultural JAS Tracts in the FNRP

Age of Frame	Total States	FNRP Total Subtracts	FNRP Total Farms from Non-agricultural Tracts	Non-agricultural Tracts w/ Potential		Non-agricultural Tracts w/ Potential Unknown		Non-agricultural Tracts w/out Potential	
				Farms	Expanded Farms	Farms	Expanded Farms	Farms	Expanded Farms
5 Years or Less	11	3,824	330	19	9,631	18	14,587	293	156,341
6 to 10 Years	4	2,799	88	8	3,500	0	0	80	46,656
11 to 15 Years	9	3,192	240	41	13,007	15	5,224	184	137,847
16+ Years	24	5,664	398	27	12,208	23	17,668	348	159,494
Totals	**48**	**15,479**	**1,056**	**95**	**38,346**	**56**	**37,479**	**905**	**500,338**

NASS maintains an area frame for each state except Alaska. Frames are constructed one state at a time and used year after year until the frame is deemed outdated, generally in about 15 to 20 years. When a frame becomes outdated, a new frame is constructed to replace it. Analysis of age of frame effect on number of farms misclassified revealed that farming operations missed were reasonably spread regardless of year in which the frame was built (See Table 15). Newer frames (5 years or less) missed about as many farms as did much older ones (16 or more years).

VII. COST ANALYSIS

An initial budget of $600,000 was allocated to FRNP for field data collection. It cost $412,000 to re-screen the 17,191 subtracts that were within the scope of the project. At the U.S. level, it cost $23.97 to screen each subtract and $161.04 for each (unexpanded) farm found. See Appendix E for cost estimates at the U.S. level and by state. In the operational 2008 JAS, the cost per tract was $43, and it cost an additional $23.97 to re-screen these tracts for FNRP.

The cost estimates for each subtract, as presented here, are an over simplification, since they assume that the re-screening cost is the same for non-agricultural and estimated tracts. However, since there was already some information on estimated tracts, it was generally less costly to attempt a re-contact of these. The same amount of information was not available on non-agricultural tracts. In fact, a reasonable amount of time was allocated to finding good name and address information on these tracts. Thus, it is safe to conclude that re-screening of non-agricultural tracts comprised a larger portion of the overall cost per subtract ($23.97). It is important to note that a direct correlation should not be made between the cost per tract for FNRP and the operational JAS, since the methodology and procedures were different. In addition to the budget allocated for direct field enumeration expenses, it cost $92,000 to purchase the real estate information from First American Spatial Solutions, which was used in identifying name and address information for the places of interest in the study.

VIII. RECOMMENDATIONS

1. ***Determine the agricultural/non-agricultural status of each tract by asking the screening questions rather than by observation.*** It is important that all tract operators be questioned whether or not there is any agricultural potential in the tract.

2. ***Add a question/column to the screening questionnaire similar to Question 1 of the Land Utilization Survey: Did you own, rent, or operate land with potential for agricultural production?*** Results of the FNRP showed that close to 63.3 percent of agricultural operations answered 'yes' to this question indicating that potentially more non-agricultural tracts could be identified as agricultural during the screening procedures. If a question is not added, emphasize to enumerators and office staff that this should not be determined by enumerator observation (if at all possible) since land operated outside the segment cannot be observed.

3. ***Retain the "½-acre rule."*** It appears that relatively few farm operations will be missed, especially given the resource savings, by continuing to employ the "½-acre rule".

4. ***Provide better training for enumerators and field office (FO) survey administrators concerning screening procedures.*** This should be a standard part of the mid-year training school. Additionally, emphasize that field enumerators and FO staff should utilize all tools available (including asking questions of the respondent's neighbors) to gain valuable information about agricultural producers in the segment. This way, enumerators will essentially ask for information about every person in the segment; thereby accounting for every farm in the segment. Training should emphasize proper drawing of boundaries and proper application of ownership/operator definition and identification. FO staff noted that when they conducted the follow-up interviews, it was obvious that several mistakes were made in June in identifying tract operators and/or tract boundaries. They found that the majority of wooded areas were not being allocated to the proper tract and/or owner/operator. Strengthen the section in the Survey Administration Manual which refers to screening and include more examples. This recommendation is already being adopted by the Survey Administration Branch.

5. ***Extend the screening period to start several weeks earlier than the usual start date to facilitate more intensive screening,*** FNRP results showed that this may reduce the number of refusal and inaccessible tracts for which data are estimated. Ensure that all FOs prescreen new segments and, for June 2010, potentially prescreen "unknown" operators from older segments. This recommendation was implemented beginning in June 2010.

6. ***Obtain better quality information on the non-agricultural tracts in order to facilitate more accurate screening.*** Results of the post re-screening showed that the information obtained from the First American Spatial Solutions (FASS) was useful to the states in accomplishing this. Additionally, states made use of Google maps as an additional source of name and address information for non-agricultural tracts. FOs were able to

categorize numerous office hold cases more accurately and many were classified as farms. The methods the FOs used in the FNRP should be reviewed to see if methods used by individual states could be used more broadly.

7. ***Conduct a yearly un-duplication review of all the June Area Survey (JAS) tracts (agricultural and non-agricultural).*** There were several FNRP subtracts which were linked to other JAS operations located in different segments or tracts. Conducting an un-duplication review could help identify the actual farm status of these non-agricultural tract operators. The name and address information on the non-agricultural tracts could be used by record linkage techniques to attempt to match them to existing agricultural tracts. If these match to other tracts on the JAS, their status could be updated from non-agricultural to agricultural.

8. ***Evaluate the results from this study for their potential impact on the Census of Agriculture number of farms estimates.***

9. ***Use FNRP to enhance regression models being developed by the NASS/NISS team to adjust for misclassification in the JAS during non-census years.*** Currently, the model uses 2007 probabilities of misclassification with a verification of the status of the non-agricultural tracts using Census of Agriculture information.

IX. REFERENCES

Abreu, D. A., J. S. McCarthy, N. J. Dickey (2009). 2007 Classification Error Survey for the United States Census of Agriculture. Research and Development Division. RDD Research Report: RDD-09-03.

Bosecker, R.R., M. S. Clark (1988). Modifying the Weighted Estimator to Eliminate Screening Interviews in Residential Areas. Research and Applications Division. NASS Research Report: SSB-88-07.

Davies, Carrie (2009). Area Frame Design for Agricultural Surveys. Research and Development Division Internal Document.

Eckman, Stephanie (2009). Coverage Rates and Coverage Bias When Interviewers Create Frames. Presented at the International Total Survey Error Workshop.

Manheimer, D. and H. Hyman (1949). Interviewer Performance in Area Sampling. Public Opinion Quarterly 13(1).

Matthews, Ralph V. (1988). Screening Residential Tracts for Agricultural Activity. Research and Applications Division. NASS Research Report: SSB-88-05.

X. ACKNOWLEDGEMENTS

The authors would like to sincerely thank all the people who provided their insightful comments and invaluable contributions to this research report and the FNRP project:
Mark Apodaca, Pam Arroway, Dale Atkinson, Kevin Barnes, Norman Bennett, Patrick Boyle, Scott Cox, Carrie Davies, Matt Deaton, Virginia Harris, Kevin Hintzman, Bill Iwig, Andrea C. Lamas, Kenneth K. Lopiano, Leslee Lohrenz, Scott Shimmin, Linda J. Young

Appendix A

Farm Numbers Operational Improvements and Research Program for 2009

(Updated 03/05/09)

The number of farms and land in farms data series are very important and sensitive ones to USDA and their production by the National Agricultural Statistics Service (NASS) is critical to the reputation of the agency as a provider of accurate, unbiased statistics. During the 2007 Census of Agriculture it was discovered that a significant number of farms were missed in the previous year's area frame estimates of number of farms. The working assumption of NASS has been that the area frame is complete and that all farms and commodities are captured in the area frame expansions. This assumption is due to the theoretical completeness of the frame, the perceived quality of the personal enumeration utilized for the data collection, and the extensive training of field enumerators collecting the data. However, the 2007 census results showed that many of the smallest farms were not captured in either the area frame or on the census mail list. In addition, the 2007 Classification Error Survey (CES) showed that a significant number of small and minority farms added to the census mail list during the 2007 mail list development cycle were not correctly classified as farms in the area frame sample.

Thus, estimates of farm numbers based on the area frame have under-estimated number of farms. This proposal is designed to test improvements to the operational process and outline research designed to improve the estimates by testing alternative procedures to measure the number of farms. The components of this are as follows:

A. Improve the Area Sample for Farm Numbers
B. Implement an Intensive Post-Survey Screening and Follow-up for Bias Reduction
C. Evaluate Capture – Recapture Estimation of the Number of Farms
D. Analyze 2007 Census data for Not-on-Mail-List (NML) Tracts to Improve Future List Development

A. Improve the Area Sample for Farm Numbers

The Area Frame Sample is allocated at the national level for farm numbers. This has resulted in many states having a smaller number of segments, with the national allocation of segments serving as a balance in support of both commodity and farm number estimation. To improve the estimation of farm numbers from the area frame, supplemental segments will be selected this year to augment the June 2009 sample size. The additional segments will be a subset of the ACES segments used in June 2007 to ensure the quality of the Not-on-Mail-List component of the 2007 Census of Agriculture results. The inclusion of the additional segments will supplement the national allocation and lower CVs for farm numbers, particularly for small farms. Many of these segments are located in the agri-urban strata where many small farms reside.

Through the efforts of the Area Frame Section and the Survey Administration Branch (in costing out sample augmentation options), the attached worksheet provides an allocation of ACES segments (181) to use for June 2009. The worksheet also shows 122 segments being added for the new Native American strata in the southwestern states, per previous discussion at the October 2008 PPC meeting.

The criteria used to determine where to add the ACES segments are as follows:

1. States with the largest difference between the "fully adjusted" Census indication and the four-year average JAS direct expansion for number of farms.

2. States with high CVs or those showing the greatest potential for improvement of CVs.

States not considered were Native American states (AZ, CO, NM, NV, and UT), new frame states (MN, SD, WA), and states with a small number of farms. Also not considered were CA, MS and TX which were part of the Minority Agricultural Coverage Evaluation Survey (MACES). The MACES segments were added primarily to improve minority number of farm estimates by creating new substrata. Adding them back for this purpose would have resulted in an overly complex sample weighting process.

Other comments:

Plans are **not** to reprint the ACES photos, meaning the segments numbers will be same as they were during June 2007. Their complete 6 digit ACES segment numbers will be unique; however, the last 4 digits will **not** be unique.

For the recommended states, only a portion of their 2007 ACES sample was allocated to minimize FO workload. Most of the allocation was distributed to the moderate to light cultivation strata (20s and 40s).

B. Implement an Intensive Post-Survey Screening and Follow-up for Bias Reduction:

No changes to normal prescreening, screening or data collection will be implemented for the 2009 June Area Survey (including any additional segments added as outlined above). FOs will follow normal times and procedures, and they should continue to utilize FSA screening where history has shown it to be advantageous.

However, an additional, post-survey screening will be conducted on all newly rotated-in 2009 June Area Survey and added ACES segments, which will focus on collecting quality names and addresses for all tracts. This additional, enhanced screening will be especially needed for tracts that were identified as non-ag, or were inaccessible or estimated for the June survey. The additional screening will be conducted through the most cost-effective way possible to collect these complete (and mailable) names and addresses for all tracts and occupied dwellings. This follow-up, intensive screening will occur after completion of the June Survey, and must be done in the most cost efficient manner possible, including coordination, where possible, with other field visits. The primary objective will be to identify quality names and addresses (and telephone numbers, as available) for all tracts not previously identified as farm operators -- not to collect data. Therefore, if mailable names and addresses can be obtained, at least in part, through other means (e.g., tax records) in lieu of door-to-door canvassing, this much-less expensive, indirect method should be used.

All tracts in the newly rotated-in June 2009 and added ACES segments previously coded as non-ag or non-ag with potential or as inaccessible or estimated will subsequently be followed-up through an Agricultural Identification Survey (AIS) to determine their true farm status. The AIS instrument will utilize enhanced screening questions developed by Research and Development Division cooperators, Drs. Don Dillman and Danna Moore of Washington State University (WSU). This subsequent screening for farm/nonfarm determination will use mail, phone, EDR

and phone follow-up to maximize response. The focus will be on a definitive determination of the farm status of each name in the most efficient manner possible. The AIS procedures being researched and developed by WSU cooperators will be designed to maximize mail and Internet response, in order to make this follow-up screening as efficient as possible.

An estimate of the number of farms missed in the current operational procedures will be made from the number of additional operations identified in the enhanced screening, and a new estimate of the number of farms from the June Area Survey will be calculated.

This component of the overall farm number improvement program for 2009 is designed to reduce the RMSE of farm number estimation (especially for small farms) by minimizing the bias from farms being missed during the screening procedures.

C. Evaluate Capture – Recapture Estimation of the Number of Farms

An estimate of the number of farms not found on the area frame will also be measured by utilizing capture-recapture methods with the 2009 Agricultural Resource Management Survey (ARMS) Phase I and the 2009 June Area Survey. The procedures will be tested in advance by utilizing the 2007 Census of Agriculture and the 2007 June Area Survey. Results for 2008 will not be available since there was no ARMS screener that year.
See the supplemental attached document (Some Thoughts On Adjusting for the Not-on-Either-Frame Component of the 2007 Census of Agriculture) by Phil Kott for a general discussion of this methodology. See also the attached document (A Capture-Recapture Indication for the 2007 Census of Agriculture and an Estimator for its Variance) by Phil for details on the capture-recapture indication and an estimator of its variance.

One of the under-pinning requirements of capture-recapture methodology is that all target population units have a non-zero probability of selection in each frame involved – here list and area.

All units innately had a non-zero probability of selection in the 2007 June Area survey, with the exception of any new operations formed after June that qualified as 2007 farms. The problem is murkier from the list side, since reduced list sampling is conducted for most surveys focusing solely on production agriculture. From the list standpoint, the requirement of non-zero probabilities of selection are probably best satisfied by the Census of Agriculture and the Agricultural Resource Management Survey (Phase I), where small and limited resource farms are of concern.

Some consideration will be given to possible additional data that could be collected on future June area surveys that would help with subsequent year capture-recapture estimation with the ARMS Phase I.

D. Analyze 2007 Census data for Not-on-Mail-List (NML) Tracts to Improve Future List Development

Appendix A

A full census questionnaire was sent to all 2007 ag and non-ag with ag potential June Area tracts not found on the census list. The data were collected primarily to ensure the accuracy of the scrubbing process in determining the correct match/non-match status between the operations on the list and area sides, but they also have potential for shedding light on the characteristics of farming operations that were missed in the census list building process. Data mining and cluster analysis will be used to identify subgroups of the NML tracts to provide guidance into future, improved list building efforts.

Appendix A

Some Thoughts On Adjusting for the Not-on-Either-Frame Component of the 2007 Census of Agriculture

The National Agricultural Statistics Service (NASS) estimated the number of farms in 2002 assuming that its area-frame survey was complete. That is to say, every 2002 farm was assumed to be either on the Census Mailing List (CML) or accounted for in the estimate of farms not on the mailing list (NML) derived for the agency's area-frame sample (AFS). In the terminology used here, both the NML and AFS estimates incorpororate the additional segments selected for the Agriculture Coverage and Evaluation Survey (ACES) and the Minority Agriculture Coverage and Evaluation Survey (MACES).

NASS will make the same assumption in 2007 but it has strong evidence the NML estimate underestimates the total number of 2007 farms not on the CML. This is because the agency's CML + NML estimate of the number of farms is signficantly larger than its AFS estimate of the number of farms.

This note will discuss methods of total-farm-count estimation that capture the farms missing from *both* the CML and the area frame (variance estimates for these methods will wait for another time). Although, in principle, the area-frame should be complete, in practice, the methods used to collect information from an area-frame sample often fail to account for all the farms in sampled segments. In this note, we will treat the expected values of the NML and AFS estimated farm counts using the 2007 sample-collection methodology as the "true" counts of the NML and area-frame populations, recognizing that they are smaller than they would be were area-frame enumeration perfect.

For simplicity, we assume there are no erroneous matches between CML and AFS records and that all true matches are uncovered, although there can be CML farms in area-frame sampled segments that do not make it into the AFS.

Appendix A

Capture-Recapture

The traditional method of total-farm-count estimation is called "capture-recapture" (the name derives from the possibility that a farm is captured in the CML and then recaptured in the AFS). It assumes that the population of interest can be broken in mutually exclusive groups such that each farm in a particular group has a fixed probability of being missing from the CML and a fixed probability of being "missing from the area frame," by that we mean if its segment were selected for the AFS, it would have this probability of failing to be accounted a farm. Both probabilities are asumed to be less than 1 so that *every farm has a positive probability of being both on the CML and the area frame*. Within a group, a farm's probability of being missing from one frame is assumed to have no effect on its probability of being missing from the other.

Restricting our attention to one of these groups, let

BOTH be the count of farms on both the CML and the area frame,

AFS be the count of farms on the area frame,

NML be the count of farms on the area frame only,

CML be the count of farms on the CML,

NAF be the count of farms on the CML only (not on the area frame), and

NEF be the count of farms not on either frame.

Note that $AFS = BOTH + NML$, while $CML = BOTH + NAF$. The total number of farms is $TOTAL = BOTH + NML + NAF + NEF = AFS + NAF + NEF = CML + NML + NEF$.

By the assumptions, the probability of farm k in the group being on the CML, call it p_{Lk}, is the same as the probability of an area-frame farm in the group being on the CML. This probability is approximately $BOTH / AFS = BOTH / (BOTH + NML)$. In fact, this ratio is an unbiased estimator of that probability. That suggest the following estimator for *TOTAL* when *NEF* is unknown but everything else is known:

Appendix A

$$\widehat{TOTAL} = CML\frac{AFS}{BOTH}$$

or

$$\widehat{TOTAL} = CML\frac{BOTH + NML}{BOTH} = CML + NML\frac{CML}{BOTH}$$
$$= CML + NML\frac{BOTH + NAF}{BOTH}$$
$$= CML + NML + NML\frac{NAF}{BOTH},$$

where NEF is apparently estimated by NML (NAF / $BOTH$).

In reality, AFS, $BOTH$, and NML need to be estimated from the area-frame sample. Even CML must be estimated (from the census) due to nonresponse (we will assume that this is done in an unbiased fashion). That means we have two alternative estimators for $TOTAL$:

$$\widehat{TOTAL}_1 = \widehat{CML}\frac{\widehat{AFS}}{\widehat{BOTH}} \quad (1)$$

and

$$\widehat{TOTAL}_2 = \widehat{CML} + \widehat{NML}\frac{\widehat{CML}}{\widehat{BOTH}}. \quad (2)$$

The latter is preferred in practice because it is closely tied to the traditional two-frame *screening estimator*,

$$\widehat{TOTAL}_0 = \widehat{CML} + \widehat{NML}, \quad (3)$$

which it will exceed when $\widehat{BOTH} < \widehat{CML}$. This *should* be the case since $BOTH < CML$ when there is a greater than zero probability that a farm in the group is missing from the area frame. Unfortunately, $\widehat{BOTH}$ is a random variable based on the area-frame sample, while $\widehat{CML}$ comes from the census, so there is no guarantee $\widehat{BOTH} < \widehat{CML}$, especially

when the group size is small. When $\widehat{BOTH} > \widehat{CML}$ in practice, one usually either uses $\widehat{TOTAL}_0 = \widehat{CML} + \widehat{NML}$ as the estimated total for the group or collapses groups until the inequality goes in the right direction.

Before proceeding, observe that $\widehat{BOTH} / \widehat{CML}$ is an estimator for the probability that a CML farm in the group is on the area frame, which makes it an estimator for the probability that *any* farm k in the group is on the area frame, call it p_{Ak}. Thus, in equation (2), $\widehat{CML} / \widehat{BOTH}$ is effectively an estimator for $1/p_{Ak}$ which is applied to all NOL farms in the group.

Determining the Groups

When the Census Bureau conducted the Census of Agriculture in 1992, it based its official tabulations strictly on the CML. There was, however, a coverage-adjusted farm count provided in the appendix computed using equation (2). For this purpose, each state was divided into two groups, one with farms having sales of $2,500 or less and one with farms having greater than $2,500 in sales.

A better way to form groups is by first fitting a sample-weighted logistic model (either by state or for the entire country with state dummies as explanatory variables) on the area-frame sample with on-the-CML (yes or no) as the dependent variable. The tract-to-farm ratios for AFS farms should be incorporated into their sampling weights. The independent variables (*e.g.*, size class, race of principal operator, type of farm, *etc.*) in the final model must be available for all AFS and Census-responding records. *This means area-frame stratum information can not be used.*

Let $\mathbf{x}_k$ be the vector of independent-variable values associated with farm k (ideally, this information should come from census forms, even for AFS farms), and $\mathbf{b}$ be the estimated logistic regression coefficients. Compute the estimate probability of being on the CML,

$$\hat{p}_{Lk} = \exp(\mathbf{x}_k'\mathbf{b})/[1 + \exp(\mathbf{x}_k'\mathbf{b})], \qquad (4)$$

for every AFS and *every Census-responding (CML) record k*. Sort the combined set of AFS farms and Census respondents by their $\hat{p}_{Lk}$ values. Create, say, G groups by collecting the farms with the smallest 1/G $\hat{p}_{Lk}$ values into one group, the farms with the next smallest 1/G $\hat{p}_{Lk}$ values into a second group, and so forth, making adjustments when there are more natural breaks in the $\hat{p}_{Lk}$ values (and collapsing if necessary so that $\widehat{BOTH} \leq \widehat{CML}$ in each group). Each group should contain enough AFS farms that are also in the CML to assure that $\widehat{BOTH}$ is stable (does not have too much variance).

I recommend creating 20 groups at the national level. If state-specific groups are preferred, I recommend no more than five groups per state.

If the groups cut across states, and a single state farm-count total is needed, one can treat the ratio $\widehat{CML}/\widehat{BOTH}$ computed for each group as an additional multiplicative factor in the area-frame weight of NML farms in the group. That is to say, the adjusted weight for NML farm k is $w_{Ak}^* = w_{Ak}(\widehat{CML}/\widehat{BOTH})_k$, where w_{Ak} is the farm's AFS weight (including its tract-to-farm ratio), and $(\widehat{CML}/\widehat{BOTH})_k$ is the $\widehat{CML}$-to-$\widehat{BOTH}$ ratio for the group containing it. Recall from the last section that this ratio effectively estimates $1/p_{Ak}$. By incorporating this factor in the NML weights, new augmented NML estimates are created that implicitly include estimates of the not-on-either-frame farms.

A Continous Approach Based on Weighting the CML Respondents
(an alternative measure for quick estimation and double checking)

Rather than arbitrarly breaking the AFS sample and Census-responding farms into groups, an alternative approach simply applies the $\hat{p}_{Lk}$ values estimated in equation (7) to every CML-responding farm. If w_{Lk} is the (pre-integerized) nonresponse weight for CML-responding farm k, then the new estimated total is

$$\widehat{TOTAL}_3 = \sum w_{Lk} / \hat{p}_{Lk} , \qquad (5)$$

where the summation is over every CML-responding farm in the domain of interest (a state, the country as a whole, etc.).

This is effectively a generalization to the estimator in equation (1) where the group is effectively all farms with the same component values of the vector **x**.

There is no guarantee this estimator will exceed the screening estimator for the domain of interest. Another drawback is that it strongly depends on the logistic form of the on-the-CML model being correct. This can be problematic in the tail where the estimated value of p_{Lk} can be very small, and as a result, $1/\hat{p}_{Lk}$ is uncomfortably larges. By contrast, the method described in the previous section only uses the logistic model to form groups. Once groups are formed, the probability of a farm being on the list is effectively estimated (some would say, "approximated") with the group-specific: $\widehat{BOTH} / (\widehat{BOTH} + \widehat{CML})$. By forming groups, we effectively smooth the impact of small estimated p_{Lk} values.

Appendix A

A Continous Approach Based on Weighting the AFS
(a cutting-edge approach that may merit research)

One fits a sample-weighted logistic regression by solving this equation for **b**:

$$\sum w_{Lk}\left(y_k - \ell(\mathbf{x}_k\,'\mathbf{b})\right)\mathbf{x}_k = \mathbf{0}, \qquad (6)$$

where $\ell(z) = \exp(z)/[1+\exp(z)]$,

the summation is over all farms in the AFS (in a state or the country), and $y_k = 1$ when when k is in the CML, 0 otherwise. In principle, an iterative seach is needed, but there are canned programs to do that. Alternatively, a truncated version of the logistic function can replace $\ell(z)$ in equation (6), one that avoids overally small values for $\ell(\mathbf{x}_k\,'\mathbf{b})$.

Taking a different tack, an iterative search routine could be undertaken to find a vector **c** that satisfies the calibration equation:

$$\sum y_k w_{Ak} f(\mathbf{x}_k\,'\mathbf{c})\mathbf{x}_k = \hat{\mathbf{X}}_{CML}, \qquad (7)$$

where $f(\mathbf{x}_k\,'\mathbf{c}) = [\ell(\mathbf{x}_k\,'\mathbf{c})]^{-1} = 1+\exp(-\mathbf{x}_k\,'\mathbf{c})$, the summation is again over all farms in the AFS (at some level), and $\hat{\mathbf{X}}_{CML}$ is the estimated CML total of the components of $\mathbf{x}_k$.

If the search is successful, the probability that farm k is on the area frame given $\mathbf{x}_k$ is implicitly estimated to be

$$\hat{p}_{Ak} = \left[1+\exp(-\mathbf{x}_k\,'\mathbf{c})\right]^{-1} = [f(\mathbf{x}_k\,'\mathbf{c})]^{-1} = \ell(\mathbf{x}_k\,'\mathbf{c}) = (\exp(\mathbf{x}_k\,'\mathbf{c})/\left[1+\exp(\mathbf{x}_k\,'\mathbf{c})\right].$$

We are implicitly assuming that for a given $\mathbf{x}_k$ the probability a farm is on the CML does not depend on whether it is on the area frame and the probability a farm is on the area

frame does not depend on whether it is on the CML. Thus, the components of $\mathbf{x}_k$ in this section are the same as in the last two sections.

Given $\hat{p}_{Ak}$ values, one can then estimate *TOTAL* with

$$\widehat{TOTAL}_4 = \sum w_{Ak} / \hat{p}_{Ak} = \sum w_{Ak} \left[1 + \exp(-\mathbf{x}_k \text{'} \mathbf{c})\right], \qquad (8)$$

where the summation is over the AFS. If $\mathbf{x}_k$ contains a constant (or the equivalent), one can show that $\widehat{TOTAL}_4$ must exceed the screening estimator (because equation (7) assures us that the summation of the AFS sample in the list frame must equal the CML-estimate of that number, while the summation of the NML farms must exceed the NML count because exp(.) is always greater than 0).

There can be alternative choices for the back-link function *f*(.) in equation (7) than the inverse of the logistic.

Appendix A

Estimating Farm Counts in Years Before 2007 (and, perhaps, after)

In principal, the estimators in equation (2) could be used in any year with a multiple-frame survey. There are two problems with that, however. Problem 1 is that the data collected on the June *Area* Survey (JAS) in non-Census years is not as extensive as in other years. This means that the components of the **x** vector will have to change and with it the groups created after fitting the logistic model to the JAS sample.

Problem 2 is that the List Frame used in non-Census years is not as extensive as the 2007 CML. As a result, it may be that farms with certain characteristics have no chance of being on the List Frame (We will assume away the parallel possibility that farms with certain characteristics have no chance of being on the Area Frame covered by the JAS.)

To get around Problem 2, I propose the following for Year *A*:

1. Fit a logistic model to the 2007 JAS-only sample (excluding ACES and MACES segments) with on-the-CML (yes of no) as the dependent variables. Use as independent variables only items that are always on the June-Area questionnaire.

2. Break the CML-responding farms and JAS-only farms in 20 groups based on their sorted new $\hat{p}_{Lk}$ values.

3. Compute $\widehat{TOTAL_2}$ in equation (2) for each group, where $\widehat{BOTH}$ and $\widehat{NML}$ are now based on the JAS-only sample in 2007.

4. Compute $\widehat{JAS} = \widehat{BOTH} + \widehat{NML}$ for each group and then $\widehat{TOTAL_2} / \widehat{JAS}$. This is a group-specific estimate of 1/the-probability-of-being-on-the-area-frame.

5. For a JAS sampled farm *k* in Year *A*, compute $\hat{p}_{Lk}$ in equation (4) using the farm's Year-*A* **x**-vector and the **b** computed for 2007.

6. Using the group definitions for 2007 in Step 2 and the Year-*A* $\hat{p}_{Lk}$ in step 5, assign each farm in the Year-*A* JAS sample to a group.

7. Create a coverage-adjusted weights for every farm in the Year-*A* JAS sample by multiplying its JAS weight (including the tract-to-farm ratio) by its group's $\widehat{TOTAL_2} / \widehat{JAS}$ in Step 4

These coverage-adjusted weights can then be used to estimate any Year-*A* farm count (either directly or by using a chain of year-to-year ratios containing only overlap segments). One assumption is that the probability of being on the Area Frame in Year *A* is the same as in 2007. This is viable only if the area-sample screening procedures in Year *A* are the same as in the 2007 JAS-only sample.

Phil Kott; 12/10/08

A Capture-Recapture Indication for the 2007 Census of Agriculture and an Estimator for its Variance

Suppose both the full JAS area sample of farms and the Census-Mail-List (CML) respondents have already been separated into G mutually exclusive capture/recapture groups. These groups can be determined using logistic regression or data-mining techniques.*

Let

C_g denote the estimate of the number for list farms in group g derived from the CML respondents,

A_g be the set of area-frame sample farms in group g, and

L_g be the subsample of A_g containing only area-frame farms on the CML.

If w_k is the area-frame weight for area-sample farm k (including the tract-to-farm adjustment), and y_k is the item value of interest for k (y_k will usually be 1 for all farms, although there are other possibilities), then the capture-recapture indication for the y-total is

$$t_y = \sum_{g=1}^{G} \left(C_g \frac{\sum_{k \in A_g} w_k y_k}{\sum_{k \in L_g} w_k} \right).$$

* The idea is to determine the characteristics that make a farm more or less likely to be on the CML *from the Census responses* of JAS area-sample farms (including ACES and MACES segments). Both area-frame and CML respondent farms are then put into capture/recapture groups by placing together farms with similar probabilities of being on the CML based solely on their characteristics.

Note that C_g and $\sum_{k\in L_g} w_k$ are estimating the number of farms on the CML in group *g,* but the latter may be subject to an undercount. In computing t_y, the adjustment ratio for the area-frame undercount of the CML in *g*, $C_g \Big/ \sum_{k\in L_g} w_k$, is applied to the full area-frame estimate in *g*: $\sum_{k\in A_g} w_k y_k$.

An extended delete-a-group jackknife estimator for the area-frame contribution to the variance of t_y is

$$v_A(t_y) = \frac{R-1}{R}\sum_{r=1}^{R}(t_y - t_y^{(r)}),$$

where $t_y^{(r)} = \sum_{g=1}^{G}\left(C_g \frac{\sum_{k\in A_g} w_k^{(r)} y_k}{\sum_{k\in L_g} w_k^{(r)}} \right)$, and the $w_k^{(r)}$ are defined in the appendix.

Since the the only "sampling" in the CML is the result of Census nonresponse, the sampling fractions are too high to use a delete-a-group jackknife to compute the likely-to-be-very-small contribution of the CML to the variance of t_y. Instead, define

$$f_k = \frac{\sum_{k\in A_g} w_k y_k}{\sum_{k\in L_g} w_k}$$

for every CML respondent in capture/recapture group *g*. Consider the CML-weighted total of the *f* values among all CML respondents. This total is exactly equal to t_y, and the CML variance tor for this total is the CML contribution to the variance of t_y. As a

result, the estimated variance of t_y is $v_A(t_y)$ plus the estimated CML variance for the total of the *f* values.

Phil Kott; 12/30/08

Appendix: *Extended Delete-a-Group Jackknife Weights*

Sort the sampled *segments* in the JAS by substratum and assign each segment systematically to one of $R = 15$ replicate groups. A replicate group is the *complement* of a replicate. It has no relationship to a capture/recapture group.

Let h denote a substratum, r a replicate group, and S^{hr} the set of *farms* from segments in the area sample (S), substratum h and replicate group r. Let n_h be the number of *segments* from substratum h in the area sample.

For each area-sample *farm* k in substratum h, 15 jackknife replicate weights are computed as follows:

$$
\begin{aligned}
w_k^{(r)} &= w_k && \text{if } S^{hr} \text{ is empty} \\
&= w_k\ 1-(n_h-1)Z && \text{if } n_h > 1 \text{ and } k \in S^{hr} \\
&= w_k\ 1+Z && \text{if } n_h > 1,\ S^{hr} \text{ is not empty, and } k \notin S^{hr} \\
&= w_k\ 1-\sqrt{\tfrac{1}{2}\tfrac{R}{R-1}} && \text{if } n_h = 1 \text{ and } k \in S^{hr},
\end{aligned}
$$

where $Z = \min\left\{\sqrt{\frac{R}{R-1}\frac{1}{n_h(n_h-1)}},\ \frac{1}{n_h-1}\right\}$.

Observe that when $n_h \geq R = 15$ (so that $Z = 1/(n_h-1)$), these replicate weights are the same as conventional delete-a-group replicate weights. In particular, $w_k^{(r)} = 0$ when $k \in S^{hr}$, and $w_k^{(r)} = [n_h/(n_h-1)]w_k$ when $k \notin S^{hr}$.

Appendix B

2009 Farm Numbers Research Project (FNRP): Segment and Tract Counts by State						
STATE	**Total Segments Rotated-In**	**Segments to Contact for FNRP**	**Non-ags Tracts w/ Potential or Potential UK**	**Non-ags Tracts w/out Potential**	**Estimated Refusal & Inaccessible Tracts**	**Total Study Tracts**
AL(1)	64	62	33	281	37	351
AZ(4)	57	38	4	59	31	94
AR(5)	102	96	11	490	37	538
CA(6)	78	73	36	123	59	218
CO(8)	60	45	4	98	48	150
CT(9)	6	6	0	11	2	13
DE(10)	3	3	1	16	2	19
FL(12)	27	25	1	135	3	139
GA(13)	48	47	11	220	25	256
ID(16)	24	23	5	72	16	93
IL(17)	79	79	12	317	70	399
IN(18)	52	51	11	301	46	358
IA(19)	96	86	8	234	56	298
KS(20)	79	63	5	45	109	159
KY(21)	37	33	13	133	17	163
LA(22)	60	56	14	293	31	338
ME(23)	24	24	0	57	6	63
MD(24)	11	11	7	48	8	63
MA(25)	23	23	0	58	10	68
MI(26)	56	56	2	104	55	161
MN(27)	78	70	16	187	45	248
MS(28)	58	53	8	202	11	221
MO(29)	75	74	21	250	97	368
MT(30)	90	58	5	118	57	180
NE(31)	74	63	7	99	81	187
NV(32)	10	4	1	4	6	11
NH(33)	5	5	0	10	3	13
NJ(34)	8	8	3	124	2	129
NM(35)	56	35	46	89	44	179
NY(36)	15	15	6	63	11	80
NC(37)	63	62	13	86	28	127
ND(38)	82	82	2	133	58	193
OH(39)	64	62	1	119	32	152
OK(40)	66	53	60	322	63	445
OR(41)	33	21	4	71	8	83

Appendix B

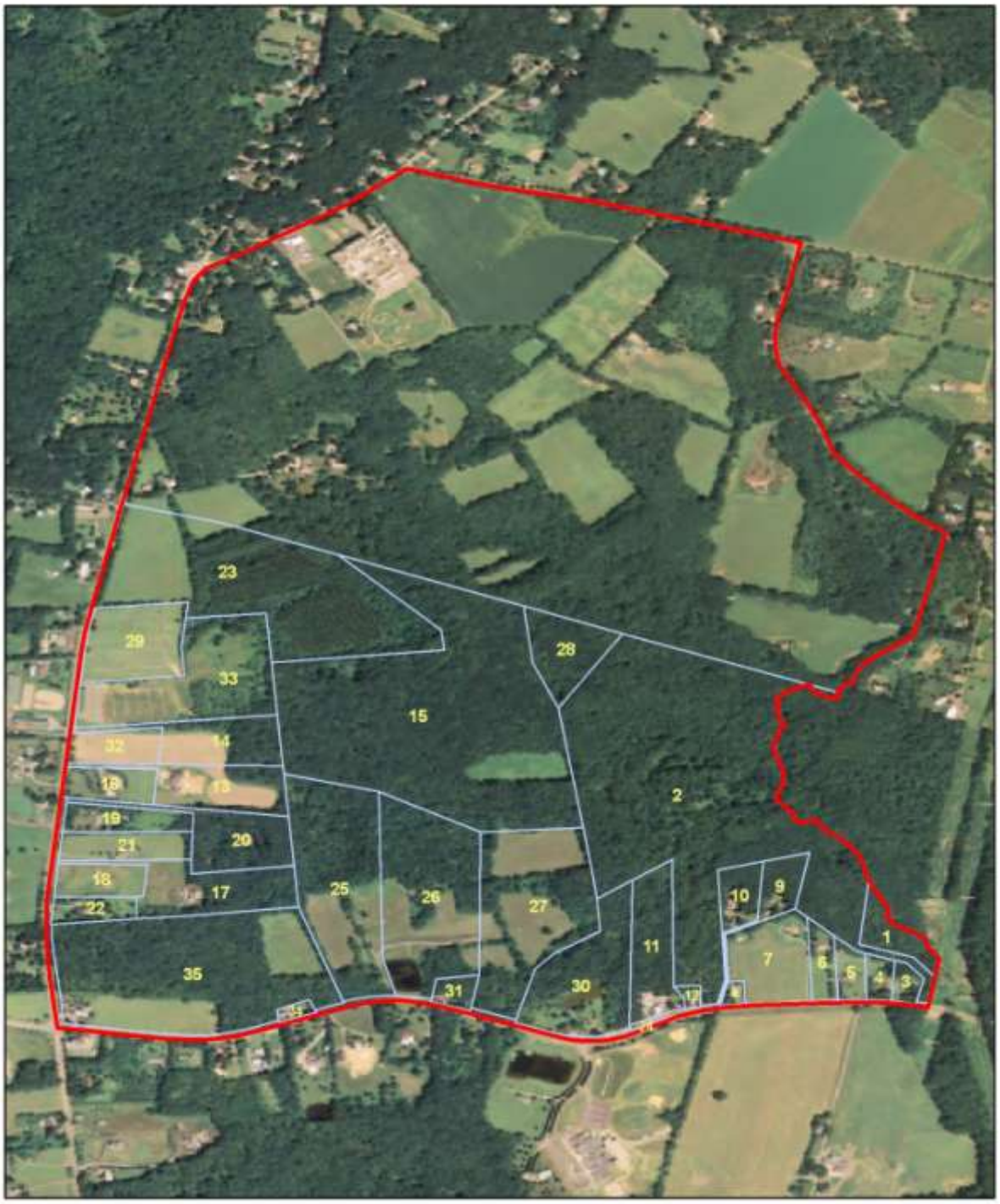
2009 Farm Numbers Research Project (FNRP): Segment and Tract Counts by State						
STATE	**Total Segments Rotated-In**	**Segments to Contact for FNRP**	**Non-ags Tracts w/ Potential or Potential UK**	**Non-ags Tracts w/out Potential**	**Estimated Refusal & Inaccessible Tracts**	**Total Study Tracts**
PA(42)	53	53	14	308	18	340
RI(44)	4	4	0	6	3	9
SC(45)	61	60	5	163	34	202
SD(46)	78	59	5	79	64	148
TN(47)	66	64	34	418	29	481
TX(48)	202	183	142	650	126	918
UT(49)	42	35	13	56	39	108
VT(50)	16	16	0	46	7	53
VA(51)	35	35	8	44	11	63
WA(53)	52	51	15	315	51	381
WV(54)	25	25	5	395	12	412
WI(55)	43	43	20	308	28	356
WY(56)	25	16	0	160	16	176
Totals	2,465	2,209	632	7,920	1,652	10,204

Appendix C

Segment 190014

County 19

New Jersey
Map Series



Appendix C

Segment 190216

County 19

Nebraska
Map Series

Appendix D

Project 132 QID OMB No. 0535-0213 Approval Expires 12/31/2010

[redacted]

LAND UTILIZATION SURVEY

NATIONAL AGRICULTURAL STATISTICS SERVICE

National Field Office
U.S. Department of Agriculture,
Rm 5030, South Building
1400 Independence Ave., S.W.
Washington, DC 20250-2000
Phone: 1-800-727-9540
Fax: 202-690-2090
Email: nass@nass.usda.gov

Please make corrections to name, address and Zip Code, if necessary.

PLEASE RETURN THIS FORM REGARDLESS OF YOUR AGRICULTURAL ACTIVITY

You have been selected for a nationwide land utilization survey. The purpose of the survey is to determine if you are involved in agriculture in any way. We need your completed form even though you may not be involved with agriculture, or have only a limited amount of agriculture. All information collected is confidential and used only in combination with similar reports from others. Response is voluntary. Thank you!

If you have any questions, call our toll-free number at 1-800-833-0867, and we will be happy to assist you.

Since January 2009…

Did you own, rent or operate land with potential for agricultural production?	1☐ Yes	2☐ No	0001
Did you own or operate cropland?	1☐ Yes	2☐ No	0002
Did you receive any federal or state agricultural payments (Include Federal Farm Program, CRP, WRP, FWP, and CREP payments)	1☐ Yes	2☐ No	0003
Did you have any hay or pasture?	1☐ Yes	2☐ No	0004
Did you produce or raise any type of livestock?	1☐ Yes	2☐ No	0005
Did you have any equine?	1☐ Yes	2☐ No	0006
Did you grow any type of poultry or birds?	1☐ Yes	2☐ No	0007
Did you own any colonies of bees regardless of location?	1☐ Yes	2☐ No	0008
Did you have any aquaculture?	1☐ Yes	2☐ No	0009
Did you have any agricultural sales?	1☐ Yes	2☐ No	0010

(**Exclude** cash rent received or share of crops received for rented out land.)

(**Include** value from Field Crops, Hay, Silage and Forage Crops, Christmas Trees, Short Rotation Woody Crops, Seed Crops, Nursery, Greenhouse, Floriculture, Sod, Mushrooms, Seeds, Bulbs, Vegetables, Melons, Fruits, Nuts, Berries, Other Crops, Maple Syrup, Hogs and Pigs, Cattle and Calves, Sheep and Lambs, Goats, Poultry, Horses, Bees and Honey, Eggs, Dairy Products, Other Animals, Livestock and Animal Products, Fish and Other Aquaculture.)

If "NO" to all of these questions, please return the questionnaire in the envelope provided. You do not need to answer the rest of the questions. You may comment on the last page.

If "YES" to any question, please complete the rest of the survey as it applies to your operation.

Acres Operated

Include the farmstead, all cropland, ponds, woodland, pastureland, wasteland, and land that is in government programs.

		MARK NONE		ACRES
1.	As of June 1, 2009, how many acres does this operation:			
a.	**Own**?	☐	+	UU11
b.	**Rent** or **Lease from Others**, or use **Rent Free**?	☐	+	UU12
c.	**Rent to Others**?	☐	-	UU13
2.	**Total Acres Operated**: [Items 1a + 1b – 1c]?	☐	=	UU14

Cropland

		MARK NONE	ACRES
3.	Of the **total acres operated** [item 2], how many acres are considered **Cropland**? (Include planted cropland, land in hay, summer fallow, cropland idle, cropland used for pasture and cropland in government programs.)	☐	UU15
4.	Of the total cropland acres [item 3], how many are used to produce:		ACRES
a.	Any hay or forage crops on this operation? (Count each acre only once, regardless of the number of cuttings or different uses.)	☐	UU16

		MARK NONE	ACRES	TENTHS
b.	Any short rotation woody crops on this operation?	☐	UU17	
c.	Any cut Christmas trees on this operation?	☐	UU18	
d.	Any berry crops on this operation?	☐	UU19	
e.	Any fruit or nut crops on this operation? (Excluding berries)	☐	UU20	.
f.	Any potato crops on this operation?	☐	UU21	.
g.	Any vegetable or melon crops on this operation? (Excluding potatoes)	☐	UU22	.

		MARK NONE	SQUARE FEET UNDER GLASS or OTHER PROTECTION	ACRES IN THE OPEN	
				ACRES	TENTHS
h.	Any nursery, greenhouse, or floriculture crops, or aquatic plants on this operation?	300	UU23	UU24	.

		MARK NONE	TOTAL ACRES
5.	Of the **total acres operated** [Item 2], how many acres are in the **Conservation Reserve Program (CRP), Wetland Reserve Program (WRP), Farmable Wetland Program (FWP), or Conservation Reserve Enhancement Program (CREP)**?	☐	UU25
			ACRES
6.	Of the **total acres operated** [Item 2], how many acres are in **Permanent Pasture, Woodland Pasture, and Rangeland**?	☐	UU26

Livestock

7. As of June 1, 2009, how many of the following **livestock species were on hand?** (Include animals of any age, including newborns.)

		MARK NONE	NUMBER OF HEAD
a.	**All Cattle and Calves**, regardless of ownership?	☐	0027
b.	**All Goats and Kids**, regardless of ownership?	☐	0028
c.	**All Horses, Ponies, Mules, Donkeys or Burros**, regardless of ownership?	☐	0029
d.	**All Sheep and Lambs owned** by this operation, regardless of location?	☐	0030
e.	**All Hogs and Pigs owned** by this operation, regardless of location?	☐	0031
f.	**All Other Animals** (Include Alpacas, Bison, Deer, Rabbits, Mink, etc.)?	☐	0032

Poultry

8. Report any poultry on the total acres operated as of June 1, 2009, regardless of ownership.

		MARK NONE	NUMBER
a.	What was the number of layers, pullets, and roosters for laying flock?	☐	0033
b.	What was the number of broilers, fryers, capons, roasters, and other chickens raised for meat production? (Exclude chickens reported in 8a.)	☐	0034

9. As of June 1, 2009, how many turkeys were on the total acres operated, regardless of ownership? ☐ 0035

Bees

10. How many colonies of bees did this operation **own**, regardless of location, on June 1, 2009? ☐ MARK NONE — NUMBER OWNED 0036 — HONEY COLLECTED In 2009 (Pounds) 0037

Sales of Agricultural Products

11. What was the **total gross value of sales of agricultural products in 2008**, including landlord's share and value of product, for all crops, livestock, and poultry?

(**Exclude** cash rent received or share of crops received for rented out land.)

(**Include** value from Field Crops, Hay, Silage and Forage Crops, Christmas Trees, Short Rotation Woody Crops, Seed Crops, Nursery, Greenhouse, Floriculture, Sod, Mushrooms, Seeds, Bulbs, Vegetables, Melons, Fruits, Nuts, Berries, Other Crops, Maple Syrup, Hogs and Pigs, Cattle and Calves, Sheep and Lambs, Goats, Poultry, Horses, Bees and Honey, Eggs, Dairy Products, Other Animals, Livestock and Animal Products, Fish and Other Aquaculture.)

None		99	☐
$	1 --- $ 999	1	☐
$	1,000 --- $ 2,499	2	☐
$	2,500 --- $ 4,999	3	☐
$	5,000 --- $ 9,999	4	☐
$	10,000 --- $ 24,999	5	☐
$	25,000 --- $ 49,999	6	☐
$	50,000 --- $ 99,999	7	☐
$	100,000 --- $ 249,999	8	☐
$	250,000 --- $ 499,999	9	☐
$	500,000 --- $ 999,999	10	☐
$	1,000,000--- $2,499,999	11	☐
$	2,500,000--- $4,999,999	12	☐
$	5,000,000 and over	13	☐

→ Code 0038

Final Comments:

12. If you have any additional comments, please write them on the lines below.

_____

_____

_____

Thank you for your cooperation.

A wide variety of agricultural statistics is available from the National Agricultural Statistics Service (NASS). NASS reports, data products, and services are also available on the Internet at www.nass.usda.gov

Completed by: _____ Area Code and Phone Number: _____ 9910 Date: _____

For Office Use Only								
Response		Respondent		Mode		R Unit	Enum.	Eval.
1-Comp 2-R 3-Inac 4-Office Hold 5-R – Est 6-Inac – Est 7-Off Hold – Est 8-Known Zero	9901	1-Op/Mgr 2-Sp 3-Acct/Bkpr 4-Partner 9-Oth	9902	1-Mail 2-Tel 3-Face-to-Face 4-CATI 5-Web 6-e-mail 7-Fax 8-CAPI 19-Other	9903	0921	098	100

State	Segment	Tract	Subtract	Number of Dwellings	Number Sampled	Tract Acres
039	040	041	044			

Appendix E

State Name	Total Expenses	Total Tracts	Tracts from estimated	Tracts from non-ags	Total Farms	Farms from non-ags	Farms from estimated tracts	Total Expenses per Tract	Total Expenses per Farms
US Level	$412,104	17,191	1,712	15,479	2,559	1,056	1,503	$23.97	$161.04
Alabama	$27,942	642	65	577	93	58	35	$43.52	$300.45
Arizona	$1,246	160	31	129	22	1	21	$7.79	$56.64
Arkansas	$19,331	716	37	679	97	60	37	$27.00	$199.29
California	$9,076	393	59	334	81	30	51	$23.09	$112.05
Colorado	$3,221	152	48	104	45	1	44	$21.19	$71.58
Florida	$7,452	190	3	187	9	6	3	$39.22	$828.00
Georgia	$18,571	353	25	328	55	34	21	$52.61	$337.65
Idaho	$5,125	128	16	112	23	8	15	$40.04	$222.83
Illinois	$11,864	522	71	451	102	33	69	$22.73	$116.31
Indiana	$6,633	583	46	537	80	35	45	$11.38	$82.91
Iowa	$7,025	403	56	347	98	42	56	$17.43	$71.68
Kansas	$4,104	195	110	85	113	6	107	$21.05	$36.32
Kentucky	$10,749	284	20	264	71	52	19	$37.85	$151.39
Louisiana	$20,127	576	35	541	90	60	30	$34.94	$223.63
Maryland	$3,312	204	10	194	17	10	7	$16.24	$194.82
Michigan	$8,236	557	61	496	66	20	46	$14.79	$124.79
Minnesota	$8,369	333	45	288	77	34	43	$25.13	$108.69
Mississippi	$19,594	461	11	450	50	40	10	$42.50	$391.88
Missouri	$11,720	450	100	350	122	29	93	$26.04	$96.07
Montana	$2,469	218	57	161	72	18	54	$11.33	$34.29
Nebraska	$2,240	224	81	143	92	11	81	$10.00	$24.35
Nevada	$111	57	6	51	6	3	3	$1.95	$18.50
New Hampshire	$15,375	744	31	713	38	24	14	$20.67	$404.61
New Jersey	$5,567	256	2	254	4	2	2	$21.75	$1,391.75
New Mexico	$16,626	231	56	175	53	20	33	$71.97	$313.70

Appendix E

New York	$3,334	175	11	164	27	17	10	$19.05	$123.48
North Carolina	$5,100	605	28	577	69	44	25	$8.43	$73.91
North Dakota	$1,488	197	58	139	64	6	58	$7.55	$23.25
Ohio	$11,401	398	32	366	52	21	31	$28.65	$219.25
Oklahoma	$7,188	565	63	502	115	53	62	$12.72	$62.50
Oregon	$2,016	111	8	103	9	2	7	$18.16	$224.00
Pennsylvania	$22,393	880	18	862	41	24	17	$25.45	$546.17
South Carolina	$11,181	408	34	374	53	24	29	$27.40	$210.96
South Dakota	$366	157	64	93	71	10	61	$2.33	$5.15
Tennessee	$20,520	907	29	878	43	15	28	$22.62	$477.21
Texas	$47,237	1,547	126	1421	248	130	118	$30.53	$190.47
Utah	$2,790	152	39	113	44	10	34	$18.36	$63.41
Virginia	$4,246	232	11	221	15	9	6	$18.30	$283.07
Washington	$3,460	439	51	388	62	26	36	$7.88	$55.81
West Virginia	$10,611	698	12	686	26	18	8	$15.20	$408.12
Wisconsin	$10,800	512	30	482	28	5	23	$21.09	$385.71
Wyoming	$1,888	176	16	160	16	5	11	$10.73	$118.00