United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Development Division
Washington DC 20250

RDD Research Report
RDD-07-01

March 2007

Supplemental Samples for the 2007 Area Frame: A Design for Estimating Numbers of NML Farms for the 2007 Census of Agriculture

Raj S. Chhikara
Floyd M. Spears
Charles R. Perry
Phillip S. Kott

SUPPLEMENTAL SAMPLES FOR THE 2007 AREA FRAME: A DESIGN FOR ESTIMATING NUMBERS OF NML FARMS FOR THE 2007 CENSUS OF AGRICULTURE by Raj S. Chhikara[1], Floyd M. Spears[2], Charles R. Perry, and Phillip S. Kott. Research and Development Division, National Agricultural Statistics Service, United States Department of Agriculture, Washington, DC 20250-2000, March, 2007, NASS Research Report No. RDD-07-01.

ABSTRACT

An area frame sample allocation is developed that will meet the survey cost and precision requirements of NASS in its estimation of the number of farms not on the mailing list (NML) for the 2007 Census of Agriculture. Direct and model-predicted estimates of the stratum standard deviations are obtained for 18 agricultural items (the 8 items used by NASS for determining the area frame sample and 10 NML items) using the 2002 area frame sample data, including the supplemental Agricultural Coverage Evaluation Survey (ACES) segments. Standard deviation estimates of for these 18 agricultural items are considered as input in the current NASS sample allocation procedure. Three different cases that utilize these NML items are considered for determining sample allocations. In each case, the previously available 2007 sample allocation is combined with that obtained for the NML items to determine a combined area frame sample allocation. The results from each case are compared to determine the ones that meet the NASS main objectives for the NML farm survey in 2007. An area frame sample design that includes a supplemental sample of 3091 segments is recommended.

KEY WORDS

NML Farm Items; Coefficient of Variation (CV); Stratum Standard Deviation; Model Fits; Multivariate Allocation; Area Frame Sample Allocation; Stratum Groups

ACKNOWLEDGEMENTS

The authors would like to thank George Hanuschak and Roberta Pense for their keen interest and support for the project. The authors would also like to thank Bill Wigton and Herb Eldridge for their contribution and providing the SAS data file for this project.

[1] Raj S. Chhikara is Professor, Division of Computing and Mathematics, University of Houston-Clear Lake, Houston, TX.

[2] Floyd M. Spears is Assistant Vice-President for Academic Affairs, Harding University, Searcy, AR.

Contents

List of Figures

List of Tables

GLOSSARY

ACES - **A**gricultural **C**overage **E**valuation **S**urvey segments

ASB Number - **A**gricultural **S**tatistical **B**oard Final Estimates (also known as Official Statistics)

CV - the ratio of the standard error of an estimate to the ASB number for an item; multiply the CV by 100 to convert to percentage

d-value - directly computed estimate of the stratum standard deviation from the 2002 area frame survey

JAS - **J**une **A**rea **S**urvey

Multivariate Allocation - a nonlinear optimization convex programming procedure used by NASS in determination of its sample allocation

m-value - estimated stratum standard deviation using the predicted value from model-fits as described in this report

NML Farm - a farm that is **n**ot on the census **m**ailing **l**ist

NML Subdomains - sub-groups of NML farms that are minority-operated or specialty farms

p-value - estimated stratum standard deviation using the predicted value from the empirical model for NML farms by Chang and Kott (2004)

R-square - measures the proportion of variability in the response data that is accounted for by the model

Sample Allocation - total number of sample segments at U.S. level and their allocation across land-use strata

Stratum - land-use stratum in NASS area frame

Stratum Group - group of land-use strata based on NASS categories of agriculture intensity or location

EXECUTIVE SUMMARY

One challenge in generating a reliable 2007 Census of Agriculture estimate is in determining that portion of the estimate represented by those farms not found on the Census mailing list (NML). To estimate this portion, NASS will supplement the regular 2007 area segment sample allocation with additional area segments. Estimates will be made for the total number of NML farms and its sub-domains of Asian-operated, Black-operated, Hispanic-operated, Native American-operated, female-operated, and vegetable, fruit, nursery and Christmas tree farms. A target of approximately 3000 additional segments will be used for this supplemental sample. This will meet the Agency target CV of 0.5 percent at the national level for the estimate of total number of NML farms and the Agency target CV of 5.0 percent for each NML sub-domain, where the target CV is defined to be one hundred times the ratio of standard error to the Agriculture Statistics Board (ASB) number for an item.

NASS uses a multivariate allocation procedure which requires stratum standard deviations for the agriculture items of interest as input to determine its sample allocation to the land use strata. The annual area frame survey data collected in the 2002 Census year included an additional 2429 Agricultural Coverage Evaluation Survey (ACES) segments. These survey data were used in determining the stratum standard deviations for the NML farm items. Three different stratum variances for each NML item were employed in this study: (1) directly computed from the 2002 survey data, (2) Chang and Kott predicted values based on logistic regression modeling of each NML item, and (3) determined from the model-fit of directly computed versus Chang and Kott estimated stratum variance. This third option was considered in an attempt to smooth the anomalies in stratum standard deviation estimates that could skew the allocation for some strata. This is to safeguard against the use of an unreliable stratum variance, particularly as a result of its estimation based on small sample size.

The NASS allocation procedure was applied using each of the ten NML items separately, and also in a multivariate approach using all ten NML items together. Another multivariate approach was to consider all eighteen agricultural items, the eight regular items and the ten NML items, together. In each case, the allocations were compared for all land strata with the regular 2007 design sample, and the maximum was taken to be the stratum sample size. The univariate NML allocations varied substantially for strata across various NML items, whereas the allocation from the multivariate procedure with 10 NML items smoothed these out and thus seemed more appropriate.

The optimum approach that met the target of adding approximately 3000 while having robust estimates of stratum standard deviations utilized the model-predicted stratum standard deviations in a single multivariate allocation for the 10 NML items. The resulting allocation was compared across land strata with the regular 2007 design sample, and the maximum was taken to be the stratum sample size.

Five stratum groups are considered on the basis of land use as described in Appendix B. These correspond to the land use strata numbered in 10's, 20's, 30's, 40's and 50's. Stratum

groups 3 and 4 correspond to urban and low cultivation areas and are expected to have relatively more NML farms than stratum groups 1 and 2 which correspond to high and moderate cultivation areas, respectively. Stratum group 5 corresponds to non-agricultural areas and is not expected to hae NML farms. The proposed sample allocation would require a total of 3091 supplemental samples, of which 287 are in stratum group 1; 699 are in group 2; 381 are in group 3, 1724 are in group 4, and 0 in group 5. There is a more than 100 percent increase in stratum group 3, and nearly a 100 percent increase in stratum group 4 when compared to the regular 2007 sample design. The increases in sample size are much smaller in stratum groups 1 and 2.

At the state level, there is a substantial increase in sample allocations in California (additional 515 segments; an increase of 127 percent), Florida (additional 119 segments; an increase of 119 percent), Michigan (additional 107 segments; an increase of 74 percent), New Mexico (additional 105 segments; an increase of 85 percent), Oregon (additional 232 segments; an increase of 120 percent), and Texas (additional 625 segments; an increase of 56 percent). There are several strata in these states that have substantial supplemental samples when compared to the regular 2007 design allocation.

The reason for the increase is determined by identifying the NML item for which the univariate allocation in a stratum was substantially higher compared to the univariate allocation for any other NML item. The substantial increases for these states are due to one or more NML items being predominant there. The primary NML item(s) for which reliable estimation would require such an increase in sample allocation is listed in Tables 5 and 6.

Total NML farms is the primary cause of most increases, especially in stratum group 4. A major exception is California, where Asian-operated farms are the cause of large increases for several strata.

It should be noted that a slightly different allocation is being used operationally for the 2007 ACES survey. The allocation takes advantage of targeted sampling in some states/strata to reduce the number of samples needed, and yet achieve the target CV for Black-operated and Asian-operated NML farms. It also takes into account the additional goal of a state-level NML target CV of 6.5 percent for states or combined states in the case of small states.

1 INTRODUCTION

The NASS list of farms will be used to mail for the 2007 Census of Agriculture. However, since every list has incomplete coverage, a reliable estimate needs to be generated for those farms not on the mailing list (NML). To accomplish this, supplemental area frame samples will be selected for the 2007 June Area Survey (JAS). From that expanded area frame, estimates will be generated for the NML farms and its sub-domains for various minority operated and specialty-type farms. These sub-domains would include the following groups: Asian-operated, Black-operated, Hispanic-operated, Native American-operated, female-operated, and vegetable, fruit, nursery and Christmas tree farms. The NML farms are usually smaller in size, operated by minorities or have specialty-type farming.

For the annual agriculture survey, NASS selects a stratified area frame sample specifically for that year based upon a multivariate allocation procedure. To do this, the stratum standard deviations must be known for eight primary agriculture items: corn, soybean, cotton, winter wheat, spring wheat, durum wheat, number of farms and number of cattle. To extend this allocation technique to include the NML farm items for Census, the stratum standard deviations for each of the NML items would be required. Hence, these are estimated for each NML item as considered and described in Section 2.

Based on its experience from the 2002 Census of Agriculture (see the technical memorandum of February 11, 2005 by NML/ACES Sample Design Team), NASS has decided to meet the following goals in estimation of NML farm items from its 2007 area frame samples. The main objective is to achieve a target CV of 0.5 percent at the national level for the estimate of total number of farms, and a target CV of 5.0 percent for each NML farm sub-domain. The target CV is defined to be 100 times the ratio of standard error to the Agriculture Statistics Board (ASB) number for an item which consists of its NML estimate plus the value deteremined from the Census. Also, it is desired to have a target CV of 6.5 percent at the state level. The exception would be for small states where a target CV of 15 percent is acceptable. Small states that are contiguous such as those in New England can be combined to achieve the 6.5 percent target CV.

Survey costs and timely acquisition of data are major concerns for the agency. NASS has set a target of approximately 3000 supplemental area segments in 2007 to achieve reliable estimates for the various NML farm items. The focus of this research was to develop a supplemental sample for the 2007 area frame design that would support these NASS goals and objectives.

This report describes the approach taken to estimate the stratum standard deviations for the NML farm items and to determine a robust and hence, a more stable sample allocation. Section 3 describes different options considered for the sample allocation. The various sample allocations obtained under these options are evaluated to determine the one that is robust and most economical to implement. The actual implementation, of course, would require selection of supplemental samples, which is beyond the scope of this research.

2 ESTIMATION OF STANDARD DEVIATION

In 2002, NASS collected data on 2,429 additional area segments, besides its 11,075 June Area Survey (JAS) segments. These supplemental samples were referred to as the Agricultural Coverage Evaluation Survey or ACES. The JAS sample was primarily designed to estimate total farm counts, major crop acreages, and livestock inventories. The ACES supplemental sample was designed to improve the overall coverage of the NML farms with a majority of these sample segments being allocated in low agricultural intensity strata. Thirty-three percent of the ACES segments reported NML farms as compared to 27 percent of the JAS segments. The ACES and JAS segments combined provided a much more useful data set for the NML area survey, and hence are used here for estimation of stratum standard deviations for the various NML farm items that are considered in this study.

One set of stratum variances consisted of directly computed survey values, called the d-values. Another set of stratum variances was made of those computed from the predicted values, called the p-values. These values were obtained from the empirically developed models for NML farms by Chang and Kott (2004). They utilized the combined 2002 area frame survey data to model the probability of an area frame farm not being on the census mailing list and used a host of covariates, including farm operator's gender, ethnicity and farm type, as outlined briefly in Section 1 of Appendix A. A logistic regression methodology was the basis for developing model-fits for the empirical NML models. For full details, see Chang and Kott (2004).

When the two estimates for stratum standard deviation are compared, d-values versus p-values, a strong linear relationship was exhibited for each NML item. For example, depicted in Figure 1 is a plot of the paired values for stratum standard deviation of the total number of NML farms across all land use strata. The d-values are on the average varying more than the p-values. This is an expected outcome since an NML model-fit is likely to predict values that have less variability than those directly computed from the survey data.

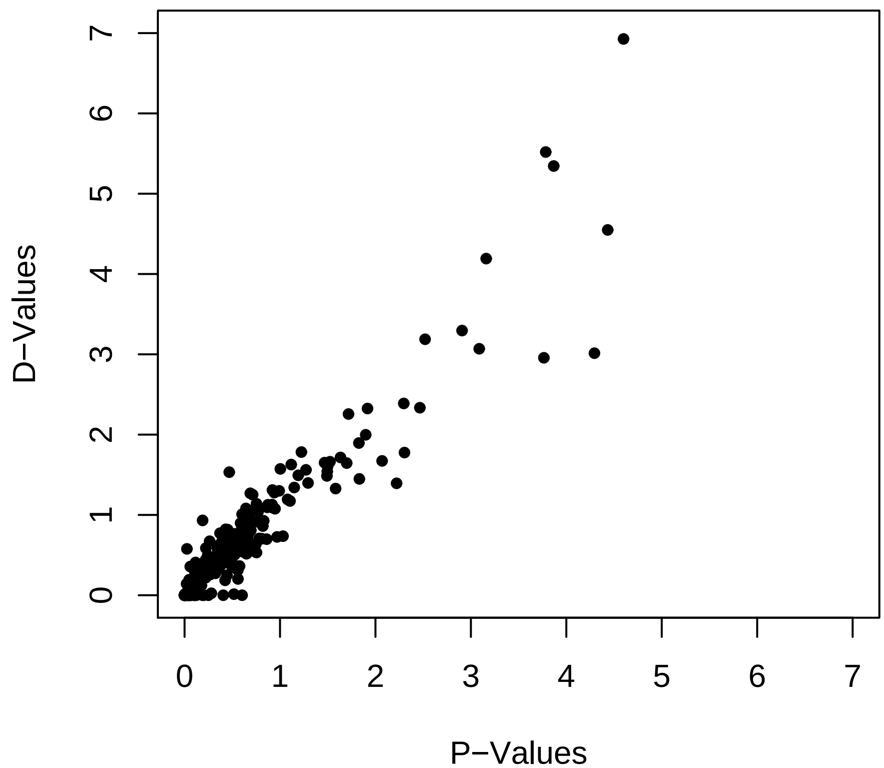
When estimated based on a small sample size, the direct survey computed stratum standard deviation is not reliable and needs to be improved upon. So we considered modeling the direct survey computed values using the NML model predicted value as a covariate. Since the covariate values are more stable and less varying than the response values, this modeling of stratum standard deviation should smooth any anomalies that exist in the directly computed survey estimates. The model-fits should result in more reliable predicted values for the stratum standard deviation. This in turn should safeguard against the use of unreliable stratum standard deviations in the determination of sample allocation to strata.

Appendix A describes the modeling of direct survey computed stratum standard deviations as a function of that obtained from the NML model predicted values.

Table 1 lists the model-fit characteristics, slope and R-square for each of the NML items. It indicates a strong linear relationship with a correlation coefficient close to 1.

The model-fits are used to predict the stratum standard deviations for the different NML items. These model-predicted stratum standard deviations, called the m-values, then provided a third set of stratum standard deviations for each NML item. Thus three different stratum standard deviations were employed in this investigative study of sample allocation. These three sets of stratum standard deviation estimates are referred to as follows:

Figure 1: Standard deviations of d-values versus p-values for the total number of NML farms across all strata in the U.S.



a. **d-values**, which are directly computed survey stratum standard deviations

b. **p-values**, which are estimated from the empirical NML models developed by Chang & Kott

c. **m-values**, which are estimated from the model-fits of d-values and p-values as described above.

3 SAMPLE ALLOCATIONS

The multivariate sample allocation for a NASS area frame design is determined using the stratum standard deviations computed from prior year survey data for the regular eight agriculture items. This multivariate allocation procedure can be carried out for any one or more agriculture items of interest by replacing one or more of the regular eight items or by simply adding more items to it. In this study, the allocation procedure considered each NML farm item individually, as univariate case, as well as all 10 NML farm items together as a single multivariate case. Also considered is a single multivariate allocation for the eighteen agriculture items of which eight are the regular items and ten are the NML items. Since the 2007 area frame design is for collection of survey data that would support the annual JAS needs as well

Table 1: Model Fits (No Intercept)

NML Item	Slope	R-Square
Total	1.12	0.931
Asian	1.84	0.934
Black	1.44	0.983
Christmas Tree	1.47	0.973
Female	1.55	0.971
Fruit	1.74	0.969
Hispanic	1.57	0.939
Native American	1.54	0.919
Nursery	1.71	0.892
Vegetable	1.66	0.918

as reliable estimation of NML farm items, the results of an allocation derived for the NML items are then combined with the sample allocations obtained for the regular eight items in order to determine a supplementary sample.

The following three ways of sample allocation were considered for the NML farm items to develop the final sample allocation.Table 2 lists the sample allocation by stratum groups. There are five stratum groups, which correspond to the land use strata numbered in 10's, 20's, 30's, 40's and 50's, as defined in Appendix B.

Case A: Univariate allocations for NML items

1. Perform univariate NML allocations for the 10 NML items.
2. Compare the univariate allocations and take the maximum of these 10 allocations stratum by stratum to obtain the combined univariate NML allocation. This corresponds to the column labeled A_1 in Table 4.

Case B: A single multivariate allocation with 10 NML items. This corresponds to the column labeled B_1 in Table 4.

Case C: A single multivariate allocation with the regular 8 agriculture items and the 10 NML items together. This corresponds to the column labeled C_1 in Table 4.

The final allocation was derived in each case by taking the maximum stratum by stratum of the regular 2007 NASS design allocation and the NML allocation. The resulting allocations for the three cases A, B and C are given in columns labeled A_2, B_2 and C_2, respectively, of Table 4.

In each of these cases, three sets of stratum standard deviations (d-values, p-values and m-values, as discussed in Section 2) were used to perform each of the three allocations described above for the NML items. Table 2 lists the sample allocations obtained when summed up at the stratum group level for each of the cases A - C using (a) d-values, (b) p-values, and (c) m-values for the NML farm item stratum standard deviations.

Each of these allocations meets the goal of an achievable target CV of 0.5 percent at the national level for the total number of NML farms and 5 percent or less at the national level for each of the NML sub-domains with some exceptions. The Black-operated farms and Christmas Tree farms have an achievable CV slightly above 5 percent, and the Asian-operated farms have the achievable CV close to 6 percent. Table 3 depicts the achievable target CVs for all cases listed in Table 2.

The NML sample allocations given in columns A_1, B_1 and C_1 of Table 2 show that at the national level Case C has the largest sample size, and that Case B has the smallest sample size. The NML sample allocations

Table 2: U.S. sample allocations by stratum group for cases A, B and C using (a) d-values, (b) p-values and (c) m-values to estimate standard deviations for NML farm items.

	Case A		Case B		Case C	
Stratum Group	Combined Univariate for 10 NML's (A_1)	Max of NASS Design & A_1 (A_2)	Multi-variate with 10 NML's (B_1)	Max of NASS Design & B_1 (B_2)	Multi-variate with 18 Items (C_1)	Max of NASS Design & C_1 (C_2)
(a) Using d-values						
10's	3653	6545	3057	6339	6038	6630
20's	3924	4294	3336	3872	3689	3883
30's	767	844	616	700	554	634
40's	4214	4431	3619	3922	3176	3435
50's	96	104	96	104	96	104
Total	12,654	16,218	10,724	14,937	13,553	14,686
(b) Using p-values						
10's	1886	6067	1808	6041	6026	6575
20's	2002	3010	1885	2987	2644	3076
30's	440	521	391	472	355	436
40's	2292	2701	2229	2664	1826	2261
50's	96	104	96	104	96	104
Total	6716	12,403	6409	12,268	10,947	12,452
(c) Using m-values						
10's	2751	6314	2303	6192	6011	6559
20's	3227	3815	2728	3474	3305	3580
30's	767	834	651	728	608	687
40's	3601	3879	3122	3491	2752	3036
50's	96	104	96	104	96	104
Total	10,442	14,946	8900	13,989	12,772	13,966

Table 3: U.S. achievable target CV for NML items using (a) d-values, (b) p-values and (c) m-values to estimate standard deviations for NML farm items.

	Case A		Case B		Case C	
Stratum Group	Combined Univariate for 10 NML's (A_1)	Max of NASS Design & A_1 (A_2)	Multi-variate with 10 NML's (B_1)	Max of NASS Design & B_1 (B_2)	Multi-variate with 18 Items (C_1)	Max of NASS Design & C_1 (C_2)
a) Using d-values						
Asian	5.63%	5.57%	5.81%	5.77%	5.79%	5.78%
Black	4.97%	4.83%	5.41%	5.19%	5.40%	5.35%
Ch. Tree	4.90%	4.88%	5.28%	5.24%	5.30%	5.29%
Female	1.73%	1.61%	1.86%	1.68%	1.83%	1.78%
Fruit	2.73%	2.64%	2.91%	2.78%	3.12%	3.08%
Hispanic	2.76%	2.57%	3.01%	2.73%	3.02%	2.96%
Nat. Amer.	4.13%	3.82%	4.96%	4.27%	4.96%	4.81%
Total	0.45%	0.41%	0.48%	0.44%	0.48%	0.47%
Nursery	3.77%	3.62%	4.63%	4.42%	4.85%	4.58%
Vegetable	3.29%	3.18%	3.61%	3.38%	3.68%	3.61%
b) Using p-values						
Asian	5.68%	5.41%	6.55%	6.25%	6.43%	6.15%
Black	4.76%	4.25%	5.51%	4.38%	5.06%	4.51%
Ch. Tree	4.51%	4.14%	5.11%	4.55%	5.17%	4.68%
Female	1.43%	1.22%	1.46%	1.23%	1.42%	1.31%
Fruit	1.91%	1.83%	1.92%	1.84%	2.19%	2.13%
Hispanic	2.21%	1.90%	2.27%	1.95%	2.26%	2.16%
Nat. Amer.	3.22%	2.54%	3.52%	2.77%	3.18%	3.03%
Total	0.47%	0.40%	0.48%	0.41%	0.47%	0.44%
Nursery	3.10%	2.86%	3.59%	3.22%	3.44%	3.20%
Vegetable	2.56%	2.31%	2.66%	2.34%	2.64%	2.53%
c) Using m-values						
Asian	5.96%	5.93%	6.11%	6.08%	6.10%	6.10%
Black	5.00%	4.81%	5.39%	5.12%	5.33%	5.24%
Ch. Tree	4.79%	4.72%	5.32%	5.21%	5.27%	5.20%
Female	1.84%	1.66%	1.98%	1.73%	1.95%	1.87%
Fruit	2.77%	2.69%	2.91%	2.80%	3.24%	3.22%
Hispanic	2.90%	2.59%	3.19%	2.79%	3.18%	3.10%
Nat. Amer.	3.91%	3.47%	4.95%	3.90%	4.12%	3.93%
Total	0.44%	0.40%	0.48%	0.42%	0.47%	0.46%
Nursery	3.79%	3.65%	5.21%	4.92%	4.68%	4.40%
Vegetable	3.39%	3.20%	3.80%	3.51%	3.90%	3.85%

using the p-values are substantially smaller than those using the d-values or m-values.

The final allocations given in columns A_2, B_2 and C_2 show that case A has the largest sample size, and cases B and C have approximately the same size, with some exception when the p-values are used. However, more samples are allocated to stratum groups 3 and 4 and fewer samples in stratum group 1 in Case B than Case C. Since it was preferred to have more samples allocated in groups 3 and 4 than group 1, case B is viewed as optimal.

Compared to the univariate NML allocations in case A which vary substantially for strata across various NML items, and their maximization which leads to higher number of samples, a single multivariate allocation for the ten NML items in case B smoothes these out and thus is a more stable allocation for the NML items considered together. Furthermore, a single multivariate for the 18 items in Case C becomes more unstable because of the optimization process involved in carrying out the sample allocation algorithm. Of the two, Case B is preferable over Case C since it allocates more samples to stratum groups 3 and 4, which are expected to have relatively more NML farms than stratum groups 1 and 2 to which Case C allocates more samples.

Next, the use of d-values in allocations for the NML items leads to a higher sample size then the survey cost would allow. The sample size obtained from the use of p-values turns out to be relatively low, especially for stratum groups 3 and 4, and hence are less likely to meet the NASS target CV requirements for the 2007 area frame survey.

It is noted that the achievable CV values as given in Table 3 cannot be used to compare the three sets of stratum standard deviations (a) - (c) for their use in the allocation procedure. When the achievable CV from a sample allocation is computed using the same stratum standard deviations as those used in determining the sample allocation, it will simply reproduce the input CV of the allocation.

4 2007 SAMPLE ALLOCATION AND SUPPLEMENTAL SAMPLE

The sample allocation obtained in Case B above using the m-values is recommended for the 2007 area frame sample design. This meets the NASS goal of survey implementation cost of about 3000 additional segments for its supplemental sample in support of achieving reliable estimates for the various NML items, as well as having a more stable stratum sample allocation. The following method is therefore implemented for developing the new 2007 sample allocation:

i. For each new NML item, carry out a model-fit of direct survey estimated stratum standard deviation versus that obtained using the NML model predicted values, and use the model-fit to obtain the NML stratum standard deviations as described in Section 2.

ii. Determine a single multivariate allocation for the 10 NML items using the stratum standard deviations obtained from the model-fits described above.

iii. Finalize the allocation as outlined in Case B by taking the stratum by stratum maximum of the regular 2007 NASS design allocation and the NML multivariate allocation in (ii).

The regular NASS 2007 design allocation and the proposed new 2007 design allocations are reported in Appendix C, by state in Table C1 and by stratum within state in Table C2. Also listed are the supplemental samples obtained as the difference between these two design allocations. The resulting sample allocations are summarized by stratum group in Table 4.

Table 4: Sample Design and Supplemental Samples by Stratum Group

Stratum Group	2007 Design	Proposed 2007 Design	Suppl. Sample
10's	5905	6192	287
20's	2775	3474	699
30's	347	728	381
40's	1767	3491	1724
50's	104	104	0
Totals:	10,898	13,989	3091

A total of 3091 supplemental samples are required, of which 287 are in the stratum group of 10's; 699 are in the 20's; 381 are in the 30's and 1724 are in the 40's. There is a 110 percent increase in stratum group 3, and a 98 percent increase in stratum group 4 when compared to the regular 2007 sample design.

5 CONCLUSIONS

Table 5 lists states that require substantial supplemental samples. It includes those with at least 80 supplemental samples or an increase of 100% or more over the regular 2007 sample design. The substantial increases in these states are due to one or more NML items predominant there. Listed next to each state is the primary NML item(s) for which reliable estimation would require such an increase in sample allocation. The reason for the increase as stated herein is determined by identifying the NML items for which the univariate allocations in a state were substantially higher compared to the univariate allocations for other NML items.

Table 5: Primary NML Items Causing Substantial Increases in Sample by State

State	Increase	(%)	NML Items
AR	110	(34)	Total, Asian, Black
CA	515	(128)	Asian
FL	119	(119)	Total, Black, Hispanic
GA	92	(32)	Total, Black
MA	25	(208)	Total, Nursery, Chr. Tree
ME	31	(97)	Total, Chr. Tree
MI	107	(74)	Total
MS	97	(33)	Black
NH	29	(290)	Total, Vegetable
NM	105	(85)	Total, Hispanic
NC	96	(30)	Total, Asian, Black
OH	88	(40)	Total, Chr. Tree
OR	232	(120)	Total, Asian, Chr. Tree
SC	109	(92)	Total, Black
TX	625	(56)	Total, Black
VT	31	(148)	Total, Hispanic

Total NML is a common cause for increase in sample allocation in all but two states. The predominance of Asian-operated farms in California and Black-operated farms in Mississippi account for the increase in those two states.

Listed in Table 6 are the specific strata that have substantial supplemental samples when compared to the regular 2007 design allocation. Total NML is the primary cause

Table 6: Strata with Substantial Supplemental Samples and their Causes

State	Stratum	Supplemental Samples	Increase due to
Arkansas	21	67	Asian
	42	43	Total, Indian
California	17	56	Asian
	21	99	Asian
	27	115	Asian
	31	162	Asian
	41	81	Asian
Florida	22	20	Total, Black
	40	28	Total
	42	32	Total
Georgia	40	92	Total, Black
Kentucky	40	24	Total
Louisiania	40	27	Total, Black, Vegetable
Maine	40	29	Christmas Tree
Massachusetts	40	25	Total, Christmas Tree, Nursery
Michigan	20	43	Total, Christmas Tree, Black, Indian
	40	46	Total, Hispanic
Mississippi	20	46	Black
	40	51	Total, Black
Missouri	40	75	Total
New Hampshire	40	29	Total, Vegetable
New Mexico	13	82	Total, Hispanic
New York	40	41	Total, Christmas Tree
North Carolina	40	88	Asian
Ohio	40	57	Total, Christmas Tree
Oklahoma	40	36	Total, Indian
Oregon	31	122	Asian
	41	65	Total, Christmas Tree
South Carolina	40	83	Total, Black
Texas	26	34	Total
	27	47	Total, Christmas Tree
	42	482	Total, Black, Christmas Tree
Virginia	40	44	Total, Black, Christmas Tree
Vermont	40	31	Total, Hispanic
Washington	31	15	Total, Vegetable
Wisconsin	12	42	Christmas Tree
Wyoming	42	19	Total
	44	11	Total

of many increases, especially in the 40's stratum group. Asian-operated farms are the cause of large increases in California for several strata.

It should be noted that the actual NASS sample allocation to be used for the 2007 ACES samples takes into account targeted sampling when a large increase in sample size is needed to estimate Asian-operated or Black-operated NML farms. The allocation also looks more closely at state-level target CV for NML farms. Thus the actual NASS allocation varies slightly from the one given here.

REFERENCES

Chang, Theodore, and Kott, Phillip S. (2004). ”Modeling NML Using the Area Frame Survey” Technical Manuscript. August 27, 2004, USDA/NASS, Washington, D.C.

Chhikara, Raj S., Spears, Floyd M. and Perry, Charles R. (2002). “Sample Allocation for Estimation of the Number of ”Not on Mail List” (NML) Farms for the 2002 Census of Agriculture,” USDA-NASS RDD Research Report No. RDD-02-01, June, 2002.

Montgomery, Douglas C. and Peck, Elizabeth A. (1992). *Introduction to Linear Regression Analysis*, *second Edition*, John Wiley, New York.

NML/ACES Sample Design Team. ”Decision Memorandum for the 2007 Agriculture Coverage Evaluation Survey (ACES) Sample Size”, Technical Memo, February 11, 2005, NASS/USDA, Washington, D.C.

APPENDIX A:

1 Chang and Kott (2004): NML Farm Modeling

A farm not on the census mailing list (NML) is likely to be small in size, and has certain characteristics as to its operation, farming, livestock or the amount of annual sales. Thus, a measure of the NML farm likelihood may be modeled by taking farm characteristics into account. If p denotes the probability of a farm being NML, then p may be expressed mathematically as a function

$$p = f(x_1, x_2, ..., x_p) \tag{1}$$

where $x_1, x_2, ..., x_p$ denotes the covariates that quantify the farm characteristics and affect the NML farm likelihood. A logistic regression model is used to specify the functional form:

$$\ln \frac{p}{1-p} = \beta_0 + \beta_1 x_1 + ... + \beta_p x_p \tag{2}$$

A model-fit is carried out to estimate the unknown coefficients $\beta_0, \beta_1, ..., \beta_p$ for the model in Equation (2).

Change and Kott (2004) conducted an extensive empirical study of modeling NML farm using the data from the June 2002 Area Frame Survey, including ACES area segments. For the covariates in the model, included were the farm sales and farm type variables, operator characteristics, farm size and stratum, among others. A stepwise regression procedure was utilized to select the variables for achieving optimal model-fit. The farm sale was found to be the most important predictor of NML farm. Certain farm operation characteristics and stratum were also significant predictors. Three population groups were determined appropriate to carry out the model-fits: (1) California, (2) Illinois, Indiana and Iowa combined together, and (3) all 48 contiguous states. Alaska and Hawaii do not participate in the Area Frame Survey.

Making use of a model-fit, an estimate of p_i, the probability of farm i being NML, was obtained from

$$\ln \frac{\hat{p}_i}{1-\hat{p}_i} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + ... + \hat{\beta}_k x_k \tag{3}$$

where $x_1, x_2, ..., x_k$ are the variables selected for model-fit.

NASS estimation procedure was applied to the model-fit determined probabilities $\hat{p}_i$ to estimate the total number of NML farms and each of its subdomains. If the tracts in a population area frame survey (AFS) are denoted by U_A, then the AFS estimate of the total number of farms in the population is given by

$$\sum_{k \in U_A} w_k f_k \hat{p}_k \tag{4}$$

where w_k is the sampling weight (expansion factor) and f_k is the adjusted tract-to-farm acreage ratio. For a subdomain, Equation (4) was applied using the $\hat{p}_k$ obtained for the subdomain item. A calibration process was involved in obtaining domain specific probabilities since a binary variable was used in the model for each subdomain of NML farms.

For full details, refer to Chang and Kott (2004).

2 Modeling Stratum Standard Deviation

NASS direct survey computed stratum variance is an unbiased estimator of the population variance for the stratum. However, we consider modeling the stratum standard deviation due to a reduced amount of variability seen in model error.

Let s denote the population standard deviation for a stratum. Two estimates of s are considered:

(a) $y =$ direct survey computed standard deviation

(b) $x =$ standard deviation obtained using predicated values from an NML model fit.

Both x and y are subject to sampling design variability. However, x is expected to be less variable than y because x is obtained using the individual values that got smoothed by an NML farm model fit described in Section 1 of Appendix B.

Considering the fact that sample standard deviation is a biased estimator of population standard deviation, we consider its modeling with errors in both variables as follows:

$$y_i = \alpha + s_i + \epsilon_i \tag{5}$$

and

$$x_i = \gamma + \phi s_i + \delta_i \tag{6}$$

where s_i is the stratum standard deviation, $E[\epsilon_i] = 0$ and $E[\delta_i] = 0$ and $\text{Cov}(\epsilon_i, \delta_i) = 0, \quad i = 1, 2, ...n$. Note that α and γ account for the bias in y and x, respectively. Further, $0 < \phi < 1$ since the NML model-smoothed value is expected to compress the actual stratum standard deviation $s_i, \ i = 1, 2, ..., n$.

There are two possible ways of modeling here; either express y in terms of x or express x in terms of y. However, there may be no difference between the two such modelings when the paired data (x_i, y_i) are very close to a straight line, which in fact is the case as discussed in Section 2. Thus, we consider y as the response and x as the regressor and write the model that follows from Equation (5) and (6) as

$$\begin{aligned} y_i &= (\alpha - \frac{\gamma}{\phi}) + \frac{1}{\phi}x_i + (\epsilon_i - \frac{\delta_i}{\phi}) \\ &= \beta_0 + \beta_1 x_i + e_i. \end{aligned} \tag{7}$$

where $\beta_1 = \phi^{-1} > 1$ and $E[e_i] = 0,\ i = 1, 2, ..., n$.

A no-intercept model is assumed. Its justification follows by recognizing that $x_i = 0$ implies $y_i = 0$. It could also be argued that $\beta_0 = (\alpha - \beta_1\gamma) \approx 0$ since both the x_i and y_i are based on the same NASS estimation method, except the x_i are computed using the NML model-smoothed values and the y_i are computed using the actual survey observed values.

The linear model in Equation (7) has error term e_i that may be correlated with the regression x_i. If so

$$\begin{aligned} \text{cov}(x_i, e_i) &= E[(x_i - E(x_i))(e_i - \beta_1\delta_i)] \\ &= E[\delta_i(e_i - \beta_1\delta_i)] \\ &= -\beta_1\sigma_\delta^2 \end{aligned} \tag{8}$$

where $E(\delta_i^2) = \sigma_\delta^2$. Since $\beta_1 > 1$, there is negative correlation between the regressor value x_i and the error term e_i. However, if σ_δ^2 is negligible, then the two are almost uncorrelated.

Table 1 on page 4 lists the model-fits obtained for various NML items. These model-fits are made by applying the standard least square methods to the data. As expected, the estimate slope $\hat{\beta}_1 > 1$ in all cases. Since the error in x_i is ignored in obtaining these model-fits, $\hat{\beta}_1$ may not be unbiased. It follows that $E[\hat{\beta}_i] = \beta_i/(1 + \theta)$, where $\theta = \sigma_\delta^2/\sigma_x^2$ and $\sigma_{\hat{x}}^2$ is the variance between the x_i in the data. However, if σ_δ^2 is negligible compared to σ_x^2, $\hat{\beta}_1$ is approximately an unbiased estimator of β_1. The x values do vary considerably across strata, as can be seen in Figure 1 in the case of total NML farms. On the other hand, it seems the empirical NML model-fit developed by Change and Kott was fairly robust to have lead to a biased yet efficient estimate of stratum standard deviation s, in which case σ_δ^2 may be negligible.

APPENDIX B: Stratum Groups

NASS land use strata are grouped together by considering similarity in their major land use, depending upon the agricultural intensity or the location. Table B1 lists the characteristics of the five groups of land use strata, the area level at which the sample allocations are reported in Table 2 and discussed in the text.

Table B1: Stratum Groups

Strata Group	Land-use Strata	Major Descriptor
1	10 - 19	Intense Cultivation
2	20 - 29	Moderate Cultivation
3	30 - 39	Urban Area
4	40 - 49	Low Cultivation
5	50 - 59	Non-Agricultural

APPENDIX C: Sample Allocation Tables

List of Tables

Table C1: Recommended Supplemental Samples by State

State FIPS	State Abbrev.	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
1	AL	236	272	36
4	AZ	118	158	40
5	AR	328	438	110
6	CA	404	919	515
8	CO	267	291	24
9	CT	8	11	3
10	DE	23	29	6
12	FL	100	219	119
13	GA	290	382	92
16	ID	148	208	60
17	IL	401	401	0
18	IN	264	264	0
19	IA	452	452	0
20	KS	487	487	0
21	KY	189	227	38
22	LA	249	276	27
23	ME	32	63	31
24	MD	61	61	0
25	MA	12	37	25
26	MI	145	252	107
27	MN	393	419	26
28	MS	298	395	97
29	MO	383	458	75
30	MT	316	316	0
31	NE	473	473	0
32	NV	26	30	4
33	NH	10	39	29
34	NJ	48	59	11
35	NM	124	229	105
36	NY	96	143	47
37	NC	319	415	96
38	ND	420	420	0
39	OH	220	308	88
40	OK	335	371	36
41	OR	194	426	232
42	PA	179	225	46
44	RI	8	8	0
45	SC	119	228	109
46	SD	395	395	0
47	TN	334	334	0

Continued on next page

Table C1 (continued)

State FIPS	State Abbrev.	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
48	TX	1120	1745	625
49	UT	69	102	33
50	VT	21	52	31
51	VA	179	223	44
53	WA	267	307	40
54	WV	66	77	11
55	WI	219	262	43
56	WY	53	83	30
Total	US	10,898	13,989	3091

Table C2: Recommended Supplemental Samples by Strata within State

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
1	AL	13	78	78	0
1	AL	20	90	90	0
1	AL	31	4	4	0
1	AL	32	2	2	0
1	AL	40	60	96	36
1	AL	50	2	2	0
4	AZ	13	52	83	31
4	AZ	14	8	8	0
4	AZ	20	12	12	0
4	AZ	21	2	2	0
4	AZ	31	4	12	8
4	AZ	32	2	2	0
4	AZ	41	23	23	0
4	AZ	44	2	2	0
4	AZ	45	2	2	0
4	AZ	46	2	2	0
4	AZ	47	2	2	0
4	AZ	48	3	3	0
4	AZ	49	2	3	1
4	AZ	50	2	2	0
5	AR	11	232	232	0
5	AR	21	32	99	67
5	AR	31	4	4	0
5	AR	32	2	2	0
5	AR	42	56	99	43
5	AR	50	2	2	0
6	CA	11	209	209	0
6	CA	17	20	76	56
6	CA	19	6	6	0
6	CA	21	63	162	99
6	CA	27	12	127	115
6	CA	31	16	178	162
6	CA	32	2	4	2
6	CA	41	54	135	81
6	CA	45	20	20	0
6	CA	50	2	2	0
8	CO	13	156	156	0
8	CO	15	6	6	0
8	CO	20	35	35	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
8	CO	24	3	3	0
8	CO	25	6	6	0
8	CO	31	3	10	7
8	CO	32	3	3	0
8	CO	34	12	12	0
8	CO	35	10	10	0
8	CO	40	3	3	0
8	CO	41	2	2	0
8	CO	42	3	9	6
8	CO	43	3	3	0
8	CO	44	3	11	8
8	CO	45	3	3	0
8	CO	47	10	13	3
8	CO	48	3	3	0
8	CO	50	3	3	0
9	CT	14	2	2	0
9	CT	31	2	2	0
9	CT	40	2	5	3
9	CT	50	2	2	0
10	DE	13	10	10	0
10	DE	20	5	11	6
10	DE	31	2	2	0
10	DE	32	2	2	0
10	DE	40	2	2	0
10	DE	50	2	2	0
12	FL	13	25	46	21
12	FL	17	6	10	4
12	FL	18	2	2	0
12	FL	21	15	18	3
12	FL	22	6	26	20
12	FL	27	6	17	11
12	FL	31	6	6	0
12	FL	32	2	2	0
12	FL	40	18	46	28
12	FL	42	12	44	32
12	FL	50	2	2	0
13	GA	13	91	91	0
13	GA	20	121	121	0
13	GA	31	2	2	0
13	GA	32	2	2	0
13	GA	40	72	164	92

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
13	GA	50	2	2	0
16	ID	10	30	30	0
16	ID	13	9	15	6
16	ID	15	48	68	20
16	ID	20	6	6	0
16	ID	22	16	16	0
16	ID	25	15	15	0
16	ID	31	3	21	18
16	ID	32	3	3	0
16	ID	40	3	18	15
16	ID	41	3	4	1
16	ID	42	3	3	0
16	ID	43	6	6	0
16	ID	50	3	3	0
17	IL	11	250	250	0
17	IL	12	70	70	0
17	IL	20	50	50	0
17	IL	31	10	10	0
17	IL	32	4	4	0
17	IL	40	15	15	0
17	IL	50	2	2	0
18	IN	11	125	125	0
18	IN	12	50	50	0
18	IN	20	55	55	0
18	IN	31	15	15	0
18	IN	32	2	2	0
18	IN	40	15	15	0
18	IN	50	2	2	0
19	IA	13	378	378	0
19	IA	20	66	66	0
19	IA	31	2	2	0
19	IA	32	2	2	0
19	IA	40	2	2	0
19	IA	50	2	2	0
20	KS	11	234	234	0
20	KS	12	112	112	0
20	KS	20	120	120	0
20	KS	31	3	3	0
20	KS	32	3	3	0
20	KS	40	12	12	0
20	KS	50	3	3	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
21	KY	13	80	80	0
21	KY	20	70	70	0
21	KY	31	5	19	14
21	KY	32	2	2	0
21	KY	40	30	54	24
21	KY	50	2	2	0
22	LA	13	187	187	0
22	LA	20	24	24	0
22	LA	31	4	4	0
22	LA	32	2	2	0
22	LA	40	30	57	27
22	LA	50	2	2	0
23	ME	14	12	14	2
23	ME	31	2	2	0
23	ME	40	16	45	29
23	ME	50	2	2	0
24	MD	13	15	15	0
24	MD	20	30	30	0
24	MD	31	2	2	0
24	MD	32	2	2	0
24	MD	40	10	10	0
24	MD	50	2	2	0
25	MA	14	2	2	0
25	MA	31	2	2	0
25	MA	32	2	2	0
25	MA	40	4	29	25
25	MA	50	2	2	0
26	MI	11	63	63	0
26	MI	12	24	42	18
26	MI	20	42	85	43
26	MI	31	2	2	0
26	MI	32	2	2	0
26	MI	40	10	56	46
26	MI	50	2	2	0
27	MN	11	210	210	0
27	MN	12	120	120	0
27	MN	20	35	35	0
27	MN	31	8	10	2
27	MN	32	2	2	0
27	MN	33	4	12	8
27	MN	40	12	28	16

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
27	MN	50	2	2	0
28	MS	11	95	95	0
28	MS	12	35	35	0
28	MS	20	80	126	46
28	MS	31	2	2	0
28	MS	32	2	2	0
28	MS	40	80	131	51
28	MS	42	2	2	0
28	MS	50	2	2	0
29	MO	11	195	195	0
29	MO	12	70	70	0
29	MO	20	70	70	0
29	MO	31	4	4	0
29	MO	32	2	2	0
29	MO	40	40	115	75
29	MO	50	2	2	0
30	MT	13	160	160	0
30	MT	20	75	75	0
30	MT	31	3	3	0
30	MT	32	3	3	0
30	MT	42	6	6	0
30	MT	43	48	48	0
30	MT	44	12	12	0
30	MT	45	3	3	0
30	MT	46	3	3	0
30	MT	50	3	3	0
31	NE	11	285	285	0
31	NE	12	77	77	0
31	NE	20	63	63	0
31	NE	31	4	4	0
31	NE	32	2	2	0
31	NE	40	40	40	0
31	NE	50	2	2	0
32	NV	13	4	4	0
32	NV	20	8	8	0
32	NV	31	2	6	4
32	NV	32	2	2	0
32	NV	41	2	2	0
32	NV	42	2	2	0
32	NV	43	2	2	0
32	NV	44	2	2	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
32	NV	50	2	2	0
33	NH	14	2	2	0
33	NH	31	2	2	0
33	NH	40	4	33	29
33	NH	50	2	2	0
34	NJ	13	5	5	0
34	NJ	20	30	30	0
34	NJ	31	5	5	0
34	NJ	32	2	2	0
34	NJ	40	2	13	11
34	NJ	42	2	2	0
34	NJ	50	2	2	0
35	NM	12	6	6	0
35	NM	13	32	114	82
35	NM	20	12	12	0
35	NM	31	2	13	11
35	NM	32	2	2	0
35	NM	40	5	8	3
35	NM	41	27	27	0
35	NM	42	10	19	9
35	NM	43	20	20	0
35	NM	44	2	2	0
35	NM	45	2	2	0
35	NM	46	2	2	0
35	NM	50	2	2	0
36	NY	13	25	25	0
36	NY	20	50	56	6
36	NY	31	3	3	0
36	NY	32	2	2	0
36	NY	40	12	53	41
36	NY	45	2	2	0
36	NY	50	2	2	0
37	NC	13	30	30	0
37	NC	20	200	200	0
37	NC	31	5	13	8
37	NC	32	2	2	0
37	NC	40	80	168	88
37	NC	50	2	2	0
38	ND	11	231	231	0
38	ND	12	90	90	0
38	ND	20	85	85	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
38	ND	31	2	2	0
38	ND	32	2	2	0
38	ND	33	2	2	0
38	ND	40	6	6	0
38	ND	50	2	2	0
39	OH	11	110	110	0
39	OH	12	35	35	0
39	OH	20	30	61	31
39	OH	31	15	15	0
39	OH	32	5	5	0
39	OH	40	20	77	57
39	OH	50	5	5	0
40	OK	11	120	120	0
40	OK	12	36	36	0
40	OK	20	80	80	0
40	OK	31	3	3	0
40	OK	32	3	3	0
40	OK	40	90	126	36
40	OK	50	3	3	0
41	OR	10	30	30	0
41	OR	13	40	40	0
41	OR	20	50	95	45
41	OR	31	2	124	122
41	OR	32	2	2	0
41	OR	41	8	73	65
41	OR	43	60	60	0
41	OR	50	2	2	0
42	PA	13	24	27	3
42	PA	20	98	121	23
42	PA	31	5	5	0
42	PA	32	2	2	0
42	PA	40	48	68	20
42	PA	50	2	2	0
44	RI	14	2	2	0
44	RI	31	2	2	0
44	RI	40	2	2	0
44	RI	50	2	2	0
45	SC	13	18	18	0
45	SC	20	60	86	26
45	SC	31	2	2	0
45	SC	32	2	2	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
45	SC	40	35	118	83
45	SC	50	2	2	0
46	SD	11	99	99	0
46	SD	12	130	130	0
46	SD	20	54	54	0
46	SD	31	2	2	0
46	SD	32	2	2	0
46	SD	33	2	2	0
46	SD	40	100	100	0
46	SD	44	4	4	0
46	SD	50	2	2	0
47	TN	13	100	100	0
47	TN	20	140	140	0
47	TN	31	10	10	0
47	TN	32	2	2	0
47	TN	40	80	80	0
47	TN	50	2	2	0
48	TX	10	88	88	0
48	TX	13	40	40	0
48	TX	14	273	273	0
48	TX	15	90	90	0
48	TX	16	16	16	0
48	TX	18	30	30	0
48	TX	20	64	64	0
48	TX	21	140	163	23
48	TX	24	70	70	0
48	TX	25	70	72	2
48	TX	26	4	38	34
48	TX	27	10	57	47
48	TX	28	25	30	5
48	TX	31	10	10	0
48	TX	32	5	5	0
48	TX	40	60	82	22
48	TX	41	49	52	3
48	TX	42	72	554	482
48	TX	43	2	9	7
48	TX	50	2	2	0
49	UT	13	28	30	2
49	UT	20	20	50	30
49	UT	31	4	4	0
49	UT	32	2	2	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
49	UT	41	5	5	0
49	UT	43	2	2	0
49	UT	44	2	2	0
49	UT	45	2	3	1
49	UT	46	2	2	0
49	UT	50	2	2	0
50	VT	14	12	12	0
50	VT	31	2	2	0
50	VT	40	5	36	31
50	VT	50	2	2	0
51	VA	13	15	15	0
51	VA	20	100	100	0
51	VA	31	10	10	0
51	VA	32	2	2	0
51	VA	40	50	94	44
51	VA	50	2	2	0
53	WA	10	119	119	0
53	WA	13	60	60	0
53	WA	20	48	64	16
53	WA	31	5	20	15
53	WA	32	2	2	0
53	WA	41	27	36	9
53	WA	44	2	2	0
53	WA	45	2	2	0
53	WA	50	2	2	0
54	WV	13	10	10	0
54	WV	20	14	14	0
54	WV	31	4	4	0
54	WV	32	2	2	0
54	WV	40	34	45	11
54	WV	50	2	2	0
55	WI	11	70	70	0
55	WI	12	30	72	42
55	WI	20	80	81	1
55	WI	31	5	5	0
55	WI	32	2	2	0
55	WI	40	30	30	0
55	WI	50	2	2	0
56	WY	11	8	8	0
56	WY	12	8	8	0
56	WY	20	8	8	0

Continued on next page

Table C2 (continued)

State FIPS	State Abbrev.	Stratum	Regular 2007 Design	Proposed 2007 Design	Supplemental Samples
56	WY	31	2	2	0
56	WY	32	2	2	0
56	WY	40	2	2	0
56	WY	42	5	24	19
56	WY	43	10	10	0
56	WY	44	2	13	11
56	WY	45	2	2	0
56	WY	46	2	2	0
56	WY	50	2	2	0
Total	US	All	10,898	13,989	3091